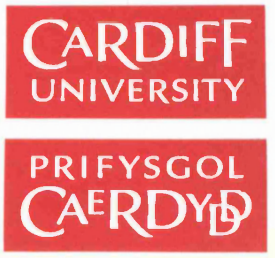


Molecular Analysis of the Diversity and Pollutant Tolerance of the *Burkholderia* Genus

Thesis presented for the Degree of Philosophiae Doctor

By

George William Payne, BSc (Hons) Biotechnology

In Candidature for the Degree of Philosophiae Doctor

Cardiff School of Biosciences

Cardiff University

May 2006

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ACKNOWLEDGEMENTS

I would like to thank the following:

- Dr Eshwar Mahenthiralingam, Professor Andrew Weightman and Dr Hefin Jones at the School of Biosciences, Cardiff University for their outstanding help, support and encouragement.
- The UK Natural Environment Research Council (NERC) for funding this PhD studentship (grant NER/S/A/2002/10347A) and additional funding from the Environmental Genomics Thematic Program (grant NER/T/S/2001/00299).
- My final year undergraduate students, Sara Morgan, Helen Rose and Daniel Brown for their assistance with the work in Chapters two, three and four respectively.
- Prof J. M. Tiedje and Dr. A Ramette for welcoming me on my placement in East Lansing, Michigan and for their help and support, especially furthering my understanding of phylogenetics.
- Dr Louise O' Sullivan for provision of the STM mutant bank and her technical guidance.
- Dr. K. E. Ashelford for assistance with *in silico* primer analysis.
- Collaborators from the International *Burkholderia cepacia* Working Group (IBCWG) for their provision of bacterial strains and methods.

SCIENTIFIC PUBLICATIONS, CONFERENCES AND PRESENTATIONS

Publications:

Payne, G. W., P. Vandamme, S. H. Morgan, J. J. Lipuma, T. Coenye, A. J. Weightman, T. H. Jones, and E. Mahenthiralingam. 2005. Development of a *recA* gene based identification approach for the entire *Burkholderia* genus. Applied & Environ Microbiol **71**:3917-27

Payne, G. W., A. Ramette, H. L. Rose, A. J. Weightman, T. H. Jones, J. M. Tiedje, and E. Mahenthiralingam. 2006. Application of a *recA* gene-based identification approach to the maize rhizosphere reveals novel diversity in *Burkholderia* species. FEMS Microbiol Lett **259**:126-32

Posters:

Payne G. W., A. J. Weightman, T. H. Jones, S. H. Morgan, P. Vandamme and E. Mahenthiralingam. Expansion of the *recA* gene based identification approach to include the entire *Burkholderia* genus. 4th NERC Environmental Genomics programme science workshop. April 18 & 19, 2005. Manchester, UK

Payne G. W., A. J. Weightman, T. H. Jones, S. H. Morgan, P. Vandamme and E. Mahenthiralingam. Expansion of the *recA* gene based identification approach to include the entire *Burkholderia* genus. 10th International Symposium on Microbial Ecology. August 22-27, 2004. **Poster 319** Cancun, Mexico

O'Sullivan L., G. W. Payne, A. Marchbank, A. Fivian, H. L. Rose, T. H. Jones, A. J. Weightman and E. Mahenthiralingam.. Environmental *Burkholderia*: fitness and phylogeny. 3rd NERC Environmental Genomics programme science workshop. April 5 & 6, 2004. Manchester, UK

Payne G. W., A. J. Weightman, T. H. Jones and E. Mahenthiralingam. Genomic analysis of diversity and mechanisms associated with bioremediation by *Burkholderia* species. 2nd NERC Environmental Genomics programme science workshop. April 14 & 15, 2003. Bristol, UK

Presentations:

Expansion of the *recA* gene based identification approach to include the entire *Burkholderia* genus. 9th International *Burkholderia cepacia* Working Group meeting. April 2-4, 2004. Vancouver, Canada

***Burkholderia* Species Diversity in an Environmental Sample. Molecular Analysis of *recA*.** 3rd All Wales Microbiology Meeting. 14-16 March, 2005. Gregynog, Aberystwyth, Wales, UK

Expansion of the *recA* gene based identification approach to include the entire *Burkholderia* genus. 2nd All Wales Microbiology Meeting. 15-17 March, 2004. Gregynog, Aberystwyth, Wales, UK

SUMMARY

Burkholderia is an important bacterial genus containing species of ecological, biotechnological and pathogenic interest which are widely distributed in the natural environment. This study describes the development of molecular methods to identify *Burkholderia* and also examines the genetic basis for their tolerance to halogenated pollutants.

The *recA* gene is a useful target for speciation of *B. cepacia* complex bacteria. Novel PCR primers were evaluated by testing their specificity against a panel of *Burkholderia* species. Nucleotide sequencing and phylogenetic analysis of 869 bp of the *recA* gene differentiated between putative and known *Burkholderia* species, including the *B. cepacia* complex. A genus-specific *recA* PCR yielding 385 bp was also able to identify all *Burkholderia* examined. Phylogenetic analysis of 188 novel *recA* genes clarified the taxonomic position of several important *Burkholderia* strains, revealing the presence of four novel *B. cepacia* complex *recA* lineages. Although the *recA* phylogeny could not differentiate *B. cepacia* complex strains recovered from clinical infection versus the natural environment, it facilitated the identification of clonal strain types of from both niches.

A *recA*-based cultivation-independent approach was used to examine *Burkholderia* diversity associated with the maize rhizosphere and cord forming fungi. Robust sequence datasets were created from the 869 bp *recA* fragments and screened to identify *recA* phylotypes matching *B. cepacia* complex species previously cultivated from the maize samples. *Burkholderia* related to known species as well as novel phylogenies were identified from the maize and fungal samples; in both cases the phylogenetic resolution of *recA* was sufficient to distinguish all individual species.

Transposon mutagenesis of *B. vietnamiensis* G4 was used to elucidate the genetic basis of resistance to halogenated phenol derivatives. Mutations resulting in increased susceptibility to 2,4-dichlorophenol included genes encoding enolase, toluene tolerance, regulators and a lipoprotein. Gene complementation restored the resistant phenotype of two mutants, a GTP-binding protein and a methyltransferase gene.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREIVATIONS

2,4,5-T	2,4,5-trichloropehnoxyacetic acid
2,4-D	2,4-dichlorophenoxyacetic acid
ANOVA	One-way analysis of variance
ATCC	American type culture collection
BCCM™	Belgian Co-ordinated Collections of Micro-organisms
BCSA	*Burkholderia cepacia* selective agar
BLAST	Basic local alignment software tool
BSM	Basal salts medium
CAP	Contig assembly program within BioEdit software
CL.DB	BUR1.2 fungal *recA* gene cone library (Chapter 4)
DDT	Dichloro-diphenyl-trichloroethane
DGGE	PCR-denaturing gradient gel electrophoresis
Dichloroprop	2-(2,4-dichlorophenoxy)-propanoic acid
DMSO	Dimethylsulfoxide
dNTPs	adenosine, cytidine, guanosine, and thymidine 5' deoxynucleotide triphosphates
EDTA	Ethylenediaminetetraacetic acid
HCl	Hydrochloric acid
IC	Inhibitory concentration
KM	Kanamycin
LB	Luria-Bertani medium agar/broth
LMG	Laboratorium voor Microbiologie, Universiteit Gent
MEGA	Molecular evolutionary genetics analysis software
MSD	Minimum significant difference
NCBI	National centre of biotechnology and information
OD	Optical density
ORF	Open reading frame
PCB	polychlorinated biphenyl
PCR	Polymerase chain reaction
Pfam	Protein family
PFGE	Pulsed field gel electrophoresis
PMX	Polymyxin B
RFLP	Restriction fragment length polymorphism
RNase	Ribonuclease
rpm	Revolutions per minute
SDS	Sodium dodecyl sulfate
STM	Signature tagged mutagenesis
Taq	*Thermus aquaticus*
TBE	Tris-borate-EDTA
TB-T	Trypan blue tetracycline medium
TCE	Trichloroethylene
TE	Tris-EDTA
TP	Trimethoprim
Tris	2-Amino-2-(hydroxymethyl)-1,3-propanediol
TSA/B	Tryptone soya agar/broth
UPGMA	Unweighted pair-group method average
X-GAL	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
$Cfu.ml^{-1}$	Colony forming units per millilitre
CL.16S	63f and 1387r fungal 16S rRNA gene cone library (Chapter 4)
CL1.2	BUR1.2 maize *recA* gene cone library (Chapter 3)
CL3.4	BUR3.4 maize *recA* gene cone library (Chapter 3)

CHAPTER 1: INTRODUCTION

1.1 BACTERIAL SYSTEMATICS

1.1.1 Introduction to classification

Bacterial classification is a means of summarising and cataloguing information about bacteria. Taxonomy is a systematic method of naming organisms according to a hierarchy and classification relies on nomenclature so that each bacterium can be identified by a universal label. Although it is impossible to classify bacterial species in terms of reproductive isolation (218), organisms may be grouped on the basis of their relationships assessed by comparison of characters to form either natural or special purpose classification schemes. Natural classification schemes may be phenetic or phylogenetic; where phylogenetic classification systems are those that infer an evolutionary relationship between taxonomic groups of organisms (216).

Phenetic based schemes, such as numerical taxonomy, involve the conversion of any number of an organism's characters into a numerical value. The increasing availability of computers in the late 1950's greatly advanced the handling of the extensive data processing required. Numerical taxonomy seeks to classify for "general", as opposed to "special", purposes and the characters are assigned unweighted values so that taxonomic groups may be assembled containing organisms with the highest proportion of similar attributes. These characters are compared and the number of differences recorded is used to score the relatedness of the taxonomic clusters within the resulting phenogram. Phenetic classification schemes are not hierarchical; instead they are static and do not imply any evolutionary relationship between clusters (216). Numerical taxonomy has recently been used where historic systems for identification by these methods have been successful and are typically compared with or contrasted to existing phylogenetic schemes (120, 134, 138). When classification criteria are chosen with the purpose of grouping bacteria according to a specific characteristic, the characters cannot be described as unweighted and the classification is not natural. Bacterial strains with special biochemical or physiological properties may be grouped according to distinguishing features with the suffix "var" and named biovars in special classification schemes; notable pathogens are pathovars; distinctive antigenic characteristics distinguish serovars; and strains with a significant degree of genetic relatedness but no corresponding phenotypic variations, genomovars. Special purpose classification schemes, using characters associated with virulence for

example, toxins, polysaccharide capsular antigens, hemolysin and fimbrial colonization factors, are particularly useful for the taxonomy of bacterial pathogens and established systems exist for several, including *Haemophilus influenzae* (153) and *Escherichia* species (72).

Approaching the height of the popularity of numerical taxonomy during the 1960s, it was considered that bacterial phylogenetics, the inference of evolutionary relationships between groups of bacteria, was impossible, and forcing phylogenentic criteria upon taxonomic groupings was logically flawed (216). Advances in this area were further hindered by the dogma that bacteria (mistakenly typified by *Escherichia coli*) and "blue-green algae" (cyanobacteria) represented a monophyletic group called the prokaryotes that were the direct ancestors of the eukaryotes. Prokaryotes were defined negatively, relying on the lack of eukaryotic features, and it was noted that such a definition was not durable in biology and in need of evolutionary support (23). At the advent of molecular evolution studies (151), eminent microbiologists such as R. Stanier and C. van Niel had nearly given up hope on a system to determine bacterial phylogenies (204, 256). It took the influence of C. Woese in the 1970s to lead a renaissance in microbial phylogenetics; using laborious oligonucleotide catalogue-based sequencing of the 16S rRNA gene, 60 kinds of bacteria were sequenced and arranged according to their genetic similarity (79, 256). The results challenged the standard, monophyletic classifications based on morphological similarities of bacteria and, more importantly, directly challenged the notion that a prokaryote-eukaryote dichotomy was a true phylogenetic distinction (204). Woese and colleagues recognised that there was a deep phylogenetic split amongst the bacteria which forced the division of the prokaryotes into two major lines of decent, the "archaebacteria" and the "eubacteria". Both lines were considered to be parallel in their phylogenetic depth and phenotypic diversity and, coupled with comparison of 18S rRNA genes from eukaryotes, it was hypothesised that an extinct simple organism, the so-called "progenote", was a likely common ancestor for all three primary lines of decent (256, 257). To emphasise the difference between these evolutionary domains, they were renamed Archaea, Bacteria and Eucarya (258).

Within the evolutionary domains, the 12 individual bacterial phyla described in 1987 (256) have since been variously referred to at the taxonomic rank of phylum, class, order, kingdom and division. The number of phyla increased to 52 by 2002, as signified by the range of publicly available 16S rRNA gene sequences; 26 of these were candidate phyla, described only from environmental clone sequences in the

absence of cultivated representatives. This abundance of uncultivated diversity was first recognised only as a result of direct 16S rRNA gene sequencing from environmental sources; at first seawater and, increasingly easy to work with since 2001, soil samples (192). By 2004 there were 80 bacterial phyla, including 54 uncultivated divisions, and the extent of known bacterial diversity is still expanding rapidly due to cultivation-independent environmental surveys (80).

The *Proteobacteria* are among well known examples of the originally described phyla. This gram-negative bacterial phylum includes two of the most studied genera of microorganisms, *Escherichia* and *Pseudomonas* (192) and also the genus *Burkholderia* (242). In some cases of taxonomy a genus may be created to accommodate newly named or identified species (39, 243), or a place may be assigned for them in an existing genus (87, 237). Use of formally recognised ranks higher than genus (families, orders, classes and divisions) are only justifiable when there is evidence of genetic relatedness. The prevalence of phylogenetic schemes to describe bacterial relatedness does, however, very often make this the case. A type strain is designated for each novel species characterised, and subsequently published, and serves as a nomenclatural type not necessarily to typify the species. Type strains may be obtained from international reference culture collections such as the American Type Culture Collection (ATCC) and the National Collection of Type Cultures (NCTC).

1.1.2 Classification by phenotypic characters

Classification of strains according to their phenotypic character encompasses the use of all techniques that are not directed towards characterising and utilising the organism's DNA or RNA. Phenotypic classification was facilitated by the advent of numerical taxonomy; the organisation and comparison of the states of several taxonomic characters using computers, data matrices and trait-related dendrograms (216). Many characteristics must be examined to yield descriptive information about genetic relatedness; between 100 and 200 different tests which evenly represent a range of categories to avoid bias. This range includes morphological, biochemical and chemotaxonomic categories (Table 1.1) (137). One of these phenotypic methods, chemotaxonomy, has been developed to become generally considered one of the essential milestones in the development of modern bacterial classification. Chemotaxonomy refers to the application of analytical methods to collect information on the various chemical constituents of the cell to classify bacteria. Amongst the many cell components, including extracellular products, which may be analysed as

Table 1.1 Categories of characters used in phenotypic classification (adapted from (137)

Category	Examples	Characteristics of method
Culture	Colonial morphology (e.g. shape, texture, elevation), pigmentation, appearance of growth in broth	Cultural characters must be evaluated under stringently standardised conditions to be of value in identification.
Morphological	Cell shape, size and arrangement. Endospore and flagella presence, staining reactions (e.g. Gram, acid fastness) and motility	Reproducibility may be attained for mean size measurements using an image analysing microscope. Morphological characters must be evaluated objectively.
Physiological	Growth temperatures, atmospheric growth requirements	Minimum and maximum temperatures permitting sustained growth may be controlled by water-bath in liquid media. Growth at different NaCl concentrations or pH. Growth under conditions of increased / decreased levels of CO_2 or O_2 is most efficient on solid media.
Biochemical	Decomposition of simple carbohydrates Metabolism of nitrogenous compounds Decomposition of large molecules (e.g. lipases, amylase). Terminal respiratory enzymes (catalase and cytochrome oxidase) Miscellaneous: coagulase, phosphatase and haemolysis	A limited number of the vast array of available enzymes and pathways is likely to be applicable to any one bacterial group. Commercially available kits (e.g. API, BBL) exist for the convenient identification of medically important organisms but may be expensive for large studies.
Nutritional	Organic acids, sole carbon and energy sources, vitamin requirements	Test procedures must be standardised to encompass defined basal media containing no growth-supporting contaminants and rigorously controlled inocula to ensure reproducibility.
Inhibitory tests	Sensitivities to antibiotics or tolerance to dyes, heavy metals and toxic or inhibitory salts. Often in combination in selective media	Sensitivities to antibiotics, chemotherapeutics and antimicrobials are constant enough to be of value for classification. Disc diffusion and E-test strips simple under standardised conditions. Dyes include crystal violet, fuchsin and brilliant green.
Serological	Agglutination by antisera to reference strains	Strict standardisation and quantification for techniques based upon agglutination, precipitation, neutralisation, immunofluorescence, and radio- and enzyme immunoassay is essential.
Chemotaxonomic	Amino acids in interpeptide bridges of cell wall, types of lipids in membranes	Components of the cell are extracted and analysed by chromatography, spectrophotometry or electrophoretic techniques to collect chemical information from bacteria. Under standard conditions these data are highly stable.

part of a chemotaxonomic study, Vandamme *et al.* (1996) reviewed several characteristics that have been used successfully in typing systems to characterise strains at the intraspecific level. These include cell wall composition, cellular fatty acids, isoprenoid quinines, whole cell proteins and polyamines, as well as the chemical analysis of DNA and RNA (242).

1.1.3 Classification by genotypic characters

Changes in the sequences of the nucleotides in nucleic acids over time can serve as a molecular chronometer. Moreover, only the *amount* of RNA and chromosomal DNA are affected by growth conditions, whilst the *composition* remains unchanged. For this reason genotypic data based on nucleotide sequences allow a more detailed allocation of taxa on a phylogenetic tree and a more accurate determination of boundaries than phenotypic data. The sequencing of entire bacterial genomes is becoming increasingly rapid and cost effective so that comparative genomic analysis is now the theoretical goal of all systematic methods. A range of methods that focus on the nucleic acid analysis of bacteria, including sequencing of individual genes, currently provides access to a rich source of evolutionary information. Various techniques may be used for classification according to genotypic characters (Table 1.2).

DNA analysis allows a more unambiguous definition of a bacterial species to be made: a group of strains, including the type strain, sharing >70% DNA-DNA relatedness with <5°C ΔT_m (252). Within a well-defined species there is a variation in moles percent G+C content of <3% and within a well-defined genus, <10%. Variation ranges between 24% and 76% in the bacterial world (242). DNA-DNA hybridisation techniques do, however, present severe disadvantages because their results are difficult to reproduce across the three methods (listed in Table 1.2) and different laboratories. Regardless of hybridisation method used, the 70% DNA hybridisation rule has proved difficult to apply and it has been recommended that two distinct genospecies that cannot be differentiated from one another on the basis of phenotype should not be named until they can be differentiated by some phenotypic property (252). A valid example is that of the genomovars of the *Burkholderia cepacia* complex, introduced in 1997 (240) and so-called until practical tests were available to endorse speciation based on DNA hybridisation values. Conversely, the genus *Bordetella* presents an example where DNA hybridisation values conflicted with the differentiation of species by conventional phenotypic methods. Such cases

Table 1.2 Categories of characters used in genotypic classification (adapted from (137, 242)

Category	Examples	Characteristics of method
DNA base composition (%mol G+C)	Caesium chloride density centrifugation Determination of thermal stability with UV spectrophotometry High-performance liquid chromatography (HPLC)	Higher G+C content is proportional to density and thermal stability. Test strains are compared to a reference (*E. coli* standard DNA).
DNA relatedness (ΔT_m)	Hydroxyapatite method Optical renumeration rates method S1 nuclease method	dsDNA denatures to ssDNA at T_m and reassociates 15-30°C lower. $T_{m(e)}$ of hetrologous duplex of test ssDNA strands is compared to T_m of homologous molecule. Low ΔT_m indicates few mismatches. Techniques require large amounts of DNA but are useful in the absence of other tests.
PFGE Typing	Restriction with rare cutting enzymes and separation of large fragments with Pulsed Field Gel Electrophoresis has replaced the use of frequent cutters in whole genome digestion	Fragments from low frequency restriction may be immobilized and probed with rRNA, elongation factor Tu, S12 ribosomal protein or flagellar proteins.
PCR Assays	10 bp oligos – arbitrarily primed PCR / randomly amplified polymorphic DNA analysis (RAPD) 5 bp oligos – DNA amplified fingerprinting	PCR may be also used to amplify repetitive elements throughout the genome of gram-positives and gram-negatives, or tRNA gene fragments.
Fragment length polymorphism	Amplified fragment length polymorphism AFLP Restriction fragment length polymorphism RFLP	Generates mostly strain-specific patterns in contrast to other DNA based methods.

were anticipated, and phenotypic characteristics are allowed to override the phylogenetic concept of species in a few exceptional situations (242, 252).

Techniques for studying rRNA, such as hybridisation studies or rRNA cataloguing of RNase T_1-resistant 16S rRNA oligonucleotides, have been replaced by direct sequencing of 16S rRNA genes using PCR and appropriate primers, and the formation of publicly available databases of partial or complete 16S rRNA gene sequences. Organisms with less than 97% 16S rRNA gene sequence homology will not undergo more than 60% DNA-DNA reassociation, which allows a general guide to be set in place for a species definition boundary based on 16S rRNA sequence identity levels. This rRNA threshold DNA-DNA reassociation value is lower to remain within the recommendation for a phylogenetic definition of a species by parameters of DNA reassociation (218).

1.1.4 Polyphasic taxonomy, a consensus approach to bacterial systematics

A polyphasic approach may include both chemotaxonomic and molecular data; large numbers of which have rendered classification systems more stable and are embedded in polyphasic classification systems. It is, however, important to consider that all methods carry information at varying levels and require different amounts of time and work. For example, pulsed-field gel electrophoresis strain typing (PFGE) is suitable for a large number of strains whereas DNA-DNA hybridisation is restricted to a minimal but representative set of strains. The inclusion of a phylogenetic dimension into bacterial taxonomy, either by the sequencing of 16S rRNA genes or other universally occurring genes, allows an isolate to be placed in a phylogenetic framework prior to the determination, where necessary, of its immediate relationships by means of a detailed polyphasic approach (242).

Recent "gold standard" polyphasic taxonomy studies are numerous in the literature (4, 230, 246), and studies to delineate species from the genus *Burkholderia* and those within the closely related *B. cepacia* complex serve as good examples. They include: expansion of the previously one-strain species *B. pyrrocinia* (222), and the descriptions of *B. anthina* (238), *B. ambifaria* (46), *B. fungorum* and *B. caledonica* (43), and *B. dolosa* (45).

1.2 MOLECULAR TAXONOMY

1.2.1 16S rRNA gene sequence analysis in bacterial systematics

The 16S rRNA gene is the primary gene that has been used in molecular taxonomy for sequence comparisons, and as a result there are large repositories of publicly-available reference sequences, for example the Ribosomal Database Project (RDP). This database contains over 205,165 aligned and annotated bacterial 16S rRNA gene sequences, and is updated frequently (http://rdp.cme.msu.edu). The 16S rRNA gene is an excellent choice for understanding the diversity and phylogenetic relationships between microorganisms: rRNA is present in all bacteria; it is comparatively easy to clone and sequence even from uncultivated or uncharacterised species; regions of conservation facilitate alignment of 16S rRNA genes from different species whilst the evolutionary substitution rate within the molecule varies sufficiently between species to allow relationships to be inferred between both close and distant relatives; and finally, it is unlikely that 16S rRNA genes have undergone lateral transfer (68). There is, however, risk associated with sole reliance on prokaryotic subdivisions described by the 16S rRNA gene so the accuracy of microbial systematics is often increased with the addition of data from another molecule complementing the 16S rRNA gene analysis (113). Historically, potential candidates for the provision of additional sequence data have included the 70 kDa heat shock proteins (HSP70, *Escherichia coli* DnaK homologs) (90), HSP60 (*E. coli* GroEL homologs) (249), elongation factor EF-Tu (141), RNA polymerases (117), glutamate dehydrogenase (16) and glutamine synthetase (229). Other systems developed to complement 16S rRNA gene data have been based on single genes that encode a functional protein which has importance within a specific group of bacteria, for example *amoA*, encoding ammonia monooxygenase (186); *gyrB* encoding DNA gyrase (262) and *mcrA* involved in methanogenesis (181). Protein-encoding functional genes can represent suitable phylogenetic markers as they have domains that must remain conserved but often their higher rate of base substitution gives them utility as a faster diverging "molecular clock" than the 16S rRNA gene.

1.2.2 The *recA* gene as a molecular chronometer in systematics

The utility of *recA* gene comparisons for phylogenetic studies has been evaluated as a suitable alternative to 16S rRNA for molecular systematic studies, applicable to all bacteria, including closely related species. The RecA protein was first identified in *E. coli* (38) and contributes to at least three main cellular processes: homologous recombination, DNA repair and the SOS response (a series of reactions triggered by

DNA damage within bacterial cells). During homologous recombination, RecA plays a critical role in catalysing the pairing of an invading strand of single stranded DNA (ssDNA) with a complementary region of a nicked double helix (dsDNA). The incoming ssDNA displaces a strand of the target DNA which in turn hybridises with the remaining ssDNA. The recombination reaction then progresses with the hydrolysis of ATP into RecA-driven branch migration and the formation of new double stranded recombinant DNA (135, 224). The function of RecA in DNA repair and the SOS response is to act as a coprotease for the LexA protein which, in a healthy cell, represses the expression of SOS genes encoding DNA repair proteins. In the event of DNA damage, ssDNA and RecA form a filament which activates self-induced catalysis of the LexA protein. With the digestion of the LexA molecules, the repression is removed and the necessary SOS proteins are produced (135, 224).

RecA is a monomer of around 352 amino acids in length which consists of three domains. The central domain contains two DNA binding sites, one each for ssDNA and dsDNA, and the binding site for ATP (223). These DNA binding regions of the molecule are highly conserved amongst bacterial species which facilitates phylogenetic alignment. Phylogenetic relationships have been inferred from RecA within robust trees that correlate well with those formed by 16S rRNA analyses and suggest that the protein is a useful genetic marker for studying bacterial evolution, with less vulnerability to factors such as non-independence of substitution patterns at different sites, variation in substitution rates between lineages and ambiguities in alignments between distantly related taxa (68, 112, 113). The bacterial *recA* gene, moreover, has phylogenetically related homologues responsible for DNA repair, DNA recombination and catalysing hybridisation between two parental molecules in all *Bacteria, Archaea* and *Eucarya*. The *RadA* homologue has been used successfully to decipher the Archaeal phylogeny. As an independent measure of Archaeal diversity it yielded results which corresponded closely to phylogenies obtained using 16S rRNA gene sequences (201). Interestingly *radA* genes from Archaeal genera were found to be more closely related to Eucaryal *RAD51* and *DMC1* genes than to Bacterial *recA* genes, suggesting that the RadA is orthologous to the common ancestor of Rad51 and Dmc1 (202).

The use of *recA* has overtaken the use of many other protein-encoding functional genes in bacterial systematics (with the potential exception of *gyrB* (211, 262) and was used with great success as the basis for a diagnostic approach that was able to discriminate between the members of the closely related *B. cepacia* complex,

providing resolution at the species level when used as a target for species-specific PCR or digested in a simple RFLP (143). It has continued to play a major part in the rapid and straightforward identification of these strains when isolated from both clinical sources and the environment (31, 32, 54, 185, 244). To discriminate *below* the species level techniques such as multi-locus restriction typing (MLRT), pulsed field gel electrophoresis (PFGE), random amplified polymorphic DNA (RAPD), and BOX-PCR are available but, although highly discriminatory, they suffer from being difficult to transfer between laboratories or require specialist equipment or use. Multi-locus restriction typing (70) offers superior strain discrimination over single-locus RFLP by analyzing multiple genes, but pattern-matching techniques based on gel banding have inherent variability and ambiguities (47).

1.2.3 Multi locus sequence typing (MLST)

A relatively recent development within the field of molecular bacterial taxonomy has combined the use of *recA* with other housekeeping genes. Multi-locus sequence typing has limited the problem that lateral transfers of genes may cause the genomes of some species to exhibit mosaic evolutionary histories (68). Much work has been published on a range of identification schemes. Protein encoding genes such as *recA, rpoA,* and *phrH* were used to identify *Vibrio* species (227); *recA, atpD, rrs* and *nifH* were used for *Sesbania* species (251); as many as seven genes, *recA, gltA, gyrB, gdhB, cpn60, gpi* and *rpoD* may be used to identify *Acinetobacter baumannii* strains (15). Other recent studies have described the identification of strains of *Lactobacillus plantarum*, *Pseudomonas aeruginosa*, *Haemophilus influenzae, Burkholderia pseudomallei* and *B. mallei,* and *Yersinia* species (27, 52, 85, 118, 157), and notably strain *and* species resolution between members of the *B. cepacia* complex (10) in a scheme utilising *atpD*, *gltB*, *gyrB*, *trpB*, *lepA*, *phaC*, and *recA*. All of these schemes utilise *recA* amongst other gene targets. MLST is fast becoming the "gold standard" of bacterial typing methods and will as such become increasingly important in the future (147).

1.3 TAXONOMIC OVERVIEW OF THE GENUS *BURKHOLDERIA* AND ITS ECOLOGY

1.3.1 Taxonomic history and phylogenetic position of the Burkholderia genus

Changes in bacterial taxonomy have been reflected in the classification scheme used to define Bergy's Manual of Systematic Bacteriology which was, up until 1986, an artificial classification based on phenotypic properties. Subsequent editions of the manual were based on genotypic properties and, for the first time, inferred phylogeny. The shift toward a phylogenetic classification system dramatically affected the biological identity of the genus *Pseudomonas,* which was a relatively large and important group of gram negative, non-spore forming oxidase-positive rods. Such basic morphological characteristics were, however, common amongst many bacterial species and hence were of limited significance in their accurate identification.

Using rRNA-DNA hybridisation methods, Palleroni *et al.* produced striking results that subdivided the *Pseudomonas* genus, which had previously consisted of about forty species not counting biovars among fluorescent pseudomonads and serovars among phytopathogenic species (175). Five well-defined rRNA homology groups were described, the first of which has remained the genus *Pseudomonas.* 16S rDNA sequencing soon confirmed the rRNA homology groups (256) at a time when the suggestion that some of the rRNA groups were closer to species of other genera than to members of the other *Pseudomonas* groups would have been flatly rejected by most bacteriologists (174).

The second rRNA homology group was composed mainly of pathogens, including *P. cepacia,* the causative agent responsible for bacterial rot of onions (29) the primary pathogens *P. mallei* and *P. pseudomallei*, the metabolically versatile *P. multivorans* (219) and *P. kingii* an opportunistic human pathogen. Preserving *Pseudomonas "cepacia"* as the type species, seven former members of the *Pseudomonas* rRNA homology group II were transferred to a novel genus, *Burkholderia,* based on a polyphasic taxonomy approach including 16S rRNA gene sequencing (260). Shortly afterwards *P. pickettii* and *P. solanacearum,* initially classified in the same rRNA homology group as *Burkholderia,* were transferred to the genus *Ralstonia* (261).

Currently, the *Burkholderia* genus consists of 34 species (83) including 10 comprising the *B. cepacia* complex (Table 1.3). Until relatively recently the members of the *B. cepacia* complex were referred to as genomovars; they were phylogenetically different and awaited separation on the basis of phenotypic characteristics before they could be named individually as species. Members of the genus are so versatile and metabolically robust they are able to occupy a wide range of ecological niches.

1.3.2 Association with plants

Species of *Burkholderia* cause plant diseases, for example *B. caryophylli* is a carnation pathogen (12) and *B. glumae* is a serious threat to crops causing rot of rice grains and seedlings (231). Some species, however, may be able to form beneficial interactions with plants, either by nodulating their roots and fixing atmospheric nitrogen (71) or by antagonising the effects of other plant pathogens, as in the case of *B. ambifaria* AMMD^T protecting pea plants against fungal infection (177, 179).

1.3.3 Growth in polluted environments

Several *Burkholderia* species are found in habitats contaminated with constituents of crude oil, agricultural pollutants and fuel additives. Most of these strains are poorly taxonomically characterised, for example only recently strain LB400, a well studied degrader of biphenyls, was assigned to a novel species *B. xenovorans* (87). Others, such as strains NF100 and CBS3, have an undetermined taxonomic status but their degradative abilities are well understood (98, 116). Species that can degrade pollutants and survive in polluted environments will be discussed further (Section 1.4).

1.3.4 Infection of susceptible individuals

Human hosts provide *Burkholderia* species with another niche in which they are well-adapted to thrive. Nine members of the *B. cepacia* complex, *B. gladioli* and *B. fungorum* have been implicated in the disease of susceptible individuals, but as yet there has been no published account of infection caused by the most recently proposed species, *B. ubonensis*. Whilst most *Burkholderia* species pose no threat to the healthy human population they have been associated with a variety of infections in persons with chronic granulomatous disease, a primary immunodeficiency, or patients who have undergone lung transplant surgery (44, 61). They have attracted

Table 1.3: Summary of the *Burkholderia* genus and its ecological distribution

Burkholderia species	Environmental Source[a]	Pathogenic Source[b]	Reference
B. cepacia	Described in 1950 as onion rot agent , isolated from soil	Cystic fibrosis	(29, 240)
B. multivorans	Strain ATCC17616^T originally isolated from Trinidad soil. Environmental distribution poorly understood	Accounts for 9-37 % of cases of *B. cepacia* complex infection in cystic fibrosis	(219, 240)
B. cenocepacia	Widely associated with the maize rhizosphere and frequently isolated from soil	Accounts for 50-80 % of cases of *B. cepacia* complex infection in Cystic fibrosis and salient cause of "*cepacia* syndrome"	(76, 239, 240)
B. stabilis	Sparsely isolated from maize samples. Environmental distribution poorly understood	Cystic fibrosis	(240, 241)
B. vietnamiensis	Able to degrade Benzene, o-cresol, m-cresol, p-cresol, phenol, toluene, TCE, naphthalene and chloroform, and fix nitrogen. Frequently associated with rice crops.	Cystic fibrosis	(84, 240)
B. dolosa	Rarely isolated from the environment	Isolated from USA cystic fibrosis patients	(45, 247)
B. ambifaria	Biocontrol Strain AMMD^T protects pea plants against disease, isolated from soil	Cystic fibrosis	(46)
B. anthina	Associated with the plant rhizosphere	Cystic fibrosis	(238)
B. pyrrocinia	Isolated from soil	Cystic fibrosis	(238)
B. ubonensis	Isolated from soil in Thailand. Reclassified from *B. thailandensis*. Maybe within *B. cepacia* complex.	Infection of cystic fibrosis patients not known	(259)
B. xenovorans	Polychlorinated biphenyl-degrading strain	Limited cystic fibrosis	(87)
B. mallei	Reservoir in chronically infected horses	Zoonotic primary pathogen, causing equestrian and human glanders	(81)
B. pseudomallei	Isolated from moist soil, rice paddies and water	Primary pathogen, causes melioidosis	(28, 55)
B. thailandensis	Similar to *B. pseudomallei*	Avirulent	(26)
B. glumae	Salient bacterial rice pathogen in Japan, Korea & Taiwan for 20 years	Rice pathogen	(231)
B. plantarii	Cause rice seedling blight; forms tropolone	Rice pathogen	(8, 231)

a; Putative "natural reservoir" or location of original isolation

b; "Cystic fibrosis" species may also pose a threat to sufferers of chronic granulomatous disease or lung transplant recipients

The *B. cepacia* complex is indicated by shading

(con/t)

Table 1.3: con/t

Burkholderia species	Environmental Source[a]	Pathogenic Source[b]	Reference
B. caryophylli	Pathogenic to carnations and cause onion rot	Plant pathogen	(12)
B. graminis	Associated with soil and isolated from wheat, corn and grasses		(250)
B. glathei	Originally isolated from fossil lateritic soils in Germany		(266)
B. phenazinium	Isolated from soils – characterised by iodinin production		(250)
B. caribensis	Isolated from tropical vertisol on Martinique, capable of root nodulation		(1, 166)
B. caledonica	Associated with soil. Originally isolated from rhizosphere of plants in Scotland		(43)
B. hospita	Originally isolated from soil in Belgium		(88)
B. terricola	Originally isolated from soil in Belgium		(88)
B. sacchari	Originally isolated from sugar-cane plantation in Brazil		(25)
B. kururiensis	Nitrogen fixer originally isolated from an aquifer polluted with trichloroethylene		(264)
B. andropogonis	Causes stripe disease in sorghum and leaf spot in velvet bean.	Plant pathogen	(42)
B. phymatum	Nitrogen fixer originally isolated from *Machaerium lunatum* root nodules in French Guiana		(237)
B. tuberum	Nitrogen fixer originally isolated from *Aspalathus carnosa* root nodules in South Africa		(237)
B. gladioli	Isolated from decayed onions, rice and human specimens	Occasionally infects cystic fibrosis patients	(40)
B. fungorum	Isolated from plant root nodules, fungi, human and animal specimens	Occasionally infects cystic fibrosis patients	(43)
B. brasiliensis	Recovered from banana and pineapple plants in South America		(71)
B. tropica	Isolated from the rhizosphere and as endophytes of sugarcane, maize and teosinte plants in Brazil, Mexico and South Africa		(193)
B. pytofirmans	Rhizosphere and endophytic populations associated with various plants, including onion roots.		(210)
B. sordidicola	Associated with the white-rot fungus *Phanerochaete sordida*		(129)
B phenoliruptrix	Well-known for its ability to degrade a variety of recalcitrant xenobiotics, including 2,4,5-T, 2,3,4,6-tetrachlorophenol and pentachlorophenol. AC1100 is a well-characterised strain.		(41)

a; Putative "natural reservoir" or location of original isolation

b; "Cystic fibrosis" species may also pose a threat to sufferers of chronic granulomatous disease or lung transplant recipients

The *B. cepacia* complex is indicated by shading

most research in relation to their ability to cause opportunistic infection in sufferers of cystic fibrosis (CF), the most common lethal inherited disorder in Caucasians, where *Burkholderia* respiratory tract infection may be asymptomatic and chronic, or associated with a rapid deterioration in pulmonary function known as *cepacia* syndrome (106). Unlike other opportunistic pathogens such as *Pseudomonas aeruginosa* and *Staphylococcus aureus*, *Burkholderia* species are not carried commensally, providing supporting evidence that these infections are acquired from nosocomial or environmental sources. Contaminated disinfectants, intravenous solutions and medical devices have all been implicated as sources of *B. cepacia* complex outbreaks (104, 173). Moreover, treatment of *Burkholderia* infection is often hindered by the range of innate antibiotic resistances that these species typically exhibit (146). Two *Burkholderia* species (Table 1.3) are primary pathogens: *B. mallei* causes glanders and is indigenous to the Far and Middle East, Northern Africa, Eastern Mediterranean, and South-eastern Europe (81); *B. pseudomallei* is the causative agent of melioidosis. These two species have attracted interest recently owing to their potential for use as bioterrorism agents (44, 55).

1.4 BIOTECHNOLOGICAL POTENTIAL OF ORGANISMS WITHIN THE *BURKHOLDERIA* GENUS

Species within the genus *Burkholderia* owe their ecological versatility to the large coding capacity afforded by their genomic size and organisation; different isolates contain from two to four (usually three) chromosomes ranging in size from 5-9 Mb (122). Their genomes are rich in insertion sequences which increase the expression of neighbouring genes and promote genomic rearrangements, contributing significantly to genomic plasticity and playing an important role in the evolution of novel catabolic functions; for example *B. multivorans* ATCC 17616 (126). Furthermore they may readily acquire and express plasmids such as pJP4 and TOM (7, 64, 77, 212, 213). The metabolic versatility of *Burkholderia* species allows them to degrade a range of recalcitrant pollutants and this ability may be exploited for biotechnology. Table 1.4 gives details of *Burkholderia* species that are able to degrade a range of recalcitrant xenobiotic compounds, as well as those that have biological control activities. Several of these *Burkholderia* strains have been particularly well studied and are described below in detail:

1.4.1 Bioremediation potential

B. vietnamiensis G4 (ATCC 53617 or R1808) can co-metabolise the abundant organic pollutant trichloroethylene (TCE) using a degradative pathway encoded by the *tom* operon on the catabolic plasmid pTOM. The first enzyme, toluene *o*-monooxygenase, must be induced by aromatic molecules such as toluene or phenol (169, 213). TCE biodegradation was recently successful on a laboratory scale in a two-stage continuous stirred reactor and trickling biofilter system in which 100% TCE removal was sustained for three months (124). An adhesion-deficient mutant of *B. vietnamiensis* G4 (ENV435), evaluated in a field-scale trial, was able to successfully remediate an aquifer contaminated with volatile organochloride compounds by 78% in two days (220). *B. xenovorans* LB400 (LMG 21463) has only recently been designated as an individual species (87) and its ability to degrade biphenyl and polychlorinated biphenyls (PCBs, used in old electrical capacitors and some oils) has been well studied (74). The complete genome sequences of strains G4 and LB400 are now available (Table 1.4).

Table 1.4: Biotechnological potential within the *Burkholderia* genus (adapted from (172, 183, 232)

***Burkholderia* species**	**Strain information**	**Relevant information; degradative or biocontrol abilities**	References
Biodegradative species and strains			
B. vietnamiensis	G4 (ATCC 53617 or R1808)	Benzene, *o*-cresol, *m*-cresol, *p*-cresol, phenol, toluene, TCE, naphthalene, chloroform; US patents 4925802 and 5543317; http://genome.ornl.gov/microbial/bcep_1808/	(77, 124, 154, 169, 212)
B. kururiensis	KP23 (JCM 10599)	Trichloroethylene (TCE)	(264)
B. xenovorans	LB400 (LMG 21463)	Biphenyl, polychlorinated biphenyls PCBs; http://genome.ornl.gov/microbial/bfun/	(74, 87)
B. phenoliruptrix	AC1100 (LMG 22037)	2,4,5-Trichlorophenoxyacetate, 2,3,4,6-tetrachlorophenol, pentachlorophenol	(41, 115)
Burkholderia sp.	JS150	TCE, benzene, phenol, toluene, chlorobenzene, ethylbenzene, cresols, naphthalene	(108)
Burkholderia sp.	CRE-7 and RP007	Polycyclic aromatic hydrocarbons (PAH)	(122, 167)
Burkholderia sp.	CSV90, EML1549, K712, RASC, TFD2 and TFD6	2,4-Dichlorophenoxyacetate (2,4-D)	(156)
Burkholderia sp.	CBS3	4-Chlorobenzoate, 2-nitrobenzoate, 3-nitrobenzoate, 4-nitrobenzoate, 3-nitrochlorobenzene, 2-nitrophenol, 3-nitrophenol, 2,4,6-trinitrotoluene.	(136)
Burkholderia sp.	Strain 8	Benzoate, 4-fluorobenzoate, 4-hydroxybenzoate	(205)
Burkholderia sp.	KZ2	2-Chlorobenzoate, 4-chlorobenzoate, 2,4-dichlorobenzoate	(263)
Burkholderia sp.	NF100	Fenitrothion	(98)
Burkholderia sp.	DNT	Dinitrotoluenes (DNT)	(127)

(con/t)

Table 1.4: con/t

Burkholderia species	Strain information	Relevant information; degradative or biocontrol abilities	References
Biocontrol species and strains			
B. cepacia	ATCC 49709	Grass seed biological control strain	(107)
B. cenocepacia	M36	Biopesticide withdrawn from commercial use	(178)
B. cenocepacia	BC-1, BC-2	Biological control strains reduce damping-off in corn	(149)
B. vietnamiensis	TVV75	Root colonising and endophytic strain increases rice yield by 13-22%	(236)
B. ambifaria	M-54 (R-5142)	Antifungal biopesticide	(178)
B. ambifaria	J82 (R-5140)	Antinematodal biopesticide	(178)
B. ambifaria	BC-F	Antifungal biological control strain	(265)
B. ambifaria	AMMD^T (LMG 19182^T)	Biological control strain	(46, 178)
B. pyrrocinia	BC11	Antifungal biological control stain	(111)
B. pyrrocinia	ATCC 39277	Production of antifungal agents	(158)
B. tropica	Ppe8^T (LMG 22274^T)	Fixes atmospheric nitrogen, endophytic	(193)

1.4.2 Biocontrol potential

B. ambifaria AMMD is one of the most studied biocontrol strains and was first isolated from *Pisum sativum* roots. It produces antifungal antibiotics and siderophores which confer resistance to *Pythium* species that are responsible for damping-off in peas (178). Two other commercially important *Burkholderia* strains are *B. ambifaria* M54 and *B. cenocepacia* M36 which were registered with the American Environmental Protection Authority (EPA) for use in the USA biopesticides Deny®, Blue Circle® and Intercept®. This registration was approved when human clinical and biological control strains were thought to be taxonomically distinct. Also of significance is *B. pyrrocinia* ATCC 39277, from which a potent antifungal agent was purified in 1987 (158).

1.4.3 Registration of *Burkholderia* species for commercial exploitation

A problem that confounds advances which exploit the degradative and biological control properties is the potential transfer of "harmless" environmental strains to clinical niches. Concerns about the commercial use of *B. cepacia* complex species (101, 131) have prompted a moratorium on the registration of new strains for use in biotechnological applications by the EPA and there is evidence to support direct environmental acquisition of *B. cenocepacia* from soil (132). Whilst some biodegradative strains are well characterised (e.g. strains G4 and LB400) further confusion surrounds those for whom the taxonomic position is poorly understood. For example *Burkholderia* species strain 8, which degrades 4-fluorobenzoate (205) and KZ2, which degrades 4-chlorobenzoate (263 149), are important strains for scientific and regulatory reasons and there is presently no way to distinguish these, or other environmental strains, from those that might cause opportunistic infection. The EPA Toxic Substances Control Act (TCSA) was claimed to act as a deterrent to the development of safer microbiological bioremediation as it would entail a lengthy and costly licence application (131), nevertheless, advances have been made in the field of transferring commercially useful features from *Burkholderia* species to non-pathogenic hosts. An example of recent success in this field is the chromosomal integration of the 2,4-dinitrotoluene pathway genes from *Burkholderia* species strain DNT into the non-pathogenic psychrotolerant rhizobacterium *Pseudomonas fluorescens* (163). For the full potential of biotechnology using *Burkholderia* species to be exploited the dilemma of opportunistic pathogenicity crossover from the environment must be understood to facilitate effective regulation.

1.5 AIMS

The aims of the current project were two-fold:

Investigate *Burkholderia* populations in the natural environment (Chapters 2, 3 & 4). The objectives of this aspect of the investigation were to develop a *recA* gene based molecular identification approach for the entire *Burkholderia* genus, and evaluate the approach and its application to environmental samples. Understanding the diversity and distribution of *Burkholderia* species within the natural environment could facilitate discovery of novel species and habitats. In turn, this understanding may lead to the exploitation of biotechnological potential or information about the origin of clinical infections.

Identify genes involved in pollutant resistance (Chapter 5). Here, the objectives were to screen a strain G4 mutant bank using 2,4-dichlorophenol as a model pollutant to identify susceptible mutants and identify them on the G4 genome in order to determine the genetic basis for the mutations. Characterising genes linked to pollutant resistance has implications for the transfer of degradative pathways to organisms that are "safe" for commercial use, as resistance is a prerequisite for degradation.

CHAPTER 2: DEVELOPMENT OF A *RECA* GENE BASED IDENTIFICATION APPROACH FOR THE ENTIRE *BURKHOLDERIA* GENUS

Published in part as Payne et al. 2005, Appendix 1

2.1 INTRODUCTION

The *Burkholderia* is a genus with complex taxonomy that currently contains 34 validly described species (48), nine of which form a closely related group, known as the *B. cepacia* complex (49). *Burkholderia* species are widely distributed in the natural environment, and although the majority appear to live either freely or as symbionts or commensals with a variety of higher organisms, several species also cause disease (48). Plant pathogenic species include *B. glumae* and *B. plantari* which are important rice pathogens. The genus also includes mammalian primary pathogens such as *B. pseudomallei*, the cause of meliodosis in humans, and *B. mallei*, which causes glanders in horses; both species have attracted recent interest as potential bioterrorism agents (140). Many other *Burkholderia* species are capable of causing opportunistic infections in humans and animals; for example the *B. cepacia* complex (49) can cause serious infections in persons with cystic fibrosis (142) and other vulnerable individuals (217). Species can also cause disease in plants (69) and animals (19). In contrast to these detrimental pathogenic properties, several *Burkholderia* species have considerable commercial and ecological importance. They have been used in agriculture as biopesticides and plant growth promoters (178), and in the bioremediation of major pollutants such as TCE (213) and PCBs (162).

The taxonomy and identification of the *Burkholderia* genus are complex, with new species being described frequently (48, 49). Closely related species such as the *B. cepacia* complex are difficult to distinguish using conventional biochemical and phenotypic tests, and species belonging to other betaproteobacterial genera (including *Pandoraea* and *Ralstonia*) may be misidentified as *Burkholderia* species (49). A polyphasic taxonomic approach (242) utilising multiple diagnostic tests is often required to identify *Burkholderia* species accurately. Although 16S rRNA gene sequence analysis forms an integral part of taxonomical analysis for many bacterial genera (242), its utility in the genus *Burkholderia* is more limited, especially within the *B. cepacia* complex where it cannot be used as means to distinguish accurately all species (130, 143).

The *recA* gene has been widely applied in bacterial systematics (113) and has proven very useful for the identification of *B. cepacia* complex species, with phylogenetic analysis of sequence variation within the gene enabling discrimination of all nine current species within the *B. cepacia* complex (143). The PCR primers designed for the original *recA*-based approach, BCR1 and BCR2, are, however, specific only to members of the *B. cepacia* complex and fail to amplify this gene from other *Burkholderia* species (143). Whilst this can be used as a positive means to confirm an isolate's assignment within the complex, it limits the application of the approach to identify other *Burkholderia* species in diverse natural habitats.

Given the ecological, biotechnological and pathogenic importance of these bacteria, there is a clear need for a molecular diagnostic scheme capable of discrimination across all *Burkholderia* species. This chapter describes the use of genome sequence data from several *Burkholderia* species, in combination with an extensive collection of *recA* sequences from *B. cepacia* complex bacteria (143), to develop and evaluate a scheme for molecular systematic analysis and identification of all *Burkholderia* based on the *recA* gene. The specific aims of the present study were to:

1) Design new PCR primers to amplify and obtain novel *recA* gene sequence data from known and unknown species.

2) Determine whether restriction fragment length polymorphism (RFLP) typing of the *Burkholderia* species *recA* PCR product is sufficient to discriminate between species within the genus.

3) Use *recA* sequence information for each *Burkholderia* species to design and test novel PCR primers specific for the entire genus.

4) Sequence and compare the *recA* genes from a large collection of *B. cepacia* complex strains to examine further the phylogenetic relationships between strains of clinical and environmental origin.

2.2 MATERIALS AND METHODS

2.2.1 Bacterial strains and cultivation conditions

Molecular identification approaches were developed and evaluated using an extensive collection of strains (Table 2.1). The panel was selected to be representative of the current species diversity of the *Burkholderia*, and contained 28 species from the genus with the exception of the primary pathogens *B. pseudomallei* and *B. mallei.* These *Burkholderia* species, and all *B. cepacia* complex strains, were obtained from the Belgian Co-ordinated Collections of Micro-organisms (BCCM^{TM}) LMG Bacteria collection, the Cardiff University collection (over 1200 isolates) (11, 143), and the US *B. cepacia* complex Research Laboratory and Repository (133) and included representatives of the published strain panels (50, 144). Nine isolates representing putative novel *Burkholderia* species, a selection of non-*Burkholderia* control species, and four isolates of the closely related genus *Pandoraea* (designated a novel genus in 2000 when it was separated from the genera *Burkholderia* and *Ralstonia*; (39), were also included in the test strain panel (Table 2.2). The *B. ubonensis* type strain, which appears to be a new species member of the *B. cepacia* complex (248), was also included in the study. Ecologically and biotechnologically relevant *B. cepacia* complex strains examined are described in Table 2.2. All *Burkholderia* species were identified as described previously (49, 143). Growth media were purchased from Fisher Scientific Ltd., Loughborough, UK. Strains from clinical or environmental sources were grown at 37°C or 30°C respectively, either for 24 h on Tyrptone soya agar (TSA) or in Tryptone soya broth (TSB) with shaking at 150-200 rpm for 18-20 h.

2.2.2 Extraction of chromosomal DNA from bacterial cells

DNA was prepared for PCR amplification from overnight cultures by one of two methods:

(i) Lysis and salting-out method: Overnight bacterial growths from 5 ml TSB cultures (grown with shaking at 37°C in 13 ml snap-top polypropylene tubes) were harvested by centrifugation (10 min, 1,200 x g). Bacterial pellets were re-suspended in 200 ml of TE (10 mM EDTA and 10 mM Tris-HCl [pH 8]) and 2.8 ml of lysis buffer (1% sodium dodecyl sulfate [SDS], 50 mM Tris-HCl [pH 8], 50 mM EDTA [pH 8]) containing 1 mg.ml^{-1}of pronase. The resulting lysate was incubated at 37°C with rocking for 4 h prior to the addition of 1 ml of saturated ammonium acetate, and vigorous mixing.

Table 2.1: Bacterial strains used for *recA* PCR primer development and evaluation

Species	Strain	*recA* GenBank accession number [a]	BUR1.2 PCR	BUR1.2 RFLP type	BUR3.4 PCR
***Pandoraea* species**					
Pandoraea pulmonicola	#40	AY619660	Positive	67	Negative
Pandoraea apista	Patient W	AY619657	Positive	04	Negative
Pandoraea sputorum	AU0012	AY619659	Positive	16	Negative
Pandoraea pnomenusa	Keulen	AY619658	Positive	15	Negative
Known *Burkholderia* species					
Burkholderia kururiensis	LMG 19447^T	AY619654	Positive	80	Positive
Burkholderia caledonica	LMG 19076^T	AY619669	Positive	78	Positive
Burkholderia phenazinium	LMG 2247^T	AY619668	Positive	—	Positive
Burkholderia plantarii	LMG 9035^T	AY619655	Positive	08	Positive
Burkholderia glathei	LMG 14190^T	AY619666	Positive	10	Positive
Burkholderia glumae	LMG 2196^T	AY619675	Positive	89	Positive
Burkholderia gladioli	LMG 2216^T	AY619665	Positive	70	Positive
Burkholderia fungorum	LMG 16225^T	AY619664	Positive	12	Positive
Burkholderia caryophylli	LMG 2155^T	AY619663	Positive	53	Positive
Burkholderia graminis	LMG 18924^T	AY619653	Positive	77	Positive
Burkholderia caribensis	LMG 18531^T	AY619662	Positive	18	Positive
Burkholderia thailandensis	LMG 20219^T	AY619656	Positive	83	Positive
Burkholderia sacchari	LMG 19450^T	AY619661	Positive	71	Positive
Burkholderia terricola	LMG 20594^T	AY619672	Positive	85	Positive
Burkholderia tuberum	LMG 21444^T	AY619674	Positive	86	Positive
Burkholderia phymatum	LMG 21445^T	AY619667	Positive	87	Positive
Burkholderia xenovorans	LB400^T	AAAJ00000000	Positive	01	Positive
Burkholderia andropogonis	LMG 2129^T	—	Negative	—	Negative
Burkholderia hospita	LMG 20598^T	—	Positive[b]	—	Positive
Indeterminate *Burkholderia* species					
"Burkholderia" sp. nov.	R-20943	AY619679	Positive	71	Positive
"Burkholderia" sp. nov.	R-8349	AY619681	Positive	84	Positive
"Burkholderia" sp. nov.	R-15273	AY619677	Positive	81	Positive
"Burkholderia" sp. nov.	R-701	AY619680	Positive	90	Positive
"Burkholderia" sp. nov.	R-15821	AY619678	Positive	01	Negative
"Burkholderia" sp. nov.	R-13392	AY619676	Positive	77	Positive
"Burkholderia" sp. nov.	LMG 21262	AY619673	Positive	02	Positive
"Burkholderia" sp. nov.	LMG 19510	AY619670	Positive	82	Positive
"Burkholderia" sp. nov.	LMG 20580	AY619671	Positive	84	Positive
***B. cepacia* complex species**					
Burkholderia vietnamiensis	LMG 10929^T	AF143793	Positive	66	Positive
Burkholderia multivorans	C1576	AF143774	Positive	73	Positive
Burkholderia multivorans	LMG 13010^T	—	Positive	—	Positive
Burkholderia cepacia	ATCC 25416^T	AF143786	Positive	88	Positive
Burkholderia cenocepacia	J2315/LMG16656^T	www.sanger.ac.uk	Positive	20	Positive
Burkholderia stabilis	LMG 14294^T	AF456031	Positive	63	Positive
Burkholderia pyrrocinia	LMG 14191^T	AF143794 BPP	Positive	75	Positive
Burkholderia ambifaria	LMG 19182^T	AF323985	Positive	48	Positive
Burkholderia dolosa	LMG 18943^T	AF323971	Positive	58	Positive

Species	Strain	*recA* GenBank accession number [a]	BUR1.2 PCR	BUR1.2 RFLP type	BUR3.4 PCR
Burkholderia anthina	LMG 20980^T	AF456059	Positive	57	Positive
Burkholderia cenocepacia	CFLG	AF456021	Positive	25	Positive
Burkholderia ubonensis	LMG 20358^T	AY780511	—	—	—
PCR control species					
Bordetella parapertussis	LMG 14449^T	—	Positive	—	Negative
Rhizobium vitis	LMG 8750^T	—	Negative	—	Negative
Xanthomonas sacchari	LMG 471^T	—	Positive	—	Negative
Ralstonia metallidurans	LMG 1195^T	—	Negative	—	Negative
Ralstonia gilardii	LMG 5886^T	—	Positive	—	Negative
Ralstonia eutropha	JMP134/LMG1197	www.jgi.doe.gov	Positive	—	Negative
Neisseria elongata	LMG 5124^T	—	Negative	—	Negative
Mycobacterium smegmatis	MC²155	—	Negative	—	Negative
Pseudomonas aeruginosa	C3719	—	Negative	—	Negative
Pseudomonas aeruginosa	PAO1	NC_002156	Negative	—	Negative

[a] —, *recA* sequence or RFLP not available or not determined.
[b] PCR amplification with BUR1.2 resulted in one product of the correct size 869 bp and another of approximately 400 bp which could not be resolved by optimisation.

Table 2.2: Relevant characteristics of selected *B. cepacia* complex strains

Species (*recA* phylogenetic cluster)	Strain (other strain designations)	Source and relevant information	Reference(s)
B. cepacia (genomovar I type strain cluster)	ATCC 49709	Grass seed biological control strain	(107)
B. cepacia (genomovar I Group K)	SAR-1	*Burkholderia* species metagenomic strain from Sargasso Sea	(245)
B. cepacia (genomovar I Group K)	383 (ATCC 17660; LMG 6991)	Forest soil, Trinidad; genome draft available (genome.jgi-psf.org/draft_microbes/bur94/bur94.home.html)	(219)
B. cepacia (genomovar I type strain cluster)	ATCC 25416^T (LMG 1222)	Onion rot; *B. cepacia* Type strain	(219, 240)
B. cepacia (genomovar I type strain cluster)	J1050	Clinical infection, USA	(187)
B. cepacia (genomovar I type strain cluster)	ATCC 17759 (LMG 2161)	Forest soil, Trinidad	(219)
B. cepacia (genomovar I type strain cluster)	LMG 14087	Wound swab UK	(240)
B. cenocepacia (III-B)	M36	Corn rhizosphere, USA; registered biopesticide withdrawn from commercial use; "Type Wisconsin" strain; encodes *B. cenocepacia* pathogenicity island	(178)
B. cenocepacia (III-B)	BC-1	Corn rhizosphere, USA; biological control strain; encodes *B. cenocepacia* pathogenicity island	(149)
B. cenocepacia (III-B)	BC-2	Corn rhizosphere, USA; biological control strain; encodes *B. cenocepacia* pathogenicity island	(149)
B. stabilis [a]	HI-2482	Veterinary shampoo contaminant, USA	This study
B. stabilis [a]	LMG 14294	Cystic fibrosis patient, Belgium	(241)
B. vietnamiensis [a]	G4 (ATCC 53617; R-1808)	Waste water, USA; capable of trichloroethylene degradation; derivative strain ENV435 effective in commercial field test on a contaminated aquifer; genome draft available (http://genome.jgi-psf.org/draft_microbes/bur08/bur08.home.html)	(212, 220)
B. ambifaria [a]	M54 (R-5142)	Corn rhizosphere, USA; registered biopesticide in commercial use; "Type Wisconsin" strain with antifungal properties	(178)
B. ambifaria [a]	J82 (R-5140)	Corn rhizosphere, USA; registered biopesticide in commercial use; "Type Wisconsin" strain with anti-nematodal properties	(178)
B. ambifaria [a]	BC-F	Corn rhizosphere, USA; US Department of Agriculture biological control strain; production of antifungal agents	(265)

Species (*recA* phylogenetic cluster)	Strain (other strain designations)	Source and relevant information	Reference(s)
B. ambifaria [a]	AMMD^T (LMG 19182^T)	Pea rhizosphere, USA; biological control strain and species type strain	(46, 178)
B. ambifaria [a]	AU212 (LMG 19466)	Cystic fibrosis strain isolated from a patient in Wisconsin, USA	(46)
B. pyrrocinia [a]	BC11	Soil; antifungal producing biological control strain	(111)
B. pyrrocinia [a]	ATCC 39277	Corn field soil; production of antifungal agents	(158)

[a]; Species group and *recA* cluster are identical

The lysate was incubated for 1 h at room temperature prior to the addition of 4 ml of chloroform ($CHCl_3$) and vigorous mixing. The resulting precipitate of protein and polysaccharide contaminants was then removed by centrifugation (10 min, 2100 x g). DNA was collected from the lysate by ethanol precipitation (the addition of 3 volumes of ethanol), washed with 70% ethanol, dried under vacuum, and dissolved in 500 µl of TE (10 mM Tris-HCl [pH 8], 1mM EDTA) containing 0.5 $\mu g.ml^{-1}$ RNase (Sigma-Aldrich Ltd., Gillingham, UK). The yield of DNA was approximately 100 mg per 2 ml of overnight bacterial growth (143).

(ii) Chelex rapid sheared DNA extraction: DNA suitable for use in PCR reactions was also obtained by boiling in the presence of a chelating resin as follows. Biomass from 1 ml of a liquid culture were harvested by centrifugation (5 min, 16,000 x g) and resuspended in 100 µl of a sterile solution containing 5% CHELEX 100 (Sigma Aldrich Ltd.). The suspension was boiled for 5 min, immediately placed on ice for a further 5 min and then this procedure repeated. The supernatant was recovered by centrifugation (5 min, 16,000 x g) and stored at -20°C. Before use DNA was diluted in sterile water and approximately 20 ng of template DNA incorporated into each PCR.

DNA was quantified for further analysis using the GeneQuant system (Biochrom Ltd., Cambridge, UK) at Cardiff University's Molecular Biology Unit according to the manufacturer's instructions. This system compared the A_{260} / A_{280} absorbance ratio of DNA solutions to calculate their concentration.

2.2.3 Agarose gel electrophoresis

DNA samples were analysed by agarose gel electrophoresis (Sigma-Aldrich Ltd. in 1 x TBE (Tris-Cl, Boric acid, EDTA; Sambrook *et al.* 1989). Gels were prepared at 0.8-1.0% to visualise DNA of >900 bp and 1.5% to separate PCR products of 300-900 bp. Low melting-temperature high-resolution agarose (MP Biomedicals, Irvine, California, USA) was used at a concentration of 2.5% to resolve restriction fragment length polymorphisms. DNA samples were mixed with bromophenol blue loading dye (containing 0.1 M EDTA pH 8.0; Sigma-Aldrich Ltd.) according to the manufacturer's instructions and loaded into the wells of the agarose gel. 5µl of a molecular weight DNA marker was included on all gels for size estimation and quantification of DNA fragments separated by electrophoresis. In this study the following markers were used, as appropriate: 1 kb Plus DNA ladder (Invitrogen Ltd., Paisley, UK) and 50 bp molecular size standards (Helena BioSciences, Sunderland, UK). DNA fragments were separated by electrophoresis in a 1 x TBE running buffer at 80-100 V (Bio-rad

power PAC 300) for up to 18 h in a Bio-rad SUB CELL® GT electrophoresis gel tank (Bio-rad Laboratories Ltd., Lewes, UK). After electrophoresis the gel was stained in a 1 x TBE buffer containing 0.5 $\mu g.ml^{-1}$ ethidium bromide for 30-45 min. DNA was visualised on a UV transilluminator and photographed using a Gene Genius Bioimaging system (Syngene Ltd., Cambridge, UK)

2.2.4 Pulsed field gel electrophoresis

Macrorestriction and pulsed field gel electrophoresis (PFGE) was performed as described (241). After overnight growth in 5 ml of LB broth the bacteria were harvested by centrifugation (10 min, 2100 x g) and were resuspended to an optical density of 0.8-0.9 at 620 nm in SE buffer (75 mM NaCl, 25 mM EDTA [pH 7.4]). The suspension was warmed to 45°C for 5 min and was mixed with an equal volume of molten 2% low-melting-point agarose (Type 7; Sigma-Aldrich Ltd.) that was kept at the same temperature, and the mixture was poured into 70-µl disposable plug molds (Bio-Rad Laboratories Ltd.). The plugs were briefly chilled to 4°C, and then three to five plugs were placed in 10 ml of PEN buffer (0.5 M EDTA [pH 9.6], 1% *N*-lauroyl sarcosine) containing 1 mg of pronase (Sigma-Aldrich Ltd.) per ml held within a 15-ml sterile tube. After 24 h of incubation with gentle rocking at 37°C, the plugs were washed with five volume changes (one per h) of TE buffer (Tris-EDTA). Slices (approximately 2 mm) were then cut from the plugs and were incubated overnight with 10 U of *Spe*I in a 150-µl digestion mixture at 37°C. The macrorestricted DNA was separated in 1.2% agarose gels made with 0.5 x TBE buffer (Tris-borate-EDTA) at 5 $V.cm^{-1}$ for 44 h, with pulse switch times ramped from 20 to 75 s according to the manufacturer's standard guidelines (CHEF-DR II apparatus; Bio-Rad). Bacteriophage lambda concatemers were included as size standards (Bio-Rad Laboratories Ltd.).

2.2.5 DNA amplification by PCR

PCR was performed using the *Taq* PCR core kit (Qiagen Ltd. Crawley, UK) according to the manufacturer's instructions. Each 25 µl PCR contained the following: 1 U of *Taq* polymerase, 250 µM of each deoxynucleoside triphosphate, 1x PCR buffer (including 1.5 mM $MgCl_2$), 10 pmol of each appropriate oligonucleotide primer, and 10 to 50 ng of template DNA. Both a positive (template DNA) and negative (sterile polished water) control was included with all PCR reactions to confirm amplification of the correct DNA sequence (143).

Table 2.3: Oligonucleotide primers used for PCR amplification and nucleotide sequencing in this study

PCR name	Primer pair (forward and reverse, 5' to 3')	Product size (bp)	Annealing temp. (°C)	Specificity (first use in this study)	Reference
UNI2.5	UNI2, GACTCCTACGGGAGGCAGCAG UNI5, CTGATCCGCGATTACTAGCGATTC	1020	60	Broad range of bacteria (Chapter 2)	(143)
BCR1.2	BCR1, TGACCGCCGAGAAGAGCAA BCR2, CTCTTCTTCGTCCATCGCCTC	1043	58	*B. cepacia* complex only (Chapter 2)	(143)
BUR1.2	BUR1, GATCGA(AG)AAGCAGTTCGGCAA BUR2, TTGTCCTTGCCCTG(AG)CCGAT	869	58	Broad specificity including *Burkholderia* species (Chapter 2)	This study
BUR3.4	BUR3, GA(AG)AAGCAGTTCGGCAA BUR4, GAGTCGATGACGATCAT	385	55 / 57 / 60	*Burkholderia* species exclusively (Chapter 2)	This study
16SPCR	63f, CAGGCCTAACACATGCAAGTC 1387r, GGGCGG(AT)GTGTACAAGGC	1,300	55	Broad range of bacteria (Chapter 4)	(150)

Novel primers were designed for the amplification of *Burkholderia* species *recA* genes; BUR1 and BUR2 (BUR1.2 PCR) and BUR3 and BUR4 (BUR3.4 PCR) (Table 2.3). Primers used for amplification of the 16S rRNA gene (UNI2.5 PCR) and the *B. cepacia* complex *recA* genes (BCR1.2 PCR) were as previously described (Table 2.3 (143). OligoCheck software ((6), http://www.cardiff.ac.uk/biosi/research/biosoft) was used to assist in primer design and analyse primers BUR1 and BUR2, by rating the likelihood of publicly available bacterial *recA* sequences being amplified. Thermal cycling was carried out in a Flexigene thermal cycler (Techgene Ltd, Cambridge, UK) as follows: Initial denaturation at 94°C for 5 min, 30 cycles of 30 s at 94°C, annealing for 30 s at 55-60°C and extension at 72°C for 45 s with a final 5 minute extension at 72°C. Approximately 2 µl of each PCR product was visualised by agarose gel electrophoresis (Section 2.2.3).

2.2.6 Restriction fragment length polymorphism (RFLP) analysis

RFLP analysis of 16S rRNA and *recA* gene sequences from *Burkholderia* species and other strains (Table 2.1) was performed on 5 µl of PCR product in a mixture containing the appropriate restriction enzyme buffer and restriction endonuclease as described by the manufacturer (Promega, Southampton, UK) and incubated at 37°C for 4 h. RFLP fingerprints of *B. cepacia* complex species were generated from UNI2.5 PCR amplicons with restriction endonuclease *Dde*I (Promega, UK) and compared to those of *Burkholderia* species by manual band-sharing pairwise comparisons. The analysis of RFLP patterns from BUR1.2 PCR *recA* amplicons from representatives of the entire *Burkholderia* genus, including those from the *B. cepacia* complex, was performed using HaeIII (Promega, UK). RFLP patterns obtained from *recA* were recorded and compared using computer software (GeneSnap, GeneTools, and GeneDirectory, Syngene, Cambridge, UK), and RFLP pattern similarity calculated as a Dice coefficient at 1% tolerance and clustered using the unweighted pair-group method average (UPGMA). RFLP patterns with a similarity index of 0.75 or higher were clustered as a single group. PFGE fingerprints (Section 2.2.4) were also compared and analysed using computer software as described above.

2.2.7 Nucleotide sequence analysis

All *recA* PCR products were sequenced directly using the primers designed in this study (BUR1.2) and previously (BCR1.2; (143). Sequencing reactions were prepared using Applied Biosystems Big Dye Terminator ready reaction mix version 3.1 and analysed using an ABI-PRISM 3100 Genetic Analyser capillary electrophoresis system running Performance Optimised Polymer 6 in accordance with the

manufacturer's instructions (Applied Biosystems, Foster City, CA., USA). Raw sequences from both strands of the PCR products were aligned, assembled and a consensus sequence was derived using the CAP contig assembly program within the BioEdit software (92). Analysis also involved the use of the National Centre for Biotechnology Information's (NCBI) Basic Local Alignment Search Tool (BLAST; www.ncbi.nlm.nih.gov, (2) to establish the correct gene identity.

2.2.8 Phylogenetic analysis

Multiple nucleotide sequence alignments spanning 760 nucleotides of the *recA* gene were constructed using CLUSTAL W (228). Phylogenetic and molecular evolutionary analyses were conducted using genetic-distance based neighbour-joining algorithms within MEGA Version 3.1 (119). Gaps and missing data were completely deleted in MEGA before trees were constructed using the Jukes and Cantor matrix model with random sequence input order and 1000 data sets examined by bootstrapping. All trees were rooted with the *Pseudomonas aeruginosa* PAO1 *recA* gene.

2.2.9 Nucleotide accession numbers and aligned sequence sets

Novel *recA* nucleotide sequences were determined for 28 *Burkholderia* strains, four *Pandoraea* strains, and the *B. ubonensis* type strain (GenBank accession numbers are listed in Table 2.1). Sequences were also determined for 106 *B. cepacia* complex strains representative of the diversity seen in the *recA* gene RFLP and submitted under the following accession numbers: AF143782, AF143797, AF143800, AF456003 to AF456015, AF456017, AF456018, AF456020, AF456023, AF456024, AF456026 to AF456028, AF456035, AF456036, AF456038, AF456039 to AF456050, AF456052, AF456054, AF456056, AF456057, AF456062 to AF456064, AF456066, AF456067, AF456069 to AF456083, AF456085 to AF456124 and AY753187. Phylogenetic analysis was performed on the latter novel sequences, 52 previously published *recA* genes (46, 143, 238, 239, 247, 248), and four sequences obtained from the following genome sequence projects: *B. cenocepacia* strain J2315 (NC_004503; www.sanger.ac.uk/Projects/B_cenocepacia), *B. vietnamiensis* strain G4 (NZ_AAEH00000000; genome.jgi-psf.org/draft_microbes/bur08/bur08.home.html), *B. cepacia* strain ATCC 17660 (strain 383; genome.jgi-psf.org/draft_microbes/bur94/bur94.home.html), *Burkholderia* Sargasso Sea Metagenome (245) strain SAR-1 (NS_000028; www.ncbi.nlm.nih.gov/genomes/static/es.html). Aligned sequence sets of all the *recA* sequences used in this study are available from the following web site: ftp://cepacia.bios.cf.ac.uk and the enclosed data CD.

2.3 RESULTS

2.3.1 PCR primer design and amplification of *Burkholderia* species *recA* gene

Data for the design of a *Burkholderia* genus-specific *recA* PCR were obtained from the alignment of ten published *recA* genes spanning the *B. cepacia* complex (143), and genomic regions spanning the *recA* genes from the *B. xenovorans* LB400^T, *B. cenocepacia* J2315 and *B. pseudomallei* K96243 genome sequence projects (Fig. 2.1). Few priming sites capable of amplification across the genus were identified *in silico* either within or just outside the *recA* coding sequence. Sites that facilitated amplification of a large (869 bp) internal fragment were selected and used to design primers BUR1 (21-mer; spanning base 72 to 92 relative to the *B. cenocepacia* J2315 genome *recA* gene [all subsequent primer positions are given relative to this sequence]) and BUR2 (20-mer; base 819 to 938) (Fig. 2.2). The 869 bases targeted by the BUR1.2 PCR (Table 2.3) accounted for 87% of the *recA* gene and the resulting amplicon was suitable in size for both RFLP and sequence analysis. Prior to laboratory testing, the OligoCheck software (6) was used to examine the performance of primers *in silico* against all available *recA* sequences. A panel of nine non-*Burkholderia* control strains, each containing fewer than five putative mismatches to the PCR primers, were selected from the OligoCheck output for testing (Table 2.1). The annealing temperature of the BUR1.2 PCR was adjusted to 58°C to optimise the reaction and obtain a single discrete band of the expected size. At this annealing temperature, the BUR1.2 PCR produced an 869 bp amplicon from all target species, including the *B. cepacia* complex and all *Burkholderia* species tested, with the exception of *B. andropogonis* (Table 2.1). Four of the control species, *Bordetella parapertussis*, *Xanthomonas sacchari*, *Ralstonia gilardii* and *Ralstonia eutropha,* also produced the same sized amplicon; the remaining control strains were negative (Table 2.1). Because of its broad specificity for *Burkholderia* species, the BUR1.2 PCR was subsequently used as a means to test a *recA*-based RFLP approach and to obtain further nucleotide sequence information for specific primer design.

2.3.2 RFLP analysis of *Burkholderia* species 16S rRNA and *recA* genes

The inability of RFLP analysis of the *Burkholderia* 16S rRNA gene to discriminate between species was immediately evident as several species possessed the same RFLP type. UNI2.5 PCR amplification and digestion with *Dde*I was performed on 35 strains representative of the *Burkholderia* genus, including the *B. cepacia* complex, and the *Pandoraea* genus (Table 2.1).

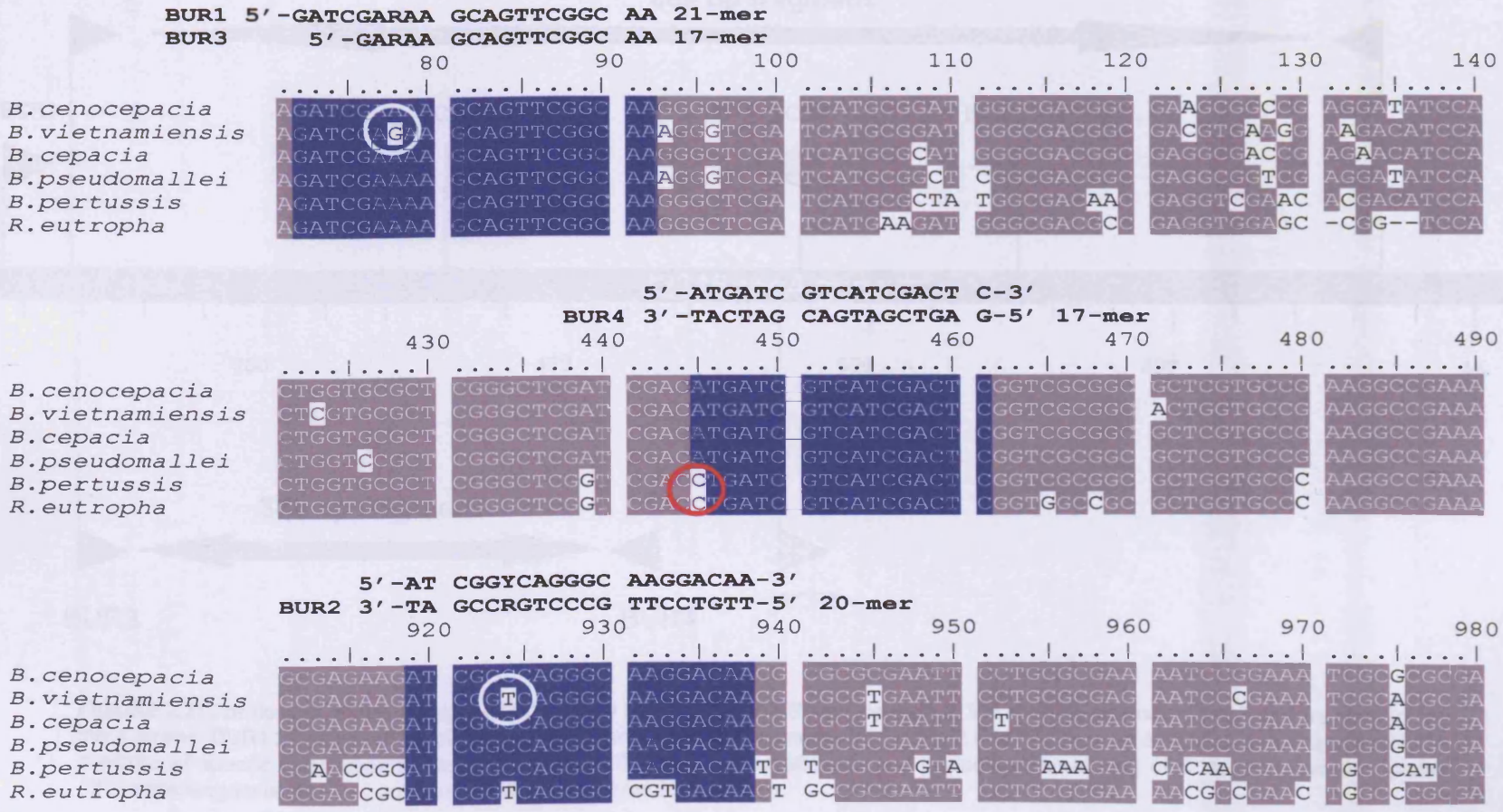


FIGURE 2.1: Nucleotide sequence alignments used to design *Burkholderia* species PCR. Multiple alignments of *recA* gene nucleotide sequences showing the regions of conservation (shaded blue) used for the design of *Burkholderia* genus specific primers. Two positions required the inclusion of a degenerate base in the primer sequence (white circles). Mismatch with closely related bacterial species at a single position facilitated *Burkholderia* genus specificity (red circle).

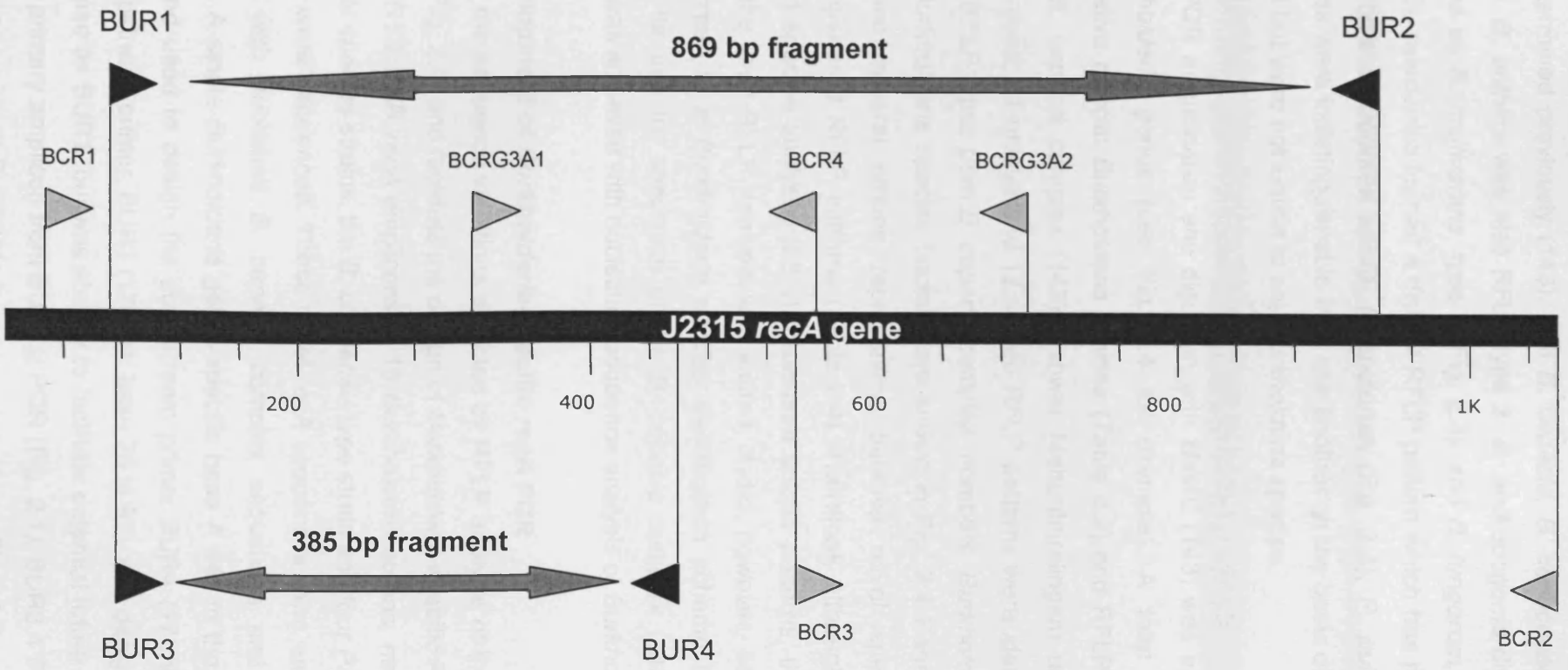


FIGURE 2.2: Positions of primers designed in this study in relation to the *B. cenocepacia* J2315 *recA* gene. Novel primer sets are marked with black arrows: BUR1.2 amplifies 87% of the total *recA* gene. BUR3.4 is nested within BUR1.2 and can be used as a nested PCR approach. The positions of specific PCR primers (designated with prefix BCR) for amplification and nucleotide sequencing of *B. cepacia* complex species (Mahenthiralingam *et al.* 2000) are marked with grey arrows for reference.

B. thailandensis, B. caryophylli and *B. pyrrocinia* shared 16S rRNA gene RFLP type 2, as determined previously (143), with *B. cepacia, B. cenocepacia* and *B. stabilis* (Fig. 2.3). *B. anthina* was also RFLP type 2. *B. andropogonis* displayed the same RFLP type as *B. multivorans*, type 3 (Fig. 2.3), and *B. fungorum, B. graminis* (Fig. 2.3) and *B. caledonica* formed a distinct RFLP pattern which has not been assigned a formal type. *Pandoraea apista, P. sputorum* (Fig. 2.3), *P. pulmonicola,* and *P. pnomenusa* were indistinguishable from one another on the basis of their novel RFLP fingerprint but were not similar to any *Burkholderia* species.

BUR1.2 PCR amplification and digestion with *Hae*III (143) was applied to analyse the *Burkholderia* genus (see Fig. 2.4 for example). A total of 103 isolates representative of both *Burkholderia* species (Table 2.2) and RFLP pattern diversity with the *B. cepacia* complex (143); Eshwar Mahenthiralingam unpublished data) were examined; 81 unique and 13 shared RFLP patterns were identified. Examples of unique RFLP types from *B. cepacia* complex members, *Burkholderia* species and two non-*Burkholderia* species bacteria are shown in Fig. 2.4 Panel B. Five known species and several strains representing putative novel species possessed overlapping/shared RFLP patterns (Table 2.4). Therefore, although the majority of strains and species analysed (86%) possessed unique patterns, the discriminatory power of the *recA*-RFLP analysis was limited. It did, however, serve as a useful primary screen in a *Burkholderia* species identification scheme in an analogous fashion to its use for speciation of the *B. cepacia* complex (143, 248). Further resolution was achieved with nucleotide sequence analysis of *Burkholderia recA.*

2.3.3 Development of *Burkholderia*-specific *recA* PCR

To confirm the sequence variations detected by RFLP analysis of the amplified *recA* fragment (Fig. 2.4) and facilitate the design of *Burkholderia* specific PCR primers, the 869 bp BUR1.2 PCR *recA* amplicons of 16 *Burkholderia* strains, nine indeterminate *Burkholderia* species strains, the *B. ubonensis* type strain and four *Pandoraea* strains (Table 2.2) were sequenced. These novel *recA* sequences were analysed following alignments with published *B. cepacia* complex sequences and genomic *recA* sequences. A single *Burkholderia* genus-specific base A 445 in the *recA* gene was identified and used to design the downstream primer BUR4 (17-mer; base 445 to 461). The upstream primer, BUR3 (17-mer; base 76 to 92), was designed to prime off the same base as BUR1, but was shorter to facilitate potential future use in a nested PCR with a primary amplicon from BUR1.2 PCR (Fig. 2.1). BUR3.4 PCR produced a 385 bp amplicon with all *Burkholderia* species examined (Figs. 2.2 and 2.4 Panel C),

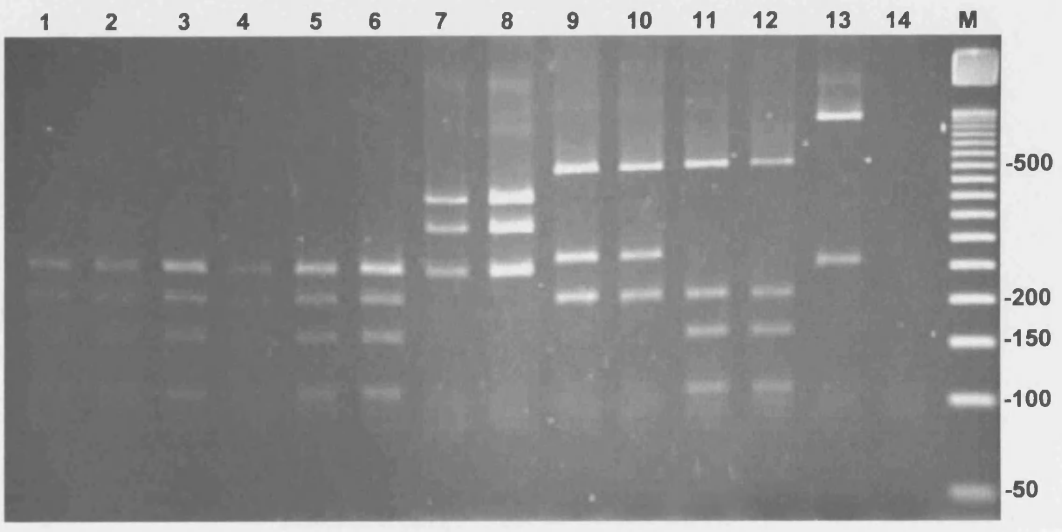


FIGURE 2.3: RFLP analysis of the *Burkholderia* 16S rRNA amplicons performed by digestion of PCR products with the enzyme *Dde*I. Samples in each lane are as follows: 1, *B. stabilis* LMG14294; 2, *B. cenocepacia* J2315; 3, *B. cepacia* LMG25416; 4, *B. thailandensis* LMG20219; 5, *B. caryophylli* LMG2155; 6, *B. pyrrocinia* LMG14191; 7, *B. fungorum* LMG16225; 8, *B. graminis* LMG18924; 9, *B. multivorans* LMG13010; 10, *B. andropogonis* LMG2129; 11, *Pandoraea apista* "W"; 12, *Pandoraea sputorum* AU0012; 13, *B. phenazinium* LMG2247; 14, H_2O negative control. Molecular size markers show size in bp in Lane M (50 bp ladder).

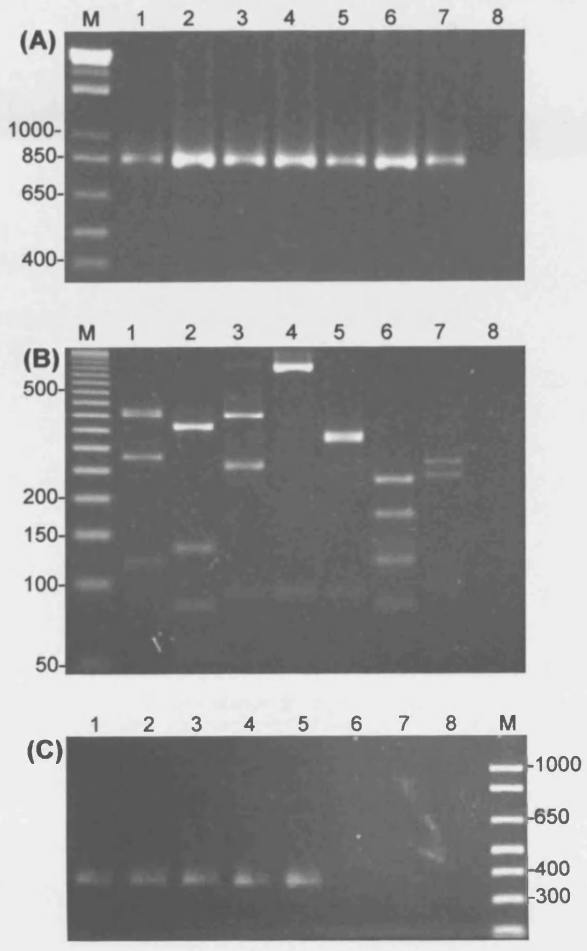


FIGURE 2.4: *Burkholderia recA* gene PCR analysis. The three steps of *Burkholderia* species identification are shown:

Panel A shows PCR products obtained with primers BUR1.2 from *Burkholderia* and control species as follows in each lane: 1, *B. stabilis* LMG14294; 2, *B. caryophylli* LMG2155^T; 3, *B. fungorum* LMG16225^T; 4, *B. graminis* LMG18924^T; 5, *B. plantarii* LMG9035^T; 6, *Bordetella parapertussis* LMG14449; 7, *Ralstonia eutropha* JMP134; 8, negative control.

Panel B shows the RFLP analysis of the BUR1.2 *recA* amplicon performed by digestion of PCR products with the enzyme *Hae* III. Samples in each lane are as follows: 1, *B. stabilis* LMG14294; 2, *B. caryophylli* LMG2155^T; 3, *B. fungorum* LMG16225^T; 4, *B. graminis* LMG18924^T; 5, *B. plantarii* LMG9035^T; 6, *Bordetella parapertussis* LMG14449; 7, *Ralstonia eutropha* JMP134; 8, negative control.

Panel C shows the specific amplification of the *Burkholderia recA* gene with primers BUR3.4 with the following species run in each lane as follows: 1, *B. stabilis* LMG14294; 2, *B. caryophylli* LMG2155^T; 3, *B. fungorum* LMG16225^T; 4, *B. graminis* LMG18924^T; 5, *B. plantarii* LMG9035^T; 6, *Bordetella parapertussis* LMG14449; 7, *Ralstonia eutropha* JMP134; 8, H_2O negative control.

Molecular size markers show size in bp in Lane M for all panels and the size of relevant bands indicated in bp.

Table 2.4: *Burkholderia* and *Pandoraea* isolates found to possess matching *recA HaeIII* RFLP types

Species	Strain name	*HaeIII* RFLP type
B. xenovorans	LB400^T	01
"Burkholderia" sp. nov.	R-15821	01
P. pulmonicola	BCC0150	15
P. pnomenusa	BCC0580	15
B. caribensis	LMG 18531	18
B. cepacia	BCC0679	18
B. cepacia complex unknown	AU0553	38
B. cepacia complex unknown	BCC0095	38
B. pyrrocinia	LMG 21823	38
B. vietnamiensis	CLO	45
B. ubonensis	R-11767	45
B. ambifaria	MVP/C1 64	46
B. cenocepacia	MDII 367	46
B. cenocepacia	MVP/C1 73	62
B. cenocepacia	BELF 2	62
B. stabilis	LMG 14294	63
B. cenocepacia	PC184	63
B. gladioli	LMG 2216	70
B. gladioli	BCC0238	70
B. sacchari	LMG 19450	71
"Burkholderia" sp. nov.	R-20943	71
B. graminis	LMG 18924	77
"Burkholderia" sp. nov.	R-13392	77
"Burkholderia" sp. nov.	R-8349	84
"Burkholderia" sp. nov.	LMG 20580	84

except *B. androprogonis* and novel "*Burkholderia*" strain R-15821. The *recA* sequence of the latter strain was subsequently shown to be more closely related to *Bordetella* (see below; Table 2.1 and Fig. 2.5). BUR3.4 PCR did not produce amplicons with non-*Burkholderia* species, including those closely related species that had originally amplified with BUR1.2 PCR (Table 2.1).

2.3.4 Phylogenetic analysis of *Burkholderia* species *recA* gene

A phylogenetic tree constructed with novel *recA* sequences from *Burkholderia* species and *Pandoraea* species (Table 2.1) illustrated the diversity of the *Burkholderia* genus (Fig. 2.5). The *B. cepacia* complex formed a distinct phylogenetic cluster. *B. thailandensis, B. pseudomallei* and *B. mallei* (the *B. pseudomallei* group) formed a cluster, which, although distinct, was more closely related to *B. cepacia* complex than other *Burkholderia* species. The *Pandoraea* genus also formed a distinct cluster consistent with their recent separation from the *Burkholderia* genus. All the indeterminate *Burkholderia* species (Table 2.2) strains clustered within the *Burkholderia* genus except strain R-15821 which was more closely related to *Bordetella* species. To test if the same cluster assignments were made using sequences obtained from the smaller BUR3.4 amplicon, the sequences were trimmed to the 300 bases within the 385 bp amplicon produced by these primers and subjected to identical phylogenetic analysis. The trimmed sequences produced a tree with the same topology and clusters that had been observed with the 800 base sequences derived from the 865 bp BUR1.2 PCR amplicon (data not shown).

2.3.5 Phylogenetic diversity of the *B. cepacia* complex

Although the *recA* gene has proven to be a valuable tool in determining the taxonomy of the *B. cepacia* complex, species such as *B. cenocepacia* and *B. cepacia* are split into distinct phylogenetic clusters when analysed in this way. Two *recA* lineages were originally observed in *B. cepacia* genomovar III, III-A and III-B (143), and subsequently clusters III-C and III-D were reported when the formal name *B. cenocepacia* was proposed for this genomovar (239). Similarly, *B. cepacia* strains also divide into two lineages by phylogenetic polymorphism in the *recA* gene; one cluster includes the type strain (ATCC 25416^T; (143) whilst the second group was named Group K based on the most common *recA* RFLP found within that cluster (248). Further novel *recA* phylogenetic groups have also been observed for other strains. The significance of these clusters was, however, difficult to interpret within a phylogeny based solely on the *B. cepacia* complex (Eshwar Mahenthiralingam, unpublished data), and so further analysis with the new data obtained in this study

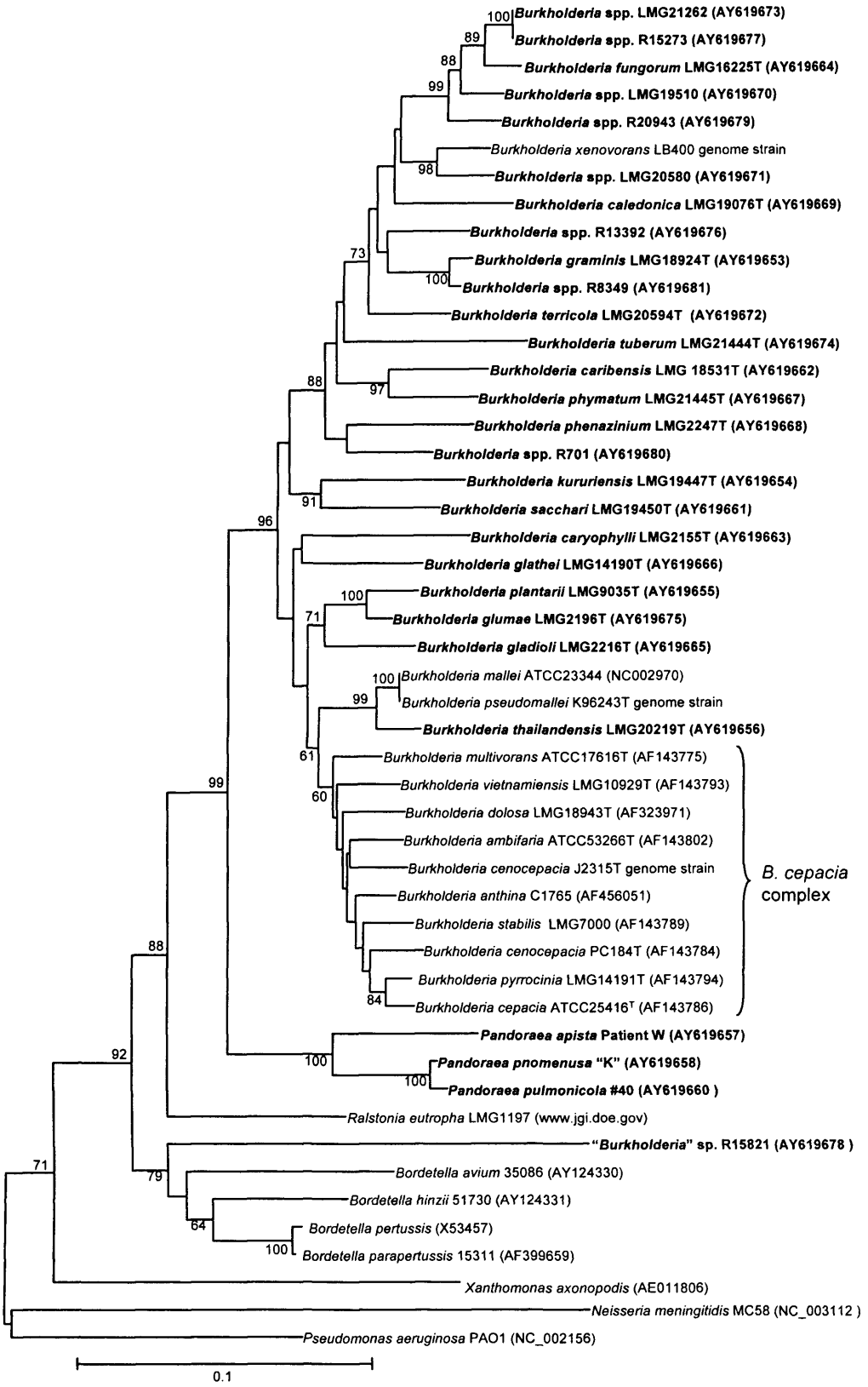


FIGURE 2.5: Phylogenetic analysis of the *Burkholderia* genus using novel *Burkholderia recA* sequences. A phylogenetic tree comparing representatives from each species is shown. 27 novel *Burkholderia recA* genes and four novel *Pandoraea recA* genes are highlighted. Bootstrap values over 70% are shown against their appropriate nodes. Number of substitutions per site is indicated by the scale bar. The species name and strain number of each isolate tested is shown, with the accession number in brackets. The tree was rooted with the *P. aeruginosa* PAO1 *recA* gene taken from the complete genome sequence as a representative member of a species outside the *Betaproteobacteria*. The following *recA* sequences were also included in the phylogenetic analysis as indicators of the ability of the analysis to differentiate amongst genera closely related to *Burkholderia*: *Ralstonia eutropha* JMP135, *Bordetella avium* AY124330, *Bordetella hinzii* AY124331, *Bordetella parapertussis* AF399659, *Bordetella pertussis* X53457, *Xanthomonas axonopodis* AE011806, *Neisseria meningitidis* NC_0031112. Sequences from the species comprising the *B. cepacia* complex were included.

was carried out to resolve phylogenetic relationships on a broader scale.

Seventeen phylogenetically distinct clusters were observed for *B. cepacia* complex species in a phylogenetic tree (Fig. 2.6) comprising a subset of 101 *B. cepacia* complex strain *recA* genes and including 28 novel *Burkholderia* and *Pandoraea* sequences (Table 2.2). As observed with the *Burkholderia* genus phylogeny (Fig. 2.5), the *B. pseudomallei* group clustered adjacent to the *B. cepacia* complex. *B. gladioli* and *B. plantarii*, and *B. glathei* and *B. carophylli* also formed distinct clusters. All the remaining *Burkholderia* species formed a diverse separate group. Within the *B. cepacia* complex, all *B. stabilis*, *B. pyrrocinia*, *B. anthina*, *B. vietnamiensis*, *B. dolosa*, *B. ambifaria*, *B. multivorans* strains and the *B. ubonensis* type strain formed single discrete phylogenetic clusters. As previously (143), *B. cepacia* strains were split into sub-lineages (Fig. 2.6) whilst *B. cenocepacia* III-A, III-B, III-C and III-D were all distinct. Three *B. cepacia* clusters were observed, one containing the type strain, ATCC 25416^T, one composed of *recA* RFLP type K strains (248) and another composed of *recA* RFLP type AW strains that previously belonged to *B. cepacia* (84); Fig. 2.6). It must, however, be noted that *B. cepacia* RFLP group K cluster was the least robust group within the phylogenetic analysis (bootstrap value 30; Fig. 2.6), suggesting that further species diversity may be present. Several strains previously difficult to assign (Mahenthiralingam, unpublished work) fell into distinct branches of the new tree. *B. cepacia* complex Group 1 strains were closely related to *B. cenocepacia* III-A, while *B. cepacia* complex Group 2 strains clustered between *B. cenocepacia* III-D and *B. anthina* (Fig. 2.6). Again all of the clusters remained intact when the *recA* sequences were trimmed to the 300 bp defining the BUR3.4 amplicon (data not shown).

2.3.6 Phylogenetic position of clinical and environmental strains

Of the 101 *B. cepacia* complex strains analysed phylogenetically (Fig. 2.6), 41 were from the environment and 60 were from human sources. No clinical strains were present in *B. cepacia* type AW, *B. pyrrocinia* and *B. cenocepacia* III-C clusters, and no environmental strains were present in the *B. cenocepacia* III-D cluster. All the remaining *B. cepacia* complex phylogenetic clusters contained strains from both sources, with the exception of *B. ubonensis* which comprised a single type strain (Fig. 2.6). The phylogenetic positions of several environmentally, biotechnologically and clinically relevant strains were also resolved (Table 2.2). Strains with documented biological control activities were found to be present in the *B. cepacia* type strain cluster (ATCC 49709), *B. cenocepacia* III-B (M36, BC-1, BC-2), *B.*

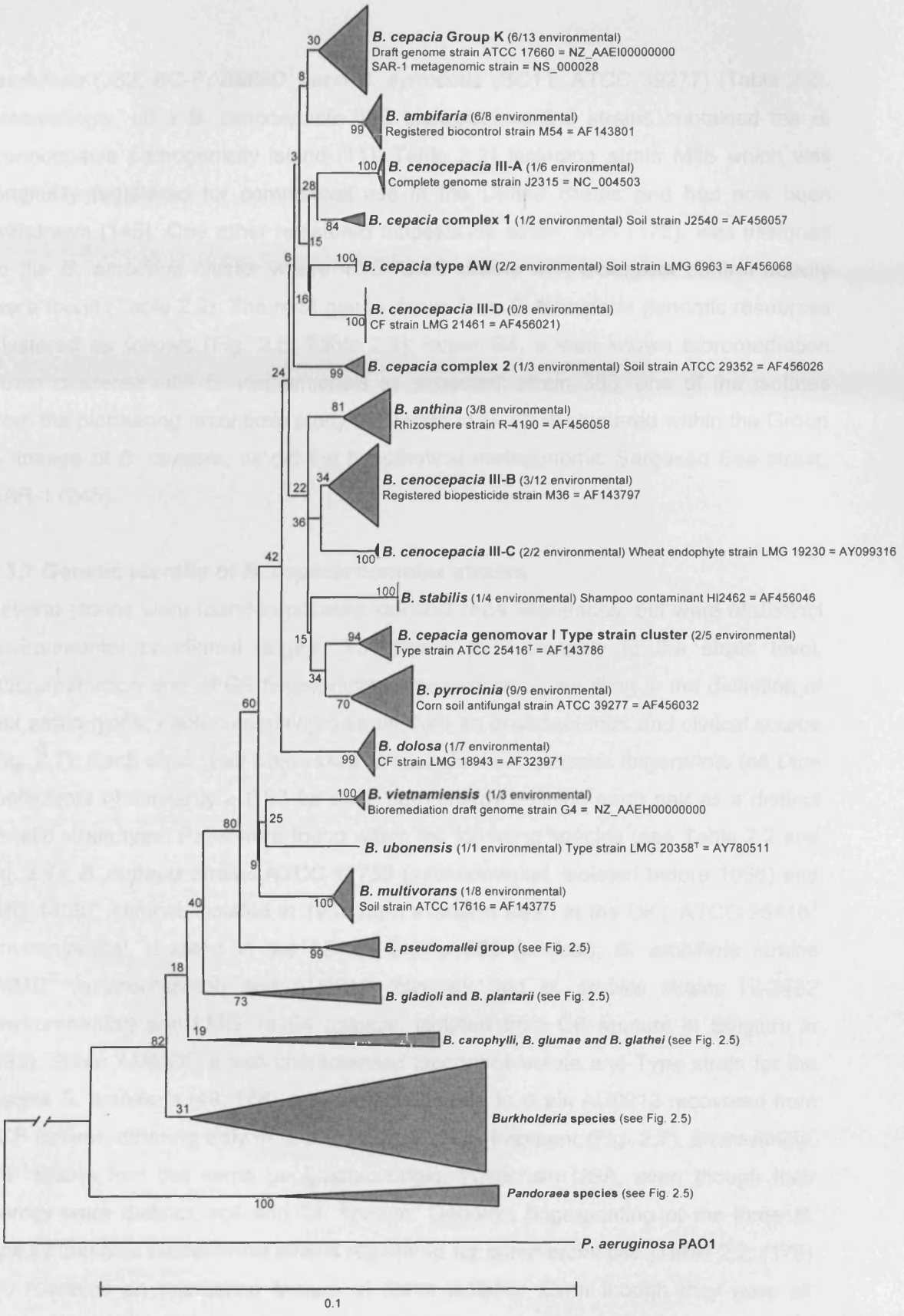


FIGURE 2.6: Phylogenetic analysis of *B. cepacia* complex strains using novel *Burkholderia recA* sequences. A phylogenetic tree comparing 101 *B. cepacia* complex strains, 28 *Burkholderia*, and four *Pandoraea recA* sequences is shown as a composite of phylogenetically distinct sequence clusters. Significant bootstrap values are shown against their appropriate nodes. Number of substitutions per site is indicated by the scale bar. The species or group name for each cluster is shown in bold. Additional information for one or more reference or interesting strains within the group is also shown. The number of sequences deriving from environmental strains versus the total number of sequences in the group is shown in brackets.

ambifaria (J82, BC-F, AMMD^T) and *B. pyrrocinia* (BC11, ATCC 39277) (Table 2.2). Interestingly, all 3 *B. cenocepacia* III-B biological control strains contained the *B. cenocepacia* pathogenicity island (11); Table 2.2) including strain M36 which was originally registered for commercial use in the United States and has now been withdrawn (145). One other registered biopesticide strain, M54 (178), was assigned to the *B. ambifaria* cluster where 4/10 listed strains with biological control activity were found (Table 2.2). The *recA* genes drawn from *Burkholderia* genomic resources clustered as follows (Fig. 2.6; Table 2.2): strain G4, a well known bioremediation strain clustered with *B. vietnamiensis* as expected; strain 383, one of the isolates from the pioneering taxonomic study of Stanier et al. (219) clustered within the Group K lineage of *B. cepacia*, as did the hypothetical metagenomic Sargasso Sea strain, SAR-1 (245).

2.3.7 Genetic identity of *B. cepacia* complex strains

Several strains were found to possess identical *recA* sequences, but were of distinct environmental or clinical origins. To resolve their clonality to the strain level, macrorestriction and PFGE fingerprinting was performed resulting in the definition of four strain types, each comprising a strain from an environmental and clinical source (Fig. 2.7). Each strain pair possessed almost identical genomic fingerprints (all Dice coefficients of similarity > 0.93 for each pair) clearly defining each pair as a distinct genetic strain type. Pairs were found within the following species (see Table 2.2 and Fig. 2.7): *B. cepacia* strains ATCC 17759 (environmental, isolated before 1966) and LMG 14087 (clinical, isolated in 1988 from a wound swab in the UK); ATCC 25416^T (environmental, isolated in the 1940s) and J1050 (clinical); *B. ambifaria* strains AMMD^T (environmental) and AU0212 (clinical); and *B. stabilis* strains HI-2482 (environmental) and LMG 14294 (clinical, isolated from CF sputum in Belgium in 1993). Strain AMMD^T, a well characterised biocontrol isolate and Type strain for the species *B. ambifaria* (46, 178), was almost identical to strain AU0212 recovered from a CF patient, differing only in one macrorestriction fragment (Fig. 2.7). Interestingly, both strains had the same geographic origin, Wisconsin USA, even though their sources were distinct, soil and CF sputum. Genomic fingerprinting of the three *B. cepacia* complex biopesticide strains registered for commercial use (Table 2.2; (178) also revealed an interesting feature of these isolates. Even though they were all registered as "Type Wisconsin," a designation derived from their phenotypic features and source, it is clear that this designation was not related to their genotype or even species identity. Strain M36 was a member of *B. cenocepacia* III-B (Fig. 2.6; Table 2.2) and possessed a macrorestriction profile clearly distinct from strains M54 and

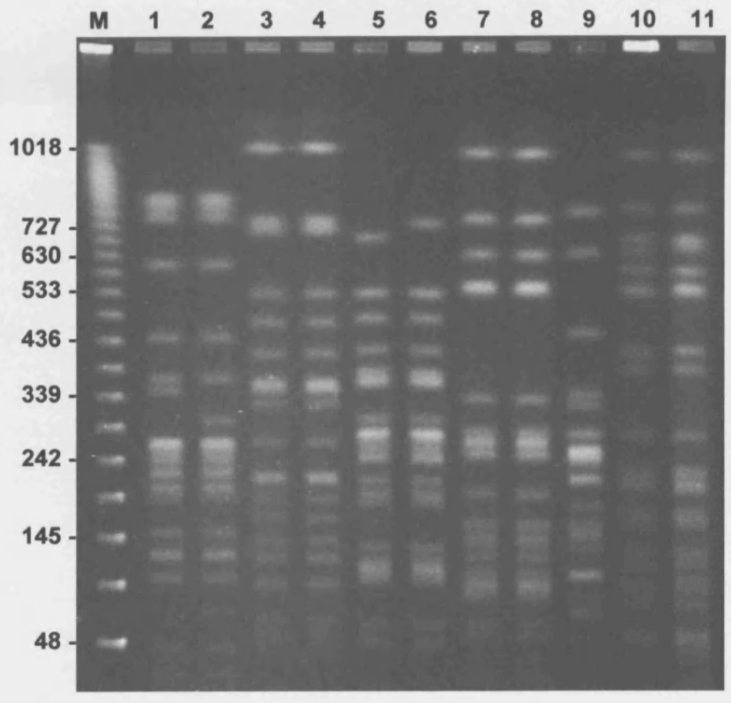


FIGURE 2.7: PFGE fingerprinting of environmental and clinical *B. cepacia* complex strains. Macrorestriction analysis with *Spe*I of the following strains is shown in each lane as follows (given respectively for each strain pair): 1 and 2, *B. cepacia* strains ATCC 17759 and LMG 14087; 3 and 4, *B. cepacia* strains ATCC 25416^T and J1050; 5 and 6, *B. ambifaria* strains AMMD^T and AU0212; 7 and 8, *B. ambifaria* strains M54 and J82; 9, *B. cenocepacia* strain M36; 10 and 11, *B. stabilis* strains HI-2482 and LMG 14294. Molecular size markers were run in Lane M and the size of relevant fragments is indicated in kb.

J82 (Fig. 2.7). Moreover, the latter biocontrol isolates were, in fact, exactly the same *B. ambifaria* strain type (Fig. 2.7), even though each possesses slightly different biopesticidal properties (Table 2.2; (178).

2.4 DISCUSSION

A genetic identification approach was developed for the entire *Burkholderia* genus that also discriminated between members of the closely related *B. cepacia* complex. Sequence polymorphisms within the *recA* gene have proven very useful in defining the taxonomy of the *B. cepacia* complex (49, 143), but the original approach could not be applied to *Burkholderia* species outside the complex. The design of new *Burkholderia* species *recA* PCRs demonstrated that RFLP analysis of the gene can be used to differentiate the majority of *Burkholderia* strains. This pattern-matching technique, however, was still limited in its ability to discriminate completely between all species, and required comparison to known reference strains for preliminary grouping of strains. Sequence comparison of *Burkholderia recA* genes ultimately provided the best means of species identification. The *Burkholderia*-specific *recA* primers and an extensive set of reference *recA* sequences provided by this study facilitated further definition of species diversity within both *Burkholderia* and the *B. cepacia* complex, and was also used to examine the phylogenetic and genotypic relationships of strains from clinical versus environmental sources.

The BUR1.2 PCR enabled amplification of an 869 bp *recA* fragment, but was not absolutely specific to *Burkholderia* and cross-reacted with strains from *Bordetella*, *Pandoraea, Ralstonia* and *Xanthomonas*. Nucleotide sequences derived from analysis of novel *Burkholderia* and *Pandoraea recA* genes did, however, enable the design of the BUR3.4 PCR primers, which were specific for the genus. Only *B. andropogonis* repeatedly failed to produce correct amplification products with both the BUR1.2 and BUR3.4 PCRs. The explanation that the sample of *B. andropogonis* was not, in fact, a *Burkholderia species* due to contamination or misidentification of the original LMG stock was unsatisfactory, because analysis of the 16S rRNA gene RFLP suggested that the isolate was a *Burkholderia* species. Moreover, Southern hybridisation of *B. andropogonis* chromosomal DNA demonstrated that DNA homologous to the *recA* gene of *B. cenocepacia* J2315 was at least partially present, suggesting that the lack of amplification had not resulted from complete gene deletion in this isolate (data not shown). The *recA* gene may not have been intact but still detectable by Southern hybridisation, hence it is most likely that there was a mismatch within the BUR1.2 and BUR3.4 primer sets. Finally, because *B. andropogonis* is an example of a one strain species (42) its amplification was not pursued.

Phylogenetic analysis of eight of the nine strains representing putative novel *Burkholderia* species (Table 2.2) corroborated the results of the phenotypic data assigning them to the genus (Fig. 2.5). Only strain R-15281 clustered outside the genus and adjacent to *Bordetella* species. Two of the novel *Burkholderia* species strains, R-15273 and LMG 21262, possessed identical *recA* genes suggesting they were members of the same novel species group that was closely related to *B. fungorum* by phylogenetic analysis (Fig. 2.5). The remaining six *Burkholderia* species strains were all highly distinct according to the bootstrap values observed in the phylogenetic analysis, suggesting they were members of novel *Burkholderia* species. These findings were consistent with the analysis of their whole cell protein profiles (Peter Vandamme, unpublished data).

The use of the 16S rRNA gene RFLP did not provide sufficient resolution to distinguish between three members of the *B. cepacia* complex (143, 209). In support of these findings, expansion of the 16S rRNA gene based RFLP approach (Section 2.3.2) confirmed that five *B. cepacia* complex species shared the same RFLP type and, moreover, could not be distinguished from two *Burkholderia* species. The previously unique RFLP type 3 of *B. multivorans* was shared with *B. andropogonis* and coupled with a lack of resolution outside the *B. cepacia* complex and within the *Pandoraea* genus; these findings illustrate the unsuitability of the 16S rRNA gene for identifying members of this genus in an RFLP approach. In contrast, greater resolution was achieved with *recA* (Table 2.4).

Although RFLP analysis of the *recA* gene did not discriminate between all species of *Burkholderia,* sequence analysis of the BUR1.2 amplicon was sufficient to separate all strains, and to differentiate them from members of the *B. cepacia* complex (Figs. 2.5 and 2.6). In particular, the degree of resolution of the *recA* phylogenetic trees for members of the *B. cepacia* complex was much greater than observed with 16S rRNA gene analysis (49). The presence of discrete *recA* lineages within species such as *B. cenocepacia* and *B. cepacia* does, however, add an additional level of complexity. A recently developed MLST scheme supported the phylogenetic placement of these *B. cepacia* complex lineages (10). A promising finding from this study was that analysis of just a 300 bp region of *recA* sequence within the BUR3.4 amplicon produced phylogenetic trees with the same topology and discrimination as the near full length trees derived from analysis of the BUR1.2 PCR. These data suggested that the BUR3.4 PCR could provide a useful and rapid cultivation-independent approach to explore the *Burkholderia* diversity in the natural environment.

Genomic resources for *Burkholderia* have increased substantially in the last few years and it has been shown that, by phylogenetic analysis of whole-genome sequence derived *recA* gene sequences, the classification of both cultivated genomic and hypothetical meta-genomic strains can be clarified. The genome strains *B. cepacia* ATCC 17760 (strain 383; (219) and *Burkholderia* Sargasso Sea strain SAR-1 reconstructed from metagenomic data; 38) were both found to be members of *B. cepacia* group K. The *recA* gene of SAR-1 was also identical to that of strain R-12710 (data not shown), cultivated from sheep with mastitis (19). Another marine isolate (obtained from the Sea of Japan) also belongs to *B. cepacia* Group K (Peter Vandamme, unpublished data). Whilst the presence of *B. cepacia* complex members in river water has already been reported (125), their isolation from marine environments is thought to be rare. The novel primers developed in this study could be applied to investigate novel environmental reservoirs for *Burkholderia* species.

Phylogenetic analysis of *Burkholderia recA* genes also demonstrated that like other taxonomic criteria, it could not be used as a means to distinguish environmental from clinical strains. Strains from both sources were found throughout the extensive phylogenetic tree derived from this study (Fig. 2.6). Although no clinical strains were found in four clusters and no environmental strains found in one *B. cepacia* complex cluster (Fig. 2.6), all of these groups contained small numbers of strains. It is likely that as more *B. cepacia* complex strains are examined and the numbers of strains in these clusters is increased, environmental and clinical strains will ultimately be identified for each cluster. Analysis of the *recA* gene of *B. cepacia* complex strains in this study resulted in the identification of several clonal pairs of strains each from distinct environmental and clinical sources (Fig. 2.7). Such pairs were found for *B. cepacia*, *B. stabilis* and *B. ambifaria*, and clearly showed that genotypically identical *B. cepacia* complex strains can be isolated from both human infections and natural or environmental sources. These findings extended the work of LiPuma and colleagues (132) who demonstrated by various genotyping methods, including macrorestriction digestion and PFGE, that epidemic CF strains from the PHDC clonal lineage, such as *B. cenocepacia* AU1107, were identical to an isolate from organic soil used for onion crops, HI2424. The *B. cenocepacia* PHDC clonal lineage accounts for the majority of CF *B. cepacia* complex infections in the mid-Atlantic region of the US (34) and the recovery of this clonal isolate from soil confirmed that human pathogenic strains were indistinct from environmental strains and initiated the comparative genomic study of *B. cenocepacia* strains HI2424 and AU1054 (www.jgi.doe.gov). *B. ambifaria* strain AMMD^T is the first strain with known biopesticidal properties (46, 178) to have an

essentially clonal relative, strain AU0212, cultivated from a patient with CF (Fig. 2.7). In conclusion, it appears that all *B. cepacia* complex bacteria are capable of colonising a wide range of habitats, and in the case of infectious niches this trait appears more dependent on the vulnerability of the host rather than the taxonomic or phylogenetic classification of the strain.

B. cepacia complex strains with useful biotechnological properties were also found throughout the *recA* phylogeny. *B. cepacia*, *B. cenocepacia*, *B. ambifaria* and *B. pyrrocinia* strains with biopesticidal or antifungal properties were identified (Table 2.2). The ability to obtain accurate species information using *recA* sequence data may prove vital in the characterisation of future biotechnologically useful strains, especially since detailed understanding of the strain taxonomy forms one of the major risk assessment criteria for commercial registration of bacteria (178). In addition, this study has shown that classification of biopesticidal strains based on phenotypic criteria alone is not adequate. The two "Type Wisconsin" strains of *B. ambifaria*, M54 and J82, which remain registered for commercial use were found to be genotypically identical (Fig. 2.7) even though they possess slightly different biopesticidal properties (Table 2.2; (178). This kind of phenotypic variability within a single clone has also been observed in isolates from CF infection (121). Although the *B. cepacia* complex has been highlighted as a group of bacterial species which require risk assessment as biological control agents (178), the case for this evaluation is not limited to these organisms.

The primary aim of this chapter was to develop a *recA* based identification approach that could be applied to the entire *B. cepacia* complex. This aim was fulfilled and in addition allowed the following to be resolved:

1) Novel PCR primers BUR1.2 were developed and amplified a near full-length *recA* gene allowing both RFLP and nucleotide sequence analysis to be performed. The BUR1.2 primer set was not, however, exclusively *Burkholderia* genus specific and amplified products from non-target *Beta* and *Alphaproteobacteria*.
2) These PCR primers formed the basis for an RFLP based approach which was insufficiently discriminatory, but nevertheless afforded greater resolution than RFLP of the 16S rRNA gene to which it was compared.

3) Novel sequence data obtained from the BUR1.2 *recA* amplicon facilitated the design of genus-specific primers suitable for the exclusive amplification of species from the *Burkholderia* genus.
4) Sequence data from the BUR1.2 region of the *recA* gene was able to differentiate between both putative and known *Burkholderia* species and all members of the *B. cepacia* complex in phylogenetic analyses.
5) The resolution of the phylogenetic analyses of a large collection of *B. cepacia* complex strains was increased by the addition of these novel sequence data, resulting in the clarification of the taxonomic position of important *Burkholderia* strains and four novel *recA* lineages. The *recA* phylogeny was not, however, a suitable basis for the differentiation between strains of clinical and environmental origin but facilitated identification of strain types capable of residing in both niches.

In summary the *recA*-based approach developed in this study provides molecular tools for the identification of *Burkholderia* species that will help to enable researchers to keep pace with the ever-increasing ecological, pathogenic and genomic interest in the genus.

CHAPTER 3: APPLICATION OF A *RECA* GENE-BASED IDENTIFICATION APPROACH TO THE MAIZE RHIZOSPHERE REVEALS NOVEL DIVERSITY IN *BURKHOLDERIA* SPECIES

Published in part as Payne et al. 2006, Appendix 2

3.1 INTRODUCTION

Bacteria from the genus *Burkholderia* occupy multiple niches of major ecological importance. Many *Burkholderia* species have been isolated from soil and have been reported to be closely associated with the plant rhizosphere (48, 178). Beneficial environmental interactions of *Burkholderia* species include their ability to facilitate both plant protection and growth-promotion as biopesticidal agents (178), the capacity of certain species to fix nitrogen (30, 161, 193), and versatile catabolic capacity that allows degradation of numerous major pollutants such as trichloroethylene (TCE) and polychlorinated biphenyls (PCB) (87, 212). In contrast, several *Burkholderia* species also cause disease in vulnerable humans (146), animals and plants (48). The *Burkholderia* genus currently comprises 34 formally described species (48), with nine species forming the *Burkholderia cepacia* complex (146). *Burkholderia* taxonomy is complex and recent studies have undertaken a polyphasic approach to characterise new species (48).

Studies on *B. cepacia* complex bacteria have typically involved the cultivation of organisms on selective media such as *Burkholderia cepacia* selective agar (BCSA; (99) and trypan blue tetracycline medium (TB-T; (91) followed by molecular analyses using the *recA* and 16S rRNA genes. Cultivation followed by RFLP analysis of the *recA* and 16S rRNA genes was used, for example, to investigate the composition of *B. cepacia* complex populations isolated from the rhizosphere of maize plants from northern, central and southern Italy (76). Similar approaches based on the *recA* and 16S rRNA genes were used to sample 91 sites in three large U.S. cities (159) and maize rhizosphere samples in conjunction with a novel colony hybridization assay (188).

The diversity of other bacterial populations associated with the plant rhizosphere has also been investigated in approaches which do not rely on cultivation and are therefore not susceptible to associated identification bias (159). The 16S rRNA gene has become an established basis for these exclusively cultivation-independent

approaches, and was used in conjunction with PCR-denaturing gradient gel electrophoresis (DGGE) to analyse the bacterial diversity in the rhizospheres of many plants. Examples include: transgenic potatoes (160), sand sedge (59) and cereal and legume crops (3). The 16S rRNA gene was also used with RAPD to characterise maize rhizosphere bacterial diversity in a study that identified a novel *Pseudomonas fulgida* strain with maize biocontrol potential (168). RFLP of the 16S rRNA gene has been used to assess bacterial diversity associated with wheat monoculture (62) and denitrifying bacteria associated with maize (35). Cultivation-independent exploration of the diversity of the *Burkholderia* genus in the environment, including rhizosphere samples, has been largely limited to 16S rRNA gene analyses. Such studies have assessed the diversity of *Burkholderia* species in the rhizosphere soil of grassland microbial communities (198) and cereal crop agricultural areas (199). They have also revealed potentially novel species ("*Candidatus* Burkholderia calva" and "*Candidatus* Burkholderia nigropunctata") as important endosymbionts that form leaf galls on African *Psychotria* plants (234) and suggested the association of novel *Burkholderia*-like species with arbuscular mycorrhizal fungi in the Gigasporaceae family (21, 22).

The *recA* gene has become established as a useful candidate for the identification of the *B. cepacia* complex species primarily because it affords greater phylogenetic discriminatory power than the 16S rRNA gene for differentiation within this closely related group (143). The *recA* gene was recently targeted in a cultivation-independent approach; this study, however, was limited to the detection of only *B. cepacia* complex species occurring in the maize rhizosphere due to the specificity of the PCR employed (185). Chapter 2 described a study that expanded the *recA*-based approach to identify the entire *Burkholderia* genus by designing two useful primer sets BUR1.2 and BUR3.4. The first primer set, BUR1.2 (Table 2.3), amplifies an almost full-length *recA* gene (869 bp) from all *Burkholderia* species, was found not to be genus-specific in that the primers amplified other *betaproteobacterial* species. However, they were useful for the generation of primary sequence data from which the specific primers BUR3.4 were designed and found to produce a 385 bp amplicon only from *Burkholderia* species (Chapter 2, Table 2.3; (183).

Recently, Ramette *et al.* (188) used a high-throughput cultivation-enrichment approach in conjunction with colony hybridisation and PCR using 16S rRNA and *recA* derived probes to examine the extent of cultivable *B. cepacia* complex species diversity present in maize-associated soil samples. The *recA* gene has not been

used as the basis for a cultivation-independent analysis of the environmental distribution of all *Burkholderia* species within this important genus.

This chapter aimed to:

1) Evaluate the application of the *recA* based identification approach to environmental samples.

2) Analyse the diversity of *Burkholderia* species as a cultivation-independent assay performed on maize rhizosphere samples from which *B. cepacia* complex species have already been cultivated (188).

3.2 MATERIALS AND METHODS

3.2.1 Environmental sampling and DNA extraction

Maize rhizosphere samples were obtained in July 2003 as described (188) from the Michigan State University, W. K. Kellogg Biological Station Long-Term Ecological Research site (Hickory Maizeers, MI) and stored at -20°C. Total DNA was extracted from 0.5 g of thawed root system samples with the Bio 101® Systems FastDNA® spin kit for soil according to the manufacturer's instructions (Qbiogene, Cambridge, UK). Four replicate DNA extractions were performed on soil aliquots from the root systems of two maize plants (Plant 1 and Plant 3; (188).

3.2.2 PCR detection of *Burkholderia* species *recA* genes

DNA extracted from the maize rhizosphere samples was subjected to PCR using primers to detect *recA* genes from the *B. cepacia* complex and other *Burkholderia* species (Table 2.3). *B. cepacia* complex *recA* genes were amplified with BCR1.2 PCR as previously described (Table 2.3 and Section 2.2.4; (143). The PCR methods described previously (Section 2.2.4) were also used to amplify *recA* genes from other *Burkholderia* species using Qiagen reagents (Qiagen Ltd. Crawley, UK) in 25 µl reactions. BUR1.2 PCR reactions (Table 2.3) were used to amplify near full length *recA* gene products from a broad range of species including *Burkholderia*. Specific PCR of *Burkholderia recA* genes was achieved using BUR3.4 PCR (Table 2.3) as described (Section 2.2.4 (183). The following modifications were included: (*i*) To amplify fragments suitable for cloning, a 20 minute final extension at 72°C was included in the thermal cycle. (*ii*) A nested BUR3.4 PCR was performed using a thousand-fold dilution (approximately 20 ng) of the BUR1.2 product as the template. (*iii*) A temperature gradient PCR with annealing temperatures between 57°C and 64°C was performed to optimise the BUR3.4 PCR for amplification of high copy number plasmid template DNA in a Dyad DNA Engine thermal cycler (Bio-Rad Laboratories Ltd, Hemel Hempstead, UK). PCR products were visualised by gel electrophoresis as described (Section 2.2.3).

3.2.3 Construction and screening of *recA* gene libraries

PCR products from BUR1.2 and BUR3.4 PCRs were cloned with pGEM®-T Easy vector and transformed into competent *Escherichia coli* JM109 cells using the manufacturer's protocols (Promega UK, Southampton). Recombinant clones were picked into 96-well plates containing Luria-Bertani broth (LB) and the antibiotic ampicillin (100 $\mu g.ml^{-1}$), grown overnight and then stored at -80°C after the addition of

8% dimethylsuphoxide to each well. Clones from the BUR1.2 library were screened individually for the presence of putative *Burkholderia* specific *recA* genes using a high throughput 96-well PCR approach as follows: after revival and growth in 96-well plates containing 150 µl LB broth with ampicillin selection, a small amount of each resulting culture was transferred with a 96-point replicator to 5 µl of lysis solution (50 mM KCl, 0.1% Tween 20, 10 mM Tris-HCl [pH 8.3]; (188) in a 96-well PCR plate and boiled at 99°C for 10 min. A cocktail of complete BUR3.4 PCR reagents was added to the boiled lysate and subjected to thermal cycling as described above and agarose gel electrophoresis.

3.2.4 Nucleotide sequence determination

Plasmid DNA from selected clones was prepared using the Wizard® *Plus* SV Minipreps (Promega) as instructed by the manufacturer. Both strands of the *recA* clones were sequenced directly with the M13F and M13R primers as described (143). Sequencing reactions were carried out as outlined (Section 2.2.6), a total of 47 representative *recA* sequences were deposited in GenBank under accession numbers DQ076251 - DQ076297. The aligned sequence set is available from ftp://cepacia.bios.cf.ac.uk/pub and on the enclosed data CD.

3.2.5 Phylogenetic analysis

A number of different phylogenetic methods had been evaluated in previous studies including neighbour-joining, maximum-likelihood and maximum-parsimony analyses (143, 183). The following phylogenetic scheme was found to be straightforward to apply and provided excellent correlation to the current species distribution in the *Burkholderia* genus (10, 143, 183, 248). Multiple nucleotide sequence alignments spanning 385 nucleotides of the *recA* gene were constructed using CLUSTAL W (228). Phylogenetic and molecular evolutionary analyses were conducted using genetic-distance based neighbour-joining algorithms within MEGA version 2.1 (http://www.megasoftware.net/). Gaps and missing data were completely deleted by MEGA2.1 before similarity matrices were constructed using the Jukes and Cantor (1969) substitution model with random sequence input order and 1000 data sets examined by bootstrapping. Trees were rooted with the *Bordetella pertussis recA* gene as an outgroup (accession number X53457). To check for the presence of chimeric *recA* genes, sequences were divided into two fragments of equal length and phylogenetic analysis performed with each half (103).

3.3 RESULTS

An overview of the experimental approach used to examine *Burkholderia* species diversity in the maize rhizosphere is shown in Fig. 3.1. Four replicate samples were obtained from two maize rhizospheres in total. Amplification of putative *recA* genes using the broad-range BUR1.2 PCR demonstrated that three of four DNA extractions from Plant 1, and four of four extractions from Plant 3 were positive for the expected 869 bp PCR product. Testing of the same samples with a BCR1.2 PCR (Table 2.3; (143) demonstrated that only two of four extractions from Plant 3 were positive for *B. cepacia* complex-specific *recA* genes. None of the maize rhizosphere samples were positive for the *Burkholderia*-specific BUR3.4 PCR product when these primers were applied directly to the extracted DNA.

To estimate *Burkholderia* diversity in the DNA extracted from Plant 3, a two-stage *recA* gene library screening strategy was applied. In the first stage, products from a BUR1.2 PCR (Table 2.3) were cloned to construct the *recA* gene library, CL1.2, and clones from this library were screened for the presence of putative *Burkholderia*-specific sequences using a high throughput 96-well BUR3.4 PCR (Table 2.3). A total of 47 putative *Burkholderia recA* positive clones were identified from 384 clones screened (Fig. 3.2) and 43 of these were successfully sequenced. Eleven (26 %) of the sequenced 869 bp *recA* clones were found to be representative of *Burkholderia* species *recA* genes using BLAST analysis (2).

28 clones from the CL1.2 library which were negative for amplification with the BUR3.4 PCR were selected at random and analysed to determine the background of non-*Burkholderia* species associated with the sample. None of these were *Burkholderia recA* sequences, but were most closely related (approximately 85% homology) to *Betaproteobacteria* such as *Herbaspirillum* species, *Rubrivivax* species and *Dechloromonas* species, and *Gammaproteobacteria* such as *Legionella* species, *Xylella* species and *Methylococcus* species.

In the second stage screen, a nested PCR with the *Burkholderia*-specific BUR3.4 primers (Table 2.3) was performed on the BUR1.2 PCR products described above (Fig. 3.1). The resulting 385 bp products were subcloned to create the CL3.4 *recA* gene library and 44 clones were sequenced, 37 (84%) of which were found to be representative of *Burkholderia* species. In total, 114 *recA* clones from the maize rhizosphere sample were sequenced.

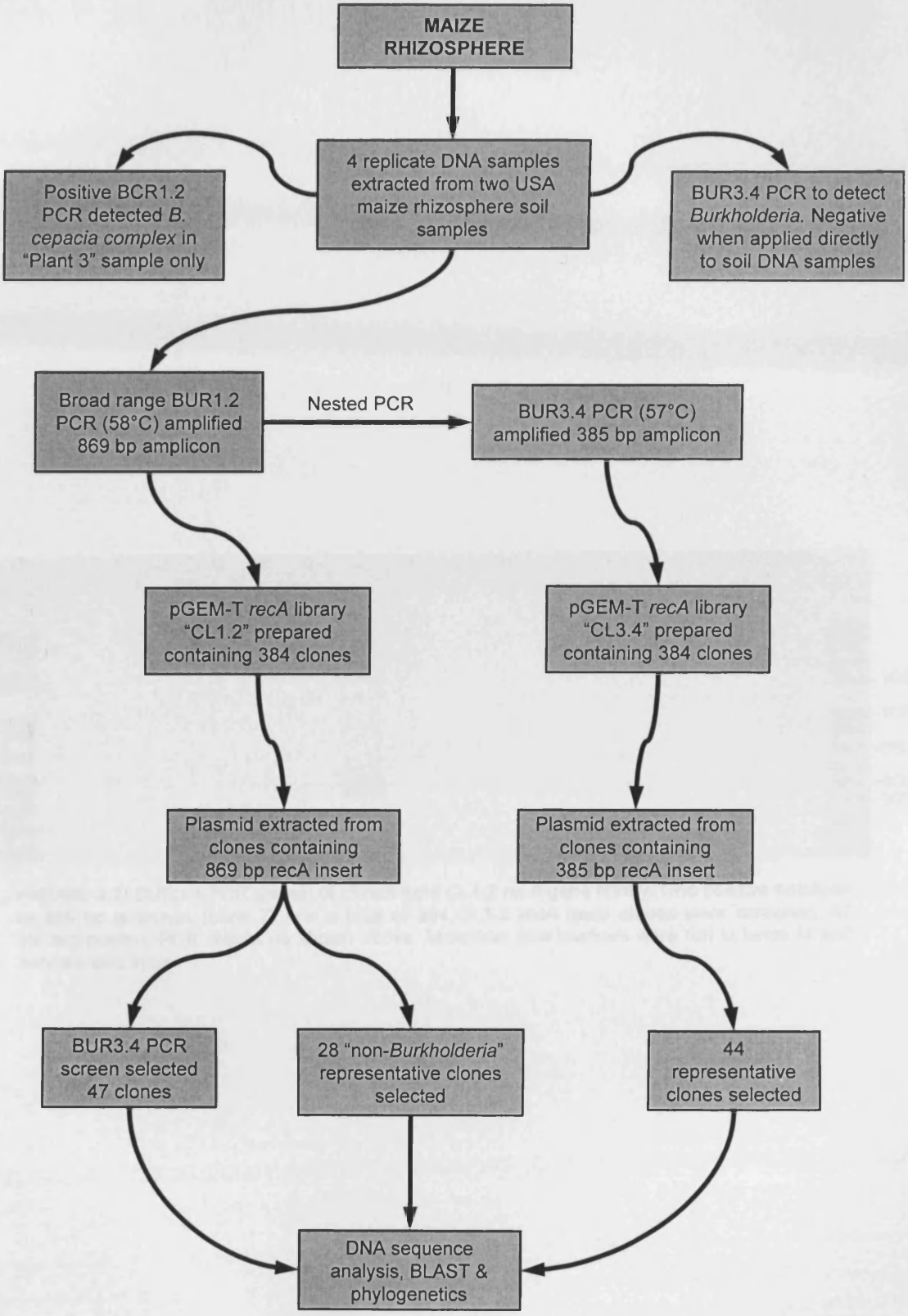


FIGURE 3.1: Flowchart of experimental strategy and overview of results obtained from the *recA* gene clone library approach.

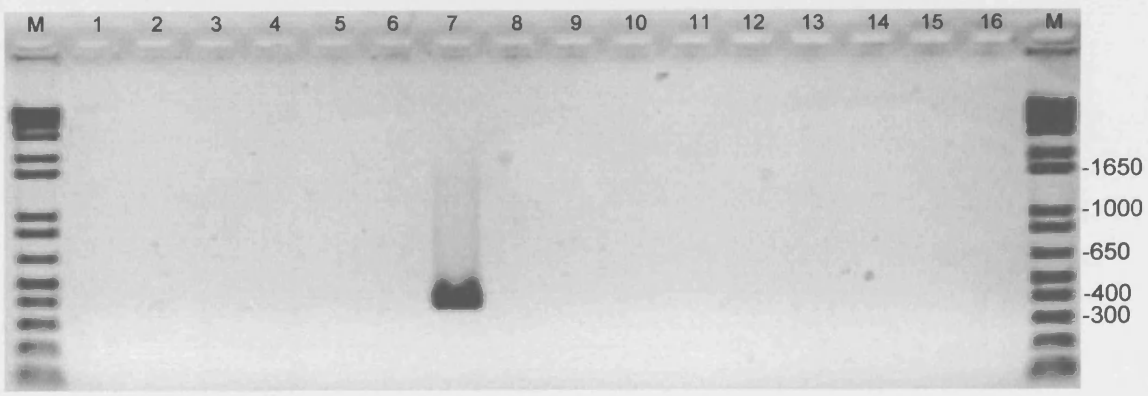


FIGURE 3.2: BUR3.4 PCR screen of clones from CL1.2 *recA* gene library. One positive amplicon of 385 bp is shown (Lane 7) and a total of 384 CL1.2 *recA* gene clones were screened; 47 yielded positive PCR results as shown above. Molecular size markers were run in lanes M and indicate size in bp.

To determine the prevalence of *Burkholderia* DNA associated with maize rhizosphere, the *recA* library clones were sequenced and subjected to phylogenetic analysis. The tree topology observed correlated well with previous studies (Chapter 2, Fig. 2.5) (143, 188) separating *Burkholderia* species from other genera, clearly defining the *B. cepacia* complex and discriminating the species within this group (Fig. 3.3). Screening *recA* gene library CL1.2 was the least sensitive means of detecting *Burkholderia* species, resulting in a final detection specificity of 26% (11/43). Eleven of the *Burkholderia* species *recA* sequences present in library CL1.2 clustered within the genus and were associated with three novel groups, only distantly related to known cultivated species, (Groups 1, 2, and 4, Fig. 3.3; bootstrap values > 50%). Only the *Burkholderia phenazinium recA* sequence clustered with these novel *Burkholderia recA* sequences, and all of the other cultivated species within the genus fell outside the cluster. Of the remaining CL1.2 *recA* sequences, one was distantly related to *Burkholderia glathei* (clone 4c11; Fig. 3.3) and three identical CL1.2 *recA* sequences clustered closely with *Burkholderia gladioli* (Group 6; Fig. 3.3).

The *recA* gene library CL3.4 derived from the nested BUR3.4 PCR (Table 2.3) proved a more effective means of detecting *Burkholderia* within the maize soil sample, with 84% (37/44) of the determined *recA* sequences falling within the genus. The majority of CL3.4 clones (35) clustered within novel phylogenetic groups, including those already defined by CL1.2 (Groups 1, 2, and 4; Fig. 3.3), as well as two further unique clusters (Groups 3 and 5; Fig. 3.3). Of the 11 BUR3.4 clones within Group 1, three were identical; while the remaining sequences were very closely related (the sequence difference across the entire group was 0.7%). The sequence variation across the four Group 2 BUR3.4 clones was 1%. Group 3 was composed of six closely related clones (2.4% variation across the group) and formed part of a novel cluster that included two discrete CL3.4 derived clones (1b1 and 2a7 Fig. 3.3). Groups 4 and 5 both contained a total of four closely related clones and the sequence variation within each of the latter groups was identical at 1.4%. The Group 5 sequences were distantly related to *Burkholderia glathei.* Only two of the CL3.4 clones, 1c5 and 2a1, were assigned to *B. cepacia* complex species, *Burkholderia ambifaria* and *Burkholderia pyrrocinia,* respectively (Fig. 3.3). The novel alignment of the majority of CL1.2 and CL3.4 clones was not due to the presence of chimeric *recA* sequences as clustering remained consistent in trees constructed from each half of the sequence.

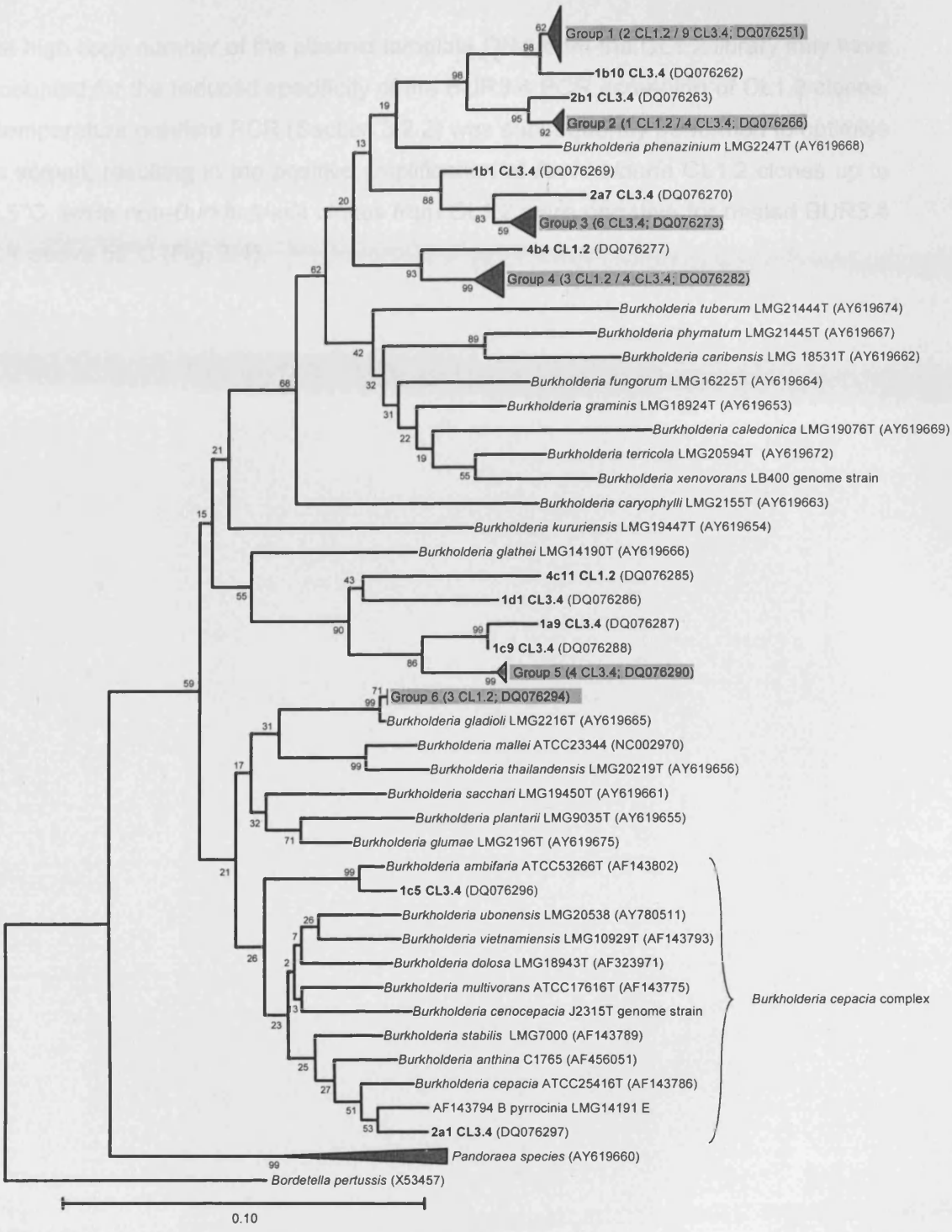


Figure 3.3: Phylogenetic analysis of the *Burkholderia recA* sequences obtained from the maize soil environment. Nucleotide sequence alignments of *recA* (385 bases) were constructed and analysed phylogenetically using genetic-distance based neighbour-joining algorithms (Jukes-Cantor matrix model; bootstrapping with 1000 replications). Six clusters (Groups 1-6) containing multiple novel *Burkholderia recA* genes are shown compressed. The clone library from which the sequence originated, the number of sequences in each cluster, and accession number for each representative *recA* gene are shown in brackets. Bootstrap values and genetic distance scale (number of substitutions per site) are indicated. Species names and strain numbers for reference *Burkholderia recA* genes are shown with accession numbers in brackets. Three *Pandoraea recA* sequences were included as a closely related genus and the *B. cepacia* complex group are also indicated.

The high copy number of the plasmid template DNA from the CL1.2 library may have accounted for the reduced specificity of the BUR3.4 PCR screening of CL1.2 clones. A temperature gradient PCR (Section 3.2.2) was subsequently performed to optimise the screen, resulting in the positive amplification of *Burkholderia* CL1.2 clones up to 60.5°C, while non-*Burkholderia* clones from CL1.2 were negative for nested BUR3.4 PCR above 59°C (Fig. 3.4).

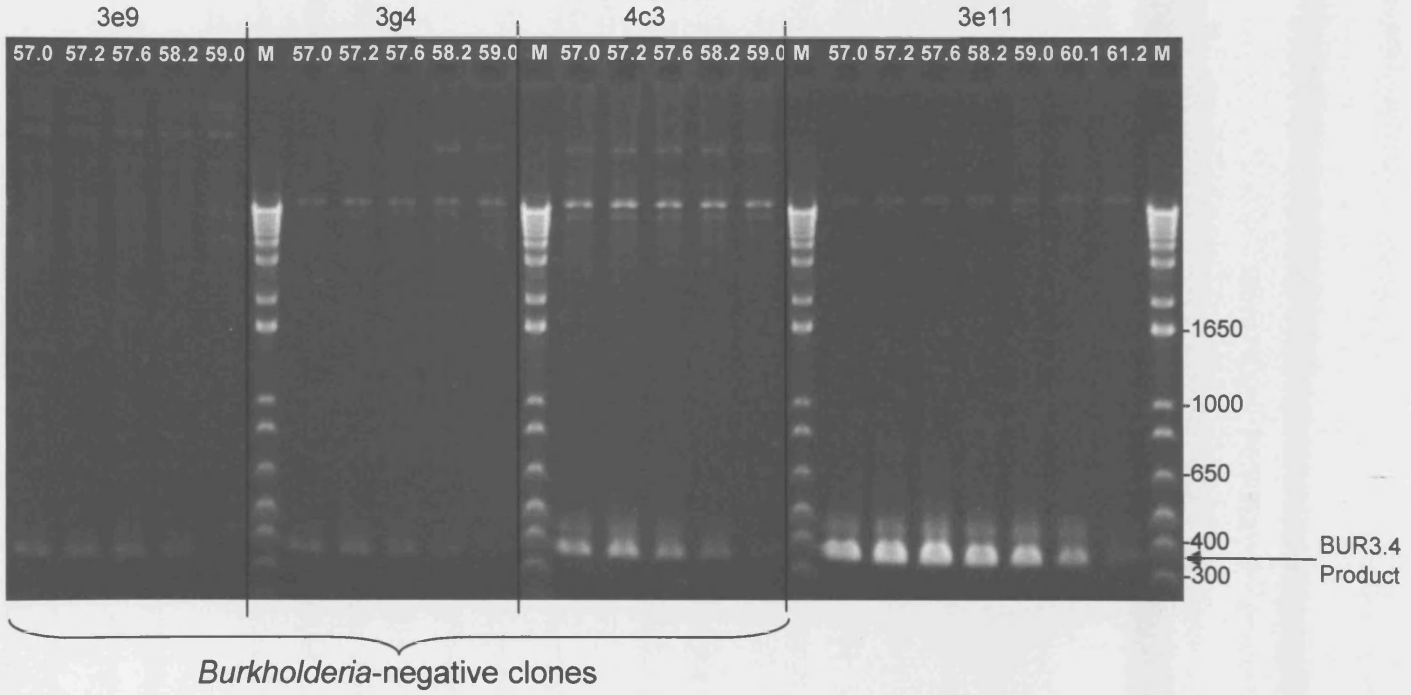


FIGURE 3.4: Thermal gradient PCR with annealing temperatures between 57°C – 64°C to optimize BUR3.4 screen of clone libraries containing high copy number plasmid encoded template DNA. *Burkholderia* CL1.2 clones were amplified up to 60.5°C but non-*Burkholderia* clones failed above 59°C.

3.4 DISCUSSION

The *recA*-based PCR approaches developed previously (Chapter 2; (183) enabled the identification and taxonomic assignment of *Burkholderia* species using DNA extracted from cultivated isolates. To test their efficacy as primers for the investigation of *Burkholderia* diversity in the natural environment BUR1.2 and BUR3.4 were used a cultivation-independent PCR-based approach to amplify *Burkholderia* species directly from maize rhizosphere soil samples known to contain cultivated members of the *B. cepacia* complex (188). The samples from Plant 3 were selected for further study because; (*i*) the positive results of the BCR1.2 PCR indicated that they contained DNA from *B. cepacia* complex species, and (*ii*) previous cultivation-based analysis had demonstrated that moderate numbers of these bacteria (3.2 x10^6 CFU/g soil) were recovered from the rhizosphere of this plant (188). The sample also provided a suitable test of discriminatory power of the *Burkholderia* genus *recA* identification approach and its ability to distinguish members of the *B. cepacia* complex from other members of this genus.

Given the success of the BUR1.2 and BCR1.2 PCRs, the failure of the BUR3.4 PCR to amplify the soil-extracted DNA is difficult to explain; lack of amplification may have occurred for a number of reasons. The BUR3.4 PCR used 17-mer primers that were shorter than the 19 to 21 base primers used in the successful *recA*-specific PCRs (Table 2.3). In combination with the annealing temperature of 57°C required to obtain the correct 385 bp product from the BUR3.4 PCR, this may have produced conditions which were too stringent to facilitate amplification of limited number of *Burkholderia recA* genes present in the soil samples. In addition, the target site of primer BUR4 was within a central region of the gene that may have a reduced capacity for amplification in comparison to the other *recA* –specific oligonucleotides which were designed to prime at each end of the 1 kb gene.

Determination of *recA* sequences present in environmental DNA samples using the described PCR approach clearly demonstrated the utility of *recA* as a cultivation-independent approach to examine *Burkholderia* diversity. It was an effective means to identify *Burkholderia* species and members of the *B. cepacia* complex in a single screen. This would not be possible using methods based on the 16S rRNA gene where limited sequence diversity precludes such discrimination within the latter group of bacteria. Although amplification with the BUR1.2 PCR was less successful at revealing the presence of *Burkholderia*, the 869 bp product constitutes >80% of the

full length of the average *recA* gene and hence has utility for the generation of sequence databases not only for *Burkholderia,* but can also target other closely related species.

Overall, nested BUR3.4 PCR showed much greater specificity for the detection of *Burkholderia* DNA within the maize rhizosphere (Fig. 3.3). The primary rationale for working with maize rhizosphere samples was that they are associated with an abundance of *B. cepacia* complex species (48, 76, 198). The *recA* approach (Chapter 2; (183) can therefore be evaluated against bacteria that have been cultivated from these samples (188). The cultivation-independent analysis correlated well with the cultivation-based approach performed previously (188), identifying *recA* clones most closely related to *B. ambifaria* and *B. pyrrocinia* (Fig. 3.3). Ramette *et al.* (188) had cultivated both of these species from soil associated with the root system of Plant 3, and while *B. ambifaria* was recovered from all seven plant samples examined, *B. pyrrocinia* was less prevalent and was only isolated from one other plant in the original survey. The abundance of these two *B. cepacia* complex species within the soil sample was $<1 \times 10^6$ CFU/g of soil (188), indicating that the sensitivity of the *recA* approach is limited with no *B. cepacia* complex species *recA* genes detected in clone library CL1.2. Only two *recA* genes corresponding to cultivated *B. cepacia* complex species were detected after nested PCR with BUR3 and BUR4.

Other studies of bacterial diversity associated with plant rhizosphere samples have used functional genes to avoid the limitations of 16S rRNA gene-based analyses and may be compared with the present study. Protein encoding genes exhibit typically faster molecular divergence in variable regions allowing greater phylogenetic discrimination within specific groups of bacteria. Additionally they may be used simultaneously to target bacteria with specific functions. The RNA polymerase encoding gene *rpoB* was used to asses the diversity of industrially and agriculturally important *Paenibacillus* species in maize rhizosphere samples (53). Antibiotic production genes have also been used; *phlD,* encoding 2,4-diacetylphloroglucinol, was used with denaturing gradient gel electrophoresis (DGGE) to identify novel antibiotic-producing *Pseudomonas* strains (18) and *prnD* was developed with RFLP to identify pyrrolnitrin-producing *Pseudomonas* and *Burkholderia* strains in wheat rhizosphere soil (60). The nitrate reductase encoding gene, *narG,* was used in a clone-library and RFLP approach which identified a dominant *Actinobacteria* population and suggested the existence of unidentified nitrate-reducing bacteria in maize rhizosphere samples (184). Use of such functional genes as phylogenetic

targets adds further resolution to cultivation-independent analyses of bacterial diversity.

In the current study, an important finding was that 90% (43 of 48 sequences) of the *recA* genes associated with the maize rhizosphere sample were assigned to putatively novel, as yet uncultivated, *Burkholderia* species. Clones from three of these novel groups (Groups 1, 2 & 4; Fig. 3.3) were identified in both the CL1.2 and CL3.4 *recA* gene libraries. The high bootstrap values for these groups indicate that they represent important uncharacterised taxa. Although simultaneous cultivation of non-*B. cepacia* complex *Burkholderia* species was not attempted in this study, the fact that the vast majority of *recA* sequences analysed were not closely related to cultivated *Burkholderia* reference species strongly indicated that uncultivated and even potentially endosymbiotic *Burkholderia* are highly abundant within the maize rhizosphere.

The aims of this study were fulfilled by the first description of the utility of the *recA* gene as a cultivation-independent means to directly examine the diversity of all *Burkholderia* species. The *recA* gene based approach has the resolution to identify and discriminate between species members of the closely related *B. cepacia* complex, and differentiate them from the rest of genus. In addition, the following points have been resolved:

1) The BUR1.2 PCR is not *Burkholderia* genus specific and detected many non-*Burkholderia* species. The production of the 869 bp amplicon, however, will prove valuable for the expansion of *recA* sequence databases to include important species abundant in maize rhizosphere samples.
2) The BUR3.4 nested PCR was most successful approach and could be considered the most efficient strategy to apply to rapid profiling of environmental samples.
3) The analysis revealed an abundance of potentially novel *Burkholderia* species associated with the maize soil system. The diversity of uncultivated *Burkholderia* species, whose potential role in the maize rhizosphere has clearly been underestimated, is clearly worthy of further analysis. Appropriate strategies would include both cultivation and cultivation-independent analyses, and metagenomics. Fluorecent *in situ* hybridisation (FISH) and microscopy could be employed to determine whether endosymbiotic relationships exist between *Burkholderia* and the maize rhizosphere.

CHAPTER 4: DETECTION OF *BURKHOLDERIA* SPECIES ASSOCIATED WITH AN ENVIRONMENTAL CORD-FORMING FUNGAL SAMPLE USING A CULTURE-INDEPENDENT APPROACH

4.1 INTRODUCTION

The association of the *Burkholderia* genus with plants has been well documented, with several species being well-recognised for their plant pathogenic, plant growth promotion or biocontrol properties (48). The most extensively studied species tend to be within the *B. cepacia* complex and are frequently implicated in interactions at the rhizosphere level. Members of this closely-related group of organisms have been associated with the rhizosphere of perennial ryegrass strain (P2; (170), pea (*B. ambifaria* AMMD; (46), cotton (93), rice (*B. vietnamiensis*; TVV75; (236) and coffee (71). In particular, maize has attracted attention as a widespread reservoir of plant-beneficial and potentially human-pathogenic strains (9, 20, 54, 76, 185, 188). Plant root environments are also habitats for *Burkholderia* species outside the complex. *B. sacchari* was named after its habitat in sugar cane (*Saccharum officinarum*) plantation soil (25) whilst *B. caledonica* was first isolated from rhizosphere soil in Scotland (43). Bacteria may compete fiercely to colonise the nutrient rich rhizospheric niche and resultantly the antibiotic concentration and diversity in the layer of soil influenced by root metabolism is comparatively high (17). The host plant, however, may gain a specific advantage in return for exudation of compounds.

Many of the bacteria associated with the plant rhizosphere exist in symbiotic relationships within root nodules. For example, members of the *Leguminoseae* form the largest plant family (approximately 18,000 species) and play host to a large diversity of soil bacteria that fix nitrogen, a process which is considered a major contributor to the global nitrogen cycle. These rhizobia commonly represent several genera; *Rhizobium*, *Bradyrhizobium*, *Azorhizobium, Sinorhizobium, Mesorhizobium* and *Methylobacterium* are all distinct alpha-proteobacterial branches. Beta-proteobacterial genera such as *Ralstonia* and *Burkholderia* have also been found to possess the *nod* genes essential for rhizobial symbiosis (166, 237). *B. caribensis* and *B. dolosa* were isolated from tropical legumes in addition to two novel nitrogen-fixing species *B. tuberum* and *B. phymatum* (237). *Burkholderia* species may also form leaf galls in the bacterial leaf nodulation process among African Rubiaceae and three novel species were identified which fulfil the definition of obligate endosymbionts as beneficial, intracellular populations dependent on the host for survival (73). As such,

they have not yet been cultivated, however, the provisional status names "*Candidatus* B. calva", "*Candidatus* B. nigropunctata" and "*Candidatus* B. kirkii" have been proposed (234, 235).

Other eukaryotes, including insects, have developed a dependence on associations with obligate endosymbionts to overcome the limitations imposed by their inability to synthesise essential amino acids and coenzymes that are necessary for their basic metabolism (73). 16S rRNA gene analysis has revealed at least two examples of endosymbiotic relationships between *Burkholderia* species and insects. *Burkholderia* was identified as part of consortium of nitrogen-fixing root-nodule bacteria closely related to *Flavobacteria, Rhizobium, Methylobacterium* and *Pseudomonas* species in the gut pouch of some *Tetraponera* ant species (233). Moreover, *Burkholderia* was recently identified as the dominant genus in the gut of Japanese broad-headed bugs (*Riptortus clavatus* and *Leptocorisa chinensis*) (114).

Endosymbionts have also been identified in fungi. 16S rRNA gene analysis and fluorescent *in-situ* hybridisation of endosymbionts from the fungus *Gigaspora margarita* revealed a cluster of novel sequences which were thought to belong to the genus *Burkholderia* (22) but were subsequently assigned to a novel proposed genus "*Candidatus* Glomeribacter gigasporum" (21). Recently the production of rhizotoxin by plant pathogenic fungi belonging to the genus *Rhizopus* was attributed to the presence of an endosymbiont. Rhizotoxin is an agriculturally important polyketide metabolite that causes rice seedling blight but its ability to inhibit mitosis and arrest the cell cycle of eukaryotic cells has also attracted considerable interest as a potential anti-tumour drug. The endosymbiont was isolated in pure culture and identified by 16S rRNA gene sequence analysis as a species of *Burkholderia* (182). The number of 16S rRNA gene sequences in this analysis was insufficient to resolve accurately the phylogenetic position of the novel strains within the *Burkholderia* genus.

This chapter centres around two types of fungi. The first, like *Gigaspora margarita* mentioned above, were arbuscluar mycorhizzal fungi. These fungi are themselves obligate symbionts and live in close association (the fungus grows within the cortex of the root) with more than 80% of vascular plants. Plants acquire inorganic nutrients (for example the poorly labile phosphate ion) from the fungus whilst the fungus, in turn, obtains carbohydrates from the plant and explores the substrata with an extensive mycelium. The fungi may confer advantage to the plant by controlling pests

(for example nematodes and fungal pathogens) and increasing the fitness of plants in polluted environments (207). Moreover, the formation of this type of fungus-autotroph symbiosis is thought to have been a fundamental step in the colonisation of land by plants (207). The second type of fungi investigated was saprophytic cord-forming fungi, in particular wood-decaying Basidiomycetes which most commonly form mycelial cords. These aggregations of longitudinally aligned hyphae sequester nutrients at the mycelial front and may translocate them over several metres, as well as antagonising or competing with woodland root pathogens, which makes them important as potential biocontrol agents (23).

On the basis that there have been no studies identifying endosymbiotic bacteria that form associations with saprophytic cord-forming fungi such as Basidiomycetes and with access to a wealth of information on these species (24), the aims of this part of the study were to:

1) Further evaluate the application of the *recA* gene based approach to environmental samples.

2) Analyse the diversity of *Burkholderia* species as a cultivation-independent assay performed on arbuscular mycorrhizal and saprophytic cord forming fungi samples from which pilot data suggesting the presence of *Burkholderia* species had already been obtained.

3) Extend the analysis of bacterial diversity in the sample with a 16S rRNA gene-based clone library and RFLP approach.

4.2 MATERIALS AND METHODS

4.2.1 Environmental sampling

Fungal samples were obtained from four individual sites within the Coed Beddick deciduous mixed woodland enclosure in Tintern, UK (NGR SO 5280001800). A pilot collection of 17 samples consisting predominantly of mycorhizzal roots was obtained on 16th May 2005 and a further set of 34, focussing on basidiomycete mycelial cords was obtained on 19th July 2005. The samples were washed vigorously in a minimum of five changes of sterile distilled water to remove surface associated bacteria. Total DNA was extracted from 0.5 g of fungal material with the Bio 101® Systems FastDNA® spin kit for soil according to the manufacturer's instructions (Qbiogene, Cambridge, UK).

Additionally, DNA was extracted from 17 pure laboratory cultures of cord forming basidiomycetes cultivated on sterile wooden blocks (65) with the Bio 101® Systems FastDNA® spin kit for soil (Qbiogene). The fungi represented the following species: *Hypholoma fasciculare, Phallus impudicus, Coprinus picaceus, Phanerochaete velutina, Serpula lacrymans, Megacollybia platyphylla,* and *Coniophora puteana.*

4.2.2 PCR detection of *Burkholderia* species *recA* genes and bacterial 16S rRNA genes

DNA extracted from the fungal samples was subjected to PCR analysis with primers BUR1.2 and BUR3.4 (Table 2.3 and Section 3.2.2), with the modification that the annealing temperature of the nested BUR3.4 PCR was reduced to 55°C. The fungal DNA samples were not amplified directly using the BUR3.4 PCR. DNA extracted from the samples was also amplified in a PCR targeting a broad range of bacterial 16S rRNA genes using primers 63f and 1387r at an annealing temperature of 55°C (16SPCR, Table 2.3) in standard PCR reactions as described (Section 3.2.2; (150)

4.2.3 Construction and screening of gene libraries

DNA extracted from Sample 30, a fungal cord sample associated with the root of nettles, was selected as the basis for clone library construction. PCR products from the BUR1.2 and 16SPCRs of Sample 30 were cloned with pGEM®-T Easy vector (Promega, Southampton, UK), selected and stored in 96-well plates (Section 3.2.3). Clones from this BUR1.2 library were individually screened with the high throughput 96-well PCR approach (Section 3.2.3) with the modification that the annealing

temperature of the BUR3.4 PCR was increased to 60°C in accordance with the results of the PCR temperature gradient to optimise the screen (Section 3.2.2).

4.2.4 Restriction fragment length polymorphism analysis

RFLP analysis was performed on 64 BUR1.2 clones that were negative in the BUR3.4 PCR screen and 76 clones from the 16S rRNA gene library. DNA was obtained from selected clones using a CHELEX extraction (Section 2.2.2). The *recA* gene clones were amplified with a BUR1.2 PCR and digested with the restriction endonuclease *Hae* III (Section 2.2.5) while the 16S rRNA gene clones were amplified in a 16SPCR (Section 4.2.2 and Table 2.3) and digested with *Dde* I. RFLP patterns were visualised by 2.5% agarose gel electrophoresis (2.2.3).

4.2.5 Nucleotide sequence determination

Selected gene clones were sequenced, aligned and identified as described (Section 3.2.4.) BLAST analysis (2) was used to determine the putative identify of the sequences derived from the 16S rRNA gene library, CL.16S.

4.2.6 Phylogenetic analysis

Using MEGA 3.1 (119), the sequences were compared to representatives of all species from the *Burkholderia* genus including those determined from the maize rhizosphere sample in Chapter 3 as previously described (Section 3.2.5).

4.3 RESULTS

4.3.1 Preliminary sampling

In a preliminary experiment, 17 pilot samples were collected from Tintern woods in May 2005. The primary focus of the sampling process were mycorrhizal roots; in fact the collection contained a range of sample types including mycorrhizal root systems, some associated with beech and oak leaf litter, a soil sample, various samples of cord forming fungi and the rhizospheres of *Urtica* (nettle), *Mercurialis* and *Vicia* species. The following results were obtained in PCRs applied directly to the extracted DNA to evaluate rapidly the distribution of putative *Burkholderia* species at the sample site: Using the BCR1.2 PCR, an amplicon of approximately 1500 bp was obtained from the *Vicia* rhizosphere. The expected size for a positive *recA* PCR from a *B. cepacia* complex species is, however, 1043 bp; this difference could have been the result of an artefactual non-specific amplification of a non-*Burkholderia* species (Fig. 4.1). The remaining samples were negative. BUR1.2; in this PCR, four of 17 samples were positive, indicating the presence of putative *Burkholderia* species. The *Vicia* rhizosphere sample yielded indiscrete products but cord-forming fungi samples yielded three positive amplicons (Fig. 4.2). BUR3.4; in a nested PCR using a 1000-fold dilution of the BUR1.2 PCR products, three amplicons of the expected size predicting *Burkholderia* were obtained from the same cord-forming fungi. The annealing temperature of this PCR was, however, lowered to 55°C for increased sensitivity (Fig. 4.3). For this preliminary survey, the only samples that produced consistent results indicative of the presence of *Burkholderia* species were mycelial cord-forming basidiomycetes. These results suggested that this sample type warranted further investigation; a change from the original focus on mycorrhizal roots.

4.3.2 *recA* gene clone library analysis

Thirty-four samples were obtained in the second sampling exercise in July 2005 and comprised exclusively of cord-forming basidiomycetes. Thirteen of these samples were positive with BUR1.2 PCR. Despite repetition, it was impossible to obtain *recA* amplicons from any sample in the BUR3.4 nested reaction. The DNA extracted from Sample 30 was selected for further study because it amplified strongly in the BUR1.2 PCR and originated from a simple mycelial cord. To estimate *Burkholderia* diversity in Sample 30, products from a BUR1.2 PCR were cloned to construct the *recA* gene library CL.DB. Clones from this library were screened for the presence of putative *Burkholderia*-specific sequences using the BUR3.4 PCR screen assuming that this

FIGURE 4.1: Application of BCR1.2 *B. cepacia* complex PCR to woodland DNA samples. Lanes 11 & 16 contain cord forming fungus and *Vicia* rhizosphere samples respectively. *B. glumae* and *B. cepacia* controls (Lanes 19 & 20) confirmed that only the *B. cepacia* complex was amplified. The remaining samples comprised mycorrhizal roots (Lanes 1, 2, 4, 6, 8, 10, 13 & 15), plant rhizosphere samples (Lanes 3 & 9), saprophytic cord-forming fungi (Lanes 5, 7, 14 & 17) and soil (Lane 18) and yielded no PCR signal. Molecular size markers are shown in Lane M and the size of relevant bands is indicated in bp.

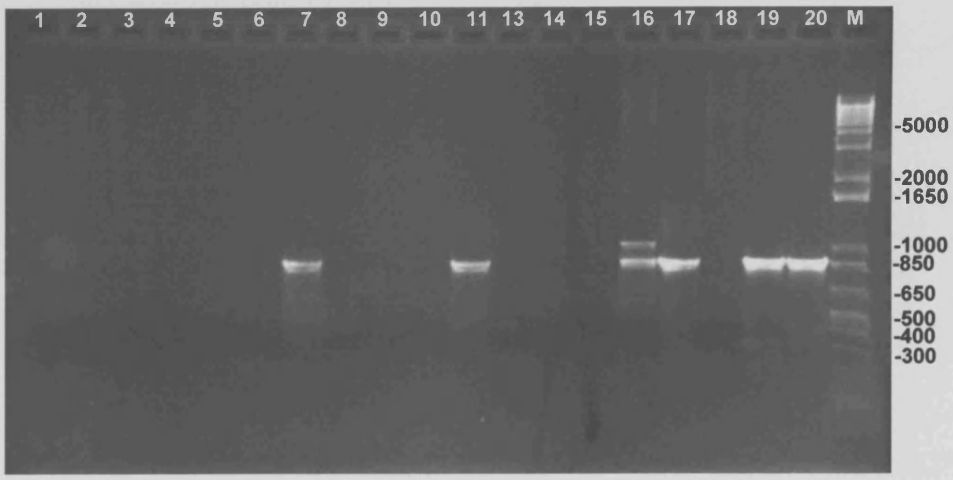


FIGURE 4.2: Application of BUR1.2 PCR to woodland DNA samples. Specificity for the *Burkholderia* genus as well as closely related *Beta-* and *Gammaproteobacterial* genera e.g. *Ralstonia*, *Bordetella,* and *Pandoraea*. Positive amplicons present in Lanes 7, 11 & 17 were all obtained from mycelial cords. Sample 16 produced non-specific amplification. Controls; 19, *B. glumae* and 20, *B. cepacia* confirmed that the PCR amplified the expected target range. The remaining samples comprised mycorrhizal roots (Lanes 1, 2, 4, 6, 8, 10, 13 & 15), plant rhizosphere samples (Lanes 3 & 9), saprophytic cord-forming fungi (Lanes 5 & 14) and soil (Lane 18) and yielded no PCR signal. Molecular size markers are shown in Lane M and the size of relevant bands is indicated in bp.

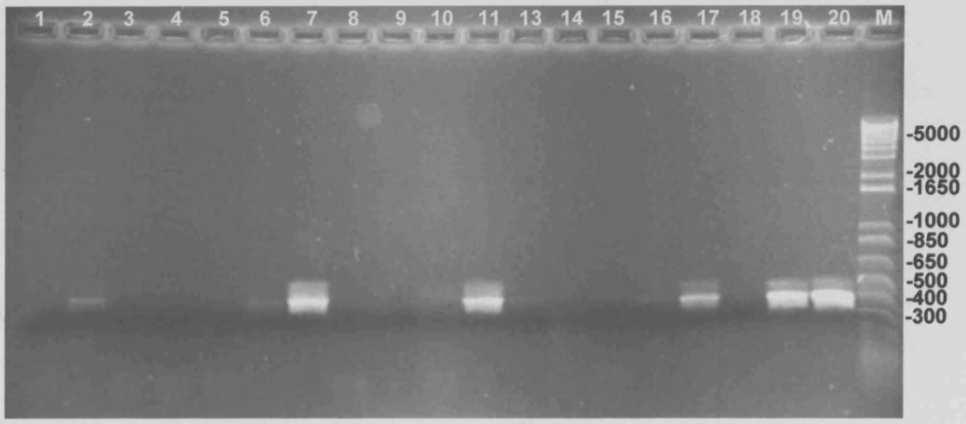


FIGURE 4.3: BUR3.4 PCR nested on BUR1.2 PCR from woodland DNA samples. Amplification conducted at a reduced annealing temperature of 55°C. Samples 7,11 & 17 showed strong amplification and were all mycelial cord fungi samples. Controls; 19, *B. glumae* and 20, *B. cepacia* confirmed that the PCR amplified the expected target range. The remaining samples comprised mycorrhizal roots (Lanes 1, 2, 4, 6, 8, 10, 13 & 15), plant rhizosphere samples (Lanes 3, 9 & 16), saprophytic cord-forming fungi (Lanes 5 & 14) and soil (Lane 18) and yielded insufficient PCR signal. Molecular size markers are shown in Lane M and the size of relevant bands is indicated in bp.

would enable detection of *Burkholderia*-like sequences. A total of 15/384 putative *Burkholderia recA* positive clones were identified (Fig. 4.4) and underwent further analysis.

The *recA* gene clone library sequences were subjected to phylogenetic analysis (Fig. 4.5) and, as in Fig. 3.3, the tree topology correlated well with previous studies (143, 183, 188). Only four of these 869 bp *recA* clones, however, were found to be representative of *Burkholderia* species resulting in a final detection specificity of 26%. Two of these sequences clustered within the genus and were associated with previously cultivated species: Sample 14 with *B. glathei,* and Sample 10 with *B. caledonica* and *B. xenovorans*. The other two clustered amongst sequences representing novel uncultivated species from the maize rhizosphere that were determined in Chapter 3. The remaining 11 *recA* genes fell outside the genus and were found by BLAST analysis (2) not to be closely related to any other *recA* sequences.

To assess the background of non-*Burkholderia* species associated with Sample 30, 64 clones which were negative for amplification with the BUR3.4 PCR were selected at random from the CL.DB library. RFLP analysis with *Hae*III was performed on this selection of clones to identify any patterns that indicated dominant groups within the mycelial cord sample. Within the *recA* analysis, two major banding patterns emerged (Fig. 4.6). The dominant pattern represented 21 of 64 (33%) samples. When subjected to sequencing and BLAST analysis no homology to was found over 84% for representatives from each pattern.

4.3.3 16S rRNA gene clone library analysis

To gain a more accurate representation of the bacterial diversity broadly represented in the sample 76 clones from the 16S rRNA gene clone library CL.16S were analysed by RFLP with *Dde*I. As expected a greater diversity of patterns was observed in this gene clone library (Fig 4.7). DNA sequence analysis was performed on ten clones representative of the dominant RFLP types amongst the 76 clones analysed, demonstrating that several unusual bacterial species were present in the sample of nettle associated cord. 16S rRNA gene sequences were discovered with close homology (≥99%) to the following sources: bacteria present in the gut of antlions (the larval stage of lacewings) (67), uncultivated *Alphaproteobacteria* from pine rhizospheres (37), legume root nodule *Bradyrhizobium* species (180, 254), methanesulfonic acid-degrading *Hyphomicrobium* species (58), uncultivated

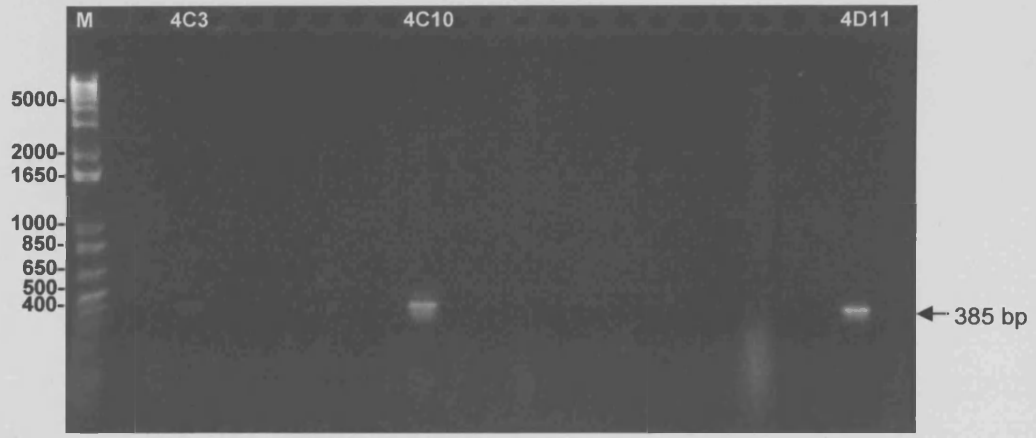


FIGURE 4.4: Screening of recA clone library CL.DB with BUR3.4 PCR screen annealing temperature was modified to 60°C. Three positive amplicons of 385 bp are shown and a total of 384 CL.DB *recA* gene clones were screened; 15 yielded positive PCR results. Molecular size markers are shown in Lane M and the size of relevant bands is indicated in bp.

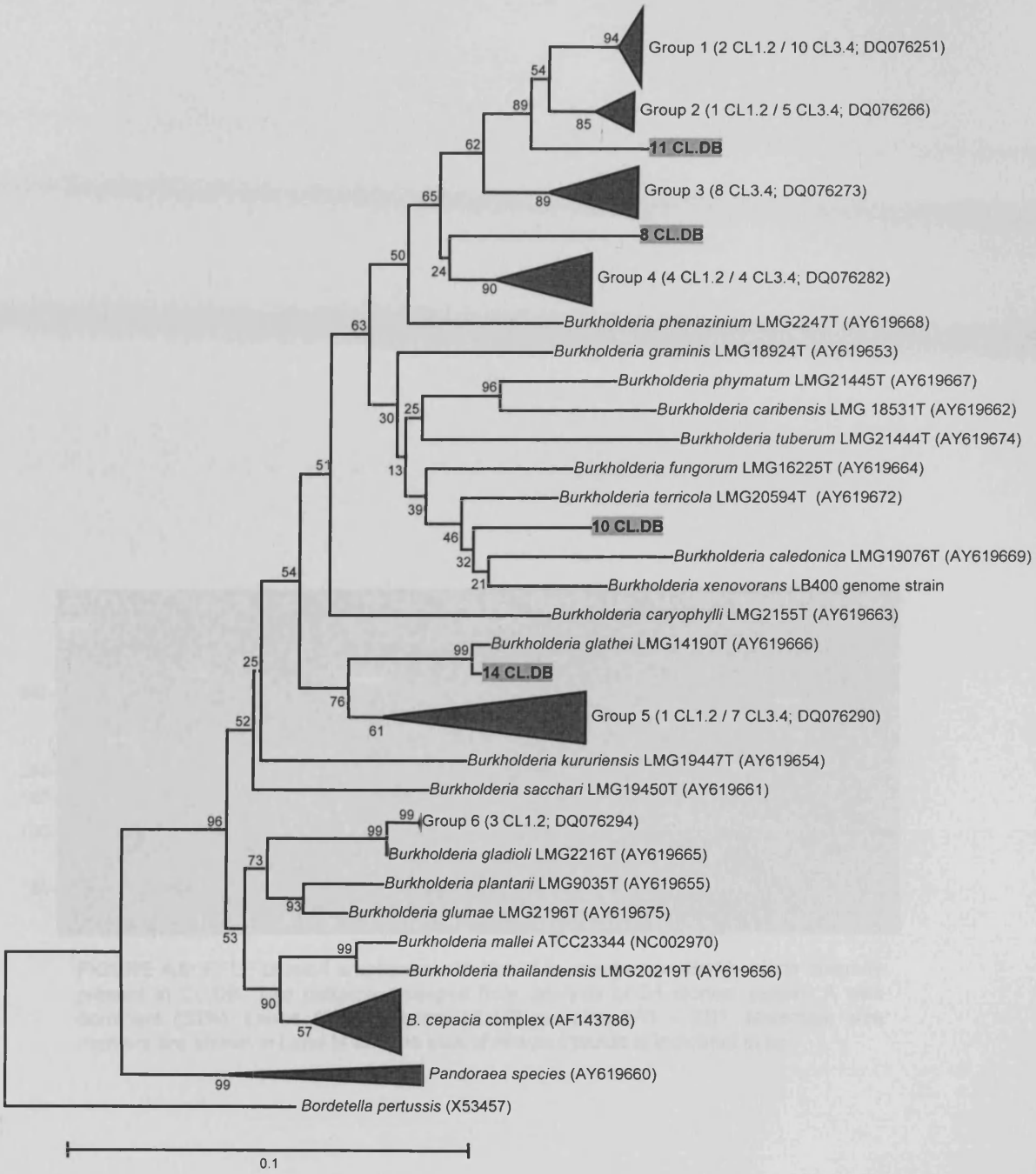


FIGURE 4.5: Phylogenetic analysis of the *Burkholderia recA* sequences from fungal samples. Nucleotide sequences were obtained from the *recA* gene clone library CL.DB, constructed from DNA Sample 30 associated with basidiomycete mycelial cord. Nucleotide sequence alignments of *recA* (385 bp) were constructed and analysed phylogenetically using genetic distance neighbour joining algorithms (Jukes-Cantor matrix model; bootstrapping with 1000 replications). Four novel *Burkholderia recA* genes are shaded. Species names and strain numbers for reference *Burkholderia recA* genes are shown with accession numbers in brackets. Bootstrap values and distance scale (number of substitutions per site) are indicated. The *B. cepacia* complex and the *Pandoraea* genera are shown compressed and are labelled with the accession number of a representative *recA* gene. Compressed groups 1-6 contain multiple novel *recA* gene sequences previously determined from a maize rhizosphere sample (Chapter 3) and are labelled with the numbers of sequences from each clone library and the accession number of a representative *recA* gene.

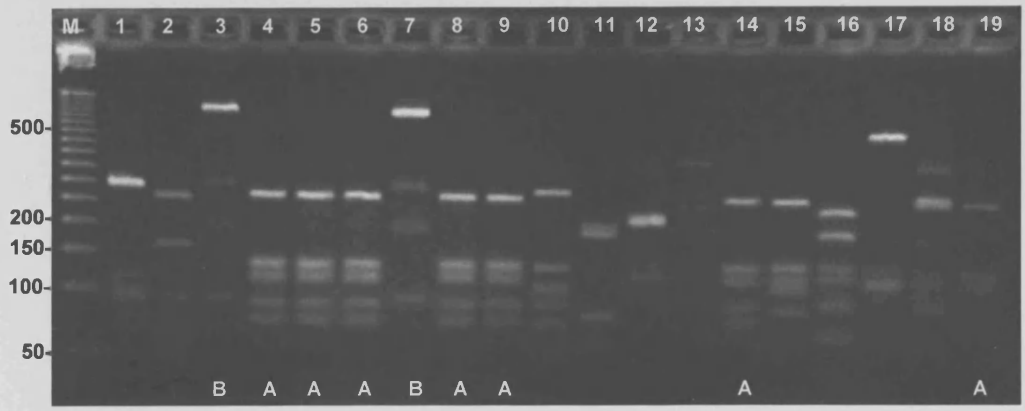


FIGURE 4.6: RFLP of *recA* amplicons with *Hae*III to assess non-*Burkholderia* diversity present in CL.DB. Two patterns emerged from analysis of 64 clones; pattern A was dominant (33%). Lanes 1-19 represent CL.DB samples 2A1 – 2B7. Molecular size markers are shown in Lane M and the size of relevant bands is indicated in bp.

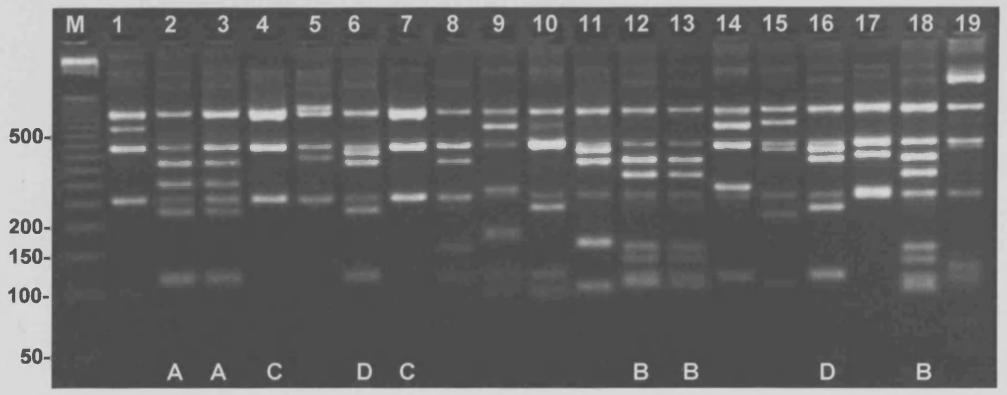


FIGURE 4.7: RFLP of the 16S rRNA gene amplicons with *Dde*I to assess non-*Burkholderia* diversity present in CL.16S. Numerous patterns emerged from analysis of 76 clones, similar ones were grouped by letter. Lanes 1-19 represent CL.DB samples 3C1 – 3D7. Molecular size markers are shown in Lane M and the size of relevant bands is indicated in bp.

Variovorax species from ectomycorrhizal fungi (14) and uncultivated bacteria from metal-contaminated soil (89).

DNA extracted from pure fungal cultures, cultivated under sterile laboratory conditions (Section 4.2.1) yielded no positive results when subjected to 16SPCR suggesting that there was little or no bacterial DNA associated with them.

4.4 DISCUSSION

Overall, low numbers of *recA* gene sequences were obtained from putative *Burkholderia* species in environmental samples from the deciduous mixed woodland of Coed Beddick Enclosure. The sample site was selected as a source of fungal material for the two sample times as it has been a previously successful source of mycorrhizal roots and basidiomycete cord formers and is well characterised by Boddy and colleagues (65, 66, 97, 255).

The commercial DNA extraction kit used (Section 4.2.1, Qbiogene) was determined the most efficient for recovery of prokaryotic DNA from low biomass sediment samples (253) and performed well on maize rhizosphere samples which contained a mixture of soil and root material (Section 3.2.1); for extracting bacterial DNA from mycelial cord samples, however, another type of commercially available kit may have been more suitable. A DNeasy Plant Mini Kit (Qiagen, Crawley, UK) was used to extract DNA from bacteria rich leaf galls (234, 235) and is also suitable for use with fungi; and Qbiogene produce a FastDNA® Spin Kit which is not specific for soil and may be more specific for eukaryotic tissue destruction.

Determination of *recA* sequences present in an environmental DNA sample using this PCR approach demonstrated, as in Chapter 3, the utility of *recA* as a cultivation-independent approach to examine *Burkholderia* diversity. A limitation of this particular study, however, was the inability to construct a library from the products of a BUR3.4 PCR nested of BUR1.2 products. When investigating the maize rhizosphere (Chapter 3) CL3.4 proved very useful, providing numerous sequences to facilitate a more detailed analysis of the *Burkholderia* diversity within a specific niche. Construction of the nested PCR clone library may have failed because the population density of *Burkholderia* species in Sample 30 was too low to facilitate this approach. Template DNA extracted from extremely low biomass samples can be susceptible to PCR bias and random amplification, and attempting to generate gene clone libraries from such samples should be approached with caution (253). Whilst a high concentration of DNA was extracted from the mycelial cord samples overall, it was impossible to quantify the ratio of bacterial to eukaryotic DNA. It was, however, likely that DNA from the fungal cords dominated the sample. Other studies have also noted this difficulty in targeting potentially endosymbiotic bacteria directly when sampling, for example, within leaf galls (235) or after enrichment and cultivation (182). Increasing the annealing temperature, and therefore the specificity, of the BUR3.4 PCR screen

applied to CL.DB had the expected effect of decreasing the number of *Burkholderia* species positive clones that were identified. It did not, however, increase the specificity of the screen as the actual percentage of *Burkholderia* species sequences recovered from the library remained around 26% of the total number of positive clones similar to the success rate observed in the maize study (Chapter 3).

The addition of the CL.16S gene clone library allowed a more accurate evaluation of the bacterial diversity in the sample. The primers 63f and 1387r were designed to amplify a wide range of 16S rRNA genes and so were well suited to this type of bacterial community analysis. It was valuable to use CL.16S in parallel with CL.DB because the number of publicly available 16S rRNA gene sequences is greater than that of *recA* sequences; hence it provides a higher resolution in terms of determining the bacterial species. The level of data provided by any gene clone library about the actual numbers of bacteria present is, at best, semi-quantitative. DGGE can assess the reliability of a clone-library by examining the diversity of sequences within a PCR product prior to the clone library construction phase (253) which contrasts to the retrospective analysis of the clone library diversity in this study with RFLP and sequencing. Adding a DGGE analysis step to check PCR reproducibility and microbial diversity in the fungal samples would yield additional information about the coverage provided by the libraries. An ideal cultivation-independent approach would quantitatively enumerate any bacteria present. For example, total bacterial counts, including differentiation between living and dead organisms may be achieved by treating fungal samples with Live/Dead *Bac*light bacterial viability stain (Invitrogen, Paisley, UK) (22, 182) Specific bacterial species may be identified by fluorescent *in-situ* hybridisation (FISH); in this case the primer sets BUR3.4 and BUR1.2 could be used to identify *Burkholderia* species and closely related species, respectively, and the broad-range 16SPCR could be used to target all bacterial species in the sample. Both of these FISH and live/dead staining techniques would be suitable to enumerate intracellular bacteria and those adhering to eukaryotic cells which would provide valuable information about a sample suspected of harbouring endosymbionts (22, 182).

Use of the broad-range bacterial 16SPCR yielded important results indicating that pure laboratory fungi cultures grown on wooden blocks did not contain any bacterial DNA. With regard to these results two hypotheses must be considered. Firstly, in the woodland samples, the bacterial DNA detected was from environmentally acquired endosymbionts. *Burkholderia* species are well known for their ability to colonise

diverse niches and are able to invade the spores of the arbuscular mycorrhizal fungi *Gigaspora decipiens* (48, 128). There is a clear distinction between permanent and cyclical endosymbioses, with the latter involving an unstable relationship with the host and regular reassociation events (152). Basidiomycetes are able to reproduce asexually which would facilitate vertical transmission of an endosymbiont, but they also produce sexual spores. Cyclical endosymbiosis supports the environmental acquisition of the bacteria detected in the fungal samples and suggests that endosymbiosis was subject to a seasonally or temporal variability that might account for the difference between the results of the pilot PCR on the samples in May and the larger sample set obtained in July.

The cord-forming basidiomycetes analysed in this study may become associated with intracellular bacteria in the same way that arbuscular mycorrhizal fungi are thought to select their bacterial symbionts from a reservoir of potential endosymbionts in their rhizosphere and hyphosphere (22). If they were environmentally acquired, it would follow that a transient invading bacterial species might be able to exist independent of a host and may inhabit other eukaryotic species causing infection or symbiosis in plants, insects or susceptible mammals. Whilst endosymbiotic relationships are host beneficial by definition (73), in some circumstances a host fungus may survive in the absence of a symbiont without any morphological changes (182). A physical constraint upon acquisition from the environment by endocytosis in fungi is the cell wall surrounding the fungal hyphae. Special circumstances allow endocytosis in some zygomycetes (22) but the interaction of *Burkholderia* species with cord-forming basidiomycetes is an area of interest which requires investigation as with arbuscular mycorrhizal species (128).

The second hypothesis is that the detected bacteria were actually contamination adhering to the fungal samples that were not washed off before the DNA extraction stage (Section 4.2.1). *B. pseudomallei* has been shown to adhere to fungal spores (128) however, this is unlikely because in the pilot data, samples containing cord forming fungi yielded a positive amplicon whilst other samples containing soil from surrounding areas were negative (Figs. 4.1-4.3).

The primary aim of this chapter was to further evaluate the application of the *recA* gene based approach to environmental samples. This aim was fulfilled with the successful gene-clone library based analysis of the bacterial diversity associated with a basidiomycete mycelial cord sample. The major conclusion, however, was that in

contrast with the investigation of the maize rhizosphere (Chapter 3), there was a lack of significant *Burkholderia* species diversity associated with fungal cords in the mixed deciduous woodland studied. The following was also resolved:

1) The different PCR results obtained in the primary and main samples suggested a temporal variation within the populations of bacteria associated with the fungal cord samples between the May and July sampling times.

2) The BUR3.4 PCR could not amplify a product directly from the main sample, possibly due to the high ratio of eukaryotic to prokaryotic DNA, despite prior success in a preliminary screen.

3) *Burkholderia* species were detected in the sample and the phylogenetic approach facilitated the placement of potentially novel bacterial diversity.

4) The addition of a 16S rRNA gene-based approach provided more information about the total bacterial diversity present in the sample. A valuable further addition would be fluorescent *in-situ* hybridisation to determine the presence of endosymbionts associated with the fungal samples.

CHAPTER 5: GENOMIC MECHANISMS OF POLLUTANT RESISTANCE IN *BURKHOLDERIA* SPECIES.

5.1 INTRODUCTION

5.1.1 Bacterial bioremediation

Many bacterial species are capable of degrading toxic pollutants to form harmless products on land and in water suffering from contamination with a range of pollutants. Bioremediation has been studied in relation to pollutants ranging from the constituents of crude oil (with a microbial consortium; (63, 226) to radioactive waste uranium with *Geobacteraceae* (5, 102). The EC Dangerous Substances Directive (76/464/EEC and daughter directives) controls the release of dangerous chemicals into water. The chemicals are divided into two groups: Discharges of List 1 Dangerous Substances, for example TCE and chloroform, have the potential to cause the most harm to aquatic life due to their persistence, toxicity or bioaccumulation and must be eliminated. List 2 Dangerous Substances are thought to be harmful, but not to the same degree, thus their discharges must be reduced (www.environment-agency.gov.uk). The majority of pollutants in these lists are halogenated and can be degraded by bacteria (13, 139, 214). Much work on the use of bacterial bioremediation has also been carried out in the field of organic solvent degradation. For example, a range of bacteria capable of metabolising the suspected carcinogen trichloroethylene (TCE), once used for processes such as metal part cleaning, dry cleaning and circuit-board manufacture and now causing problems as a recalcitrant groundwater contaminant, has been studied (213). Toluene is found in kerosene fuel and petrol, is abundant in the manufacture of phenol, dyes and adhesives and can be degraded by a range of bacterial species, the best characterised of which is *Pseudomonas putida* (51, 190). Another notable group of pollutants includes herbicides such as 2,4-dichlorophenoxyacetate (2,4-D) and 2,4,5-trichlorophenoxyacetic acid (2,4,5-T) and pesticides such as DDT. Studies on the bioremediation of these compounds have focussed on their degradation by bacteria such as *Pseudomonas fluorescens* (203) and *Burkholderia* species (36, 56).

Bacteria with the ability to degrade toxic compounds should also possess resistances to these same chemicals; to allow survival upon initial exposure and during degradation when the toxic substrate may permeate the cell envelope. Whilst a great deal is understood about the pathways that allow bacteria to degrade pollutants,

there is, however, limited understanding of the genetic basis of intrinsic resistance to pollutants.

5.1.2 Solvent tolerance and intrinsic resistance to antimicrobial agents

Chemical toxicity can obstruct the use of bacteria in the removal of pollutants from contaminated land and water. This has a serious impact on the efficiency of bioremediation in biofilters, soils and reactors (189), and even bacteria degrading radioactive actinide are found to suffer the effects of chemical toxicity rather than radiological poisoning (195). Ramos *et al.* (189) provide an extensive review of potential cellular responses to toxic chemicals, listing: *(a)* alteration of the cell membrane phospholipid composition; *(b)* reduction in cell permeability; *(c)* efflux or extrusion of the toxin in an energy dependent process; and *(d)* formation of vesicles that remove the solvent from the cell surface. The identification of bacterial strains that are tolerant to certain pollutants but that lack (105) or have lost (165) the ability to metabolise them suggests that degradation of toxic chemicals is only of minor importance to protecting cell viability (189). Pollutant tolerance mechanisms may be likened to the intrinsic resistance found across the *Burkholderia cepacia* complex to polymyxin B, in which alteration of membrane lipopolysaccharide reduces permeability to the antibiotic (164, 171).

5.1.3 *Burkholderia vietnamiensis* strain G4 as a model organism

The genus *Burkholderia* contains more than 30 characterised species which display remarkable metabolic versatility. *Burkholderia* species are found in soil and plant rhizosphere environments, where they may act either as plant pathogens or plant growth promoters. They also cause concerns in the clinical context as they are able to colonise the cystic fibrosis lung and typically exhibit broad-spectrum antimicrobial resistance, for example to polymyxin B, chloramphenicol and trimethoprim (44, 49, 171, Coenye, 2003 #17). *Burkholderia* species owe this ecological versatility to a huge coding capacity, afforded by genomes of up to 9 Mb, and an array of insertion sequences that promote genomic plasticity and general adaptability. It is this trait that has made members of this genus candidates for exploitation as degraders of crude oil constituents, herbicides, fuel additives and organic solvents (48).

Several *Burkholderia* species have already had their genomes sequenced; these include the PCB degrader *B. xenovorans* LB400, the opportunistic pathogen *B. cenocepacia* J2315, two primary pathogens *B. mallei* ATCC 23344 and *B. pseudomallei* K96243, the rhizosphere coloniser and bioremediation strain *B.*

ambifaria AMMD, and *B. vietnamiensis* G4. The potential to study the mechanisms of pollutant resistance and tolerance is considerably expanded with access to such genetic resources. *B. vietnamiensis* strains have been released in commercial field-scale environmental trials: derivatives of the TCE-degrading strain ATCC 53617 were used to bioremediate a chlorinated solvent-contaminated aquifer and a TCE groundwater plume (220). They have also been associated with plant growth promotion (178) and may also degrade benzene, *o*-cresol, *m*-cresol, *p*-cresol, phenol, toluene, naphthalene and chloroform (172, 232). *B. vietnamiensis* strain G4 has been extensively studied in relation to the degradation of TCE where a co-metabolic pathway requires a primary substrate such as toluene or phenol to generate the enzyme toluene *ortho*-monooxygenase that is necessary for aerobic degradation (169, 213). The genes that allow an organism to thrive in habitats polluted with any of these compounds are of considerable interest; for these reasons strain G4 has been selected for use as a model in this study.

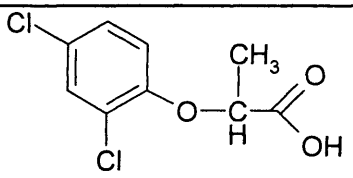
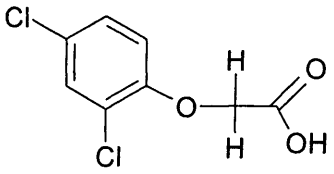
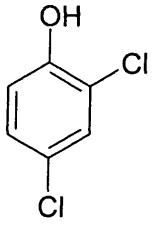
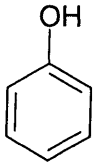
5.1.4 Selection of model pollutants

Of the substances used in the current study as model pollutants (Table 5.1), 2,4-D and 2,4-dichlorophenol are both EC Dangerous Substances Directive List 2 substances. 2,4-D is a herbicide and 2,4-dichlorophenol is a compound released during the manufacture and degradation of herbicides. Phenol is of special interest because it is an inducer of toluene *ortho*-monooxygenase in *B. vietnamiensis* strain G4 and 2-(2,4-dichlorophenoxy)-propanoic acid (dichloroprop) is an agricultural herbicide (www.speclab.com). None of these chemicals require the implementation of safety precautions above those afforded by a laboratory chemical fume extraction hood and as such are relatively convenient for use in screening experiments.

5.1.5 Signature-tagged transposon mutagenesis (STM)

Mini-transposon mutagenesis was developed in 1990 as a way of inserting a single transposon into the genome of an organism to create a random mutation. Central to the operation of Mini-*Tn*5 mutagenesis are two factors: (*a*) they contain antibiotic selection that allows successful insertion mutants to be selected, and (*b*) they encode transcriptional terminators from phage T4 and stop codons in all three reading frames, which generate strong polar mutations at their point of insertion (57). The mutant bank used in this study was signature tagged (197) to allow simultaneous identification of each mutant after a negative selection event; however, it was possible to screen individual mutants separately.

Table 5.1: Chemical structure and concentration of model pollutants used in this study

Pollutant Chemical Abstracts Service (CAS) nomenclature "common name"	**Structure**	**G4 IC** (mM)	**G4 SC** (mM)
2-(2,4-dichlorophenoxy) propanoic acid "dichlorprop"		7.0	3.0
(2,4-dichlorophenoxy) acetic acid "2,4-D"		2.0	1.0
2,4-dichlorophenol		1.0	0.7
Phenol		10.0	8.0

G4 IC: Inhibitory concentration at which there was a significant reduction in growth compared to G4 wild type. G4 SC: Estimated concentration permitting G4 wild-type growth but allowing the identification of mutants showing attenuated growth in the pollutants.

5.1.6 Aims

Since little is known about the intrinsic resistance of bacteria to pollutants the aim of this particular study was to elucidate the genetic basis for this phenotype. The following aims were fulfilled:

1) Screen a library of *B. vietnamiensis* strain G4 mutants against a panel of halogenated phenol derivatives ("model pollutants") and identify mutants in which growth in the presence of these agents was reduced.

2) Identify the genes in which expression has been interrupted by the insertion of a mini-Tn5 transposon and locate these genes on the strain G4 genome.

3) Restore the pollutant resistance phenotype of mutants by gene complementation.

4) Link mutated gene to function in order to determine the molecular basis for intrinsic pollutant resistance in *B. vietnamiensis* strain G4.

5.2 MATERIALS AND METHODS:

5.2.1 Bacterial strains and plasmids

5.2.1.1 Bacterial growth conditions

The strains and plasmids used in this study are described in detail in Table 5.2. *B. vietnamiensis* strains were grown at 37°C, either for 24 h on tryptone soya agar (TSA, Fisher Scientific Ltd., Loughborough, UK) or in tryptone soya broth (TSB) with shaking at 150-200 rpm for 18-20 h. *E. coli* strains were grown at 37°C on Luria-Bertani (LB) agar or in LB broth with shaking at 150-200 rpm for 18-20 h. Antibiotics (MP Biomedicals London, UK) were added to growth media as follows: kanamycin (KM) >200 $\mu g.ml^{-1}$ to select for the mini-Tn*5Km2* inserts or maintain the pRK2013 helper plasmid; polymyxin B (PMX) >300 $Units.ml^{-1}$ to select for *B. vietnamiensis*; and trimethoprim (TP) >100 $\mu g.ml^{-1}$ to select for the plasmid vector pASE101. Antibiotic stock solutions were prepared in sterile deionised water and filter sterilised using a 0.45 µm minisart filter (Sartorius AG, Goettingen, Germany) with the exception of TP, which is insoluble in water and was prepared in dimethylsulfoxide (DMSO).

5.2.1.2 Storage and maintenance of bacterial strains

All bacterial stock cultures were stored at -80°C in liquid growth media containing 8% DMSO. *B. vietnamiensis* G4 mutants were also archived into 96-well plates at -80°C. When required, strains were revived on TSA or LB agar plates and maintained for no longer than a week to limit subculture and ensure optimal viability of the organism. Strain purity was confirmed using the standard streak culture method. Individual bacterial strains were transferred directly from the frozen stock culture to the surface of an agar plate using a sterile swab and streaked to single pure colonies in four directions with a sterile loop.

5.2.1.3 Enumeration and isolation of bacteria

Cell suspensions were diluted as follows: a series of 10-fold dilutions were performed by diluting 100 µl of the bacterial suspension in 900 µl of a suitable broth medium. The dilution series were vortexed and 10 µl of each was plated in triplicate using a drop method onto the surface of a dry agar plate. After incubation at 37°C to allow visible colonies to appear the dilution yielding between 1 to 30 colonies was counted to calculate the number of colony forming units per ml ($cfu.ml^{-1}$). If isolation of single bacterial colonies was needed, 100 µl of each appropriate serial dilution was

Table 5.2: Bacterial Strains and Plasmids used in this study

Species and strain	Plasmid	Relevant characteristics [a]	Comments	Source / reference
Burkholderia vietnamiensis				
G4 (ATCC 53617)		Phe$^+$, PMX^R	Creation of transposon mutant bank and wild-type controls	Cardiff collection
Escherichia coli				
S17.1λpir	pUT-mini-Tn*5 Km2*	Apr; tnp* gene of TnS-IS50R inserted in SalI site of pGP704	Transfer of mini-Tn*5* to *B. vietnamiensis*	Cardiff collection
ED8654	pASE101	IncW, AP^r, GM^r, Mob$^+$, *lacZα*$^+$, Par$^+$, TP^r	Plasmid used as cloning vector for complementation	(69)
HB101	pRK2013	KM^r, Tra$^+$, Mob$^+$, ColE1 replicon	Mating helper plasmid	(75)
OmniMAX™ 2 T1 Phage-Resistant (T1^R)		F′ {*pro*AB *lacI*^{q} *lacZ*ΔM15 *Tn*10 (TET^R) Δ *(ccdA*B)} *mcr*A Δ*(mrr hsd*RMS-*mcr*BC) Φ80(*lacZ*) ΔM15 Δ(*lacZ*YA-*arg*F)U169 *end*A1 *rec*A1 *sup*E44 *thi*-1 *gyr*A96 *rel*A1 *ton*A *pan*D	Cloning strain used for transformation with pASE101 and subsequent transfer to *B. vietnamiensis* mini-Tn*5* mutants.	Invitrogen Corporation, Paisley, UK

[a] Antibiotic resistances. AP^r, ampicillin; GM^r, gentamycin; KM^r, kanamyacin; TET^r, tetracycline; TP^r, trimethoprim; PMX^r, polymyxin B; Phe$^+$ Phenol degradation phenotype.

dispensed onto the surface of a dry agar plate. A glass spreader, flamed in 100% ethanol, was used to distribute the cells evenly over the surface of the agar.

5.2.2 Mutagenesis of *Burkholderia vietnamiensis* G4 to create a mutant bank

B. vietnamiensis G4 was mated with *E. coli* S17λpir::pUTmini-Tn*5* in order to transfer the mini-Tn*5* as follows: 5 ml starter cultures were grown in SOB medium containing 50 $\mu g.ml^{-1}$ kanamycin (KM) for *E. coli* only. Cells were harvested by centrifugation and washed in SOB, before re-suspension in 5 ml SOB. 100 µl *E. coli* and 100 µl *B. vietnamiensis* G4 were mixed on a 0.2 µm filter on a dry TSA plate containing 10 mM $MgSO_4$. After incubation for 6 h at 25°C, the filter was vortexed in 1 ml TSB. Individual bacterial colonies were purified and isolated as described in Section 5.2.1.3 on TSA containing 180 $Units.ml^{-1}$ polymyxin B and 200 $\mu g.ml^{-1}$ KM, and incubated for 3-4 days at 25°C to allow growth of single transformant colonies. Polymyxin B selected against donor *E. coli* and KM selected against untransformed recipient *B. vietnamiensis* G4. The original donor and recipient were plated onto the selective media as controls. The transformant colonies were robotically picked from the selective TSA into 150 µl TSB in 96-well plates and incubated at 25°C for 72 h with shaking at 150 rpm. This bank of transconjugant *B. vietnamiensis* G4 was frozen at -80 °C as described in Section 5.2.1.2. Random single insertion of the transposon was confirmed by Southern hybridisation with the transposon sequence used as a probe.

5.2.3 Development of a screen to identify *Burkholderia vietnamiensis* G4 transposon mutants with reduced resistance to model pollutants:

5.2.3.1 Preparation of model pollutants

The compounds used as model pollutants in this study, obtained from Sigma-Aldrich (Poole, UK), are shown in Table 5.1. All stock solutions were prepared in sterile deionised water and filter sterilised using a 0.45 µm minisart filter (Sartorius AG, Goettingen, Germany) with the exception of 2,4-dichlorophenol. At the screening concentration of 2,4-dichlorophenol, the addition of 1.4% DMSO had no significant effect on bacterial growth. Stock solutions of all pollutants were prepared freshly as required and adjusted to pH 7 with NaOH or HCl.

5.2.3.2 Basal salts medium for screening

Basal salts medium (BSM) containing (per litre) $K_2HPO_4 \cdot 3H_2O$, 4.25 g; $NaH_2PO_4 \cdot H_2O$, 1.0 g; NH_4Cl, 2.0 g; $MgSO_4 \cdot 7H_2O$, 0.2 g; nitrilotriacetic acid, 0.1 g; $FeSO_4 \cdot 7H_2O$, 0.012 g; $MnSO_4 \cdot H_2O$, 0.003 g; $ZnSO_4 \cdot 7H_2O$, 0.003 g; $CoSO_4 \cdot 7H_2O$,

0.001g was prepared at 2x strength and diluted with water before steam sterilisation at 121°C for 15 min. Sterile CAS amino acids (0.05%), yeast extract (0.05%) and glucose (4 $g.l^{-1}$) were added to provide growth nutrients (96). Pollutants (Table 5.1) were prepared as described in Section 5.2.1 before addition to BSM.

5.2.3.3 Determination of screening concentration of model pollutants for *B. vietnamiensis* G4

To develop a screen to identify pollutant-susceptible mutants from the entire mutant bank it was necessary to determine the pollutant concentration at which wild-type growth was inhibited (IC). This allowed the estimation of a screening concentration at which the wild-type would grow successfully but mutants showing attenuated growth in the pollutants would be identified.

All screening was performed in 96-well, 200 µl polycarbonate micro-titre plates (Fisher Scientific Ltd., Loughborough, UK). For consistent initial IC determination, a fresh *B. vietnamiensis* G4 overnight broth culture was adjusted to an optical density (OD) of 1 at 620 nm. Pollutant broth BSM medium was inoculated from a dilution series at a final ratio of 1:1000 before addition of each pollutant. To confirm extensive growth of 94 individual colonies at the selected screening concentration, single colonies of *B. vietnamiensis* G4 wild type were isolated (Section 5.2.2.4), picked into 96-well format and then frozen with 8% DMSO at -80°C. Prior to test plate inoculation with a 96-pin replicator the 94 colony grid was revived in TSB overnight at 37°C. As DMSO was required as a solvent for 2,4-dichlorophenol, a BSM control containing DMSO in the corresponding concentrations was included to confirm that it caused no significant effect on the *B. vietnamiensis* G4.

Each IC test included: (*i*) un-inoculated control wells to confirm sterility of the test medium and provide an OD blank for spectrophotometer calibration, and (*ii*) inoculated wells without selection to measure the uninhibited growth of the organisms and provide a reference maximum OD. Plates were sealed with gas-permeable adhesive seals (ABgene, Epsom, UK) and incubated at 37°C with shaking at 150 rpm. For each pollutant, the groups of mean optical density were compared with one-way analysis of variance (ANOVA) and the minimum significant difference (MSD) calculated for individual comparisons between means.

5.2.4 Identification of pollutant susceptible mutants.

5.2.4.1 Screen of mutant bank against model pollutants

A total of 1920 STM strain G4 mutants were separately inoculated into BSM medium containing pollutants at concentrations indicated in Table 5.1 and screened for growth. The STM bank was revived overnight in TSB containing 50 $\mu g.ml^{-1}$ KM to select for the transposon and used to inoculate twenty 96 well plates for each pollutant. These were sealed with gas-permeable adhesive seals and incubated at 37°C, shaking at 150 rpm. Following this screen 184 mutants with attenuated resistance were subsequently subjected to a duplicate pollutant screen as described above, and a control screen using BSM containing no pollutant to confirm that all putative mutants could grow with glucose as a sole carbon source. The OD of each well was recorded at 630 nm after 24 h with a Dynex MRX Revelation 96-well plate absorbance reader (Dynex Labsystems, Middlesex, UK)

5.2.4.2 Extraction of OD data from Revelation output

A software pipeline was developed to identify the reduced OD readings within the raw data files produced by the plate reader as follows: Wells with an optical density of below a threshold 0.7 were identified from each list of 96 and re-labelled with the appropriate well position using a short Perl script. Data were then imported into Microsoft Excel as a comma delimited input file. Finally, a Microsoft Access database was constructed to record the results. The 184 mutants selected for further analysis were revived from freezer stocks onto TSA and re-frozen in 8% DMSO in both 96 well format and as individual 2 ml stocks.

5.2.4.3 Sequence analysis of pollutant susceptible mutants

The sites of mini-Tn*5* transposon insertion were determined by direct PCR sequence analysis. DNA was obtained from the mutants of interest using the Chelex extraction method as follows: cells from 1 ml of a liquid culture were harvested by centrifugation (5 min, 16,000 x g) and resuspended in 100 µl of a sterile solution containing 5% chelex 100 (Sigma Aldrich, Gillingham, UK). The suspension was boiled for 5 min, immediately placed on ice for a further 5 min, and then the procedure was repeated. The final supernatant was recovered after centrifugation and stored at -20°C. The resulting DNA extracts, which included transposon-chromosome junctions, were diluted with sterile polished water prior to use to allow approximately 20 ng of template DNA to be incorporated into each PCR.

A two stage PCR was adapted from a previously described method (148) to amplify the transposon-chromosome junctions. The reactions were prepared essentially as described, but utilised two modified primers 1 & 3, which were re-designed (109) to

target the primers to mini-Tn*5Km2*. The method was also altered to include all three second primers (2a, 2b and 2c) *simultaneously*, in a multiplex PCR reaction, to increase the chance of amplifying the target sequence. The primer sequences are listed in Table 5.3, and a schematic of the primer binding and PCR program is shown in Figs. 5.1 and 5.2. The amplification products were purified with the QIAquick PCR purification kit according to the manufacturer's (Qiagen Ltd., Crawley, UK) instructions. Sequencing reactions were prepared using 1.6 pmol of Primer 3 with Applied Biosystems Big Dye Terminator ready reaction mix version 3.1 and analysed using an ABI-PRISM 3100 Genetic Analyser capillary electrophoresis system running Performance Optimised Polymer 6 (POP-6) in accordance with the manufacturer's (Applied Biosystems, Foster City, CA., USA) instructions. The transposon sequences were removed to leave only the chromosome at the point of insertion.

5.2.4.4 Bioinformatic analysis of genes on G4 genome

The sequences derived from the transposon inserts were used to localise and mark sites of mutations onto the *B. vietnamiensis* G4 genome within the sequence viewing software Artemis (196). Fig. 5.3 shows a screen shot from Artemis genome viewer. This software allowed all open reading frames (ORFs) surrounding each insertion to be identified and compared to the strain G4 genome annotation with the Basic Local Alignment Search Tool (BLAST; www.jgi.doe.gov). In addition, the ORFs were searched against GenBank (www.ncbi.nlm.nih.gov) to measure similarity to genes from other bacterial sequences.

5.2.4.5 Determination of the growth rate of mutants

A selection of 18 different mutants of interest, the mutated genes of which had been successfully identified, was screened in quadruplicate in TSB. These mutants were tested simultaneously in a 96-well plate in addition to one sterile blank and one strain G4 growth control per plate. The OD at 630 nm of each well was read after 10 s shaking every 10 min for 24 h at 37°C. The mean culture doubling time was calculated during the exponential phase of each mutant in order to determine whether there was a difference between the growth rates of the individual mutants by comparison with the wild type. Six randomly selected mutants from the original bank of 1920 that had not been identified as susceptible to any of the model pollutants were also tested as controls to determine whether the presence of the inserted transposon alone caused any significant effect on the growth rate. Both groups of mean growth rates (random and test groups) were compared using ANOVA and a Fisher's *a priori* test was performed to compare individual means.

Table 5.3: Oligonucleotide primers used in this study

Primer	Primer sequence 5' – 3' [b]	Specificity / Restriction site tail	Reference or source
Sequencing the transposon – chromosome junction			
1	TTTTTACACTGATGAATGTTCCG	Transposon sequence	(109)
2a	GGCCACGCGTCGACTAGTACNNNNNNNNNNAGAG	Targets multiple sites in the genome and introduces a unique sequence at 5' end	(148)
2b	GGCCACGCGTCGACTAGTACNNNNNNNNNNACGCC		
2c	GGCCACGCGTCGACTAGTACNNNNNNNNNNGATAT		
3	CGGATTACAGCCGGATCCCCG	Transposon sequence	(109)
4	GGCCACGCGTCGACTAGTAC	Hybridises to unique sequence of primer 2	(148)
Amplifying fragments for gene complementation			
*typA*_for	TTATTAGGATCCTAGCGGATTGCGAATCAAGT	*Bam*H I	This study
*typA*_rev	TATTATGGATCCTGTTAATCGAACAGGCAACG	*Bam*H I	This study
*grtA*_for	TTATTAGGATCCCCGGCTTGATGTTATGAAGG	*Bam*H I	This study
*grtA*_rev	TATTATGGATCCGGTGCCGTTCACGTTCTC	*Bam*H I	This study
*enoL*_for	TTATTAGGATCCCGTCTGTAGCGACAGCAGGT	*Bam*H I	This study
*enoL*_rev	TATTATGGATCCTCTGCAGATCCTGCACCTC	*Bam*H I	This study
*strE*_for	TTATTAGGATCCTACTCGGATCTGCTCGTCGT	*Bam*H I	This study
*strE*_rev	TATTATGGATCCCTTTGGCATAAAAGCCGTGA	*Bam*H I	This study
*traR*_for	TTATTAGGATCCGCTCGTTTGCAGACATTTCA	*Bam*H I	This study
*traR*_rev	TATTATGGATCCGAGCACCGCTCTAGCTTCC	*Bam*H I	This study
*vacJ*_for	TTATTAGGATCCAGTACATCGGCCTCGATCC	*Bam*H I	This study
*vacJ*_rev	TATTATGGATCCACGGGGATCAGGAACAGTTT	*Bam*H I	This study
*tolU*_for	TTATTAGGATCCGCTGCCGAAGTACGACCTG	*Bam*H I	This study
*tolU*_rev	TATTATGGATCCGATCGGACAGCGTCAATGT	*Bam*H I	This study
*lysM*_for	TTATTAGGATCCGCACACGAATACACGCTGTC	*Bam*H I	This study
*lysM*_rev	TATTATGGATCCCGAAACACGTACGACACGAA	*Bam*H I	This study
*hspM*_for	TTATTAGGATCCCGACGTGGTTCTTCTTCCAC	*Bam*H I	This study
*hspM*_rev	TATTATGGATCCTTCGCCATATCGATTCACTG	*Bam*H I	This study
*mtrA*_for	TTATTAAAGCTTCGTATTTTTCTCGGCAACG	*Hind* III	This study
*mtrA*_rev	TATTATAAGCTTGTGATCGGCTGCATCTTCA	*Hind* III	This study
*forM*_for	TTATTAAAGCTTTGATCGACTTCATCCACAGC	*Hind* III	This study
*forM*_rev	TATTATAAGCTTACGAGTGCTCCGACAGGT	*Hind* III	This study

[b] Restriction sites underlined. AT tag at 5' end to facilitate restriction site recognition.

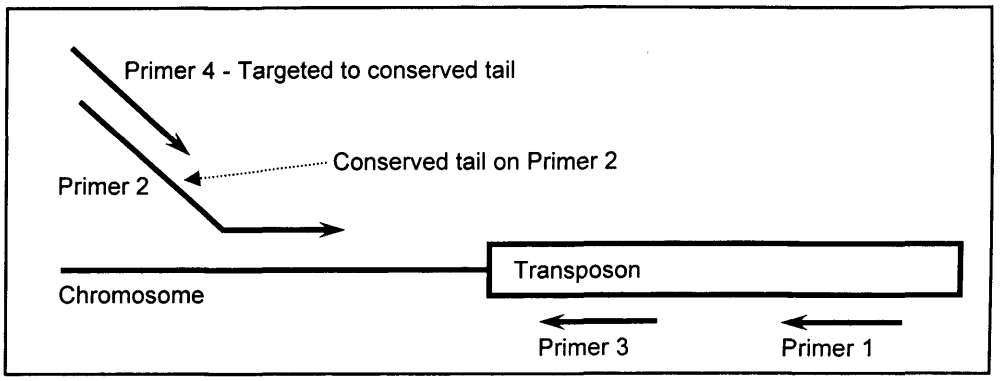


Figure 5.1: A schematic diagram illustrating PCR amplification of G4 chromosomal DNA adjacent to the mini-Tn*5Km2* insertion by a 2-step PCR protocol (modified from 148)

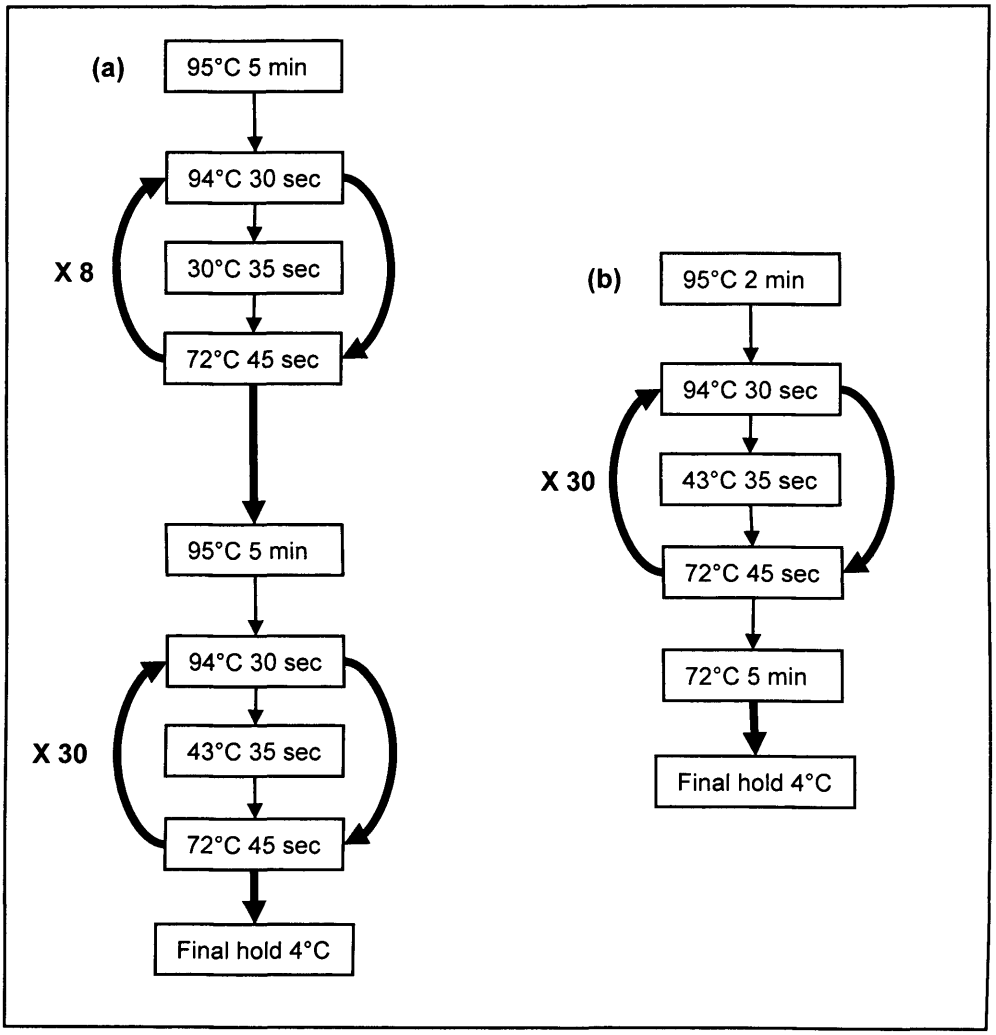


Figure 5.2: 2-step PCR thermal cycling conditions used to amplify DNA adjacent to the mini-Tn*5Km2* insertion (modified from 148): (a) 1st PCR, (b) 2nd PCR

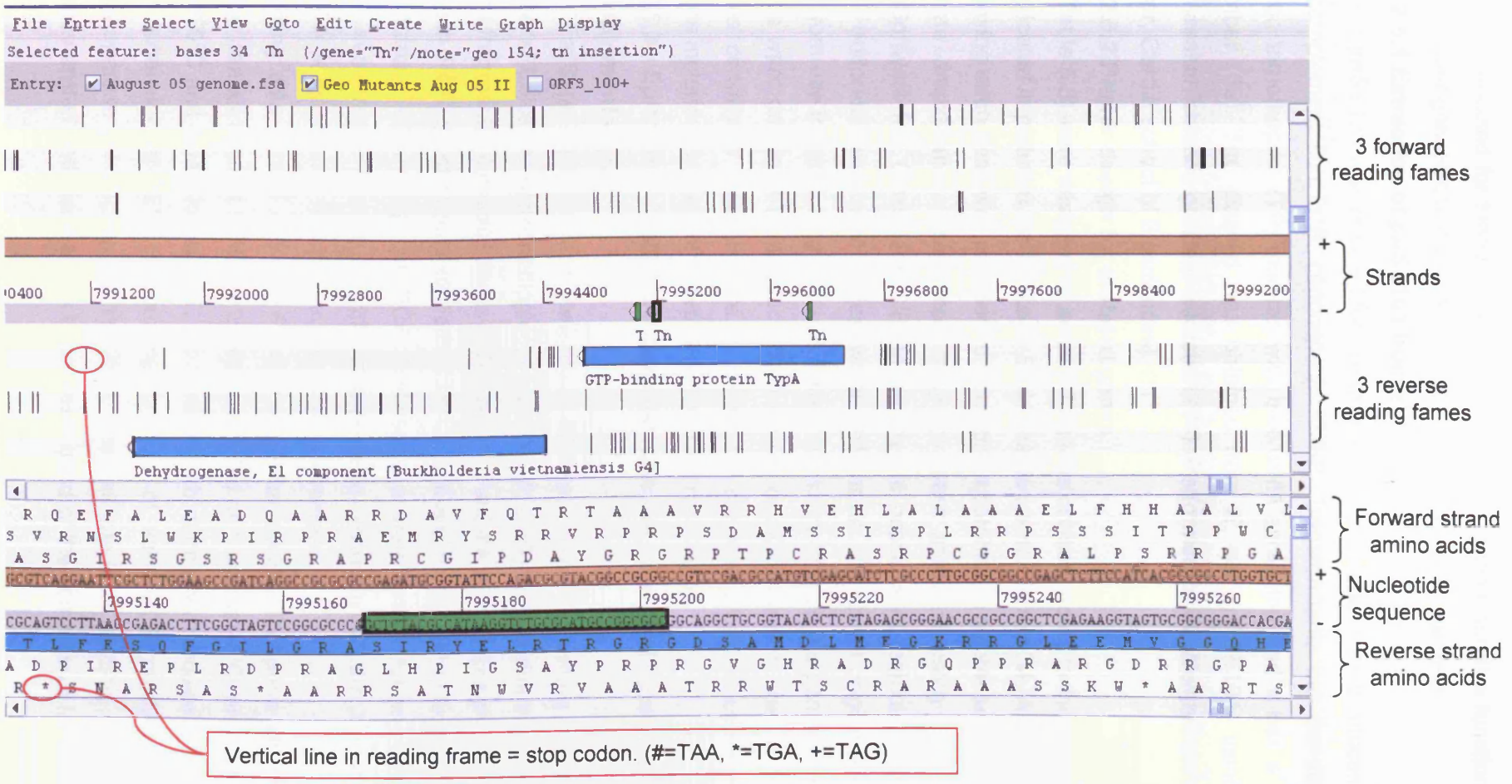


Figure 5.3: The identification of open reading frames in the *B. vietnamiensis* G4 genome sequence using Artemis software (196). Screen-shot of the Artemis sequence viewer window indicating two putative coding sequences, *typA*107 and an E1 dehydrogenase component downstream. The short green sequences represent the points of transposon insertion. Also marked are the nucelotide and amino acid sequences, and stop codons.

5.2.5 Complementation of mutants

Mutants were selected for further analysis on the basis of their putative function as identified by their genetic comparison to the *B. vietnamiensis* G4 genome.

5.2.5.1 Extraction of pASE101 from *E. coli*

The plasmid pASE101 was selected for use as a vector on the basis of its successful use for complementation in *B. cenocepacia* K56-2. Transformants can be stably selected on trimethoprim (TP) medium with blue/white screening and it has a low copy number (69). The plasmid was obtained from *E. coli* (pASE101) using a modified caesium chloride gradient extraction (94) (Performed by A. Weightman and G. Webster, Cardiff school of Biosciences).

5.2.5.2 Primer design for mutated genes

Primers (Table 5.3) were designed to amplify 11 genes corresponding to the mutated genes implicated in pollutant resistance in *Burkholderia vietnamiensis* G4. A region of 100 bp flanking each gene was included to incorporate potential promoter sequences and facilitate amplification. Sequences of genes were analysed using DSGene (Accelrys Software Inc. Cambridge, UK) to identify the positions of restriction sites. Restriction endonuclease cutting sites which fell inside the target region were precluded from use in the cloning steps as it was necessary to keep the genes intact. The primers were designed to introduce *Bam*HI tails, or *Hin*dIII tails in the case of those genes containing a *Bam*HI site. A six bp tail containing only adenine and thiamine was included to assist effective endonuclease restriction.

5.2.5.3 Epicentre "FailSafe" amplification of mutated genes for complementation

Epicentre "FailSafe PCR premix selection kit" was used according to the manufacturer's (Epicentre Biotechnologies, Madison, WI, USA) instructions to amplify the selected fragments at optimal levels of $MgCl_2$ and FailSafe™ PCR Enhancer with Betaine. Each PCR reaction was replicated in 12 FailSafe™ PCR "PreMixes" that covered a matrix of enzyme-specific PCR conditions that were optimal for amplifying problematic sequences. The nucleotide incorporation accuracy of the FailSafe Enzyme Mix is at least thrice higher than *Taq* DNA polymerase

Thermal cycling was carried out in a Flexigene thermal cycler (Techgene Ltd) as follows: Initial denaturation at 94°C for 2 min, 30 cycles of 1 min at 94°C, annealing for 1 min at 50°C and extension at 72°C for 1 min per kb of fragment required with a final 5 min extension at 72°C. Approximately 2 µl of each PCR product was visualised by agarose gel electrophoresis (200). The FailSafe PCR reaction yielding the most concentrated and discreet electrophoresis band of the correct size for each

gene was purified with QIAquick PCR purification kit (Qiagen Ltd., Crawley, UK) into 50 µl of sterile polished water.

5.2.5.4 Ligation of amplified fragments into pASE101

Restriction endonuclease digests, using either *Bam*HI or *Hin*dIII as appropriate, were performed according to the manufacturer's (Promega Corporation Inc, Southampton) instructions in a 50 µl reaction for the vector plasmid pASE101, and a 10 µl reaction on the fragments from the FailSafe PCR reaction. The digests were incubated at 37°C overnight and heat inactivated at 65°C for 2 min. The DNA fragments were ligated using a 10 x ligation kit (Promega Corporation Inc.). A 20 µl reaction was prepared containing 2 µl 10 x ligation buffer (10% v/v), 1 µl T4 DNA-ligase, 4 µl vector DNA and 8 µl of substrate DNA. The ligation mixture was incubated overnight at 16°C unless otherwise stated. A vector control was constructed in a control reaction where the substrate PCR was replaced with H_2O. The complementing constructs based on the vector pASE101 are described in Table 5.4.

5.2.5.5 Transformation of One Shot OmniMAX™ 2 T1 Phage-Resistant *E. coli*

OmniMAX™ 2 $T1^R$ *E. coli* strain cells (Invitrogen Corporation, Paisley, UK) were thawed on ice and 2.5 µl of ligation mix was gently mixed with each 50 µl aliquot of cells. The cells were incubated for 30 min on ice then heat-shocked at 42°C for 30 sec followed by a further 2 min on ice. 250 µl of SOC medium (Invitrogen) was added to the transformation mixture and incubated with shaking (225 rpm) at 37°C for 1 h. The reaction volume was increased to 1 ml by the addition of 700 µl of SOC medium and 100 µl aliquots were spread onto LB agar plates containing TP 100 $\mu g.ml^{-1}$ to select for the plasmid and XTRA-Blue Plus X-Gal/IPTG Solution (MP Biomedicals, London, UK) at 2.8 $\mu l.ml^{-1}$ to enable the growth of individual clones to produce visible colonies.

5.2.5.6 Confirmation of insert size by restriction endonuclease digests

The pASE101 plasmid containing each insert was extracted using an alkaline-sodium dodecyl sulphate lysis method (200) and dissolved in 30 µl of TE containing 0.5 $\mu g.ml^{-1}$ of RNase (Sigma-Aldrich Ltd., Poole, UK). The insert was excised from the vector by digestion with the appropriate restriction endonuclease (Section 5.2.5.4) and visualised on a 1.8% concentration agarose gel in 1 x TBE. Plasmids not shown to contain the fragment were discarded as unsuitable for restoration of phenotype complementation experiments.

Table 5.4: Complementing constructs used in this study

Complementing gene	Size & restriction site of complementing insert into pASE101	Characteristics of interrupted gene	Reference
*typA*107	2363 bp; *Bam*HI	GTP-binding protein TypA. The *typA* gene in *E. coli* is sometimes tyrosine phosphorylated and thought to be a virulence regulator which has homologues in several pathogens including *Bordetella pertussis* and *Pseudomonas aeruginosa*.	(208)
*gtrA*90	1758 bp; *Bam*HI	Family 2 glycosyltransferase. The biosynthesis of disaccharides, oligosaccharides and polysaccharides involves the action of hundreds of different glycosyltransferases, which catalyse the transfer of sugar moieties to form glycosidic bonds. Family 2 glycosyltransferases are classified as inverting enzymes according to the stereochemistries of their reaction substrates and products.	(215)
*enoL*91	1649 bp; *Bam*HI	Enolase is a glycolytic enzyme that is classically cytoplasmic and catalyses the formation of phosphoenol-pyruvate. This multifunctional protein has recently been identified on the surface of streptococci, invoking interest in its potential role in pathogenesis.	(78, 224)
*strE*80	1527 bp; *Bam*HI	Homology to *E. coli* ribosomal protein L25, which binds specifically to 5S rRNA and also the general stress protein Ctc which has been well studied in *Bacillus subtilis*.	(206, 221)
*traR*146	711 bp; *Bam*HI	Closest homology to a transcriptional regulator in the cyanobacteria species *Nostoc* strain PCC 7120.	(110)
*vacJ*71	1416 bp; *Bam*HI	VacJ like lipoprotein. VacJ itself is a surface lipoprotein required for intercellular spreading in the gram negative pathogen *Shigella flexneri*, however, generally lipoproteins function at the bacterial outer membrane-aqueous interface as structural proteins, receptors or transporters.	(95, 225)
*tolU*167	1317 bp; *Bam*HI	Toluene-tolerance. Homology to part of an ABC transport system or periplasmic binding protein-dependent transport system. Typically, genes that encode the three components are arranged in an operon.	(100)
*lysM*67	2269 bp; *Bam*HI	Predicated peptidoglycan-binding LysM domain, found in a variety of enzymes involved in bacterial cell wall degradation.	
*hspM*49	2552 bp; *Bam*HI	ATP dependent 70-kDa heat shock protein Hsp70. Hsp70 chaperones help to fold proteins.	(155)
*mtrA*38	1434 bp; *Hin*dIII	Putative methyltransferase gene, potentially coding for an enzyme that catalyzes the transfer of methyl groups. It was not possible to determine the specific methyltransferase type because the coding sequence is predicted in the annotation of G4.	
*forM*31	3037 bp; *Hin*dIII	Homology to the alpha subunit of formate dehydrogenase, from the prokaryotic molybdopterin-containing oxidoreductase family. It allows the use of formate as a major electron donor during anaerobic respiration and the alpha subunit possibly forms the active site.	(224)

[a] Compared to yield of strain G4 growing on medium containing 2,4-dichlorophenol

5.2.5.7 Tri-parental mating to transfer construct to the host by conjugation

The three types of strain to be used in the mating were revived from -80°C freezer stocks on the following media: (*i*) Donor *E. coli* (pASE101) and gene of interest (also control plasmid with no gene insert) on LB agar containing TP 100 $\mu g.ml^{-1}$, (*ii*) "Helper plasmid" *E. coli* (pRK2013) on LB agar containing KM 50 $\mu g.ml^{-1}$ and (*iii*) Recipient *B. vietnamiensis* G4 transposon mutant on TSA containing KM 50 $\mu g.ml^{-1}$. The strains were grown in 3 ml broth cultures with the appropriate selection overnight at 37°C and then harvested by centrifugation for 10 min at 1,200 x g. They were washed once in TSB containing 10 mM $MgSO_4$, centrifuged again and resuspended in 2 ml TSB containing 10 mM $MgSO_4$. At this stage 10 µl of all three parental strains were plated, in triplicate, on TSA containing TP 100 $\mu g.ml^{-1}$ and PMX 300 $Units.ml^{-1}$ to determine the viable count of each strain on the "final selection" medium. 50 µl aliquots of each strain were combined and 100 µl of this mixture was applied to a sterile nitrocellulose filter on a dry TSA plate containing 10 mM $MgSO_4$. After incubation at 37°C for 4 h the bacteria were resuspended from the filter by vortex-mixing in 1 ml TSB containing 10mM $MgSO_4$. Serial dilutions of were plated on TSA containing TP 100 $\mu g.ml^{-1}$ and PMX 300 $Units.ml^{-1}$ after overnight growth at 37°C as described in Section 5.2.2.4. Transformant colonies were transferred with sterile toothpicks into TSB containing KM 50 $\mu g.ml^{-1}$ and TP 50 $\mu g.ml^{-1}$, and stored at -80°C as described in Section 5.2.2.2.

5.2.6 Characterisation of mutants and complemented mutants

The phenotypic characteristics of the individual mutants were compared with the corresponding mutants containing corresponding complementing recombinant pASE101 derivatives ("complement"). For each test, a mutant containing the plasmid vector (pASE101) alone ("vector control") and the un-attenuated *B. vietnamiensis* G4 wild type ("wild type") were included. In each test, means were compared by ANOVA and used to calculate MSDs to determine the significance of individual results.

5.2.6.1 Determination of inhibitory concentration (IC) of 2,4-dichlorophenol for *B. vietnamiensis* G4 in Tryptone Soya Broth (TSB)

As the mutant that displayed the most acute and easily distinguishable effect, 2,4-dichlorophenol was selected as the model pollutant against which to characterise the mutants. The inhibitory concentration was determined in a 96-well microtitre plate, essentially as described in Section 5.2.3.3, with concentrations of 2,4-dichlorophenol ranging between 0 mM - 2.5 mM in 0.25 mM increments. The inoculum was adjusted

to an OD of 1 at 620 nm and then diluted serially to allow the pollutant medium to be inoculated at a dilution of 10^{-3}. Eight replicates of each test were performed.

5.2.6.2 Screening to determine the effects of complementation

To examine the success of the complementation experiments the optical densities of each mutant was compared to that of the corresponding complemented mutant, its vector control and the *B. vietnamiensis* wild type at 0.75 mM and 1.0 mM 2,4-dichlorophenol concentrations in TSB. There were eight replicates of each strain in each 96-well plate, and five plates corresponding to the five individual mutants that had been transformed with complementing constructs. Each plate also included eight wells containing the wild type growing in TSB and eight blank TSB wells. The plates were incubated with shaking at 37°C and the OD read after 48 h at 630 nm.

5.2.6.3 Screening to determine the final susceptibility of the identified mutants

All mutants for which the interrupted gene had been identified and located on the *B. vietnamiensis* genome were screened as described in Section 5.2.4.1 in TSB containing 1 mM 2,4-dichlorophenol. A total of 52 mutants were screened in 8x replicates alongside wild type controls and media blanks in 96-well plates. The plates were incubated with shaking at 37°C and the OD read after 48 h to estimate susceptibility to 2,4-dichlorophenol.

5.2.6.4 Suspension test to determine 2,4-dichlorophenol lethal effects

To distinguish between bacterial *killing* and bacterial *inhibition* the lethality of 2,4-dichlorophenol was determined by cell suspension tests. The method used was adapted from one that had previously determined the lethal effect of the biocide Triclosan, a halogenated phenol, and considered suitable for use with 2,4-dichlorophenol (123). Bacterial cell cultures of *B. vietnamiensis* G4 wild-type, derivative transposon mutants and complemented mutants were prepared as described in Section 5.2.2.1 and washed twice by centrifugation at 1,200 x g for 10 min to remove antibiotics. Suspensions were prepared in TSB and adjusted to an OD of 1 at 620 nm. A 100 µl volume of the washed cell suspension (OD 1 at 620 nm) was transferred to 900 µl of TSB containing 2mM 2,4-dichlorophenol. The suspension was incubated at 37°C for 5 min after which 100 µl of the suspension was expelled into 900 µl of neutraliser (0.75% w/v azolectin in 5% v/v Tween 80, filter sterile) terminating any further activity of the pollutant. Bacteria were left in contact with the neutraliser for 5 min before enumeration of viable cells on TSA as described (Section 5.2.1.3). All plates were incubated at 37°C for up to 48 h to allow growth of viable cells to produce visible colonies. Controls were included as follows: (*i*) 100 µl of the washed cell suspension was added to 900 µl of sterile deionised water and

incubated at 37 °C for 5 min prior to enumeration of viable cell number on TSA. (*ii*) To ensure that the neutraliser being used was effective in terminating the activity of pollutant, 100 µl of 2mM 2,4-dichlorophenol was added to 800 µl of neutraliser. Immediately 100 µl of washed cells was added and incubated at 37°C for 5 min prior to the enumeration of viable cells on TSA. The number of viable cells was compared to the sterile deionised water control. (*iii*) To ensure the neutraliser was not toxic to the organisms subjected to testing, 100 µl of the washed bacterial suspension was added to 900 µl of neutraliser and incubated at 37°C for 5 min prior to the enumeration of viable cells. The number of viable cells was compared to a control in which 900 µl of sterile deionised water replaced the neutraliser.

5.3 RESULTS:

The experimental approach in this study was to screen 1920 mutants initially and then focus finally on complementing and characterising five mutants (Fig. 5.4).

5.3.1 Mutagenesis of *Burkholderia vietnamiensis* G4 to produce a mutant bank

The transposon was shown by (L. A. O'Sullivan, unpublished results) to insert into the strain G4 genome to produce single, random mutations by Southern hybridisation.

5.3.2 Inhibitory concentration (IC) of "model pollutants" for initial screen of mutant bank

The IC was recorded (Table 5.1, Figs. 5.5-5.8) as the concentration at which the mean growth was significantly below the growth of strain G4 in the absence of pollutant as determined by ANOVA at 95% and calculation of the MSD. The pollutant screening concentrations were qualitatively selected as the *last* increments of concentration at which every replicate of *B. vietnamiensis* G4 grew (Figs. 5.5-5.8). A BSM control containing DMSO across a range of concentrations corresponding to those of its solute (2,4-dichlorophenol) revealed no significant effect on the growth of *B. vietnamiensis* G4 at either the IC or the screening concentration. DMSO significantly reduced the growth of G4 above 10% when 5 mM 2,4-dichlorophenol was used (Fig. 5.7). As an additional test 94 individually isolated *B. vietnamiensis* G4 colonies were grown in the presence of each pollutant at its selected screening concentration. All of the wells showed strong growth, confirming that the concentrations were suitable for screening the STM bank.

5.3.3 Screening of the transposon mutant bank against model pollutants

Of 1920 mutants screened for susceptibility to the four pollutants (Table 5.1), 184 mutants failed to grow beyond an OD of 0.7 at 630 nm after 24 h and qualified for further analysis. The susceptibilities were confirmed in repeat screens on these 184 individuals, additionally, the glucose control screen (Section 5.2.4.1) revealed 35 of the 184 mutants that were unable to grow on glucose as a sole carbon source (Fig. 5.4). 59 sequences of a suitable quality for analysis were obtained and located on the *B. vietnamiensis* G4 genome (www.jgi.doe.gov). Of the 52 open reading frames (Table 5.5) into which the transposon inserted all but seven were susceptible to 2,4-dichlorophenol; these were amongst 18 others identified by the phenol screen. There were 15 2,4-D mutants; all of these were susceptible to both 2,4-dichlorophenol and Dichloroprop, however, three Dichloroprop mutants were found that were not

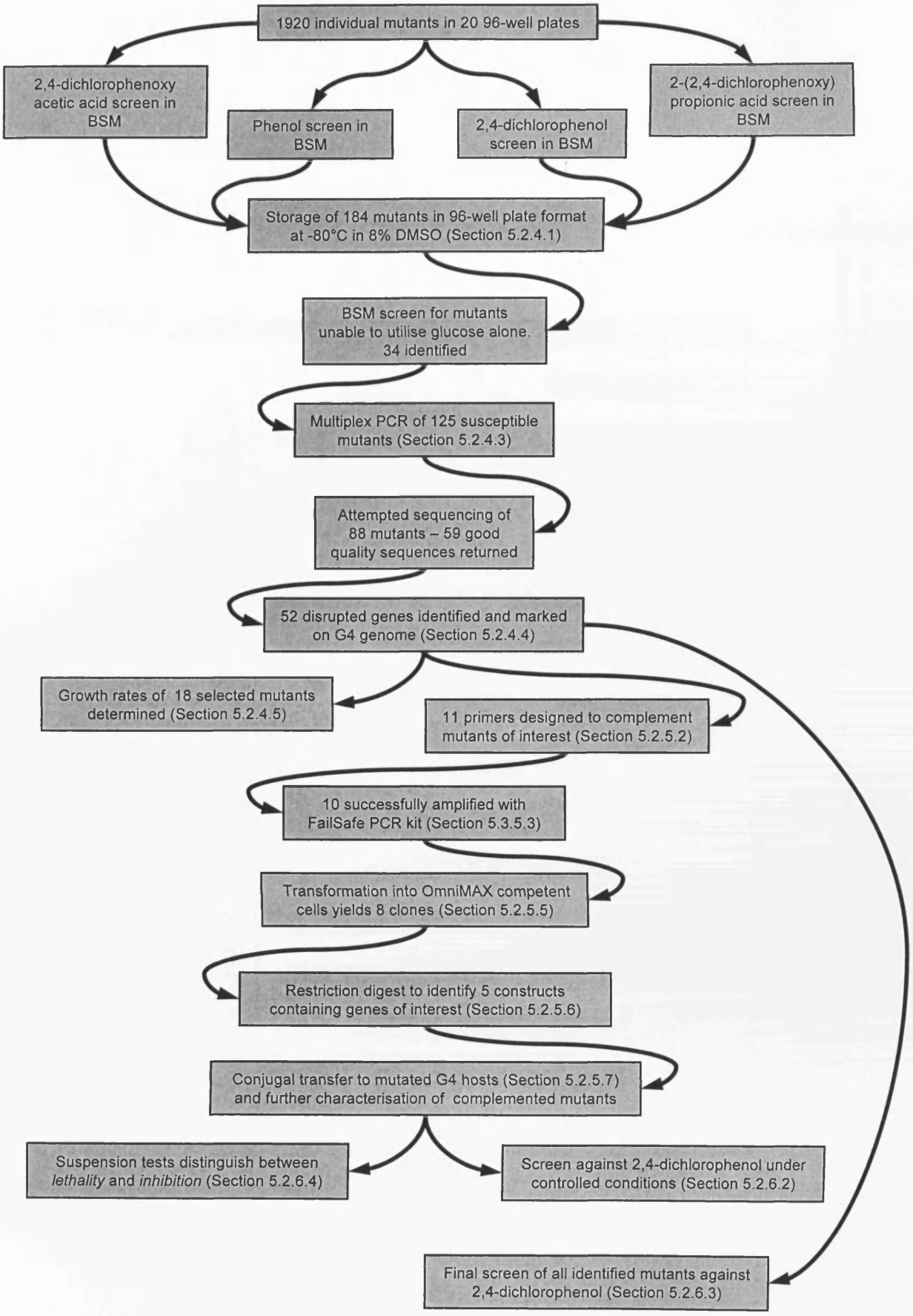


Figure 5.4: Flowchart showing experimental progression. 1920 original mutants were screened; five mutants of interest were finally focussed upon.

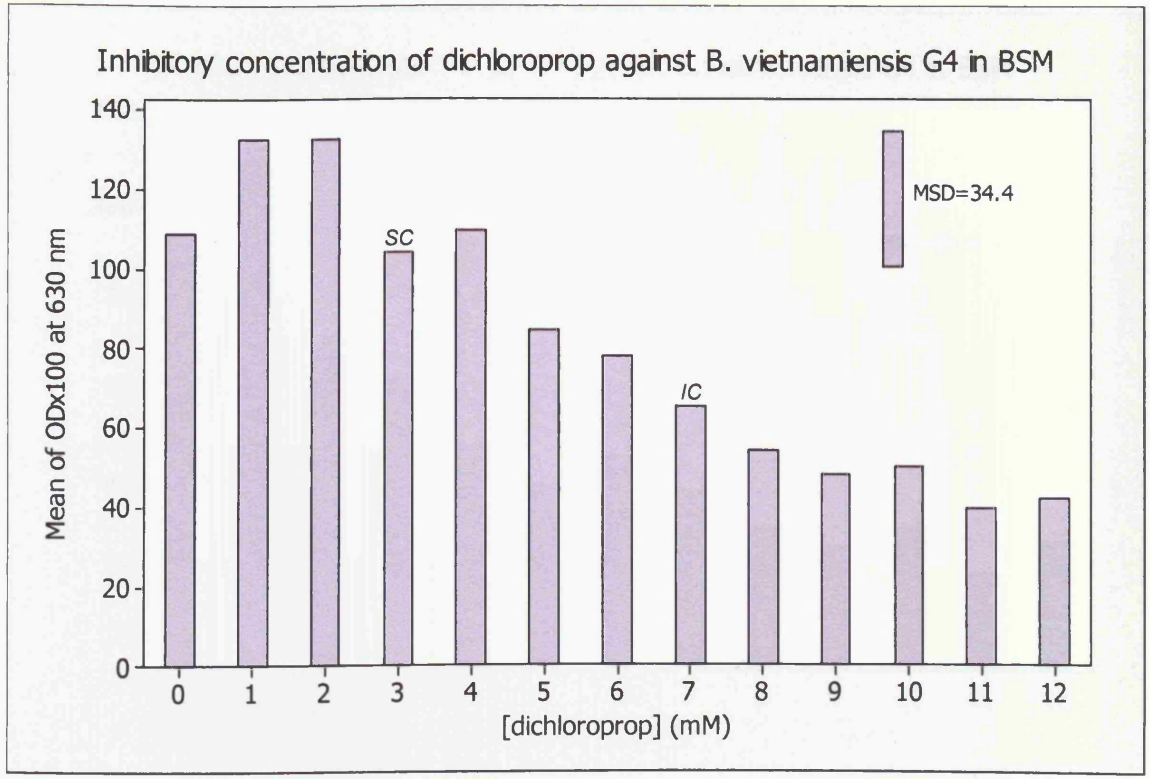


Figure 5.5: Mean optical density of *Burkholderia vietnamiensis* G4 at 24 h in 2-(2,4-dichlorophenoxy) propanoic acid (Dichloroprop). The estimated inhibitory concentration *(IC)* was 7.0 mM and the screening concentration *(SC)* was set at 3.0 mM.

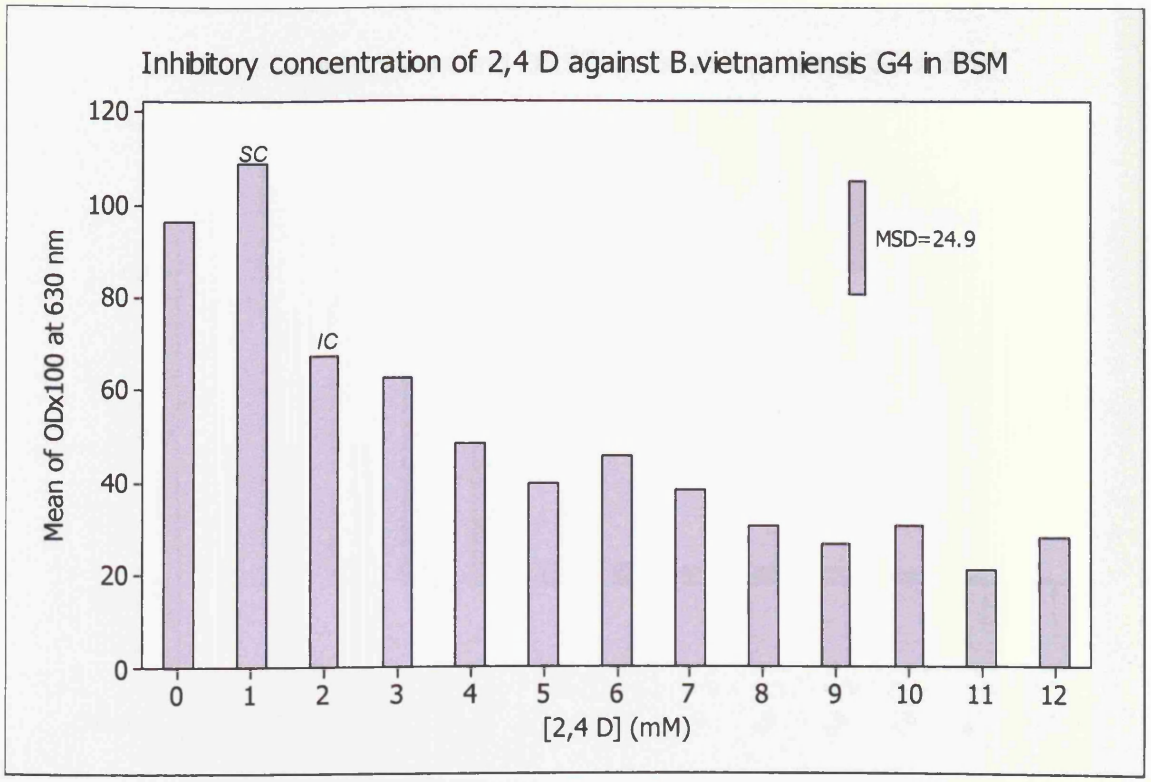


Figure 5.6: Mean optical density of *Burkholderia vietnamiensis* G4 at 24 h in (2,4-dichlorophenoxy) acetic acid (2,4-D). The inhibitory concentration was 2.0 mM, therefore the screening concentration was set at 1.0 mM.

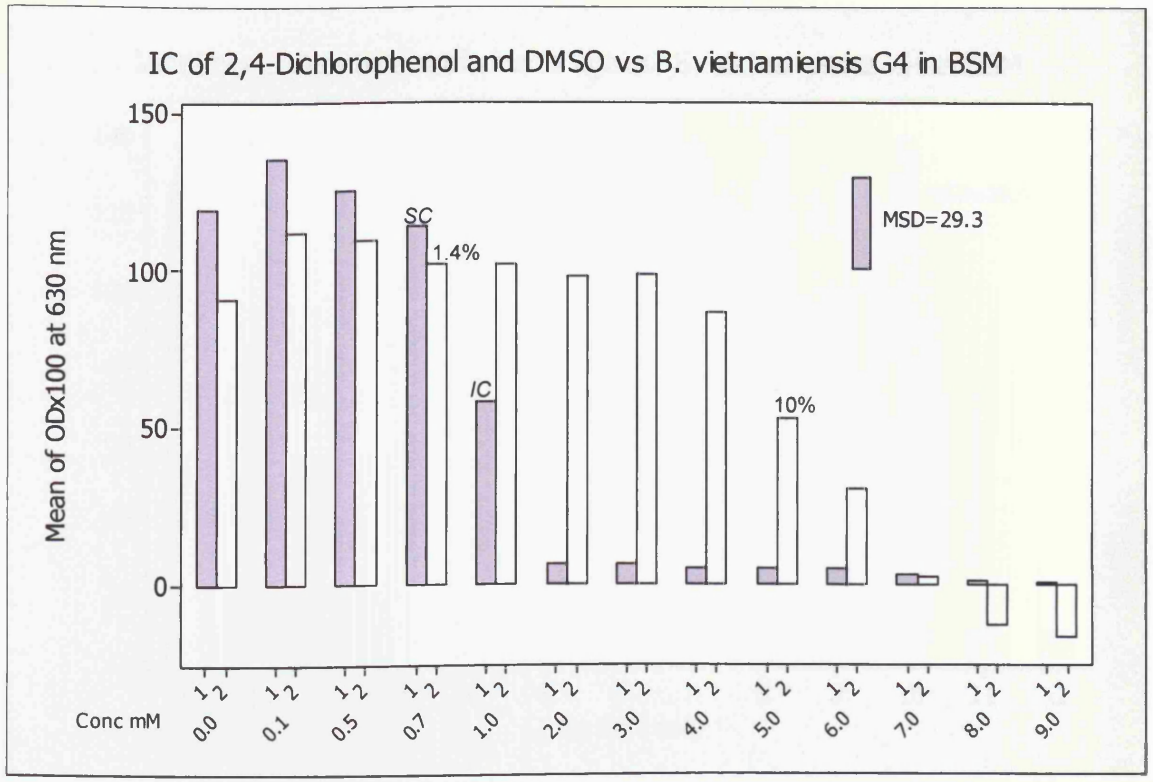


Figure 5.7: Mean optical density of *Burkholderia vietnamiensis* G4 at 24 h in *(1)* 2,4-dichlorophenol and *(2)* DMSO solvent. The estimated inhibitory concentration *(IC)* of 2,4-dichlorophenol was 1.0 mM, therefore the screening concentration was set at 0.7 mM. DMSO concentration is shown at the level required for screening, and the level at which it had a significant effect on growth.

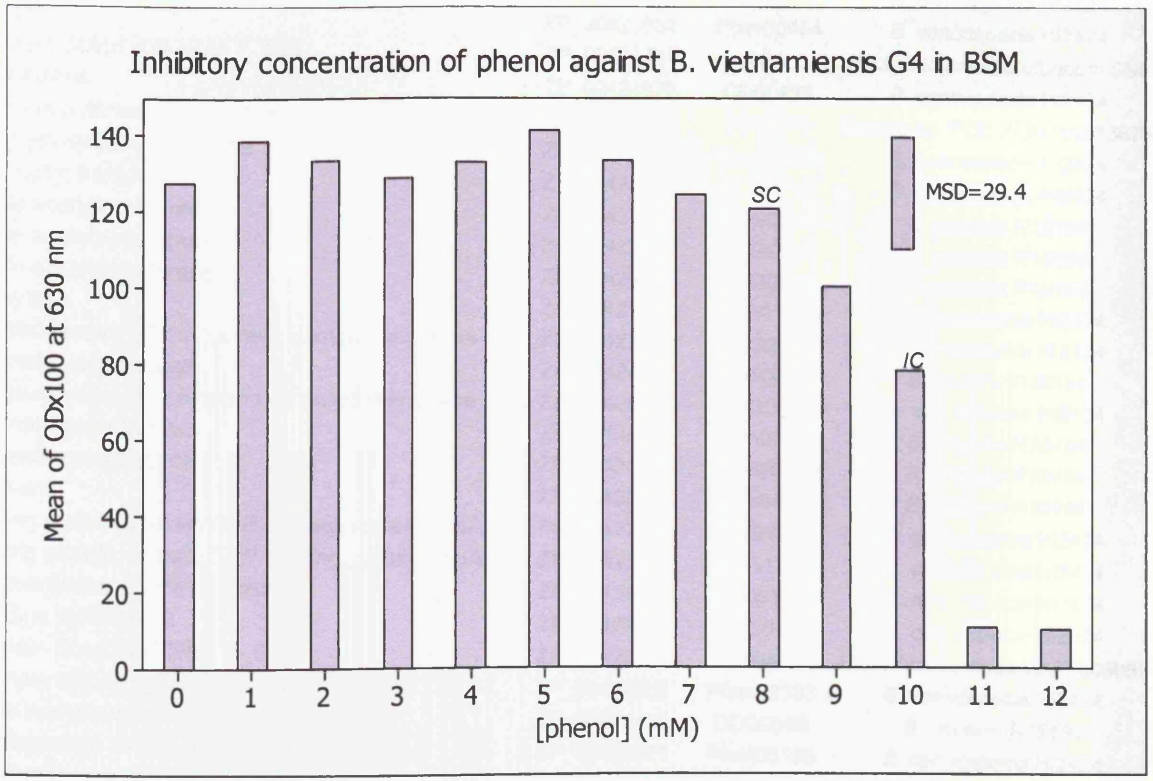


Figure 5.8: Mean optical density of *Burkholderia vietnamiensis* G4 at 24 h in phenol. The estimated inhibitory concentration *(IC)* was 10.0 mM and the screening concentration *(SC)* was set at 8.0 mM.

Table 5.5: Summary of STM insertions identified on the G4 genome ordered by % G4 WT growth [b]

Mutant ID [a]	Protein Function	G4 Accession	Protein ID	Closest match with homology	Homolog Accession	% of G4 OD [b]
7	2-oxo-acid dehydrogenase E1 component homodimeric type	ZP_00423460	COG2609	*B. cenocepacia* HI2424	ZP_00464635	26.17
42	Tryptophan synthase, alpha chain	ZP_00425264	Pfam00290	*B. cenocepacia* HI2424	ZP_00464301	26.17
*enoL*91	Enolase	ZP_00427562	Pfam00113	*B. cenocepacia* HI2424	ZP_00463627	27.87
136	Enolase	ZP_00427562	Pfam00113	*B. cenocepacia* HI2424	ZP_00463627	33.84
138	Glyceraldehyde-3-phosphate dehydrogenase, type I	ZP_00424087	COG0057	*B. cenocepacia* HI2424	ZP_00464205	36.53
*tolU*167	Toluene tolerance	ZP_00423664	Pfam05494	*B. cenocepacia* HI2424	ZP_00460565	38.31
54	Hypothetical protein Bcep1808DRAFT_4697	ZP_00423431	N/A	*Photobacterium profundum* SS9	YP_130032	40.61
37_G	Leucyl aminopeptidase	ZP_00424970	CD00433	*B. cenocepacia* HI2424	ZP_00459092	40.61
*traR*146	*B. vietnamiensis* G4 putative transcriptional regulator	ZP_00425376	CD00093	*Nostoc* sp. PCC 7120 BAB73675	BAB73675	41.99
152_G	Phosphoribosylglycinamide synthetase	ZP_00422952	Pfam01071	*B. cenocepacia* HI2424	ZP_00459261	44.52
151_G	Amidophosphoribosyl transferase	ZP_00425259	CD00715	*B. cenocepacia* HI2424	ZP_00464306	45.90
117_G	Dihydrolipoamide acetyltransferase	ZP_00423459	COG0508	*B. cepacia R18194*	ZP_00212748	47.84
87_G	Dihydrolipoamide acetyltransferase	ZP_00423459	COG0508	*B. cepacia R18194*	ZP_00212748	49.73
118_G	Dihydrolipoamide acetyltransferase	ZP_00423459	COG0508	*B. cepacia R18194*	ZP_00212748	50.36
155_G	Lysine decarboxylase	ZP_00427852	Pfam01276	*B. cenocepacia* HI2424	ZP_00459127	53.47
127_G	2-oxo-acid dehydrogenase E1 component homodimeric type	ZP_00423460	COG2609	*B. cenocepacia* HI2424	ZP_00464635	54.28
102	6-phosphogluconate dehydratase	ZP_00424101	Pfam00920	*B. cepacia R18194*	ZP_00216635	55.51
39	2-oxo-acid dehydrogenase E1 component homodimeric type	ZP_00423460	COG2609	*B. cenocepacia* HI2424	ZP_00464635	60.61
2	6-phosphogluconate dehydratase	ZP_00424101	Pfam00920	*B. cepacia R18194*	ZP_00216635	60.61
88	6-phosphogluconate dehydratase	ZP_00424101	Pfam00920	*B. cepacia R18194*	ZP_00216635	63.93
*vacJ*71	VacJ-like lipoprotein	ZP_00423665	Pfam04333	*B. cepacia R18194*	ZP_00211840	64.13
*typA*107	Small GTP-binding protein domain:GTP-binding protein TypA	ZP_00427539	COG1217	*B. cenocepacia* HI2424	ZP_00464145	74.38
154	Small GTP-binding protein domain:GTP-binding protein TypA	ZP_00427539	COG1217	*B. cenocepacia* HI2424	ZP_00464145	84.34
48	Peptidase U62, modulator of DNA gyrase	ZP_00424681	Pfam01523	*B. cenocepacia* HI2424	ZP_00462606	85.64
18	tRNA pseudouridine synthase B	ZP_00427544	COG0130	*B. cenocepacia* HI2424	ZP_00464150	85.64
43	Hypothetical protein Bcep1808DRAFT_6821	ZP_00420823	N/A	*Mesorhizobium loti* MAFF303099	NP_085706	85.95
8	6-phosphogluconate dehydrogenase, decarboxylating	ZP_00424035	Pfam00393	*B. cenocepacia* HI2424	ZP_00462313	85.95
41	Dihydrolipoamide acetyltransferase	ZP_00423459	COG0508	*B. cepacia R18194*	ZP_00212748	87.13
6	Isocitrate dehydrogenase NADP-dependent, prok type	ZP_00425021	Pfam00180	*B. cenocepacia* HI2424	ZP_00465381	87.13
175_P	Histidine biosynthesis protein HisF	ZP_00423651	COG0107	*B. multivorans*	BAC65276	89.07
183_P	Cytochrome bd ubiquinol oxidase, subunit II	ZP_00420131	COG1294	*B. cepacia R18194*	ZP_00218419	95.15
*mtrA*38 $_G$	Methyltransferase	ZP_00426792	Pfam01795	*B. cenocepacia* HI2424	ZP_00459343	97.21
57	Unidentified		COG5475	*Azotobacter vinelandii* AvOP	ZP_00415768	97.21
182_P	Glu/Leu/Phe/Val dehydrogenase, C terminal:Glu/Leu/Phe/Val	ZP_00424107	COG0334	*B. cenocepacia* HI2424	ZP_00464178	97.50

Mutant ID [a]	Protein Function	G4 Accession	Protein ID	Closest match with homology	Homolog Accession	% of G4 OD [b]
	dehydrogenase, dimerisation region					
*gtrA*90	Glycosyl transferase, family 2 [*B. vietnamiensis* G4]	ZP_00420186	Pfam00535	*B. cenocepacia* HI2424	ZP_00464361	98.12
*strE*80	Ribosomal protein L25 (general stress protein Ctc)	ZP_00424464	COG1825	*B. cenocepacia* HI2424	ZP_00462937	98.28
149	Cytochrome bd ubiquinol oxidase, subunit I	ZP_00420130	COG1271	*B. cenocepacia* HI2424	ZP_00462437	99.06
173 $_P$	Putative membrane protein	ZP_00427592	COG0348	*B. cenocepacia* HI2424	ZP_00465319	99.31
156	Tryptophan synthase, beta chain	ZP_00425266	COG0133	*B. multivorans*	BAC65264	100.09
59	Hypothetical protein Bcep1808DRAFT_4697	ZP_00423431	N/A	*Photobacterium profundum* SS9	YP_130032	100.94
120	Isocitrate dehydrogenase NADP-dependent, monomeric type	ZP_00425020	Pfam03971	*B. cenocepacia* HI2424	ZP_00465380	102.15
180 $_P$	tRNA pseudouridine synthase B	ZP_00427544	COG0130	*B. cenocepacia* HI2424	ZP_00464150	102.83
165	3-isopropylmalate dehydratase large subunit	ZP_00425274	CD01583	*B. multivorans*	BAC65257	103.26
*lysM*67	Peptidoglycan-binding LysM:SLT:MLTD_N	ZP_00420270	Pfam01464	*B. cenocepacia* HI2424	ZP_00460478	104.76
169 $_P$	FeS cluster assembly scaffold IscU	ZP_00423448	Pfam01592	*B. cenocepacia* HI2424	ZP_00464623	105.29
*hspM*49	Heat shock protein Hsp70	ZP_00425517	Pfam00012	*B. multivorans*	BAD82894	105.52
20	Cytidyltransferase-related	ZP_00423886	Pfam01467	*B. cenocepacia* HI2424	ZP_00463005	105.52
40	2-oxo-acid dehydrogenase E1 component homodimeric type	ZP_00423460	COG2609	*B. cenocepacia* HI2424	ZP_00464635	106.22
4	Hypothetical protein Bcep1808DRAFT_4599	ZP_00423333	Pfam01051	*B. pseudomallei* K96243	YP_112354	106.22
*forM*31	Formate dehydrogenase, alpha subunit, anaerobic	ZP_00425245	Pfam01568	*B. cenocepacia* HI2424	ZP_00464321	112.03
51	Peptidase S41A, C-terminal protease	ZP_00424510	Pfam03572	*B. cenocepacia* HI2424	ZP_00464553	112.03
184 $_P$	tRNA synthetases, class-II (G, H, P and S)	ZP_00421650	Pfam00587	*B. cepacia R18194*	ZP_00216101	118.95

[b] Mutants are ordered according to the percentage of the *B. vietnamiensis* G4 WT OD they were able to grow (Section 5.2.6.3) [a] Specific names are designated in the mutant ID column to identify mutants for which complementation was attempted; $_G$ Mutants that were unable to grow with glucose as a sole carbon source. $_P$ Mutants that were identified by the phenol screen but were resistant to 2,4-dichlorophenol.

susceptible to 2,4-D.

5.3.4 Comparison of growth rates of mutants

The mean growth rates of 18 individual mutants and that of wild type G4 were compared by ANOVA at the 95% confidence level and Fishers *a priori* test (Section 5.2.4.5). Only mutants *mtrA*38, *enoL*91 and 136 were significantly different from the wild type G4. The growth rates of two of the six random control mutants were significantly different from that of the wild type G4.

5.3.5 Bioinformatic characterisation and selection of mutants

Of the 52 mutants with mini-Tn*5* insertions into identified genes (Section 5.2.4.4), 11 were initially selected for complementation on the basis of their putative function. These mutants were assigned a name according to the putative gene function designated by the DOE Joint Genome Institute's annotation of the strain G4 genome (Tables 5.4 & 5.5). For example, *enoL*91 was a putative enolase, previously numbered 91; *tolU*167 was identified as a toluene resistance gene, and *traR*146 showed homology with a strain G4 transcriptional regulator. The 11 mutants, their names and their corresponding gene homologs are listed in Table 5.5. One of the mutants that was unable to grow on glucose as a sole carbon source but had a significantly increased growth rate in TSB, *mtrA*38, was included amongst those selected for further study as a control; all of the other selected mutants grew strongly in BSM containing glucose only (Section 5.2.4.1). The putative protein functions of the genes selected for further study are detailed in Table 5.4 and Figs. 5.9-5.13 show the locations of the mini-Tn*5* insertions in relation to the surrounding genes.

5.3.6 Characterisation of selected STM mutants by complementation

A discreet product of the expected size was amplified by PCR for all but one (*tolU*167) of the 11 primer sets designed to target complementing genes (Section 5.2.5.3; Table 5.3; Fig. 5.14). The purified fragments were ligated into the vector pASE101 (Section 5.2.5.4; Table 5.4) and transformed into One Shot OmniMAX™ 2 T1 Phage-Resistant *E. coli.* The reactions to transform *strE*80 and *vacJ*71 did not yield any recombinant clones and further analysis of these mutations was thus precluded. The eight recombinant pASE101 plasmids successfully transformed were extracted and subjected to restriction digest to excise each gene fragment from its complementing construct and confirm its size (Section 5.2.5.6). Three constructs failed to yield discreet product of the correct size (Fig. 5.15). These negative template plasmid preps were diluted by 1:4 and the restriction was repeated but it was not

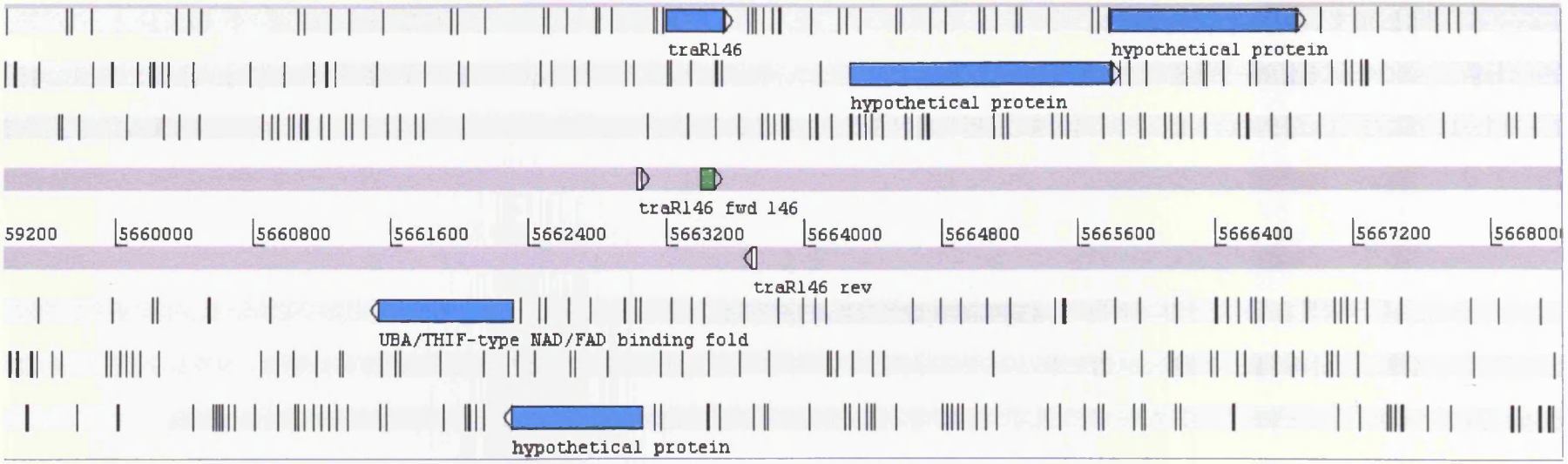


Figure 5.9: Artemis (196) screen view of the putative transcriptional regulator mutant *traR*146 and the surrounding genes marked in blue. Full details of the mutant are given in Table 5.5. Primer sites allocated for gene complementation are shown in white (Table 5.3) and the position of mini-Tn*5* insertions into the strain G4 genome is indicated in green. Six reading frames are visible with vertical lines indicating stop codons, and blue rectangles representing open reading frames derived from automatic annotation of the strain G4 genome by JGI (Fig. 5.3).

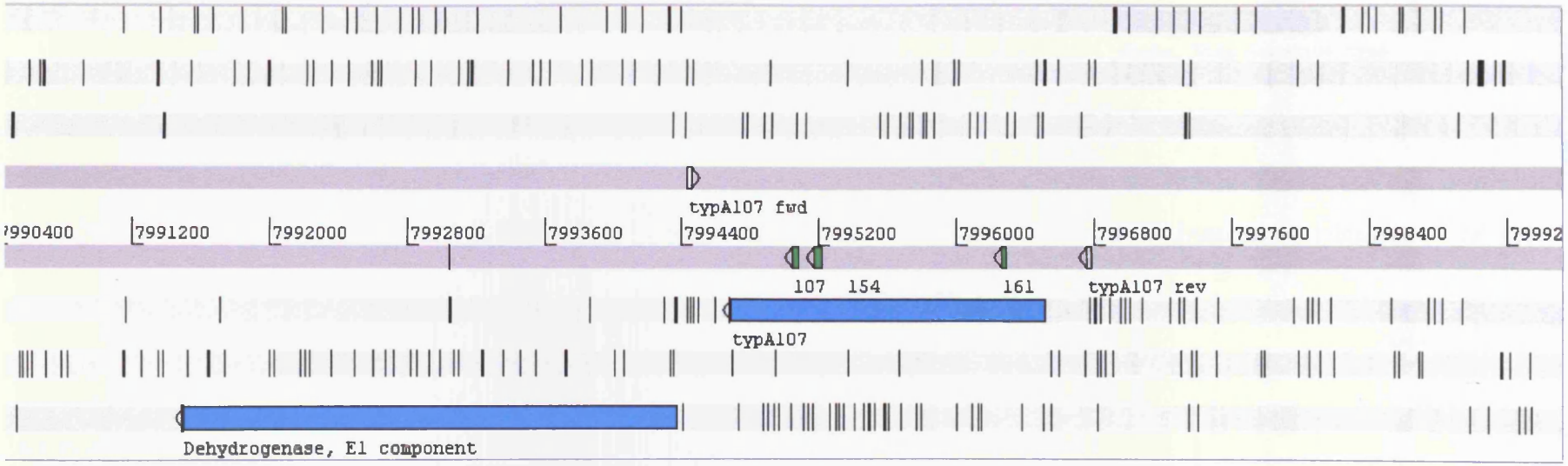


Figure 5.10: Artemis (196) screen view of the small GTP-binding protein domain mutant *typA*107 and the surrounding genes marked in blue. Full details of the mutant are given in Table 5.5. Primer sites allocated for gene complementation are shown in white (Table 5.3) and the position(s) of mini-Tn*5* insertions into the strain G4 genome is indicated in green.

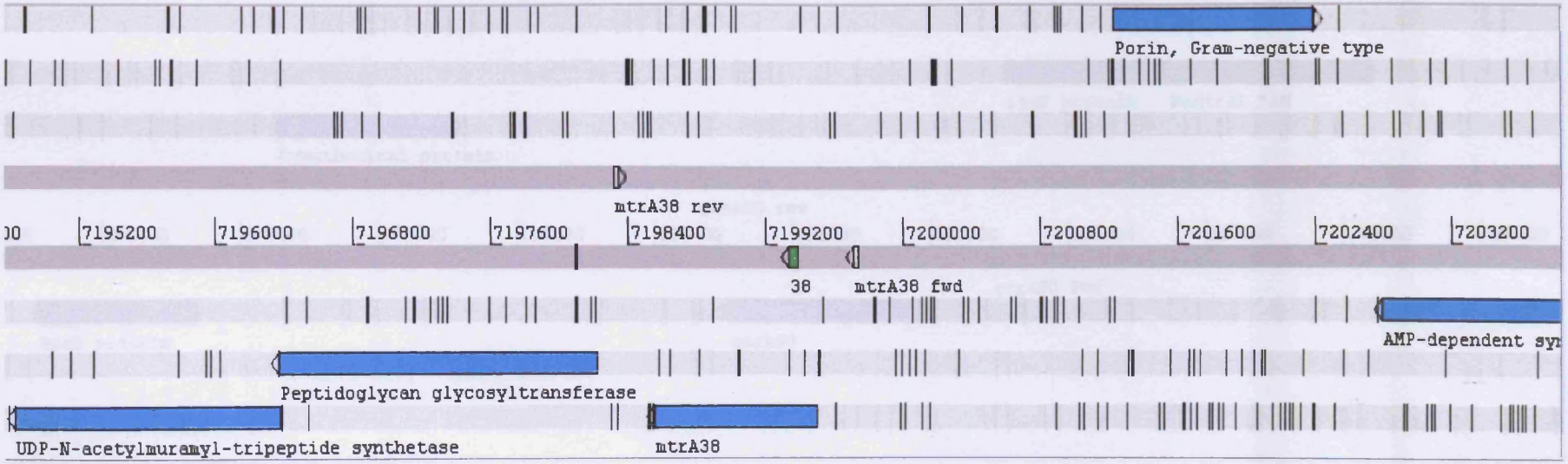


Figure 5.11: Artemis (196) screen view of the methyltransferase mutant *mtrA*38 and the surrounding genes marked in blue. Full details of the mutant are given in Table 5.5. Primer sites allocated for gene complementation are shown in white (Table 5.3) and the position of mini-Tn*5* insertions into the strain G4 genome is indicated in green.

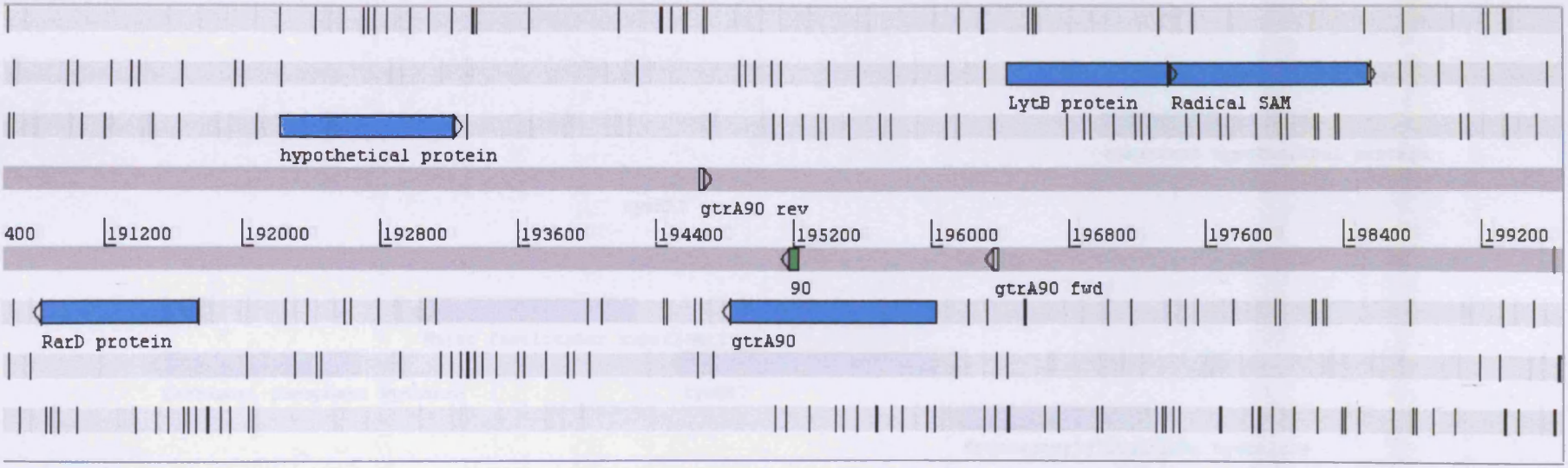


Figure 5.12: Artemis (196) screen view of the glycosyl transferase mutant *gtrA*90 and the surrounding genes marked in blue. Full details of the mutant are given in Table 5.5. Primer sites allocated for gene complementation are shown in white (Table 5.3) and the position of mini-Tn*5* insertions into the strain G4 genome is indicated in green.

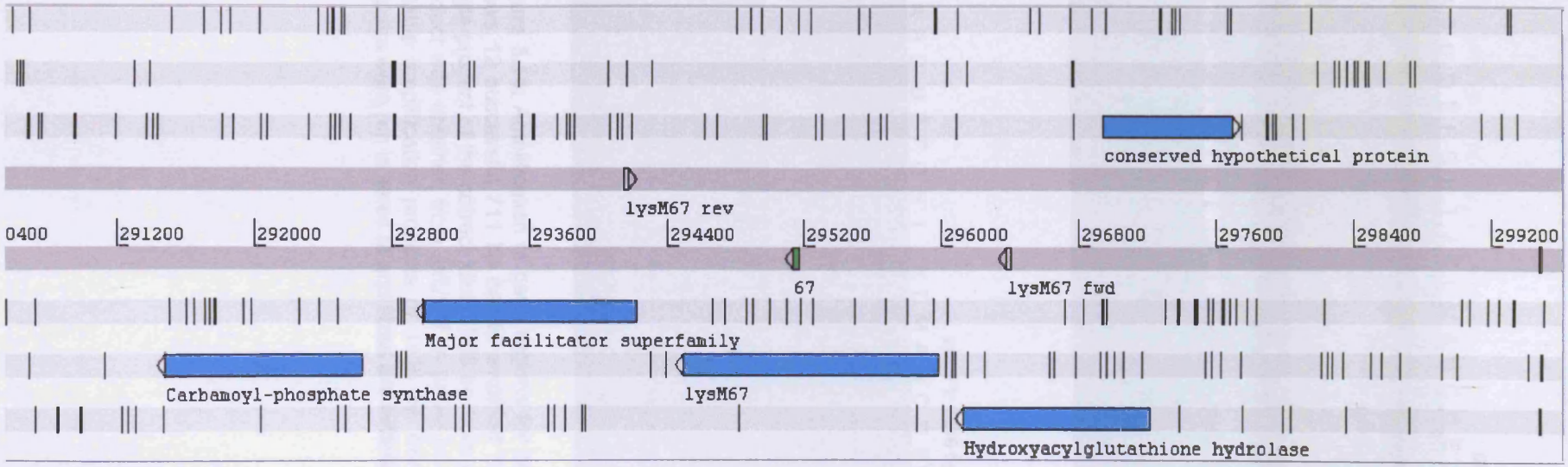


Figure 5.13: Artemis (196) screen view of the peptidoglycan binding LysM mutant *lysM*67 and the surrounding genes marked in blue. Full details of the mutant are given in Table 5.5. Primer sites allocated for gene complementation are shown in white (Table 5.3) and the position of mini-Tn*5* insertions into the strain G4 genome is indicated in green.

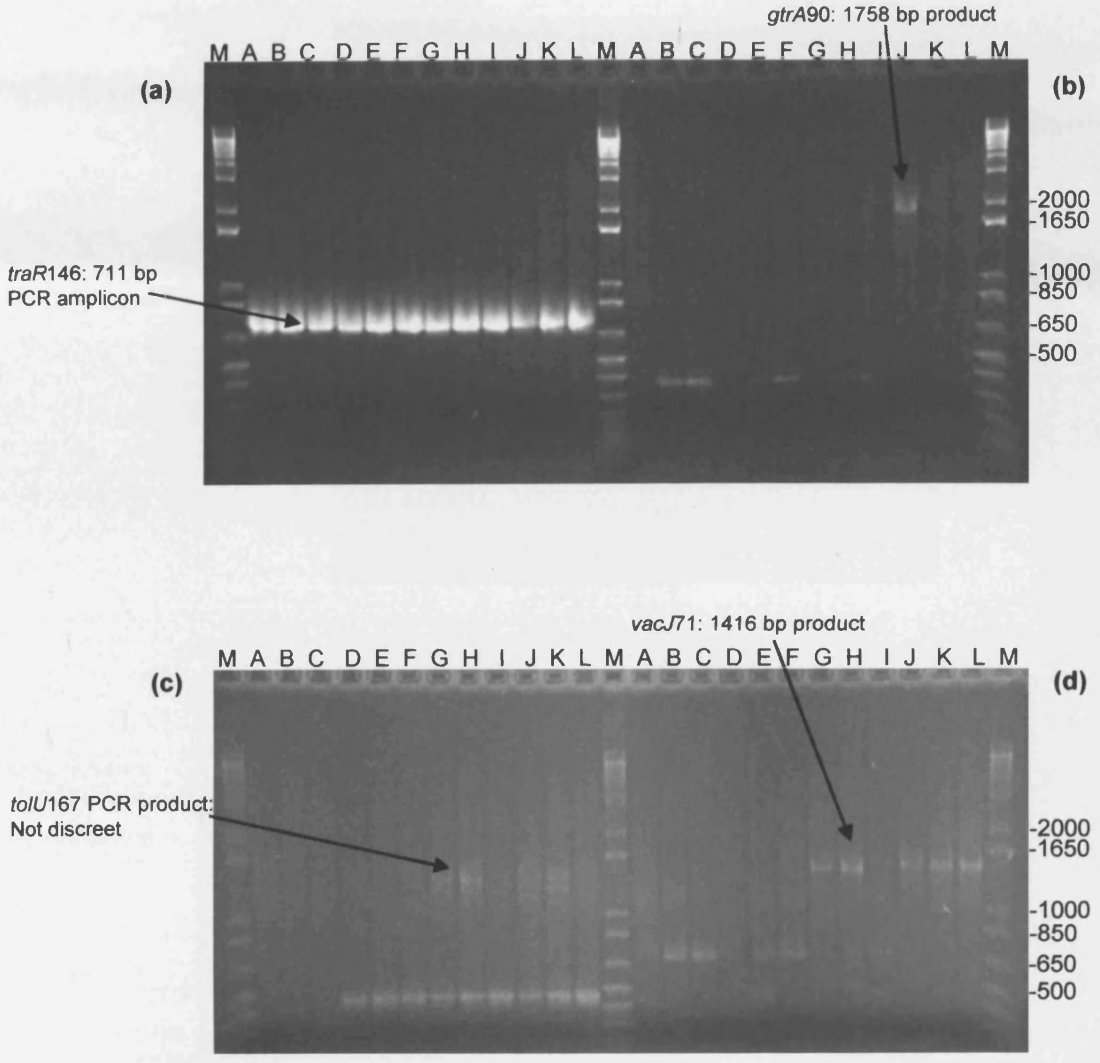


Figure 5.14: Amplification of gene regions for complementation. Panel (*a*) shows 12 successful 711 bp bands amplified from *traR*146, (*b*) shows a single product of the correct size, 1758 bp from *gtrA*90. No single, discreet product was obtained from *tolU*167 in panel (*c*) and panel (*d*) shows five suitable amplification products of 1416 bp from *vacJ*71. Molecular size markers were run in lanes M and sizes are indicated in bp.

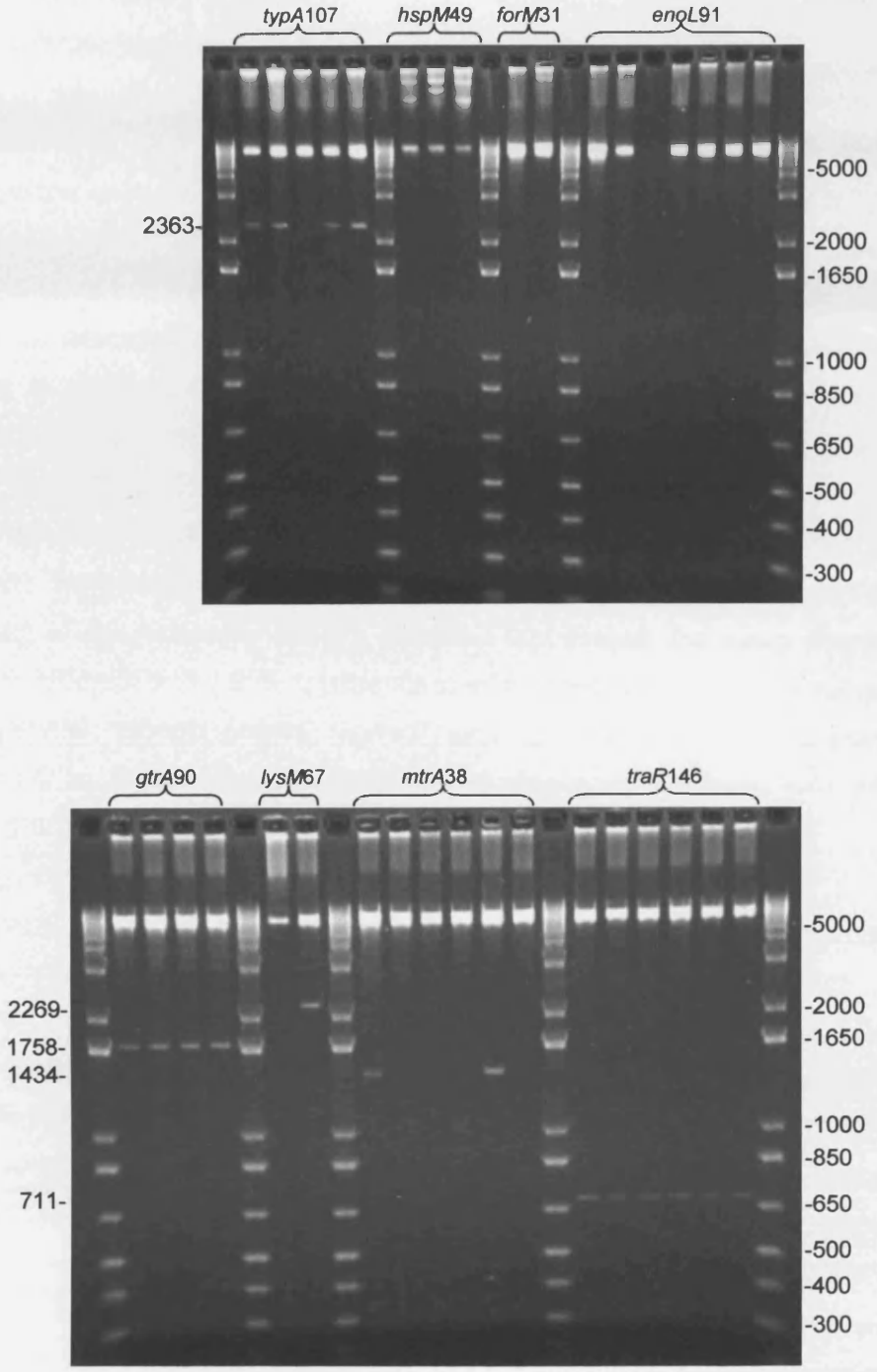


Figure 5.15: Excision of gene fragments from pASE101 recombinant plasmids prepared by alkaline SDS lysis. Products of the correct sizes are shown from five out of eight complementing constructs. Molecular size markers were included to indicate size in bp.

possible to excise the appropriate gene fragments. The five constructs containing gene fragments were transferred to their respective mutants in a tri-parental mating as described in Section 5.2.5.7.

The test to determine the inhibitory concentration (IC) of 2,4-dichlorophenol against *B. vietnamiensis* G4 in TSB showed a significant decrease in growth at concentrations over 0.5 mM and a marked decrease 1.25 mM after 48 h Fig. 5.16. Two screening concentrations were set at 0.75 mM and 1.00 mM. The mean OD for five test mutants and their corresponding complements and controls were measured after 48 h as described in Section 5.2.6.2. Complementation, defined as a significant difference in pollutant resistance between the mutant and the complement, was observed at the 0.75 mM 2,4-dichlorophenol level in mutants *typA*107, *gtrA*90 and *mtrA*38 (Figs. 5.17-5.19). No complementation was observed in mutants *traR*146 and *lysM*67 (Figs. 5.20 & 5.21). To evaluate the lethal effects of 2,4-dichlorophenol viable counts were determined after the suspension tests as described in Section 5.2.6.4. Comparison of the means by ANOVA indicated that overall, the mean counts were significantly different at the 95% confidence level with respect to the different mutants (*traR*146, *gtrA*90, *mtrA*38, *lysM*67, *typA*107) and also with respect to the treatments (water control or 2 mM 2,4-dichlorophenol). No significant difference was detected within the groups of means for different "type" of each mutant (Fig. 5.22).

5.3.7 Screen to determine susceptibility of all identified mutants to 2,4-Dicholorphenol in TSB.

During the characterisation of selected complemented mutants the pollutant screening method was re-optimised for use with complex medium and focused on 2,4-dichlorophenol as a single pollutant as described in Sections 5.2.6.1-5.2.6.3. The results obtained by the original screen were therefore revisited. The mutants that had been sequenced and their interrupted genes marked on the genome were re-tested to determine their susceptibility to 1 mM 2,4-dichlorophenol in TSB after 48 h. 14 of the mutants grew to over 100% of the optical density to which *B. vietnamiensis* G4 grew under the same conditions (Table 5.5), and it should be noted that amongst the mutants showing growth over 89% were seven phenol mutants that were not identified as 2,4-dichlorophenol susceptible in the initial screen.

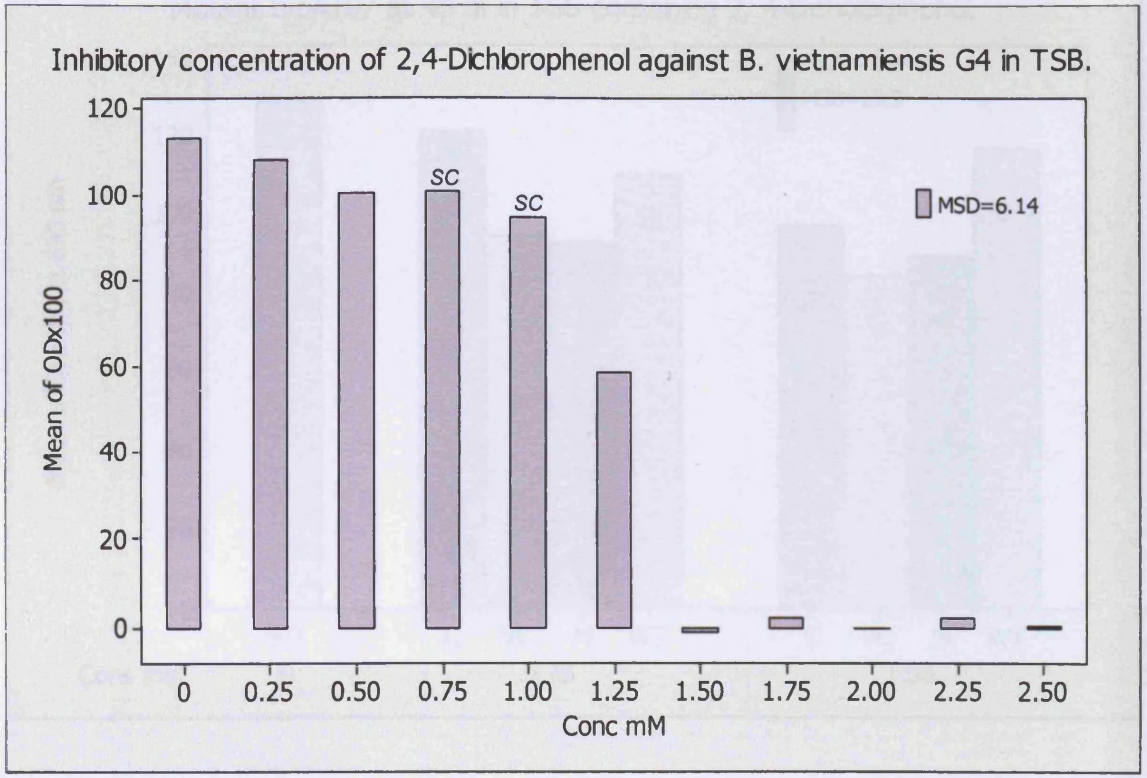


Figure 5.16: The decrease in optical density at 630 nm (as a measure of *B. vietnamiensis* growth) as concentration of 2,4-dichlorophenol increased in 0.25 mM increments. *SC*: Screening concentration. Each concentration was replicated eight times. ANOVA indicated a highly significant difference between the mean values at the 99% confidence level. Minimum significant difference (MSD), indicated individual comparisons, shows significant decrease in growth above 0.50 mM 2,4-dichlorophenol.

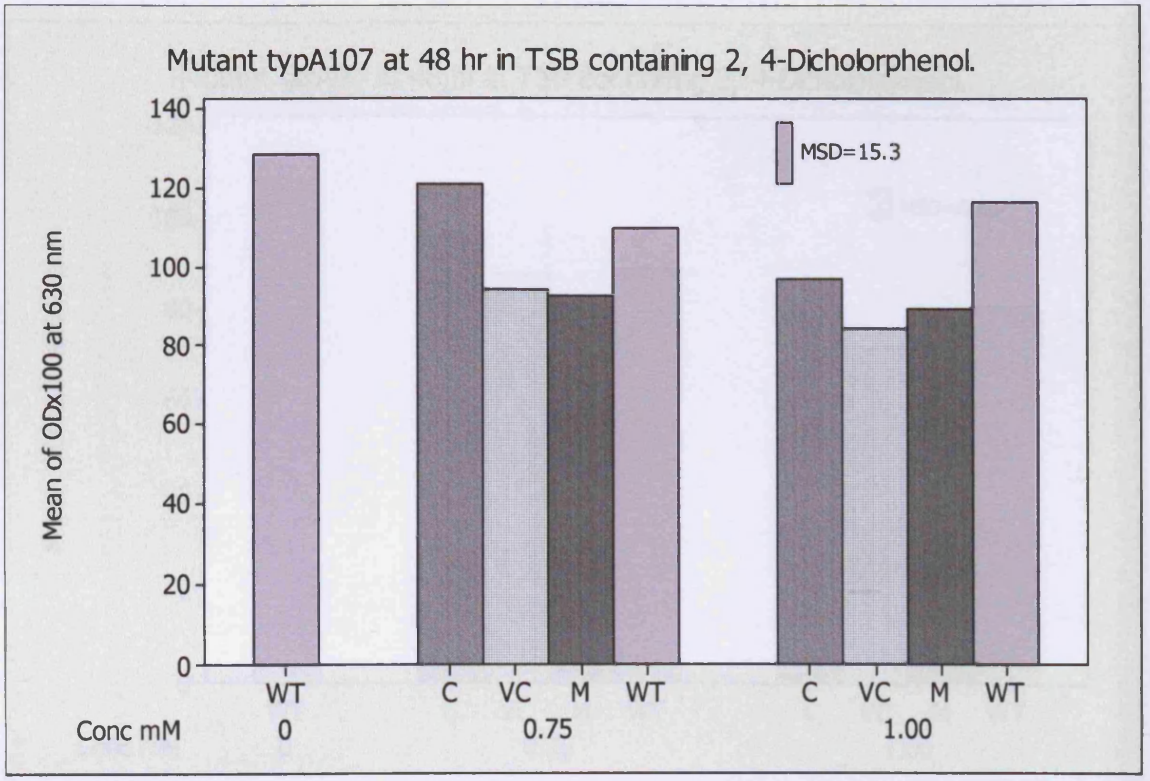


Figure 5.17: Mean optical density after 48 h of: C, mutant *typA*107 carrying complementing construct; VC, mutant *typA*107 vector control carrying pASE101 with no complementing gene region insert; M, pollutant susceptible mutant *typA*107. WT indicates *B. vietnamiensis* strain G4 wild type.
At the 0.75 mM concentration level the VC and mutant were significantly lower than the complement and WT. There was no significant difference between the complement, WT 0.75 mM and WT 0 mM control. At the 1.0 mM level there was no significant difference between the complement and VC. However the difference between the mutant and WT still indicated a significant knockdown at this level. MSD bar indicated for individual comparisons.

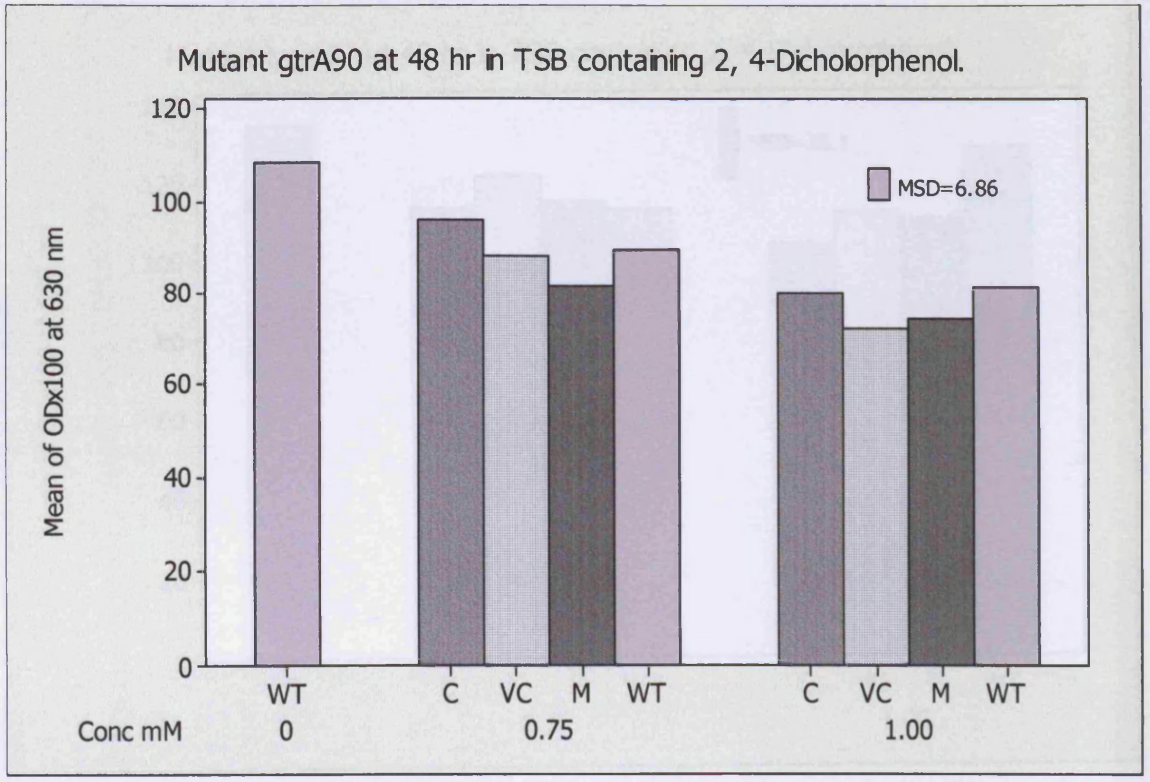


Figure 5.18: Mean optical density after 48 h of: C, mutant *gtrA*90 carrying complementing construct; VC, mutant *gtrA*90 vector control carrying pASE101 with no complementing gene region insert; M, pollutant susceptible mutant *gtrA*90. WT indicates *B. vietnamiensis* strain G4 wild type.
At the 0.75 mM level there was no significant difference between the WT and complement but they were both significantly higher than mutant *gtrA*90. However, the WT was not higher than the VC. At the 1.0 mM level, mutant *gtrA*90 was lower than the WT, indicating significant knockdown, but not different to the complement. MSD bar indicated for individual comparisons.

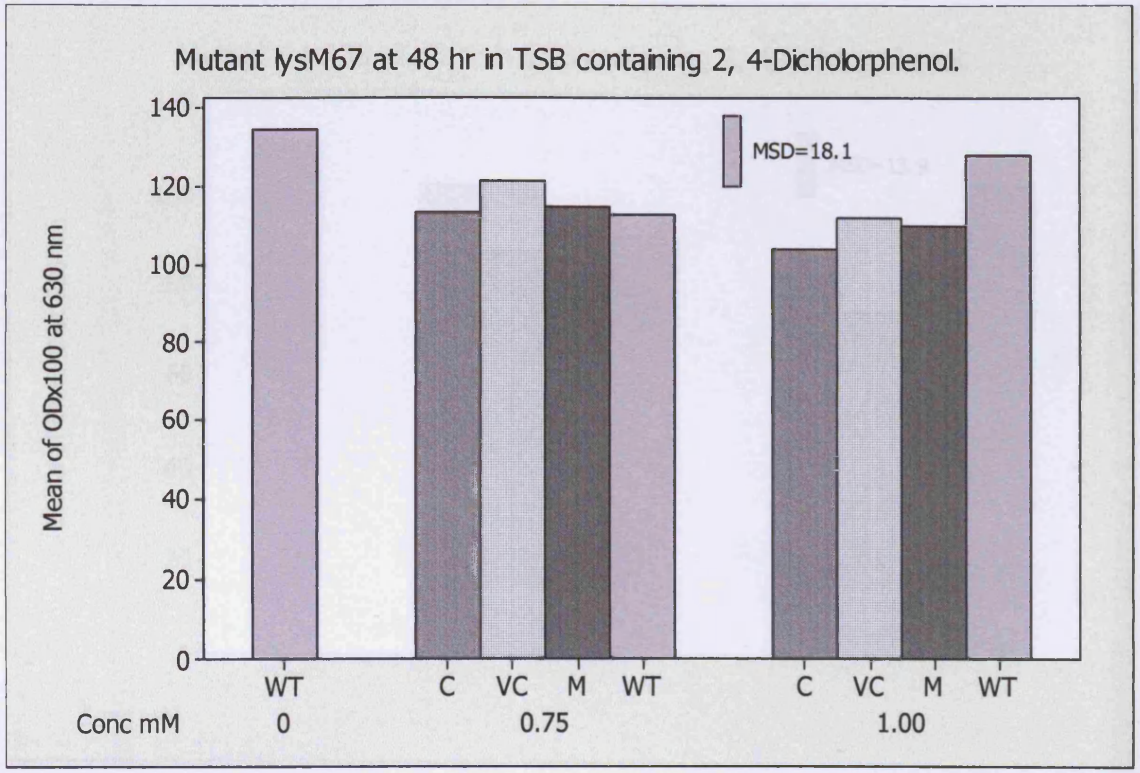


Figure 5.19: Mean optical density after 48 h of: C, mutant *lysM*67 carrying complementing construct; VC, mutant *lysM*67 vector control carrying pASE101 with no complementing gene region insert; M, pollutant susceptible mutant *lysM*67. WT indicates *B. vietnamiensis* strain G4 wild type.
Significant differences between the wild types and the complements and no significant differences between mutant and the complements indicated that complementation had not been successful. MSD bar indicated for individual comparisons.

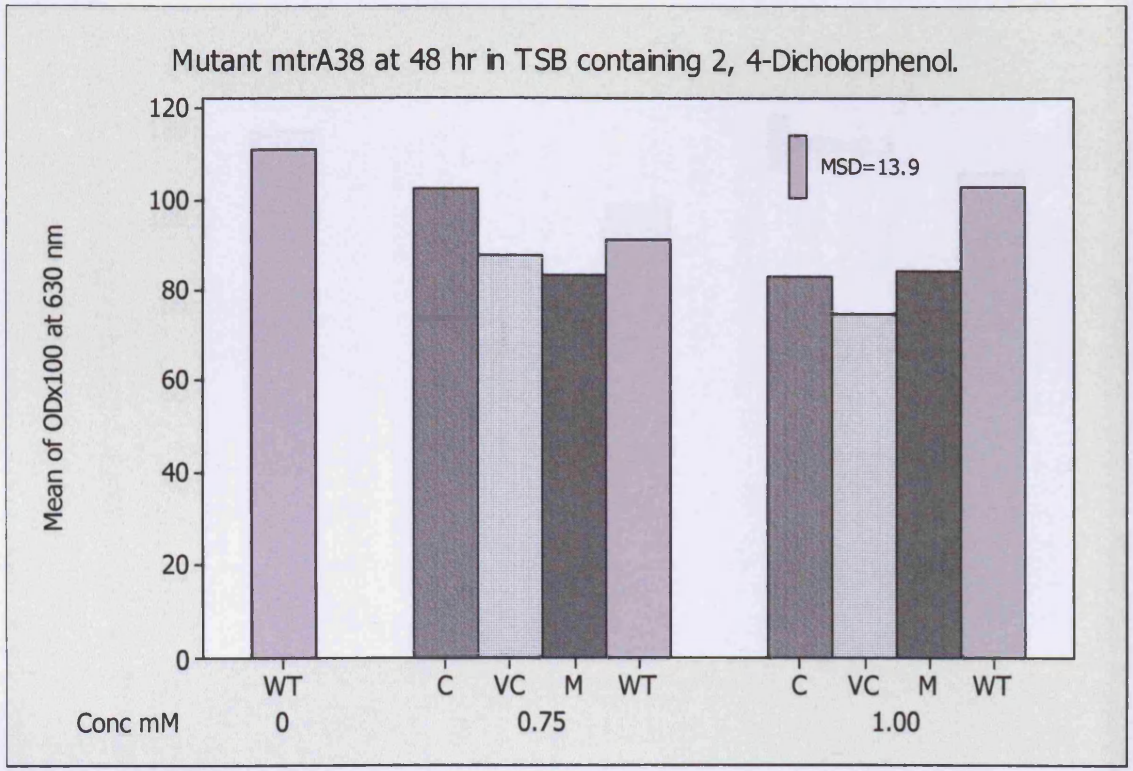


Figure 5.20: Mean optical density after 48 h of: C, mutant *mtrA*38 carrying complementing construct; VC, mutant *mtrA*38 vector control carrying pASE101 with no complementing gene region insert; M, pollutant susceptible mutant *mtrA*38. WT indicates *B. vietnamiensis* strain G4 wild type.
At the 0.75 mM concentration level there was a significant difference between the complement and the VC and mutant *mtrA*38. However no significant difference was observed between the mutant and the WT indicating no knockdown in this test. Significant knockdown was observed, but there was no complementation at the 1.0 mM level. MSD bar indicated for individual comparisons.

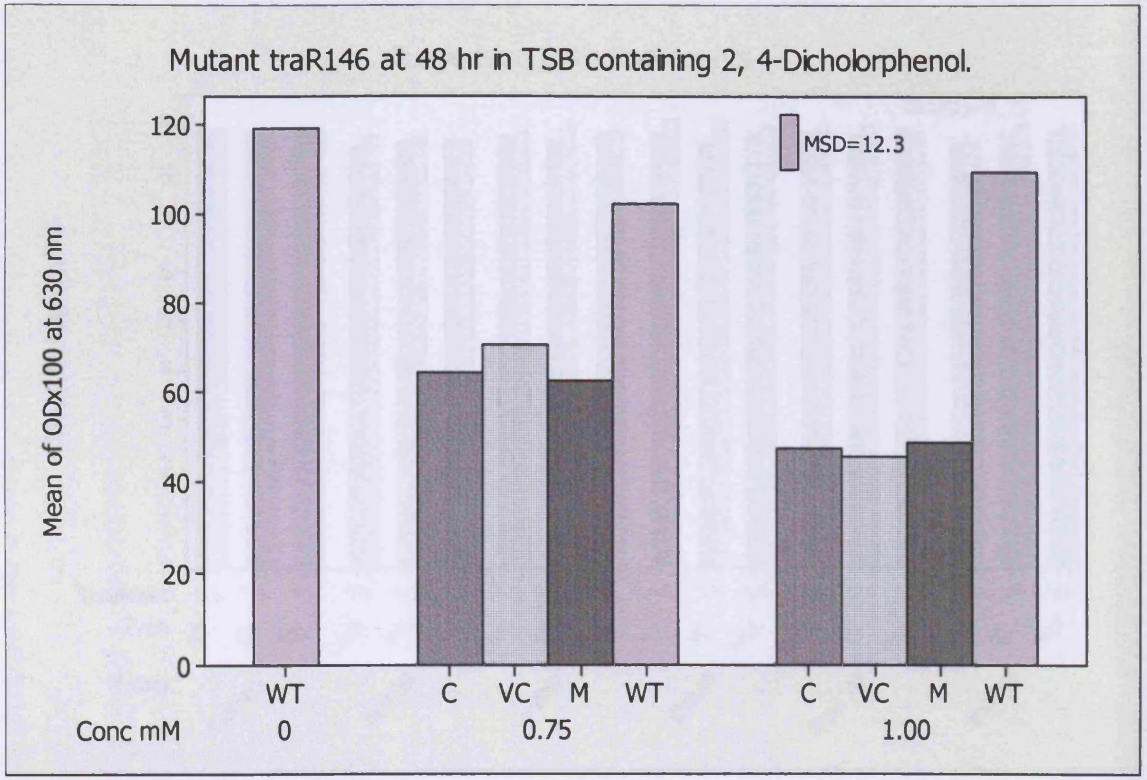


Figure 5.21: Mean optical density after 48 h of: C, mutant *traR*146 carrying complementing construct; VC, mutant *traR*146 vector control carrying pASE101 with no complementing gene region insert; M, pollutant susceptible mutant *traR*146. WT indicates *B. vietnamiensis* strain G4 wild type.
No significant difference was observed between the VC, mutant or complements at either concentration, indicating no successful complementation. However there was a highly significant knockdown observed between mutant *traR*146 and the WT. MSD bar indicated for individual comparisons.

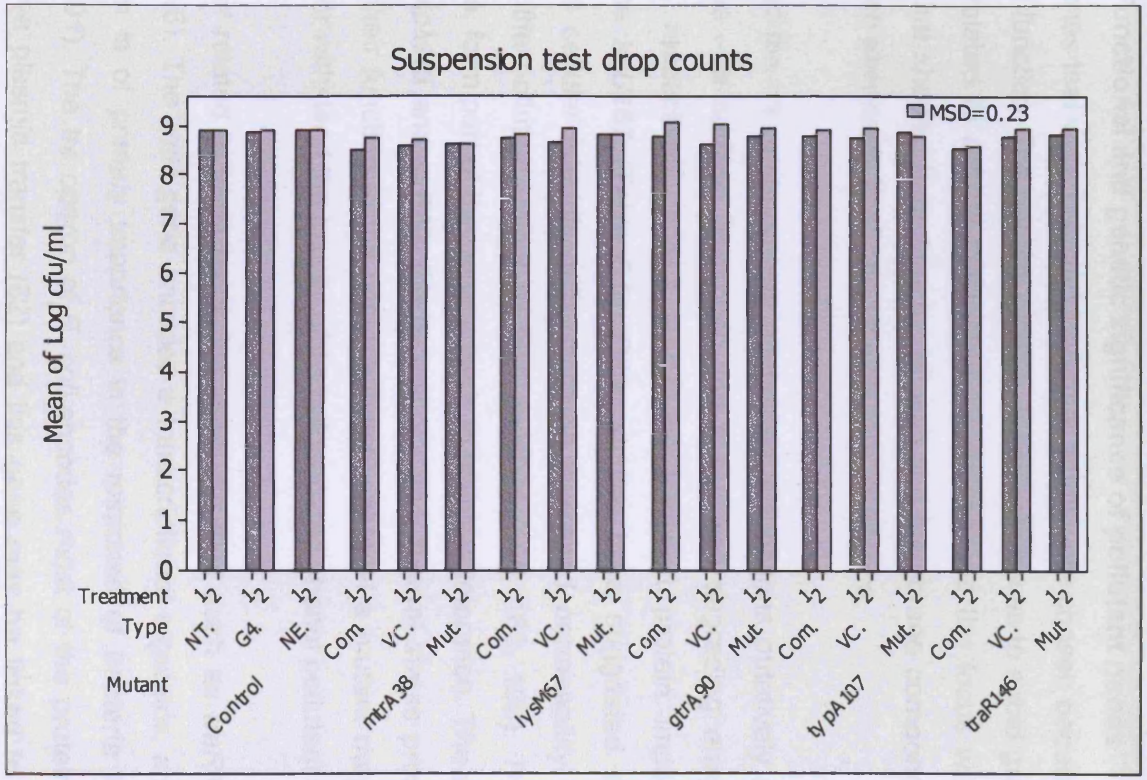


Figure 5.22: Viable counts of pollutant susceptible mutants after suspension tests in 2 mM 2,4-dichlorophenol. MSD bar indicated for individual comparisons.
Treatments: 1=2,4-dichlorophenol or neutraliser test, 2=water control
Types: NT=Neutraliser toxicity, G4= *B. vietnamiensis* G4, NE=Neutraliser efficacy, Com=mutant and complementing construct, VC=Vector control, Mut=pollutant susceptible mutant.

5.4 DISCUSSION:

This study set out to (*i*) identify novel mutants of *B. vietnamiensis* G4 that were susceptible to a range of model pollutants and (*ii*) identify their genetic basis of resistance. A bank of mutants showing pollutant susceptibility was successfully isolated and the corresponding genes which were implicated in resistance were identified. This novel set of genes facilitates speculation about how they might interact to confer pollutant resistance in *B. vietnamiensis* G4.

5.4.1 Functional and genetic significance of pollutant genes

The genes that were selected for further study were chosen because of their putative protein function. The general criterion for selection was to avoid genes that appeared to be related to central metabolism processes and the focus was directed toward those that showed homology to cell wall and membrane components or genes that had been previously implicated in pollutant resistance.

Three of the 11 mutants were attenuated in functions putatively related to the cell envelope – a surface lipoprotein and a cell wall degrading enzyme, *vacJ*71 and *lysM*67, respectively, and a membrane transport protein implicated in toluene tolerance *tolU*167 (Table 5.5). Other studies have suggested that bacteria with impaired cellular integrity, either through increased permeability or an inability to excrete the pollutant, might be less resistant (105, 165, 190). These genes may, therefore, form part of the genetic basis to pollutant tolerance. The same can be said about *hspM*49 and *strE*80, the heat shock and general stress proteins (Table 5.5); without their function at the post-transcriptional level a mutant may be less able to reverse or withstand the toxicity of the halogenated phenol pollutant (191).

Mutations related to gene regulation were identified such as *traR*146 and *typA*107 (Table 5.5). The *traR* gene encodes a transcriptional regulator, and transcriptional regulation is of primary importance in the response of bacteria to environmental stress (191). The tra operon of *E. coli* encodes most of the proteins necessary for conjugative plasmid transfer (82) and this gene may be linked to horizontal gene transfer although its function was only predicted in the JGI annotation of the strain G4 genome. Alternatively a transcriptional regulator may confer resistance in a similar way to *typA*107, a virulence regulator. Transcriptional regulators may either repress or enhance the activity of RNA polymerase at the promoter sequences of prokaryotic genes or operons, and may be encoded several kb away from their target sequence. The discovery of the virulence regulator may imply that the ability of G4 to survive in

polluted environments might be related to the versatility of *Burkholderia* species as a pathogen.

In the context of 2,4-dichlorophenol resistance, methylation of nucleic acids or proteins by methyltransferase might constitute an alteration of the pollutant target site, a suggested mechanism of resistance which is missing in *mtrA*38 (Table 5.5). The loss of a glycosyltransferase in mutant *gtrA*90 may have simply impaired the ability of the bacteria to grow as that class of enzymes catalyses the formation of oligosaccharides and polysaccharides. The remaining mutants, *forM*31 and *enoL*91 were interrupted in their genes for formate dehydrogenase and enolase and may have simply been at a disadvantage growing in conditions where they would not normally have been affected by their mutation. The enolase mutant may, however, have had an aspect of membrane function attenuated because it has recently been implicated in the pathogenesis of streptococci (176).

5.4.2 Metabolism mutants

A screen of the original 192 mutants, against BSM with no pollutant added, revealed that 35 were unable to grow to an optical density of 0.7 at 630 nm with glucose as a sole carbon source. These mutants had their mini-Tn*5Km2* insertions in genes coding for central metabolism enzymes such as pyruvate dehydrogenase and pyruvate/2-oxoglutarate dehydrogenase. Leucyl aminopeptidase, involved in the degradation of intracellular proteins (86), was also attenuated. 26 of the 35 "metabolism mutants" were not sequenced. This may be attributed to the method used during the creation of the mutant library (Section 5.2.2). After the filter mating the transconjugants were recovered on TSA containing 180 Units.ml^{-1} PMX and 200 μg.ml^{-1} KM. At this early stage, auxotrophic mutants and central metabolism mutants could have been excluded with the use of minimal media (197). This would have made the subsequent screens more efficient as they would have had an increased chance of exclusively identifying pollutant mutants. This study, however, utilised a library of signature-tagged mutants that was assembled as part of an ongoing affiliated study (L. A. O' Sullivan, unpublished results) of which it was a requirement that auxotrophs were included during the creation of the mutant bank. The fact that the mutant screens were able to detect this type of metabolism mutant is an indication that the method was working correctly.

5.4.3 Multiple insertions into the same gene

It was possible to rank the mutants for which the interrupted genes were identified according to the percentage of the strain G4 WT growth that they showed in the final 2,4-dichlorophenol screen in TSB (Table 5.5). On several occasions there were multiple insertions of the mini-Tn*5Km2* into the same gene. For example, *enoL*91 and mutant 136 both carried a mutation in the same enolase gene, and mutants 117, 87 and 118 were all mutated in the same dihydrolipidamide acetyltransferase gene. Interestingly, these sibling mutants tended to grow to a similar optical density (Table 5.5), supporting the reliability of the re-optimised final optical density screen in 2,4-dicholorphenol.

There are three possible answers as to why there are multiple insertions into the same gene on several occasions: (*i*) The mini-Tn*5Km2* inserted into the genes preferentially – that is to say there were several "insertion hotspots". It was confirmed by Southern hybridisation, however, that the insertion was random (L. A. O'Sullivan, unpublished results) making this an unlikely explanation; (*ii*) The genes were essential for pollutant tolerance and the screen was working efficiently, however, only 1920 mutants were screened in the first place and it seems rather unlikely that that proportion of duplicates would have existed; and (*iii*) the PCR of the transposon-chromosome junction and the subsequent sequencing introduced bias into the genetic identification procedure. This effect would still have to exist in tandem with either option (*i*) or (*ii*) – many of the 185 mutants that were flagged by the initial screens did not return a PCR product or, if they did, failed to yield sequence suitable for analysis. For this reason interesting mutants may have been "lost" from the initial results and others may have been over-represented.

5.4.4 Restoration of wild-type phenotype by gene complementation

Mutants were selected for further characterisation; growth rate analysis and complementation (Section 5.2.4.5-5.2.6.2) from those for which sequence was available, on the general basis of their putative function. When, however, the screen was re-optimised it became clear that it would have been more meaningful to make the selection on the basis of final pollutant susceptibility. Of the five mutants for whom complementation was sought, presence of a complementing construct was seen to restore the resistance of mutants *typA*107 and *mtrA*38 to within a significant similarity of the un-attenuated wild-type at 0.75 mM whilst the vector controls and the mutants were significantly lower than the complemented mutant. However, *gtrA*90, *lysM*67 and *traR*146 showed no successful complementation effect. The least

efficient complementation was in *traR*146 which showed no significant difference between the vector control, the mutant or the complemented mutant (Fig. 5.21).

The five mutants that were characterised were not shown by the final screen against 2,4-dichlorophenol to be those with the lowest optical density. Positive effects of complementation may be more pronounced upon these mutants with a lower resistance to 2,4-dichlorophenol, for example 7, 42, *enoL*91, 136, 138 and *tolU*167. All of these grew to less than 40% of the wild type and it is these mutants that should certainly be considered in further analyses (Table 5.5).

5.4.5 Design of pollutant screen experiment

Retrospectively, several aspects of the design of the initial mutant bank screen (Section 5.2.4.1) were modified before the final screen (Section 5.2.6.3) as follows: (*i*) Inconsistencies in preparing pollutant stocks; 2,4-D, and Dichloroprop required a very high pH in order to dissolve in water. When preparing 100 mM stock solutions it was difficult to re-adjust the pH to seven because these pollutants were halogenated phenols and incompatible with an electronic pH meter. In the initial screen the stock of 2,4-Diclorophenol was prepared at 750mM in DMSO in order to reduce the final amount of DMSO in the growth medium. In the second screen, however, the stock was pre-diluted in sterile water to allow the addition of a greater volume to the medium to achieve the same concentration. This afforded greater accuracy and reproducibility with regards final 2,4-dichlorophenol screening concentration. (*ii*) Numbers of samples; The first screen was designed to identify mutants with susceptibly to four pollutants from a bank of 1920 mutants. Twenty 96 well plates containing mutants were subjected to four pollutant screens so 80 plates were processed in total. In the second screen, however, only pollutant-susceptible mutants with identified genes were tested against a single pollutant and each result was replicated eight times. (*iii*) Inoculation; The primary screen was inoculated with a 96 point replicator. The inoculum for the second screen was standardised to OD 1 and diluted to 1:1000 which made the results much more robust. (*iv*) Length of screen; The second screen was run for 48 h as opposed to 24 h in the first screen. This caused the ODs to be higher overall. Mutants that were marked as susceptible but then grew to an OD above that of the wild type may have been "slow growers" instead of pollutant susceptible mutants.

5.4.6 Further work

There are several ways in which this study might be improved in the future: The initial screens must be repeated with all the pollutants, 2,4-D, Dichloroprop, 2,4-dichlorophenol and phenol but under the optimal conditions elucidated for the final screen. This would include using a rich medium, TSB, and screening the mutants for longer, 48 h. More focus could be placed on resistance to phenol and the study could be extended to include toluene since more is known about the catabolism of these compounds by strain G4. From the current study, an improvement would be to select the mutants with the lowest resistance and growth, and ensure that good sequence is obtained from as many as possible. The Epicentre failsafe PCR kit for amplifying problem templates (Section 5.2.5.3) was successful and it could be applied to obtain high-quality nucleotide sequence data that would give a balanced representation of the genes that were actually mutated without the possibility of any bias.

CHAPTER 6: GENERAL DISCUSSION, CONCLUSIONS AND FUTURE WORK

The *Burkholderia* genus contains strains with a number of ecologically important functions (179, 220). In order for these functions to be exploited for future bacteriological biotechnology applications, two gaps in our knowledge must be addressed. The first is the accurate identification of novel and existing strains. Correct identification is especially important for the *Burkholderia* genus because of the strict requirements restricting the registration of new biotechnologically useful strains (178). Despite containing a few very well-studied examples (48, 77, 87) the genus remains poorly characterised. Secondly, it is necessary to gain an understanding of the mechanisms by which biotechnologically interesting strains are resistant to and able to degrade compounds that are harmful to other, susceptible strains. Understanding these pollutant resistance genes and pathways offers the potential genetically engineer more proficient or safer (with regard to those capable of human opportunistic infections) biotechnological strains. In addition, if any of the pollutant resistance genes are conserved in *Burkholderia* or other bioremediation strains they may offer a means to examine the diversity of microorganisms in contaminated sites using functional gene probes as well as recA or the 16S rRNA gene. This study, therefore, aimed to simultaneously explore two important questions about the *Burkholderia* genus by developing an approach to investigate the taxonomic diversity of *Burkholderia* populations within the natural environment and also to identify the genes involved in the resistance to pollutants of the model organism *B. vietnamiensis* G4.

The *recA* gene based molecular identification approach developed and evaluated in Chapters 2, 3 & 4 provided a successful solution to the problem of limited phylogenetic resolution within the traditional target for bacterial identification, the 16S rRNA gene. The choice of the *recA* gene as a target built upon existing work with the *Burkholderia cepacia* complex (143) and the novel sequence data obtained in Chapter 2 allowed the resolution of the phylogenetic analyses of a large collection of *B. cepacia* complex strains that had previously not been assigned to a taxonomic group. The *recA* phylogeny, however, was not a suitable basis for the differentiation between strains of clinical and environmental origin.

The novel BUR1.2 PCR primers developed in (Chapter 2) were not exclusively *Burkholderia* genus specific and detected many non-*Burkholderia* species when applied to the maize rhizosphere sample (Chapter 3). The production of the 869 bp

amplicon will prove valuable for expansion of the *recA* gene sequence databases to include new species that are important in maize rhizosphere samples. The novel sequence data obtained from these primers also allowed the design of the BUR3.4 primer set.

The BUR3.4 primers developed (Chapter 2) and evaluated (Chapters 3 & 4) in this study were suitable for the rapid profiling of the *Burkholderia* species diversity in a nested PCR approach applied to screen a clone library. The BUR3.4 primers could also have utility when used with DGGE to assess the diversity associated with an environmental sample before the construction of a clone library. This would be particularly useful when it is necessary to determine the population density of various species within a sample as construction of a clone library directly from an environmental DNA sample may be susceptible to PCR bias (253).

Cultivation independent methods are valuable for evaluation of environmental populations as analyses that rely on cultivation underestimate the abundance of uncultivable species, such as endosymbionts (159). The question of whether bacteria exist in genuine endosymbiotic relationships within the maize rhizosphere or cord forming fungal species is an interesting one, because such relationships have been described previously in other host organisms (21, 182, 234). Aside from their utility as a molecular identification system, the primers developed in this study may be developed for use with FISH and microscopy to examine the distribution of *Burkholderia* species (BUR3.4) or closely related *Burkholderia*-like species (BUR1.2). This technique could provide important information about bacteria that are closely associated with or even contained within host cells.

Metagenomics describes the functional and sequence-based analysis of the collective microbial genomes contained in an environmental sample. Metagenomic methods to capture entire bacterial communities are becoming more common with further advances in sequencing technology and have focused on a range of environments including soil, the oral cavity aquatic habitats and the hospital metagenome (194). A metagenomic approach could be applied to investigate microbial communities, including those of *Burkholderia* species, within environmental samples of interest such as the maize rhizosphere. The *recA* gene could be used as a target to identify *Burkholderia* clones from the metagenomic library, and then the entire clone would be sequenced to search for genes of interest amongst the genes flanking the phylogenetic anchor. Both novel primer sets developed in this study

would be suitable for this, either BUR1.2 to target *Burkholderia*-like species or BUR3.4 to identify more exclusively *Burkholderia* species. The 16S rRNA gene is the most commonly used phylogenetic anchor, but other genes that contain phylogenetic information such as *recA* would also be suitable. Metagenomic studies conducted in this way often result in the identification of a large number of ORFs of unknown function or hypothetical ORFs flanking the phylogenetic marker. These sequences, whilst initially of little use to describe the niche the organism fills in the environment, serve to enrich the databases and describe characteristics that may be of use when the function is determined for a homologue of the same gene family subsequently (194). The primers developed in this study might also have utility as probes in microarrays of phylogenetically important markers to analyse bacterial populations in natural environment.

In the second part of this study of diversity and function in *Burkholderia* species, the ability of *B. vietnamiensis* to survive in the presence of important hydrocarbon pollutants was examined. Several *Burkholderia* strains have been shown to be excellent bioremediation agents and *B. vietnamiensis* strain G4 was investigated as a model species since it is known to degrade a variety of pollutants and its genome sequence is available. Specifically, genes were identified that were implicated in the resistance of the model pollutant 2,4-dichlorophenol. The application of the results of such a study might ultimately involve the removal of genes that are involved in pollutant resistance from *B. vietnamiensis* G4 to a biotechnologically useful host strain with no record of pathogenicity. Taking advantage of the increased *recA* gene dataset using a metagenomic approach in conjunction with these newly-identified resistance genes might identify homologues in uncultivated *Burkholderia* species that may, in the future, represent suitable candidates for biotechnological registration. Molecular probes targeting these genes could be used, with the *recA* gene as a phylogenetic marker where necessary, to screen a metagenomic library and identify clones encoding homologous genes. Again, the genes flanking the genes of interest on the metagenomic clone may provide valuable sequence data leading to the identification of further pollutant genes, or regulators of the gene of interest.

Further work that must be carried out in relation to the 2,4-dichlorophenol resistance genes (Chapter 5) includes complementation experiments, which are required to clarify the relationship between loss of phenotype and transposon mutation. The five mutants that were the final focus of the study (*traR*146, *gtrA*90, *mtrA*38, *lysM*67, *typA*107) and others exhibiting decreased growth at the screening concentration of

2,4-dichlorophenol are worthy of further characterisation and complementation. Such genes, for which the phenotype is known, provide valuable options when searching for key genes within environmental settings. For example, there are 22 *Burkholderia* species genomes sequencing projects either completed or in progress (33). A BLASTp search (2) of these genomes using the amino-acid sequence of the *Burkholderia vietnamiensis* G4 genes implicated in 2,4-Dichlorophenol resistance (Table 5.5; amino acid sequence data can be obtained from the accession numbers) reveals different levels of conservation of the genes between various *Burkholderia* species. The most susceptible mutant, "mutant 7" (Table 5.5), putatively attenuated in 2-oxo-acid dehydrogenase expression, has sequenced homologues in *B. cenocepacia, B. cepacia, B. dolosa, B. ambifaria, B. thailandensis, B. mallei, B. pseudomallei* and *B. xenovorans.* The least similar *Burkholderia* species, *B. xenovorans,* shares 90% homologous sequence with *B. vietnamiensis* at the amino acid level. The next BLASTp 2-oxo-acid dehydrogenase homologue is in *Ralstonia eutropha*, with 70% identity. These results indicate that the 2-oxo-acid dehydrogenase gene is highly conserved within the *Burkholderia* genus. Another example of conservation within the putative pollutant resistance genes is that of the toluene tolerance ABC transporter mutant, *tolU*167 (Table 5.5), with shared identity of over 97% within the same range of species but only 49% with the next hit, *Ralstonia metallidurans.* The extensive level of amino acid homology of both the latter functional pollution genes suggests that it would be easy to design environmental diversity probes using similar strategies developed for recA as described in this study (Chapters 2-4). These approaches could then be used for more rapid and accurate identification of cultivatable strains capable of survival in the presence of pollutants well as culture independent analysis or metagenomic studies of polluted environments to detect the resident *Burkholderia* populations. A different phylogenetic perspective could be gained from exploiting such novel conserved functional genes in the place of the well studied 16S rRNA gene and *recA* genes targeted for most diversity studies.

In contrast to pollutant resistance genes with a high level of amino acid conservation within the *Burkholderia* genus, the mutant *enoL*91, attenuated in *B. vietnamiensis* G4 (Chapter 5) displayed more limited distribution amongst *Burkholderia* species with homologues in *B. cepacia, B. pseudomallei, B. xenovorans* and *Ralstonia eutropha* only. That this gene appears to be absent from other sequenced *Burkholderia* genomes allows the hypothesis that it might represent a more specific high-level

pollutant resistance gene. This potential correlation with high resistance might make it more useful in a specific phenotypic investigation as described above.

Finally, in the present study, the investigation into pollutant resistance was conducted with each individual mutant in pure culture. A logical extension of the method would be to repeat the screen in a controlled microcosm environment where signature tagged transposon mutants were tested for their ability to compete in a population as well as for survival in the presence of a pollutant. This same experiment might be repeated, but with several model pollutants simultaneously to narrow the range of mutants that were obtained. It would be possible to screen a greater number of mutants from the original library in this way, and an opportunity to exploit the full potential of the signature tagged mutants.

In summary, two biotechnologically relevant aspects of *Burkholderia* biology have been explored in this study. Firstly, genetic tools based on the *recA* gene have been developed to enable better characterisation of species diversity within the *Burkholderia* genus. Secondly, the genetic basis for pollutant tolerance has been examined and uncovered a diversity of putative hydrocarbon resistance genes. Together, both aspects of this study have furthered our knowledge of how to exploit and harness the biotechnological potential within the genus *Burkholderia*.

REFERENCES:

1. **Achouak, W., R. Christen, M. Barakat, M. H. Martel, and T. Heulin.** 1999. *Burkholderia caribensis* sp. nov., an exopolysaccharide-producing bacterium isolated from vertisol microaggregates in Martinique. Int J Syst Bacteriol **49 Pt 2**:787-94.

2. **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**:3389-402.

3. **Alvey, S., C. H. Yang, A. Buerkert, and D. E. Crowley.** 2003. Cereal/legume rotation effects on rhizosphere bacterial community structure in West African soils. Biol Fertil of Soils **37**:73-82.

4. **Anderson, A. S., and E. M. Wellington.** 2001. The taxonomy of *Streptomyces* and related genera. Int J Syst Evol Microbiol **51**:797-814.

5. **Anderson, R. T., H. A. Vrionis, I. Ortiz-Bernad, C. T. Resch, P. E. Long, R. Dayvault, K. Karp, S. Marutzky, D. R. Metzler, A. Peacock, D. C. White, M. Lowe, and D. R. Lovley.** 2003. Stimulating the *in situ* activity of *Geobacter* species to remove uranium from the groundwater of a uranium-contaminated aquifer. Appl Environ Microbiol **69**:5884-5891.

6. **Ashelford, K. E., N. Banning, J. C. Fry, and A. Weightman.** 2004. Assessing the effectiveness of commonly used 16S rRNA universal primer sets to target intended taxa; an *in silico* study. FEMS Microbiol Ecol: manuscript in preparation.

7. **Aspray, T. J., S. K. Hansen, and R. G. Burns.** 2005. A soil-based microbial biofilm exposed to 2,4-D: bacterial community development and establishment of conjugative plasmid pJP4. FEMS Microbiol Ecol **54**:317-27.

8. **Azegami, K., K. Nishiyama, Y. Watanabe, I. Kadota, A. Ohuchi, and C. Fukazawa.** 1987. *Pseudomonas plantarii* sp nov, the causal agent of rice seedling blight. Int J Syst Bacteriol **37**:144-152.

9. **Balandreau, J., V. Viallard, B. Cournoyer, T. Coenye, S. Laevens, and P. Vandamme.** 2001. *Burkholderia cepacia* genomovar III is a common plant-associated bacterium. Appl Environ Microbiol **67**:982-985.

10. **Baldwin, A., E. Mahenthiralingam, K. M. Thickett, D. Honeybourne, M. C. Maiden, J. R. Govan, D. P. Speert, J. J. Lipuma, P. Vandamme, and C. G. Dowson.** 2005. Multilocus sequence typing scheme that provides both species and strain differentiation for the *Burkholderia cepacia* complex. J Clin Microbiol **43**:4665-73.

11. **Baldwin, A., P. A. Sokol, J. Parkhill, and E. Mahenthiralingam.** 2004. The *Burkholderia cepacia* epidemic strain marker is part of a novel genomic island encoding both virulence and metabolism-associated genes in *Burkholderia cenocepacia*. Infect Immun **72**:1537-1547.

12. **Ballard, R. W., N. J. Palleroni, M. Doudoroff, R. Y. Stanier, and M. Mandel.** 1970. Taxonomy of the aerobic pseudomonads: *Pseudomonas*

cepacia, *P. marginata*, *P. alliicola* and *P. caryophylli.* J Gen Microbiol **60:**199-214.

13. **Barac, T., S. Taghavi, B. Borremans, A. Provoost, L. Oeyen, J. V. Colpaert, J. Vangronsveld, and D. van der Lelie.** 2004. Engineered endophytic bacteria improve phytoremediation of water-soluble, volatile, organic pollutants. Nat Biotechnol **22:**583-8.

14. **Barbieri, E., L. Bertini, I. Rossi, P. Ceccaroli, R. Saltarelli, C. Guidi, A. Zambonelli, and V. Stocchi.** 2005. New evidence for bacterial diversity in the ascoma of the ectomycorrhizal fungus *Tuber borchii* Vittad. FEMS Microbiol Lett **247:**23-35.

15. **Bartual, S. G., H. Seifert, C. Hippler, M. A. Luzon, H. Wisplinghoff, and F. Rodriguez-Valera.** 2005. Development of a multilocus sequence typing scheme for characterization of clinical isolates of *Acinetobacter baumannii.* J Clin Microbiol **43:**4382-90.

16. **Benachenhou-Lahfa, N., B. Labedan, and P. Forterre.** 1994. PCR-mediated cloning and sequencing of the gene encoding glutamate dehydrogenase from the archaeon *Sulfolobus shibatae*: identification of putative amino-acid signatures for extremophilic adaptation. Gene **140:**17-24.

17. **Berg, G., L. Eberl, and A. Hartmann.** 2005. The rhizosphere as a reservoir for opportunistic human pathogenic bacteria. Environ Microbiol **7:**1673-1685.

18. **Bergsma-Vlami, M., M. E. Prins, M. Staats, and J. M. Raaijmakers.** 2005. Assessment of genotypic diversity of antibiotic-producing *Pseudomonas* species in the rhizosphere by denaturing gradient gel electrophoresis. Appl Environ Microbiol **71:**993-1003.

19. **Berriatua, E., I. Ziluaga, C. Miguel-Virto, P. Uribarren, R. Juste, S. Laevens, P. Vandamme, and J. R. Govan.** 2001. Outbreak of subclinical mastitis in a flock of dairy sheep associated with *Burkholderia cepacia* complex infection. J Clin Microbiol **39:**990-994.

20. **Bevivino, A., C. Dalmastri, S. Tabacchioni, L. Chiarini, M. L. Belli, S. Piana, A. Materazzo, P. Vandamme, and G. Manno.** 2002. *Burkholderia cepacia* complex bacteria from clinical and environmental sources in Italy: Genomovar status and distribution of traits related to virulence and transmissibility. J Clin Microbiol **40:**846-851.

21. **Bianciotto, V., E. Lumini, P. Bonfante, and P. Vandamme.** 2003. '*Candidatus* glomeribacter gigasporarum' gen. nov., sp. nov., an endosymbiont of arbuscular mycorrhizal fungi. Int J Syst Evol Microbiol **53:**121-4.

22. **Bianciotto, V., E. Lumini, L. Lanfranco, D. Minerdi, P. Bonfante, and S. Perotto.** 2000. Detection and identification of bacterial endosymbionts in arbuscular mycorrhizal fungi belonging to the family *Gigasporaceae.* Appl Environ Microbiol **66:**4503-9.

23. **Bisset, K. A.** 1973. Do bacteria have a nuclear membrane? Nature **241:**45.

24. **Boddy, L.** 1993. Saprotrophic cord-forming fungi - warfare strategies and other ecological aspects. Mycol Res **97:**641-655.

25. **Bramer, C. O., P. Vandamme, L. F. da Silva, J. G. Gomez, and A. Steinbuchel.** 2001. Polyhydroxyalkanoate-accumulating bacterium isolated from soil of a sugar-cane plantation in Brazil. Int J Syst Evol Microbiol **51:**1709-13.

26. **Brett, P. J., D. DeShazer, and D. E. Woods.** 1998. *Burkholderia thailandensis* sp. nov., a *Burkholderia pseudomallei*-like species. Int J Syst Bacteriol **48 Pt 1:**317-20.

27. **Bringel, F., A. Castioni, D. K. Olukoya, G. E. Felis, S. Torriani, and F. Dellaglio.** 2005. *Lactobacillus plantarum* subsp. argentoratensis subsp. nov., isolated from vegetable matrices. Int J Syst Evol Microbiol **55:**1629-34.

28. **Brook, M. D., B. Currie, and P. M. Desmarchelier.** 1997. Isolation and identification of *Burkholderia pseudomallei* from soil using selective culture techniques and the polymerase chain reaction. J Appl Microbiol **82:**589-96.

29. **Burkholder, W.** 1950. Sour skin, a bacterial rot of onion bulbs. Phytopathology **64:**468-475.

30. **Caballero-Mellado, J., L. Martinez-Aguilar, G. Paredes-Valdez, and P. E. Santos.** 2004. *Burkholderia unamae* sp. nov., an N2-fixing rhizospheric and endophytic species. Int J Syst Evol Microbiol **54:**1165-72.

31. **Campana, S., G. Taccetti, N. Ravenni, F. Favari, L. Cariani, A. Sciacca, D. Savoia, A. Collura, E. Fiscarelli, G. De Intinis, M. Busetti, A. Cipolloni, A. d'Aprile, E. Provenzano, I. Collebrusco, P. Frontini, G. Stassi, M. Trancassini, D. Tovagliari, A. Lavitola, C. J. Doherty, T. Coenye, J. R. Govan, and P. Vandamme.** 2005. Transmission of *Burkholderia cepacia* complex: evidence for new epidemic clones infecting cystic fibrosis patients in Italy. J Clin Microbiol **43:**5136-42.

32. **Carvalho, G. M., A. P. Carvalho, T. W. Folescu, L. Higa, L. M. Teixeira, M. C. Plotkowski, V. L. Merquior, R. M. Albano, and E. A. Marques.** 2005. Transient isolation of *Burkholderia multivorans* and *Burkholderia cenocepacia* from a Brazilian cystic fibrosis patient chronically colonized with *Burkholderia vietnamiensis.* J Cyst Fibros **4:**267-270.

33. **Chain, P. S., V. J. Denef, K. T. Konstantinidis, L. M. Vergez, L. Agullo, V. L. Reyes, L. Hauser, M. Cordova, L. Gomez, M. Gonzalez, M. Land, V. Lao, F. Larimer, J. J. LiPuma, E. Mahenthiralingam, S. A. Malfatti, C. J. Marx, J. J. Parnell, A. Ramette, P. Richardson, M. Seeger, D. Smith, T. Spilker, W. J. Sul, T. V. Tsoi, L. E. Ulrich, I. B. Zhulin, and J. M. Tiedje.** 2006. *Burkholderia xenovorans* LB400 harbors a multi-replicon, 9.73-Mbp genome shaped for versatility. Proc Natl Acad Sci USA **103:**15280-7.

34. **Chen, J. S., K. A. Witzmann, T. Spilker, R. J. Fink, and J. J. LiPuma.** 2001. Endemicity and inter-city spread of *Burkholderia cepacia* genomovar III in cystic fibrosis. J Pediatr **139:**643-9.

35. **Cheneby, D., S. Perrez, C. Devroe, S. Hallet, Y. Couton, F. Bizouard, G. Iuretig, J. C. Germon, and L. Philippot.** 2004. Denitrifying bacteria in bulk

and maize-rhizospheric soil: diversity and N2O-reducing abilities. Can J Microbiol **50:**469-474.

36. **Cho, Y. S., H. Y. Kahng, C. K. Kim, J. J. Kukor, and K. H. Oh.** 2002. Physiological and cellular responses of the 2,4-D degrading bacterium, *Burkholderia cepacia* YK-2, to the phenoxyherbicides 2,4-D and 2,4,5-T. Curr Microbiol **45:**415-22.

37. **Chow, M. L., C. C. Radomski, J. M. McDermott, J. Davies, and P. E. Axelrood.** 2002. Molecular characterization of bacterial diversity in Lodgepole pine (*Pinus contorta*) rhizosphere soils from British Columbia forest soils differing in disturbance and geographic source. FEMS Microbiol Ecol **42:**347-357.

38. **Clark, A. J., and A. D. Margulies.** 1965. Isolation and characterization of recombination-deficient mutants of *Escherichia coli* K12. Proc Natl Acad Sci USA **53:**451-9.

39. **Coenye, T., E. Falsen, B. Hoste, M. Ohlen, J. Goris, J. R. Govan, M. Gillis, and P. Vandamme.** 2000. Description of *Pandoraea* gen. nov. with *Pandoraea apista* sp. nov., *Pandoraea pulmonicola* sp. nov., *Pandoraea pnomenusa* sp. nov., *Pandoraea sputorum* sp. nov. and *Pandoraea norimbergensis* comb. nov. Int J Syst Evol Microbiol **50 Pt 2:**887-99.

40. **Coenye, T., M. Gillis, and P. Vandamme.** 2000. *Pseudomonas antimicrobica* Attafuah and Bradbury 1990 is a junior synonym of *Burkholderia gladioli* (Severini 1913) Yabuuchi et al. 1993. Int J Syst Evol Microbiol **50:**2135-2139.

41. **Coenye, T., D. Henry, D. P. Speert, and P. Vandamme.** 2004. *Burkholderia phenoliruptrix* sp. nov., to accommodate the 2,4,5-trichlorophenoxyacetic acid and halophenol-degrading strain AC1100. Syst Appl Microbiol **27:**623-7.

42. **Coenye, T., S. Laevens, M. Gillis, and P. Vandamme.** 2001. Genotypic and chemotaxonomic evidence for the reclassification of *Pseudomonas woodsii* (Smith 1911) Stevens 1925 as *Burkholderia andropogonis* (Smith 1911) Gillis et al. 1995. Int J Syst Evol Microbiol **51:**183-185.

43. **Coenye, T., S. Laevens, A. Willems, M. Ohlen, W. Hannant, J. R. Govan, M. Gillis, E. Falsen, and P. Vandamme.** 2001. *Burkholderia fungorum* sp. nov. and *Burkholderia caledonica* sp. nov., two new species isolated from the environment, animals and human clinical samples. Int J Syst Evol Microbiol **51:**1099-107.

44. **Coenye, T., and J. J. LiPuma.** 2003. Molecular epidemiology of *Burkholderia* species. Front Biosci **8:**e55-67.

45. **Coenye, T., J. J. LiPuma, D. Henry, B. Hoste, K. Vandemeulebroecke, M. Gillis, D. P. Speert, and P. Vandamme.** 2001. *Burkholderia cepacia* genomovar VI, a new member of the *Burkholderia cepacia* complex isolated from cystic fibrosis patients. Int J Syst Evol Microbiol **51:**271-9.

46. **Coenye, T., E. Mahenthiralingam, D. Henry, J. J. LiPuma, S. Laevens, M. Gillis, D. P. Speert, and P. Vandamme.** 2001. *Burkholderia ambifaria* sp. nov., a novel member of the *Burkholderia cepacia* complex including

biocontrol and cystic fibrosis-related isolates. Int J Syst Evol Microbiol **51**:1481-90.

47. **Coenye, T., T. Spilker, A. Martin, and J. J. LiPuma.** 2002. Comparative assessment of genotyping methods for epidemiologic study of *Burkholderia cepacia* genomovar III. J Clin Microbiol **40**:3300-7.

48. **Coenye, T., and P. Vandamme.** 2003. Diversity and significance of *Burkholderia* species occupying diverse ecological niches. Environ Microbiol **5**:719-29.

49. **Coenye, T., P. Vandamme, J. R. Govan, and J. J. LiPuma.** 2001. Taxonomy and identification of the *Burkholderia cepacia* complex. J Clin Microbiol **39**:3427-36.

50. **Coenye, T., P. Vandamme, J. J. LiPuma, J. R. Govan, and E. Mahenthiralingam.** 2003. Updated version of the *Burkholderia cepacia* complex experimental strain panel. J Clin Microbiol **41**:2797-8.

51. **Cruden, D. L., J. H. Wolfram, R. D. Rogers, and D. T. Gibson.** 1992. Physiological properties of a *Pseudomonas* strain which grows with p-xylene in a two-phase (organic-aqueous) medium. Appl Environ Microbiol **58**:2723-9.

52. **Curran, B., D. Jonas, H. Grundmann, T. Pitt, and C. G. Dowson.** 2004. Development of a multilocus sequence typing scheme for the opportunistic pathogen *Pseudomonas aeruginosa*. J Clin Microbiol **42**:5644-9.

53. **da Mota, F. F., E. A. Gomes, E. Paiva, and L. Seldin.** 2005. Assessment of the diversity of *Paenibacillus* species in environmental samples by a novel *rpoB*-based PCR-DGGE method. FEMS Microbiol Ecol **53**:317-328.

54. **Dalmastri, C., L. Pirone, S. Tabacchioni, A. Bevivino, and L. Chiarini.** 2005. Efficacy of species-specific *recA* PCR tests in the identification of *Burkholderia cepacia* complex environmental isolates. FEMS Microbiol Lett **246**:39-45.

55. **Dance, D. A.** 1991. Melioidosis: the tip of the iceberg? Clin Microbiol Rev **4**:52-60.

56. **Daubaras, D. L., C. E. Danganan, A. Hubner, R. W. Ye, W. Hendrickson, and A. M. Chakrabarty.** 1996. Biodegradation of 2,4,5-trichlorophenoxyacetic acid by *Burkholderia cepacia* strain AC1100: evolutionary insight. Gene **179**:1-8.

57. **de Lorenzo, V., M. Herrero, U. Jakubzik, and K. N. Timmis.** 1990. Mini-Tn*5* transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. J Bacteriol **172**:6568-72.

58. **De Marco, P., J. C. Murrell, A. A. Bordalo, and P. Moradas-Ferreira.** 2000. Isolation and characterization of two new methanesulfonic acid-degrading bacterial isolates from a Portuguese soil sample. Arch Microbiol**173**:146-153.

59. **de Ridder-Duine, A. S., G. A. Kowalchuk, P. J. A. K. Gunnewiek, W. Smant, J. A. van Veen, and W. de Boer.** 2005. Rhizosphere bacterial

community composition in natural stands of *Carex arenaria* (sand sedge) is determined by bulk soil community composition. Soil Biol Biochem **37**:349-357.

60. **de Souza, J. T., and J. M. Raaijmakers.** 2003. Polymorphisms within the *prnD* and *pltC* genes from pyrrolnitrin and pyoluteorin-producing *Pseudomonas* and *Burkholderia* spp. FEMS Microbiol Ecol **43**:21-34.

61. **De Soyza, A., L. Archer, A. McDowell, J. Moore, J. H. Dark, S. Elborn, E. Mahenthiralingam, P. A. Corris, and K. Gould.** 2001. Lung transplantation for cystic fibrosis; the effect of *B. cepacia* genomovars on post transplant outcomes. J Heart Lung Transplant **20**:158.

62. **Depret, G., S. Houot, M. R. Allard, M. C. Breuil, R. Nouaim, and G. Laguerre.** 2004. Long-term effects of crop management on *Rhizobium leguminosarum* biovar viciae populations. FEMS Microbiol Ecol **51**:87-97.

63. **Diez, S., J. Sabatte, M. Vinas, J. M. Bayona, A. M. Solanas, and J. Albaiges.** 2005. The prestige oil spill. I. Biodegradation of a heavy fuel oil under simulated conditions. Environ Toxicol Chem **24**:2203-17.

64. **Don, R. H., and J. M. Pemberton.** 1981. Properties of six pesticide degradation plasmids isolated from *Alcaligenes paradoxus* and *Alcaligenes eutrophus*. J Bacteriol **145**:681-6.

65. **Donnelly, D. P., and L. Boddy.** 1998. Developmental and morphological responses of mycelial systems of *Stropharia caerulea* and *Phanerochaete velutina* to soil nutrient enrichment. New Phytologist **138**:519-531.

66. **Donnelly, D. P., and L. Boddy.** 2001. Mycelial dynamics during interactions between *Stropharia caerulea* and other cord-forming, saprotrophic basidiomycetes. New Phytologist **151**:691-704.

67. **Dunn, A. K., and E. V. Stabb.** 2005. Culture-independent characterization of the microbiota of the ant lion *Myrmeleon mobilis* (Neuroptera: Myrmeleontidae). Appl Environ Microbiol **71**:8784-94.

68. **Eisen, J. A.** 1995. The RecA protein as a model molecule for molecular systematic studies of bacteria: comparison of trees of RecAs and 16S rRNAs from the same species. J Mol Evol **41**:1105-23.

69. **Engledow, A. S., E. G. Medrano, E. Mahenthiralingam, J. J. LiPuma, and C. F. Gonzalez.** 2004. Involvement of a plasmid-encoded type IV secretion system in the plant tissue watersoaking phenotype of *Burkholderia cenocepacia*. J Bacteriol **186**:6015-24.

70. **Enright, M. C., and B. G. Spratt.** 1999. Multilocus sequence typing. Trends Microbiol **7**:482-7.

71. **Estrada-De los Santos, P., R. Bustillos-Cristales, and J. Caballero-Mellado.** 2001. *Burkholderia*, a genus rich in plant-associated nitrogen fixers with wide environmental and geographic distribution. Appl Environ Microbiol **67**:2790-2798.

72. **Evans, D. J., Jr., and D. G. Evans.** 1983. Classification of pathogenic *Escherichia coli* according to serotype and the production of virulence factors, with special reference to colonization-factor antigens. Rev Infect Dis **5 Suppl 4**:S692-701.

73. **Fares, M. A., A. Moya, and E. Barrio.** 2004. *GroEL* and the maintenance of bacterial endosymbiosis. Trends Genetics **20**:413-416.

74. **Ferrer, M., P. Golyshin, and K. N. Timmis.** 2003. Novel maltotriose esters enhance biodegradation of Aroclor 1242 by *Burkholderia cepacia* LB400. World J Microbiol Biotech **19**:637-643.

75. **Figurski, D. H., and D. R. Helinski.** 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. Proc Natl Acad Sci USA **76**:1648-52.

76. **Fiore, A., S. Laevens, A. Bevivino, C. Dalmastri, S. Tabacchioni, P. Vandamme, and L. Chiarini.** 2001. *Burkholderia cepacia* complex: distribution of genomovars among isolates from the maize rhizosphere in Italy. Environ Microbiol **3**:137-43.

77. **Folsom, B. R., and P. J. Chapman.** 1991. Performance characterization of a model bioreactor for the biodegradation of trichloroethylene by *Pseudomonas cepacia* G4. Appl Environ Microbiol **57**:1602-1608.

78. **Fontan, P. A., V. Pancholi, M. M. Nociari, and V. A. Fischetti.** 2000. Antibodies to streptococcal surface enolase react with human alpha-enolase: implications in poststreptococcal sequelae. J Infect Dis **182**:1712-21.

79. **Fox, G. E., E. Stackebrandt, R. B. Hespell, J. Gibson, J. Maniloff, T. A. Dyer, R. S. Wolfe, W. E. Balch, R. S. Tanner, L. J. Magrum, L. B. Zablen, R. Blakemore, R. Gupta, L. Bonen, B. J. Lewis, D. A. Stahl, K. R. Luehrsen, K. N. Chen, and C. R. Woese.** 1980. The phylogeny of prokaryotes. Science **209**:457-63.

80. **Fox, J. L.** 2005. Ribosomal gene milestone met, already left in dust, p. 6-7, ASM News, vol. 71.

81. **Fritz, D. L., P. Vogel, D. R. Brown, and D. M. Waag.** 1999. The hamster model of intraperitoneal *Burkholderia mallei* (glanders). Vet Pathol **36**:276-91.

82. **Frost, L. S., K. Ippen-Ihler, and R. A. Skurray.** 1994. Analysis of the sequence and gene products of the transfer region of the F sex factor. Microbiol Rev **58**:162-210.

83. **Garrity, G. M., J. A. Bell, and T. G. Lilburn.** 2004. Taxonomic outline of the prokaryotes Bergey's Manual of systematic bacteriolgy, Second Edition. Springer New York.

84. **Gillis, M., T. Vanvan, R. Bardin, M. Goor, P. Hebbar, A. Willems, P. Segers, K. Kersters, T. Heulin, and M. P. Fernandez.** 1995. Polyphasic taxonomy in the genus *Burkholderia* leading to an emended description of the genus and proposition of *Burkholderia vietnamiensis* sp nov for N2-fixing isolates from rice in vietnam. Int J Syst Bacteriol **45**:274-289.

85. **Godoy, D., G. Randle, A. J. Simpson, D. M. Aanensen, T. L. Pitt, R. Kinoshita, and B. G. Spratt.** 2003. Multilocus sequence typing and evolutionary relationships among the causative agents of melioidosis and glanders, *Burkholderia pseudomallei* and *Burkholderia mallei*. J Clin Microbiol **41**:2068-79.

86. **Gonzales, T., and J. RobertBaudouy.** 1996. Bacterial aminopeptidases: Properties and functions. FEMS Microbiol Rev **18**:319-344.

87. **Goris, J., P. De Vos, J. Caballero-Mellado, J. Park, E. Falsen, J. F. Quensen, 3rd, J. M. Tiedje, and P. Vandamme.** 2004. Classification of the biphenyl- and polychlorinated biphenyl-degrading strain LB400T and relatives as *Burkholderia xenovorans* sp. nov. Int J Syst Evol Microbiol **54**:1677-81.

88. **Goris, J., W. Dejonghe, E. Falsen, E. De Clerck, B. Geeraerts, A. Willems, E. M. Top, P. Vandamme, and P. De Vos.** 2002. Diversity of transconjugants that acquired plasmid pJP4 or pEMT1 after inoculation of a donor strain in the A- and B-horizon of an agricultural soil and description of *Burkholderia hospita* sp. nov. and *Burkholderia terricola* sp. nov. Syst Appl Microbiol **25**:340-52.

89. **Gremion, F., A. Chatzinotas, and H. Harms.** 2003. Comparative 16S rDNA and 16S rRNA sequence analysis indicates that Actinobacteria might be a dominant part of the metabolically active bacteria in heavy metal-contaminated bulk and rhizosphere soil. Environ Microbiol **5**:896-907.

90. **Gupta, R. S., G. B. Golding, and B. Singh.** 1994. Hsp70 phylogeny and the relationship between archaebacteria, eubacteria, and eukaryotes. J Mol Evol **39**:537-40.

91. **Hagedorn, C., W. D. Gould, T. R. Bardinelli, and D. R. Gustavson.** 1987. A selective medium for enumeration and recovery of *Pseudomonas cepacia* biotypes from soil. Appl Environ Microbiol **53**:2265-8.

92. **Hall, T. A.** 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids Symp. Ser. **41**:95-98.

93. **Hallmann, J., R. Rodriguez-Kabana, and J. W. Kloepper.** 1999. Chitin-mediated changes in bacterial communities of the soil, rhizosphere and within roots of cotton in relation to nematode control. Soil Biol Biochem **31**:551-560.

94. **Hansen, J. B., and R. H. Olsen.** 1978. Isolation of large bacterial plasmids and characterization of the P2 incompatibility group plasmids pMG1 and pMG5. J Bacteriol **135**:227-38.

95. **Hantke, K., and V. Braun.** 1973. Covalent binding of lipid to protein - diglyceride and amide-linked fatty-acid at N-terminal end of murein-lipoprotein of *Escherichia coli* outer membrane. Euro J Biochem **34**:284-296.

96. **Hareland, W. A., R. L. Crawford, P. J. Chapman, and S. Dagley.** 1975. Metabolic function and properties of 4-hydroxyphenylacetic acid 1-hydroxylase from *Pseudomonas acidovorans*. J Bacteriol **121**:272-85.

97. **Harold, S., G. M. Tordoff, T. H. Jones, and L. Boddy.** 2005. Mycelial responses of *Hypholoma fasciculare* to collembola grazing: effect of inoculum age, nutrient status and resource quality. Mycol Res **109:**927-935.

98. **Hayatsu, M., M. Hirano, and S. Tokuda.** 2000. Involvement of two plasmids in fenitrothion degradation by *Burkholderia* sp. strain NF100. Appl Environ Microbiol **66:**1737-1740.

99. **Henry, D. A., M. E. Campbell, J. J. LiPuma, and D. P. Speert.** 1997. Identification of *Burkholderia cepacia* isolates from patients with cystic fibrosis and use of a simple new selective medium. J Clin Microbiol **35:**614-9.

100. **Higgins, C. F.** 1992. ABC transporters: from microorganisms to man. Annu Rev Cell Biol **8:**67-113.

101. **Holmes, A., J. Govan, and R. Goldstein.** 1998. Agricultural use of *Burkholderia* (*Pseudomonas*) *cepacia*: a threat to human health? Emerg Infect Dis **4:**221-7.

102. **Holmes, D. E., K. T. Finneran, R. A. O'Neil, and D. R. Lovley.** 2002. Enrichment of members of the family *Geobacteraceae* associated with stimulation of dissimilatory metal reduction in uranium-contaminated aquifer sediments. Appl Environ Microbiol **68:**2300-2306.

103. **Hugenholtzt, P., and T. Huber.** 2003. Chimeric 16S rDNA sequences of diverse origin are accumulating in the public databases. Int J Syst Evol Microbiol **53:**289-93.

104. **Hutchinson, G. R., S. Parker, J. A. Pryor, F. Duncan-Skingle, P. N. Hoffman, M. E. Hodson, M. E. Kaufmann, and T. L. Pitt.** 1996. Home-use nebulizers: a potential primary source of *Burkholderia cepacia* and other colistin-resistant, gram-negative bacteria in patients with cystic fibrosis. J Clin Microbiol **34:**584-7.

105. **Inoue, A., and K. Horikoshi.** 1989. A *Pseudomonas* thrives in high concentrations of toluene. Nature **338:**264-266.

106. **Isles, A., I. Maclusky, M. Corey, R. Gold, C. Prober, P. Fleming, and H. Levison.** 1984. *Pseudomonas cepacia* infection in cystic fibrosis: an emerging problem. J Pediatr **104:**206-10.

107. **Jayaswal, R. K., M. Fernandez, R. S. Upadhyay, L. Visintin, M. Kurz, J. Webb, and K. Rinehart.** 1993. Antagonism of *Pseudomonas cepacia* against phytopathogenic fungi. Curr Microbiol **26:**17-22.

108. **Johnson, G. R., and R. H. Olsen.** 1997. Multiple pathways for toluene degradation in *Burkholderia* sp. strain JS150. Appl Environ Microbiol **63:**4047-4052.

109. **Jones, B. V., R. Young, E. Mahenthiralingam, and D. J. Stickler.** 2004. Ultrastructure of *Proteus mirabilis* swarmer cell rafts and role of swarming in catheter-associated urinary tract infection. Infect Immun **72:**3941-50.

110. **Kaneko, T., Y. Nakamura, C. P. Wolk, T. Kuritz, S. Sasamoto, A. Watanabe, M. Iriguchi, A. Ishikawa, K. Kawashima, T. Kimura, Y. Kishida,**

M. Kohara, M. Matsumoto, A. Matsuno, A. Muraki, N. Nakazaki, S. Shimpo, M. Sugimoto, M. Takazawa, M. Yamada, M. Yasuda, and S. Tabata. 2001. Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120. DNA Res **8:**205-13; 227-53.

111. **Kang, Y., R. Carlson, W. Tharpe, and M. A. Schell.** 1998. Characterization of genes involved in biosynthesis of a novel antibiotic from *Burkholderia cepacia* BC11 and their role in biological control of *Rhizoctonia solani.* Appl Environ Microbiol **64:**3939-47.

112. **Karlin, S., and L. Brocchieri.** 1996. Evolutionary conservation of *recA* genes in relation to protein structure and function. J Bacteriol **178:**1881-94.

113. **Karlin, S., G. M. Weinstock, and V. Brendel.** 1995. Bacterial classifications derived from *recA* protein sequence comparisons. J Bacteriol **177:**6881-93.

114. **Kikuchi, Y., X. Y. Meng, and T. Fukatsu.** 2005. Gut symbiotic bacteria of the genus *Burkholderia* in the broad-headed bugs *Riptortus clavatus* and *Leptocorisa chinensis* (Heteroptera: Alydidae). Appl Environ Microbiol **71:**4035-4043.

115. **Kilbane, J. J., D. K. Chatterjee, J. S. Karns, S. T. Kellogg, and A. M. Chakrabarty.** 1982. Biodegradation of 2,4,5-Trichlorophenoxyacetic acid by a pure culture of *Pseudomonas cepacia.* Appl Environ Microbiol **44:**72-78.

116. **Klages, U., A. Markus, and F. Lingens.** 1981. Degradation of 4-Chlorophenylacetic acid by a *Pseudomonas* species. J Bacteriol **146:**64-68.

117. **Klenk, H. P., and W. Zillig.** 1994. DNA-dependent RNA polymerase subunit-B as a tool for phylogenetic reconstructions - branching topology of the archaeal domain. J Mol Evol **38:**420-432.

118. **Kotetishvili, M., A. Kreger, G. Wauters, J. G. Morris, Jr., A. Sulakvelidze, and O. C. Stine.** 2005. Multilocus sequence typing for studying genetic relationships among *Yersinia* species. J Clin Microbiol **43:**2674-84.

119. **Kumar, S., K. Tamura, and M. Nei.** 2004. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform **5:**150-63.

120. **Lai, S., and L. N. Manchester.** 2000. Numerical phenetic study of the genus *Carnobacterium.* Antonie Van Leeuwenhoek **78:**73-85.

121. **Larsen, G. Y., T. L. Stull, and J. L. Burns.** 1993. Marked phenotypic variability in *Pseudomonas cepacia* isolated from a patient with cystic fibrosis. J Clin Microbiol **31:**788-92.

122. **Laurie, A. D., and G. Lloyd-Jones.** 1999. Conserved and hybrid meta-cleavage operons from PAH-degrading *Burkholderia* RP007. Biochem Biophys Res Comm **262:**308-314.

123. **Lear, J. C., J. Y. Maillard, P. W. Dettmar, P. A. Goddard, and A. D. Russell.** 2002. Chloroxylenol- and triclosan-tolerant bacteria from industrial sources. J Ind Microbiol Biotechnol **29:**238-42.

124. **Lee, E. Y.** 2003. Continuous treatment of gas-phase trichloroethylene by *Burkholderia cepacia* G4 in a two-stage continuous stirred tank reactor/trickling biofilter system. J Biosci Bioeng **96:**572-574.

125. **Leff, L. G., R. M. Kernan, J. V. McArthur, and L. J. Shimkets.** 1995. Identification of aquatic *Burkholderia* (*Pseudomonas*) *cepacia* by hybridization with species-specific rRNA gene probes. Appl Environ Microbiol **61:**1634-6.

126. **Lessie, T. G., W. Hendrickson, B. D. Manning, and R. Devereux.** 1996. Genomic complexity and plasticity of *Burkholderia cepacia*. FEMS Microbiol Lett **144:**117-28.

127. **Leungsakul, T., B. G. Keenan, M. A. Mori, M. D. Morton, J. D. Stuart, B. F. Smets, and T. K. Wood.** 2005. Oxidation of aminonitrotoluenes by 2,4-DNT dioxygenase of *Burkholderia* sp. strain DNT. Biotechnol Bioeng.

128. **Levy, A., B. J. Chang, L. K. Abbott, J. Kuo, G. Harnett, and T. J. J. Inglis.** 2003. Invasion of spores of the arbuscular mycorrhizal fungus *Gigaspora decipiens* by *Burkholderia* spp. Appl Environ Microbiol **69:**6250-6256.

129. **Lim, Y. W., K. S. Baik, S. K. Han, S. B. Kim, and K. S. Bae.** 2003. *Burkholderia sordidicola* sp. nov., isolated from the white-rot fungus *Phanerochaete sordida*. Int J Syst Evol Microbiol **53:**1631-6.

130. **LiPuma, J. J., B. J. Dulaney, J. D. McMenamin, P. W. Whitby, T. L. Stull, T. Coenye, and P. Vandamme.** 1999. Development of rRNA-based PCR assays for identification of *Burkholderia cepacia* complex isolates recovered from cystic fibrosis patients. J Clin Microbiol **37:**3167-70.

131. **LiPuma, J. J., and E. Mahenthiralingam.** 1999. Commercial use of *Burkholderia cepacia*. Emerg Infect Dis **5:**305-6.

132. **LiPuma, J. J., T. Spilker, T. Coenye, and C. F. Gonzalez.** 2002. An epidemic *Burkholderia cepacia* complex strain identified in soil. Lancet **359:**2002-3.

133. **LiPuma, J. J., T. Spilker, L. H. Gill, P. W. Campbell, 3rd, L. Liu, and E. Mahenthiralingam.** 2001. Disproportionate distribution of *Burkholderia cepacia* complex species and transmissibility markers in cystic fibrosis. Am J Respir Crit Care Med **164:**92-6.

134. **Litzner, B. R., T. M. Caton, and M. A. Schneegurt.** 2006. Carbon substrate utilization, antibiotic sensitivity, and numerical taxonomy of bacterial isolates from the Great Salt Plains of Oklahoma. Arch Microbiol

135. **Lodish, H., D. Baltimore, A. Berk, S. L. Zipursky, P. Matsudaira, and J. Darnell.** 1995. Molecular Cell Biology, Third ed. Scientific American Books, New York.

136. **Loffler, F., F. Lingens, and R. Muller.** 1995. Dehalogenation of 4-Chlorobenzoate - Characterization of 4-Chlorobenzoyl-coenzyme a dehalogenase from *Pseudomonas* sp Cbs3. Biodegradation **6:**203-212.

137. **Logan, N. A.** 1994. Bacterial Systematics. Blackwell Scientific Publications, Oxford.

138. **Lonni, A. A., S. Scarminio, L. M. Silva, and D. T. Ferreira.** 2005. Numerical taxonomy characterization of *Baccharis* genus species by ultraviolet-visible spectrophotometry. Anal Sci **21**:235-9.

139. **Lovley, D. R.** 2003. Cleaning up with genomics: applying molecular biology to bioremediation. Nat Rev Microbiol **1**:35-44.

140. **Lowe, P., C. Engler, and R. Norton.** 2002. Comparison of automated and nonautomated systems for identification of *Burkholderia pseudomallei*. J Clin Microbiol **40**:4625-7.

141. **Ludwig, W., J. Neumaier, N. Klugbauer, E. Brockmann, C. Roller, S. Jilg, K. Reetz, I. Schachtner, A. Ludvigsen, M. Bachleitner, U. Fischer, and K. H. Schleifer.** 1993. Phylogenetic relationships of bacteria based on comparative sequence analysis of elongation factor Tu and ATP synthase Beta-subunit genes. Antonie Van Leeuwenhoek **64**:285-305.

142. **Mahenthiralingam, E., A. Baldwin, and P. Vandamme.** 2002. *Burkholderia cepacia* complex infection in patients with cystic fibrosis. J Med Microbiol **51**:533-8.

143. **Mahenthiralingam, E., J. Bischof, S. K. Byrne, C. Radomski, J. E. Davies, Y. Av-Gay, and P. Vandamme.** 2000. DNA-Based diagnostic approaches for identification of *Burkholderia cepacia* complex, *Burkholderia vietnamiensis*, *Burkholderia multivorans*, *Burkholderia stabilis*, and *Burkholderia cepacia* genomovars I and III. J Clin Microbiol **38**:3165-73.

144. **Mahenthiralingam, E., T. Coenye, J. W. Chung, D. P. Speert, J. R. Govan, P. Taylor, and P. Vandamme.** 2000. Diagnostically and experimentally useful panel of strains from the *Burkholderia cepacia* complex. J Clin Microbiol **38**:910-3.

145. **Mahenthiralingam, E., and P. Sayre.** 2003. *Burkholderia cepacia* complex, p. 41-44. *In* M. D. Licker (ed.), McGraw-Hill Yearbook of Science and Technology. McGraw-Hill, New York.

146. **Mahenthiralingam, E., T. A. Urban, and J. B. Goldberg.** 2005. The multifarious, multireplicon *Burkholderia cepacia* complex. Nat Rev Microbiol **3**:144-56.

147. **Maiden, M. C., J. A. Bygraves, E. Feil, G. Morelli, J. E. Russell, R. Urwin, Q. Zhang, J. Zhou, K. Zurth, D. A. Caugant, I. M. Feavers, M. Achtman, and B. G. Spratt.** 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci USA **95**:3140-5.

148. **Manoil, C.** 2000. Tagging exported proteins using *Escherichia coli* alkaline phosphatase gene fusions. Methods Enzymol **326**:35-47.

149. **Mao, W., R. D. Lumsden, J. A. Lewis, and P. K. Hebbar.** 1998. Seed treatment using pre-infiltration and biocontrol agents to reduce damping-off of corn caused by species of *Pythium* and *Fusarium*. Plant Dis **82**:294-299.

150. **Marchesi, J. R., T. Sato, A. J. Weightman, T. A. Martin, J. C. Fry, S. J. Hiom, and W. G. Wade.** 1998. Design and evaluation of useful bacterium-

specific PCR primers that amplify genes coding for bacterial 16S rRNA. Appl Environ Microbiol **64:**795-799.

151. **Margoliash, E.** 1963. Primary structure and evolution of cytochrome C. Proc Natl Acad Sci USA **50:**672-9.

152. **Margulis, L., and M. J. Chapman.** 1998. Endosymbioses: cyclical and permanent in evolution. Trends Microbiol **6:**342-5; discussion 345-6.

153. **Marrs, C. F., G. P. Krasan, K. W. McCrea, D. L. Clemans, and J. R. Gilsdorf.** 2001. *Haemophilus influenzae* - human specific bacteria. Front Biosci **6:**E41-60.

154. **Mars, A. E., J. Houwing, J. Dolfing, and D. B. Janssen.** 1996. Degradation of toluene and trichloroethylene by *Burkholderia cepacia* G4 in growth-limited fed-batch culture. Appl Environ Microbiol **62:**886-891.

155. **Mayer, M. P., and B. Bukau.** 2005. Hsp70 chaperones: cellular functions and molecular mechanism. Cell Mol Life Sci **62:**670-84.

156. **McGowan, C., R. Fulthorpe, A. Wright, and J. M. Tiedje.** 1998. Evidence for interspecies gene transfer in the evolution of 2,4-dichlorophenoxyacetic acid degraders. Appl Environ Microbiol **64:**4089-4092.

157. **Meats, E., E. J. Feil, S. Stringer, A. J. Cody, R. Goldstein, J. S. Kroll, T. Popovic, and B. G. Spratt.** 2003. Characterization of encapsulated and noncapsulated *Haemophilus influenzae* and determination of phylogenetic relationships by multilocus sequence typing. J Clin Microbiol **41:**1623-36.

158. **Meyers, E., G. S. Bisacchi, L. Dean, W. C. Liu, B. Minassian, D. S. Slusarchyk, R. B. Sykes, S. K. Tanaka, and W. Trejo.** 1987. Xylocandin: a new complex of antifungal peptides. Taxonomy, isolation and biological activity. J Antibiot (Tokyo) **40:**1515-9.

159. **Miller, S. C., J. J. LiPuma, and J. L. Parke.** 2002. Culture-based and non-growth-dependent detection of the *Burkholderia cepacia* complex in soil environments. Appl Environ Microbiol **68:**3750-8.

160. **Milling, A., K. Smalla, F. X. Maidl, M. Schloter, and J. C. Munch.** 2004. Effects of transgenic potatoes with an altered starch composition on the diversity of soil and rhizosphere bacteria and fungi. Plant and Soil **266:**23-39.

161. **Minerdi, D., R. Fani, R. Gallo, A. Boarino, and P. Bonfante.** 2001. Nitrogen fixation genes in an endosymbiotic *Burkholderia* strain. Appl Environ Microbiol **67:**725-32.

162. **Mondello, F. J.** 1989. Cloning and expression in *Escherichia coli* of *Pseudomonas* strain LB400 genes encoding polychlorinated biphenyl degradation. J Bacteriol **171:**1725-32.

163. **Monti, M. R., A. M. Smania, G. Fabro, M. E. Alvarez, and C. E. Argarana.** 2005. Engineering *Pseudomonas fluorescens* for biodegradation of 2,4-dinitrotoluene. Appl Environ Microbiol **71:**8864-72.

164. **Moore, R. A., and R. E. Hancock.** 1986. Involvement of outer membrane of *Pseudomonas cepacia* in aminoglycoside and polymyxin resistance. Antimicrob Agents Chemother **30:**923-6.

165. **Mosqueda, G., M. I. Ramos-Gonzalez, and J. L. Ramos.** 1999. Toluene metabolism by the solvent-tolerant *Pseudomonas putida* DOT-T1 strain, and its role in solvent impermeabilization. Gene **232:**69-76.

166. **Moulin, L., A. Munive, B. Dreyfus, and C. Boivin-Masson.** 2001. Nodulation of legumes by members of the beta-subclass of Proteobacteria. Nature **411:**948-950.

167. **Mueller, J. G., R. Devereux, D. L. Santavy, S. E. Lantz, S. G. Willis, and P. H. Pritchard.** 1997. Phylogenetic and physiological comparisons of PAH-degrading bacteria from geographically diverse soils. Antonie Van Leeuwenhoek **71:**329-343.

168. **Naureen, Z., S. Yasmin, S. Hameed, K. A. Malik, and F. Y. Hafeez.** 2005. Characterization and screening of bacteria from rhizosphere of maize grown in Indonesian and Pakistani soils. J Basic Microbiol **45:**447-459.

169. **Nelson, M. J. K., S. O. Montgomery, E. J. Oneill, and P. H. Pritchard.** 1986. Aerobic metabolism of trichloroethylene by a bacterial isolate. Appl Environ Microbiol **52:**383-384.

170. **Nijhuis, E. H., M. J. Maat, I. W. E. Zeegers, C. Waalwijk, and J. A. Vanveen.** 1993. Selection of bacteria suitable for introduction into the rhizosphere of grass. Soil Biol Biochem **25:**885-895.

171. **Nzula, S., P. Vandamme, and J. R. Govan.** 2002. Influence of taxonomic status on the in vitro antimicrobial susceptibility of the *Burkholderia cepacia* complex. J Antimicrob Chemother **50:**265-9.

172. **O'Sullivan, L. A., and E. Mahenthiralingam.** 2005. Biotechnological potential within the genus *Burkholderia*. Lett Appl Microbiol **41:**8-11.

173. **Oie, S., and A. Kamiya.** 1996. Microbial contamination of antiseptics and disinfectants. Am J Infect Control **24:**389-95.

174. **Palleroni, N. J.** 2003. Prokaryote taxonomy of the 20th century and the impact of studies on the genus *Pseudomonas*: a personal view. Microbiol **149:**1-7.

175. **Palleroni, N. J., R. Kunisawa, R. Contopou, and M. Doudorof.** 1973. Nucleic acid homologies in the genus *Pseudomonas*. Int J Syst Bacteriol **23:**333-339.

176. **Pancholi, V., and V. A. Fischetti.** 1998. Alpha-enolase, a novel strong plasmin(ogen) binding protein on the surface of pathogenic streptococci. J Biol Chem **273:**14503-15.

177. **Parke, J. L.** 1990. Population dynamics of *Pseudomonas cepacia* in the pea spermosphere in relation to biocontrol of pythium. Phytopathol **80:**1307-1311.

178. **Parke, J. L., and D. Gurian-Sherman.** 2001. Diversity of the *Burkholderia cepacia* complex and implications for risk assessment of biological control strains. Annu Rev Phytopathol **39:**225-58.

179. **Parke, J. L., R. E. Rand, A. E. Joy, and E. B. King.** 1991. Biological control of pythium damping off and aphanomyces root rot of peas by application of *Pseudomonas cepacia* or *P. fluorescens* to seed. Plant Dis **75:**987-992.

180. **Parker, M. A., J. L. Doyle, and J. J. Doyle.** 2004. Comparative phylogeography of *Amphicarpaea* legumes and their root-nodule symbionts in Japan and North America. J Biogeog **31:**425-434.

181. **Parkes, R. J., G. Webster, B. A. Cragg, A. J. Weightman, C. J. Newberry, T. G. Ferdelman, J. Kallmeyer, B. B. Jorgensen, I. W. Aiello, and J. C. Fry.** 2005. Deep sub-seafloor prokaryotes stimulated at interfaces over geological time. Nature **436:**390-4.

182. **Partida-Martinez, L. P., and C. Hertweck.** 2005. Pathogenic fungus harbours endosymbiotic bacteria for toxin production. Nature **437:**884-8.

183. **Payne, G. W., P. Vandamme, S. H. Morgan, J. J. Lipuma, T. Coenye, A. J. Weightman, T. H. Jones, and E. Mahenthiralingam.** 2005. Development of a *recA* gene-based identification approach for the entire *Burkholderia* genus. Appl Environ Microbiol **71:**3917-27.

184. **Philippot, L., S. Piutti, F. Martin-Laurent, S. Hallet, and J. C. Germon.** 2002. Molecular analysis of the nitrate-reducing community from unplanted and maize-planted soils. Appl Environ Microbiol **68:**6121-6128.

185. **Pirone, L., L. Chiarini, C. Dalmastri, A. Bevivino, and S. Tabacchioni.** 2005. Detection of cultured and uncultured *Burkholderia cepacia* complex bacteria naturally occurring in the maize rhizosphere. Environ Microbiol **7:**1734-42.

186. **Purkhold, U., A. Pommerening-Roser, S. Juretschko, M. C. Schmid, H. P. Koops, and M. Wagner.** 2000. Phylogeny of all recognized species of ammonia oxidizers based on comparative 16S rRNA and *amoA* sequence analysis: Implications for molecular diversity surveys. Appl Environ Microbiol **66:**5368-5382.

187. **Rabkin, C. S., W. R. Jarvis, R. L. Anderson, J. Govan, J. Klinger, J. LiPuma, W. J. Martone, H. Monteil, C. Richard, S. Shigeta, and et al.** 1989. *Pseudomonas cepacia* typing systems: collaborative study to assess their potential in epidemiologic investigations. Rev Infect Dis **11:**600-7.

188. **Ramette, A., J. J. LiPuma, and J. M. Tiedje.** 2005. Species abundance and diversity of *Burkholderia cepacia* complex in the environment. Appl Environ Microbiol **71:**1193-201.

189. **Ramos, J. L., E. Duque, M. T. Gallegos, P. Godoy, M. I. Ramos-Gonzalez, A. Rojas, W. Teran, and A. Segura.** 2002. Mechanisms of solvent tolerance in gram-negative bacteria. Annu Rev Microbiol **56:**743-68.

190. **Ramos, J. L., E. Duque, M. J. Huertas, and A. Haidour.** 1995. Isolation and expansion of the catabolic potential of a *Pseudomonas putida* strain able to

grow in the presence of high concentrations of aromatic hydrocarbons. J Bacteriol **177**:3911-6.

191. **Ramos, J. L., M. T. Gallegos, S. Marques, M. I. Ramos-Gonzalez, M. Espinosa-Urgel, and A. Segura.** 2001. Responses of Gram-negative bacteria to certain environmental stressors. Curr Opin Microbiol **4**:166-71.

192. **Rappe, M. S., and S. J. Giovannoni.** 2003. The uncultured microbial majority. Annu Rev Microbiol **57**:369-94.

193. **Reis, V. M., P. Estrada-de los Santos, S. Tenorio-Salgado, J. Vogel, M. Stoffels, S. Guyon, P. Mavingui, V. L. Baldani, M. Schmid, J. I. Baldani, J. Balandreau, A. Hartmann, and J. Caballero-Mellado.** 2004. *Burkholderia tropica* sp. nov., a novel nitrogen-fixing, plant-associated bacterium. Int J Syst Evol Microbiol **54**:2155-62.

194. **Riesenfeld, C. S., P. D. Schloss, and J. Handelsman.** 2004. Metagenomics: genomic analysis of microbial communities. Annu Rev Genet **38**:525-52.

195. **Ruggiero, C. E., H. Boukhalfa, J. H. Forsythe, J. G. Lack, L. E. Hersman, and M. P. Neu.** 2005. Actinide and metal toxicity to prospective bioremediation bacteria. Environ Microbiol **7**:88-97.

196. **Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajandream, and B. Barrell.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16**:944-5.

197. **Saenz, H. L., and C. Dehio.** 2005. Signature-tagged mutagenesis: technical advances in a negative selection method for virulence gene identification. Curr Opin Microbiol **8**:612-9.

198. **Salles, J. F., F. A. De Souza, and J. D. van Elsas.** 2002. Molecular method to assess the diversity of *Burkholderia* species in environmental samples. Appl Environ Microbiol **68**:1595-603.

199. **Salles, J. F., J. A. van Veen, and J. D. van Elsas.** 2004. Multivariate analyses of *Burkholderia* species in soil: effect of crop and land use history. Appl Environ Microbiol **70**:4012-20.

200. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Press, Cold Spring Harbor, N.Y.

201. **Sandler, S. J., P. Hugenholtz, C. Schleper, E. F. DeLong, N. R. Pace, and A. J. Clark.** 1999. Diversity of *radA* genes from cultured and uncultured *archaea*: comparative analysis of putative RadA proteins and their use as a phylogenetic marker. J Bacteriol **181**:907-15.

202. **Sandler, S. J., L. H. Satin, H. S. Samra, and A. J. Clark.** 1996. *recA*-like genes from three archaean species with putative protein products similar to Rad51 and Dmc1 proteins of the yeast *Saccharomyces cerevisiae*. Nucleic Acids Res **24**:2125-32.

203. **Santacruz, G., E. R. Bandala, and L. G. Torres.** 2005. Chlorinated pesticides (2,4-D and DDT) biodegradation at high concentrations using immobilized *Pseudomonas fluorescens*. J Environ Sci Health B **40:**571-83.

204. **Sapp, J.** 2005. The prokaryote-eukaryote dichotomy: meanings and mythology. Microbiol Mol Biol Rev **69:**292-305.

205. **Schlomann, M., E. Schmidt, and H. J. Knackmuss.** 1990. Different types of dienelactone hydrolase in 4-Fluorobenzoate-utilizing bacteria. Journal of Bacteriology **172:**5112-5118.

206. **Schmalisch, M., I. Langbein, and J. Stulke.** 2002. The general stress protein Ctc of *Bacillus subtilis* is a ribosomal protein. J Mol Microbiol Biotechnol **4:**495-501.

207. **Schussler, A., D. Schwarzott, and C. Walker.** 2001. A new fungal phylum, the Glomeromycota: phylogeny and evolution. Mycol Res **105:**1413-1421.

208. **Scott, K., M. A. Diggle, and S. C. Clarke.** 2003. *TypA* is a virulence regulator and is present in many pathogenic bacteria. Br J Biomed Sci **60:**168-70.

209. **Segonds, C., T. Heulin, N. Marty, and G. Chabanon.** 1999. Differentiation of *Burkholderia* species by PCR-restriction fragment length polymorphism analysis of the 16S rRNA gene and application to cystic fibrosis isolates. J Clin Microbiol **37:**2201-8.

210. **Sessitsch, A., T. Coenye, A. V. Sturz, P. Vandamme, E. A. Barka, J. F. Salles, J. D. Van Elsas, D. Faure, B. Reiter, B. R. Glick, G. Wang-Pruski, and J. Nowak.** 2005. *Burkholderia phytofirmans* sp. nov., a novel plant-associated bacterium with plant-beneficial properties. Int J Syst Evol Microbiol **55:**1187-92.

211. **Shen, F. T., H. L. Lu, J. L. Lin, W. S. Huang, A. B. Arun, and C. C. Young.** 2005. Phylogenetic analysis of members of the metabolically diverse genus *Gordonia* based on proteins encoding the *gyrB* gene. Res Microbiol.

212. **Shields, M. S., S. O. Montgomery, S. M. Cuskey, P. J. Chapman, and P. H. Pritchard.** 1991. Mutants of *Pseudomonas cepacia* G4 defective in catabolism of aromatic compounds and trichloroethylene. Appl Environ Microbiol **57:**1935-1941.

213. **Shields, M. S., M. J. Reagin, R. R. Gerger, R. Campbell, and C. Somerville.** 1995. TOM, a new aromatic degradative plasmid from *Burkholderia* (*Pseudomonas*) *cepacia* G4. Appl Environ Microbiol **61:**1352-6.

214. **Singer, A. C., C. J. van der Gast, and I. P. Thompson.** 2005. Perspectives and vision for strain selection in bioaugmentation. Trends Biotechnol **23:**74-7.

215. **Sinnott, M. L.** 1990. Catalytic mechanisms of enzymatic glycosyl transfer. Chem Rev **90:**1171-1202.

216. **Sneath, P. H., and R. R. Sokal.** 1962. Numerical taxonomy. Nature **193:**855-60.

217. **Speert, D. P., M. Bond, R. C. Woodman, and J. T. Curnutte.** 1994. Infection with *Pseudomonas cepacia* in chronic granulomatous disease: role of nonoxidative killing by neutrophils in host defense. J Infect Dis **170:**1524-31.

218. **Stackebrandt, E., and B. M. Goebel.** 1994. Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. Int J Syst Bacteriol **44:**846-849.

219. **Stanier, R. Y., N. J. Palleroni, and M. Doudoroff.** 1966. The aerobic pseudomonads: a taxonomic study. J Gen Microbiol **43:**159-271.

220. **Steffan, R. J., K. L. Sperry, M. T. Walsh, S. Vainberg, and C. W. Condee.** 1999. Field-scale evaluation of in situ bioaugmentation for remediation of chlorinated solvents in groundwater. Environ Sci Tech **33:**2771-2781.

221. **Stoldt, M., J. Wohnert, M. Gorlach, and L. R. Brown.** 1998. The NMR structure of *Escherichia coli* ribosomal protein L25 shows homology to general stress proteins and glutaminyl-tRNA synthetases. Embo Journal **17:**6377-6384.

222. **Storms, V., N. Van Den Vreken, T. Coenye, E. Mahenthiralingam, J. J. LiPuma, M. Gillis, and P. Vandamme.** 2004. Polyphasic characterisation of *Burkholderia cepacia*-like isolates leading to the emended description of *Burkholderia pyrrocinia*. Syst Appl Microbiol **27:**517-26.

223. **Story, R. M., I. T. Weber, and T. A. Steitz.** 1992. The structure of the *E. coli recA* protein monomer and polymer. Nature **355:**318-25.

224. **Stryer, L.** 1995. Biochemistry, Fourth ed. W. H. Freeman and Company, New York.

225. **Suzuki, T., T. Murai, I. Fukuda, T. Tobe, M. Yoshikawa, and C. Sasakawa.** 1994. Identification and characterization of a chromosomal virulence gene, *vacJ*, required for intercellular spreading of *Shigella flexneri*. Mol Microbiol **11:**31-41.

226. **Swannell, R. P., K. Lee, and M. McDonagh.** 1996. Field evaluations of marine oil spill bioremediation. Microbiol Rev **60:**342-65.

227. **Thompson, F. L., D. Gevers, C. C. Thompson, P. Dawyndt, S. Naser, B. Hoste, C. B. Munn, and J. Swings.** 2005. Phylogeny and molecular identification of vibrios on the basis of multilocus sequence analysis. Appl Environ Microbiol **71:**5107-15.

228. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res **22:**4673-80.

229. **Tiboni, O., P. Cammarano, and A. M. Sanangelantoni.** 1993. Cloning and sequencing of the gene encoding glutamine synthetase I from the archaeum *Pyrococcus woesei:* anomalous phylogenies inferred from analysis of archaeal and bacterial glutamine synthetase I sequences. J Bacteriol **175:**2961-9.

230. **Uilenberg, G., F. Thiaucourt, and F. Jongejan.** 2004. On molecular taxonomy: what is in a name? Exp Appl Acarol **32:**301-12.

231. **Urakami, T., C. Itoyoshida, H. Araki, T. Kijima, K. I. Suzuki, and K. Komagata.** 1994. Transfer of *Pseudomonas plantarii* and *Pseudomonas glumae* to *Burkholderia* as *Burkholderia* spp and description of *Burkholderia vandii* sp nov. Int J Syst Bacteriol **44:**235-245.

232. **Urbance, J. W., J. Cole, P. Saxman, and J. M. Tiedje.** 2003. BSD: the Biodegradative Strain Database. Nucleic Acids Res **31:**152-5.

233. **van Borm, S., A. Buschinger, J. J. Boomsma, and J. Billen.** 2002. Tetraponera ants have gut symbionts related to nitrogen-fixing root-nodule bacteria. Proc R Soc Lon Ser Biol Sci **269:**2023-2027.

234. **Van Oevelen, S., R. De Wachter, P. Vandamme, E. Robbrecht, and E. Prinsen.** 2004. '*Candidatus* Burkholderia calva' and '*Candidatus* Burkholderia nigropunctata' as leaf gall endosymbionts of African *Psychotria*. Int J Syst Evol Microbiol **54:**2237-9.

235. **Van Oevelen, S., R. De Wachter, P. Vandamme, E. Robbrecht, and E. Prinsen.** 2002. Identification of the bacterial endosymbionts in leaf galls of *Psychotria* (Rubiaceae, angiosperms) and proposal of '*Candidatus* Burkholderia kirkii' sp nov. Int J Syst Evol Microbiol **52:**2023-2027.

236. **Van, V. T., O. Berge, S. N. Ke, J. Balandreau, and T. Heulin.** 2000. Repeated beneficial effects of rice inoculation with a strain of *Burkholderia vietnamiensis* on early and late yield components in low fertility sulphate acid soils of Vietnam. Plant and Soil **218:**273-284.

237. **Vandamme, P., J. Goris, W. M. Chen, P. de Vos, and A. Willems.** 2002. *Burkholderia tuberum* sp. nov. and *Burkholderia phymatum* sp. nov., nodulate the roots of tropical legumes. Syst Appl Microbiol **25:**507-12.

238. **Vandamme, P., D. Henry, T. Coenye, S. Nzula, M. Vancanneyt, J. J. LiPuma, D. P. Speert, J. R. Govan, and E. Mahenthiralingam.** 2002. *Burkholderia anthina* sp. nov. and *Burkholderia pyrrocinia*, two additional *Burkholderia cepacia* complex bacteria, may confound results of new molecular diagnostic tools. FEMS Immunol Med Microbiol **33:**143-9.

239. **Vandamme, P., B. Holmes, T. Coenye, J. Goris, E. Mahenthiralingam, J. J. LiPuma, and J. R. Govan.** 2003. *Burkholderia* cenoc*epacia* sp. nov.-a new twist to an old story. Res Microbiol **154:**91-6.

240. **Vandamme, P., B. Holmes, M. Vancanneyt, T. Coenye, B. Hoste, R. Coopman, H. Revets, S. Lauwers, M. Gillis, K. Kersters, and J. R. Govan.** 1997. Occurrence of multiple genomovars of *Burkholderia cepacia* in cystic fibrosis patients and proposal of *Burkholderia multivorans* sp. nov. Int J Syst Bacteriol **47:**1188-200.

241. **Vandamme, P., E. Mahenthiralingam, B. Holmes, T. Coenye, B. Hoste, P. De Vos, D. Henry, and D. P. Speert.** 2000. Identification and population structure of *Burkholderia stabilis* sp. nov. (formerly *Burkholderia cepacia* genomovar IV). J Clin Microbiol **38:**1042-7.

242. **Vandamme, P., B. Pot, M. Gillis, P. de Vos, K. Kersters, and J. Swings.** 1996. Polyphasic taxonomy, a consensus approach to bacterial systematics. Microbiol Rev **60:**407-38.

243. **Vaneechoutte, M., P. Kampfer, T. De Baere, E. Falsen, and G. Verschraegen.** 2004. *Wautersia* gen. nov., a novel genus accommodating the phylogenetic lineage including *Ralstonia eutropha* and related species, and proposal of *Ralstonia* [*Pseudomonas*] *syzygii* (Roberts et al. 1990) comb. nov. Int J Syst Evol Microbiol **54:**317-27.

244. **Vanlaere, E., T. Coenye, E. Samyn, C. Van den Plas, J. Govan, F. De Baets, K. De Boeck, C. Knoop, and P. Vandamme.** 2005. A novel strategy for the isolation and identification of environmental *Burkholderia cepacia* complex bacteria. FEMS Microbiol Lett **249:**303-7.

245. **Venter, J. C., K. Remington, J. F. Heidelberg, A. L. Halpern, D. Rusch, J. A. Eisen, D. Wu, I. Paulsen, K. E. Nelson, W. Nelson, D. E. Fouts, S. Levy, A. H. Knap, M. W. Lomas, K. Nealson, O. White, J. Peterson, J. Hoffman, R. Parsons, H. Baden-Tillson, C. Pfannkoch, Y. H. Rogers, and H. O. Smith.** 2004. Environmental genome shotgun sequencing of the Sargasso Sea. Science **304:**66-74.

246. **Ventura, M., D. van Sinderen, G. F. Fitzgerald, and R. Zink.** 2004. Insights into the taxonomy, genetics and physiology of *bifidobacteria*. Antonie Van Leeuwenhoek **86:**205-23.

247. **Vermis, K., T. Coenye, J. J. LiPuma, E. Mahenthiralingam, H. J. Nelis, and P. Vandamme.** 2004. Proposal to accommodate *Burkholderia cepacia* genomovar VI as *Burkholderia dolosa* sp. nov. Int J Syst Evol Microbiol **54:**689-91.

248. **Vermis, K., T. Coenye, E. Mahenthiralingam, H. J. Nelis, and P. Vandamme.** 2002. Evaluation of species-specific *recA*-based PCR tests for genomovar level identification within the *Burkholderia cepacia* complex. J Med Microbiol **51:**937-40.

249. **Viale, A. M., A. K. Arakaki, F. C. Soncini, and R. G. Ferreyra.** 1994. Evolutionary relationships among eubacterial groups as inferred from *GroEL* (chaperonin) sequence comparisons. Int J Syst Bacteriol **44:**527-33.

250. **Viallard, V., I. Poirier, B. Cournoyer, J. Haurat, S. Wiebkin, K. Ophel-Keller, and J. Balandreau.** 1998. *Burkholderia graminis* sp, nov., a rhizospheric *Burkholderia* species, and reassessment of [*Pseudomonas*] *phenazinium*, [*Pseudomonas*] *pyrrocinia* and [*Pseudomonas*] *glathei* as *Burkholderia*. Int J Syst Bacteriol **48:**549-563.

251. **Vinuesa, P., C. Silva, M. J. Lorite, M. L. Izaguirre-Mayoral, E. J. Bedmar, and E. Martinez-Romero.** 2005. Molecular systematics of rhizobia based on maximum likelihood and Bayesian phylogenies inferred from *rrs*, *atpD*, *recA* and *nifH* sequences, and their use in the classification of *Sesbania* microsymbionts from Venezuelan wetlands. Syst Appl Microbiol **28:**702-16.

252. **Wayne, L. G., D. J. Brenner, R. R. Colwell, P. A. D. Grimont, O. Kandler, M. I. Krichevsky, L. H. Moore, W. E. C. Moore, R. G. E. Murray, E. Stackebrandt, M. P. Starr, and H. G. Truper.** 1987. Report of the ad hoc

committee on reconciliation of approaches to bacterial systematics. Int J Syst Bacteriol **37:**463-464.

253. **Webster, G., C. J. Newberry, J. C. Fry, and A. J. Weightman.** 2003. Assessment of bacterial community structure in the deep sub-seafloor biosphere by 16S rDNA-based techniques: a cautionary tale. J Microbiol Methods **55:**155-64.

254. **Weir, B. S., S. J. Turner, W. B. Silvester, D. C. Park, and J. A. Young.** 2004. Unexpectedly diverse *Mesorhizobium* strains and *Rhizobium leguminosarum* nodulate native legume genera of New Zealand, while introduced legume weeds are nodulated by *Bradyrhizobium* species. Appl Environ Microbiol **70:**5980-5987.

255. **Wells, J. M., M. J. Harris, and L. Boddy.** 1998. Encounter with new resources causes polarized growth of the cord-forming basidiomycete *Phanerochaete velutina* on soil. FEMS Micro Ecol **36:**372-382.

256. **Woese, C. R.** 1987. Bacterial evolution. Microbiol Rev **51:**221-71.

257. **Woese, C. R.** 2000. Interpreting the universal phylogenetic tree. Proc Natl Acad Sci USA **97:**8392-6.

258. **Woese, C. R., O. Kandler, and M. L. Wheelis.** 1990. Towards a natural system of organisms: proposal for the domains *Archaea*, *Bacteria*, and *Eucarya*. Proc Natl Acad Sci USA **87:**4576-9.

259. **Yabuuchi, E., Y. Kawamura, T. Ezaki, M. Ikedo, S. Dejsirilert, N. Fujiwara, T. Naka, and K. Kobayashi.** 2000. *Burkholderia uboniae* sp. nov., L-arabinose-assimilating but different from *Burkholderia thailandensis* and *Burkholderia vietnamiensis*. Microbiol Immunol **44:**307-17.

260. **Yabuuchi, E., Y. Kosako, H. Oyaizu, I. Yano, H. Hotta, Y. Hashimoto, T. Ezaki, and M. Arakawa.** 1992. Proposal of *Burkholderia* gen. nov. and transfer of seven species of the genus *Pseudomonas* homology group II to the new genus, with the type species *Burkholderia cepacia* (Palleroni and Holmes 1981) comb. nov. Microbiol Immunol **36:**1251-75.

261. **Yabuuchi, E., Y. Kosako, I. Yano, H. Hotta, and Y. Nishiuchi.** 1995. Transfer of two *Burkholderia* and an *Alcaligenes* species to *Ralstonia* gen. Nov.: Proposal of *Ralstonia pickettii* (Ralston, Palleroni and Doudoroff 1973) comb. Nov., *Ralstonia solanacearum* (Smith 1896) comb. Nov. and *Ralstonia eutropha* (Davis 1969) comb. nov. Microbiol Immunol **39:**897-904.

262. **Yamamoto, S., and S. Harayama.** 1996. Phylogenetic analysis of *Acinetobacter* strains based on the nucleotide sequences of *gyrB* genes and on the amino acid sequences of their products. Int J Syst Bacteriol **46:**506-511.

263. **Zaitsev, G. M., T. V. Tsoi, V. G. Grishenkov, E. G. Plotnikova, and A. M. Boronin.** 1991. Genetic control of degradation of chlorinated benzoic acids in *Arthrobacter globiformis, Corynebacterium sepedonicum* and *Pseudomonas cepacia* strains. FEMS Microbiol Lett **81:**171-176.

264. **Zhang, H., S. Hanada, T. Shigematsu, K. Shibuya, Y. Kamagata, T. Kanagawa, and R. Kurane.** 2000. *Burkholderia kururiensis* sp. nov., a trichloroethylene (TCE)-degrading bacterium isolated from an aquifer polluted with TCE. Int J Syst Evol Microbiol **50 Pt 2:**743-9.

265. **Zhou, H., F. Yao, D. P. Roberts, and T. G. Lessie.** 2003. AHL-deficient mutants of *Burkholderia ambifaria* BC-F have decreased antifungal activity. Curr Microbiol **47:**174-9.

266. **Zolg, W., and J. C. G. Ottow.** 1975. *Pseudomonas glathei* sp nov a new nitrogen scavenging rod isolated from acid lateritic relicts in Germany. Zeitschrift Fur Allgemeine Mikrobiologie **15:**287-299.

APPLIED AND ENVIRONMENTAL MICROBIOLOGY, July 2005, p. 3917–3927 Vol. 71, No. 7
0099-2240/05/$08.00+0 doi:10.1128/AEM.71.7.3917–3927.2005

Development of a *recA* Gene-Based Identification Approach for the Entire *Burkholderia* Genus

George W. Payne,[1] Peter Vandamme,[2] Sara H. Morgan,[1] John J. LiPuma,[3] Tom Coenye,[2] Andrew J. Weightman,[1] T. Hefin Jones,[1] and Eshwar Mahenthiralingam[1]*

Cardiff School of Biosciences, Cardiff University, Cardiff, Wales CF10 3TL, United Kingdom[1]; Laboratorium voor Microbiologie, Faculteit Wetenschappen Vakgroep Biochemie, Fysiologie en Microbiologie, Universiteit Gent, B-9000 Ghent, Belgium[2]; and Department of Pediatrics, University of Michigan, Ann Arbor, Michigan 48109-0646[3]

Received 15 November 2004/Accepted 4 February 2005

***Burkholderia* is an important bacterial genus containing species of ecological, biotechnological, and pathogenic interest. With their taxonomy undergoing constant revision and the phenotypic similarity of several species, correct identification of *Burkholderia* is difficult. A genetic scheme based on the *recA* gene has greatly enhanced the identification of *Burkholderia cepacia* complex species. However, the PCR developed for the latter approach was limited by its specificity for the complex. By alignment of existing and novel *Burkholderia recA* sequences, we designed new PCR primers and evaluated their specificity by testing a representative panel of *Burkholderia* strains. PCR followed by restriction fragment length polymorphism analysis of an 869-bp portion of the *Burkholderia recA* gene was not sufficiently discriminatory. Nucleotide sequencing followed by phylogenetic analysis of this *recA* fragment differentiated both putative and known *Burkholderia* species and all members of the *B. cepacia* complex. In addition, it enabled the design of a *Burkholderia* genus-specific *recA* PCR that produced a 385-bp amplicon, the sequence of which was also able to discriminate all species examined. Phylogenetic analysis of 188 novel *recA* genes enabled clarification of the taxonomic position of several important *Burkholderia* strains and revealed the presence of four novel *B. cepacia* complex *recA* lineages. Although the *recA* phylogeny could not be used as a means to differentiate *B. cepacia* complex strains recovered from clinical infection versus the natural environment, it did facilitate the identification of clonal strain types of *B. cepacia*, *B. stabilis*, and *B. ambifaria* capable of residing in both niches.**

Burkholderia is a genus with complex taxonomy that currently contains 34 validly described species (5), nine of which are a closely related group, known as the *Burkholderia cepacia* complex (6). *Burkholderia* species are widely distributed in the natural environment, and although the majority appear to live either freely or as symbionts or commensals with a variety of higher organisms, several species also cause disease (5). Plant-pathogenic species include *Burkholderia glumae* and *Burkholderia plantari*, which are important rice pathogens. The genus also includes mammalian primary pathogens such as *Burkholderia pseudomallei*, the cause of meliodosis in humans, and *Burkholderia mallei*, which causes glanders in horses; both species have attracted recent interest as potential bioterrorism agents (19).

Many other *Burkholderia* species are capable of causing opportunistic infections in humans and animals; for example, the *B. cepacia* complex (6) can cause serious infections in persons with cystic fibrosis (20) and other vulnerable individuals (31), as well as disease in plants (8) and animals (3). In contrast to these detrimental pathogenic properties, several *Burkholderia* species have considerable commercial and ecological importance. They have been used in agriculture as biopesticides and plant growth promoters (27), and in the bioremediation of major pollutants such as trichloroethylene (30) and polychlorinated biphenyls (26).

The taxonomy and identification of the genus *Burkholderia* are complex, with new species being described rapidly (5, 6). Closely related species such as the *B. cepacia* complex are difficult to identify using conventional biochemical and phenotypic tests, and species belonging to other betaproteobacterial genera (including *Pandoraea* and *Ralstonia*) may be misidentified as *Burkholderia* species (6). A polyphasic taxonomic approach (39) utilizing multiple diagnostic tests is often required to identify *Burkholderia* species accurately.

Although 16S rRNA gene sequence analysis forms an integral part of taxonomical analysis for many bacterial genera (39), its utility in the genus *Burkholderia* is more limited, especially within the *B. cepacia* complex, where it cannot be used as a means to accurately distinguish all species (16, 21). The *recA* gene has been widely applied in bacterial systematics (13) and has proven very useful for the identification of *B. cepacia* complex species, with phylogenetic analysis of sequence variation within the gene enabling discrimination of all nine current species within the *B. cepacia* complex (21). However, the PCR primers designed for the original *recA*-based approach, BCR1 and BCR2, are specific only to members of the *B. cepacia* complex and fail to amplify this gene from other *Burkholderia* species (21). While this can be used as a positive means to confirm an isolate's position within the complex, it limits the application of the approach to identify other *Burkholderia* species in diverse natural habitats.

Given the ecological, biotechnological, and pathogenic im-

* Corresponding author. Mailing address: Cardiff School of Biosciences, Main Building, Museum Avenue, PO Box 915, Cardiff University, Cardiff, Wales CF10 3TL, United Kingdom. Phone: 44(0)29 20875875. Fax: 44 (0)29 20874305. E-mail: MahenthiralingamE @cardiff.ac.uk.

portance of these bacteria, there is a clear need for a molecular diagnostic scheme capable of discrimination across all *Burkholderia* spp. This paper describes the use of genome sequence data from several *Burkholderia* species genome sequencing projects, in combination with an extensive collection of *recA* sequences from *B. cepacia* complex bacteria (21), to develop and evaluate a scheme for identification of all *Burkholderia* spp. based on the *recA* gene. New PCR primers were designed to amplify 87% of the *recA* gene, and novel *recA* sequence data were obtained from known and unknown species. The large size of this *recA* PCR product enabled the development of a PCR-restriction fragment length polymorphism method. Once *recA* sequence information for each *Burkholderia* species had been obtained, PCR primers specific for the genus were designed and tested. Finally, to further examine the phylogenetic relationships between strains of clinical and environmental origin, the *recA* genes from a large collection of *B. cepacia* complex strains were sequenced and compared.

MATERIALS AND METHODS

Bacterial strains, identification, and culture. Molecular identification approaches were developed and evaluated on the collection of strains listed in Table 1. The panel was selected to be representative of the current species diversity of *Burkholderia* and contained 28 species from the genus, with the exception of the primary pathogens *B. pseudomallei* and *B. mallei*. These *Burkholderia* species and all *B. cepacia* complex strains were obtained from the Belgian Co-ordinated Collections of Microorganisms (BCCM), LMG Bacteria collection, the Cardiff University collection (over 800 isolates) (2, 21), and U.S. *B. cepacia* Complex Research Laboratory and Repository (18) and included representatives of the published strain panels (7, 22). Nine isolates representing putative novel *Burkholderia* species, a selection of non-*Burkholderia* control species for PCR, and four isolates of the closely related genus *Pandoraea* were also included in the test strain panel (Table 1). The *Burkholderia ubonensis* type strain, which appears to be a new species member of the *B. cepacia* complex (42), was also included in the study. The ecologically and biotechnologically relevant *B. cepacia* complex strains examined are described in Table 2. All *Burkholderia* species were cultured and identified as described previously (6, 21).

Chromosomal DNA extraction. DNA was prepared for PCR amplification from overnight cultures as described previously (21). Rapid DNA extraction by boiling in the presence of a chelating resin was also carried out as follows. Bacteria from 1 ml of a liquid culture were harvested by centrifugation and resuspended in 100 μl of a sterile solution containing 5% Chelex 100 (Sigma Aldrich, Gillingham, United Kingdom). The suspension was boiled for 5 min and immediately placed on ice for a further 5 min, and then this procedure was repeated. The final supernatant was recovered after centrifugation and stored at −20°C. Before use DNA was diluted in sterile water, and approximately 20 ng of template DNA was incorporated into each PCR.

PCR analysis. PCR was performed as described previously (21) using QIAGEN reagents (QIAGEN Ltd., Crawley, United Kingdom). Each 25-μl PCR contained the following: 1 U *Taq* polymerase, 250 μM of each deoxynucleoside triphosphate, 1x PCR buffer (including 1.5 mM $MgCl_2$), 10 pmol of each appropriate oligonucleotide primer, and 10 to 50 ng of template DNA.

PCR of *B. cepacia* complex *recA* genes was carried out using primers BCR1 and BCR2 as previously described (21). New primers for specific amplification and sequencing of *Burkholderia* species *recA* were designed: BUR1, GATC GA(AG)AAGCAGTTCGGCAA, and BUR2, TTGTCCTTGCCCTG(AG)C CGAT, amplifying an 869-bp fragment; and BUR3, GA(AG) AAG CAG TTC GGC AA, and BUR4, GAG TCG ATG ACG ATC AT, amplifying a 385-bp fragment. OligoCheck (1; http://www.cf.ac.uk/biosi/research/biosoft) was used to assist in primer design and analyze primers BUR1 and BUR2 by rating the likelihood of publicly available bacterial *recA* sequences being amplified. Thermal cycling was carried out in a Flexigene thermal cycler (Techgene Ltd., Cambridge, United Kingdom) for 30 cycles of 30 s at 94°C, annealing for 30 s at 60°C, and extension at 72°C for 45 s, with a final 5-min extension at 72°C. Approximately 2 μl of each PCR product was visualized by agarose gel electrophoresis as described previously (21).

Restriction fragment length polymorphism analysis. Restriction fragment length polymorphism (RFLP) fingerprinting of the *recA* amplicons of *B. cepacia* complex bacteria were performed as previously described (21). RFLP analysis of *Burkholderia* species and other strains (Table 1) was performed on 5 μl of PCR product in a mixture containing the appropriate restriction enzyme buffer and restriction endonuclease as outlined by the manufacturer (Promega, Southampton, United Kingdom) and incubated at 37°C for 4 h. Analysis of the BUR1 and BUR2 *recA* amplicons was performed with HaeIII. RFLP patterns obtained from *recA* were recorded and compared using computer software (GeneSnap, GeneTools, and GeneDirectory, Syngene, Cambridge, United Kingdom). RFLP pattern similarity was calculated as a Dice coefficient at 1% tolerance and clustered using the unweighted pair-group method average. RFLP patterns with a similarity index of 0.75 or higher were clustered as a single group. Macrorestriction and pulsed-field gel electrophoresis (PFGE) analysis was performed as described previously (38); PFGE fingerprints were also compared and analyzed using computer software as described above.

Nucleotide sequence analysis. All *recA* PCR products were sequenced directly using the appropriate primers as described previously (21) as well as those developed in this study (BUR1 and BUR2). Sequencing reactions were prepared using Applied Biosystems Big Dye Terminator ready reaction mix version 3.1 in accordance with the manufacturer's instructions and analyzed using Applied Biosystems ABI-Prism 3100 genetic analyzer capillary electrophoresis system running Applied Biosystems Performance Optimized Polymer 6 (POP-6). Raw sequences from both strands of the PCR products were aligned, and a consensus sequence was derived using the CAP contig assembly program within the BioEdit software (10). Analysis also involved the use of Basic Local Alignment Search Tool (BLAST; www.ncbi.nlm.nih.gov) to establish the correct gene identity.

Phylogenetic analysis. Multiple nucleotide sequence alignments spanning 760 nucleotides of the *recA* gene were constructed using CLUSTAL W (34). Phylogenetic and molecular evolutionary analyses were conducted using genetic-distance-based neighbor-joining algorithms within MEGA version 2.1 (http://www.megasoftware.net/). Gaps and missing data were completely deleted in MEGA before trees were constructed using the Jukes and Cantor matrix model with random sequence input order and 1,000 data sets examined by bootstrapping. All trees were rooted with the *Pseudomonas aeruginosa* PAO1 *recA* gene.

Nucleotide sequence accession numbers and aligned sequence sets. Novel *recA* nucleotide sequences were determined for 28 *Burkholderia* strains, four *Pandoraea* strains, and the *B. ubonensis* type strain (GenBank accession numbers are listed in Table 1). Sequences were also determined for 106 *B. cepacia* complex strains representative of the diversity seen in the *recA* gene RFLP, and these have been submitted under accession numbers AF143782, AF143797, AF143800, AF456003 to AF456015, AF456017, AF456018, AF456020, AF456023, AF456024, AF456026 to AF456028, AF456035, AF456036, AF456038, AF456039 to AF456050, AF456052, AF456054, AF456056, AF456057, AF456062 to AF456064, AF456066, AF456067, AF456069 to AF456083, AF456085 to AF456124, and AY753187. Phylogenetic analysis was performed on the latter novel sequences, 52 previously published *recA* genes (4, 21, 35, 36, 41, 42), and four sequences obtained from the following genome sequence projects: *B. cenocepacia* strain J2315 (NC_004503; www.sanger.ac.uk/Projects/B_cenocepacia/), *B. vietnamiensis* strain G4 (NZ_AAEH00000000; http://genome.jgi-psf.org/draft_microbes/bur08/bur08.home.html), *B. cepacia* strain ATCC 17660 (strain 383; http://genome.jgi-psf.org/draft_microbes/bur94/bur94.home.html), and the *Burkholderia* Sargasso Sea Metagenome (40) strain SAR-1 (NS_000028; www.ncbi.nlm.nih.gov/genomes/static/es.html). Aligned sequence sets of all the *recA* sequences used in this study are available from ftp://cepacia.bios.cf.ac.uk

RESULTS

Design of PCR probes and amplification of *Burkholderia* species *recA* gene. Data for the design of a *Burkholderia* genus-specific *recA* PCR were obtained from the alignment of 10 published *recA* genes spanning the *B. cepacia* complex (21) and genomic regions spanning the *recA* genes from the *B. xenovorans* LB400^T, *B. cenocepacia* J2315, and *B. pseudomallei* K96243 genome sequence projects. Few primer sites capable of amplification across the genus were identified in silico either within or just outside the *recA* coding sequence. Sites that facilitated amplification of a large (869-bp) internal fragment suitable for both RFLP and sequence analysis were selected and used to design primers BUR1 (21-mer; spanning bases 72 to 92 relative to the *B. cenocepacia* J2315 genome *recA* gene

TABLE 1. Bacterial strains used for *recA* PCR primer development and evaluation

Species	Strain	*recA* GenBank accession no.[a]	BUR1 and BUR2 PCR	BUR1 and BUR2 RFLP type	BUR3 and BUR4 PCR
Pandoraea species					
Pandoraea pulmonicola	#40	AY619660	Positive	67	Negative
Pandoraea apista	Patient W	AY619657	Positive	04	Negative
Pandoraea sputorum	AU0012	AY619659	Positive	16	Negative
Pandoraea pnomenusa	Keulen	AY619658	Positive	15	Negative
Known *Burkholderia* species					
Burkholderia kururiensis	LMG 19447^T	AY619654	Positive	80	Positive
Burkholderia caledonica	LMG 19076^T	AY619669	Positive	78	Positive
Burkholderia phenazinium	LMG 2247^T	AY619668	Positive	—	Positive
Burkholderia plantarii	LMG 9035^T	AY619655	Positive	08	Positive
Burkholderia glathei	LMG 14190^T	AY619666	Positive	10	Positive
Burkholderia glumae	LMG 2196^T	AY619675	Positive	89	Positive
Burkholderia gladioli	LMG 2216^T	AY619665	Positive	70	Positive
Burkholderia fungorum	LMG 16225^T	AY619664	Positive	12	Positive
Burkholderia caryophylli	LMG 2155^T	AY619663	Positive	53	Positive
Burkholderia graminis	LMG 18924^T	AY619653	Positive	77	Positive
Burkholderia caribensis	LMG 18531^T	AY619662	Positive	18	Positive
Burkholderia thailandensis	LMG 20219^T	AY619656	Positive	83	Positive
Burkholderia sacchari	LMG 19450^T	AY619661	Positive	71	Positive
Burkholderia terricola	LMG 20594^T	AY619672	Positive	85	Positive
Burkholderia tuberum	LMG 21444^T	AY619674	Positive	86	Positive
Burkholderia phymatum	LMG 21445^T	AY619667	Positive	87	Positive
Burkholderia xenovorans	LB400^T	AAAJ00000000	Positive	01	Positive
Burkholderia andropogonis	LMG 2129^T	—	Negative	—	Negative
Burkholderia hospita	LMG 20598^T	—	Positive[b]	—	Positive
Indeterminate *Burkholderia* species					
	R-20943	AY619679	Positive	71	Positive
	R-8349	AY619681	Positive	84	Positive
	R-15273	AY619677	Positive	81	Positive
	R-701	AY619680	Positive	90	Positive
	R-15821	AY619678	Positive	01	Negative
	R-13392	AY619676	Positive	77	Positive
	LMG 21262	AY619673	Positive	02	Positive
	LMG 19510	AY619670	Positive	82	Positive
	LMG 20580	AY619671	Positive	84	Positive
B. cepacia complex species					
Burkholderia vietnamiensis	LMG 10929^T	AF143793	Positive	66	Positive
Burkholderia multivorans	C1576	AF143774	Positive	73	Positive
Burkholderia multivorans	LMG 13010^T	—	Positive	—	Positive
Burkholderia cepacia	ATCC 25416^T	AF143786	Positive	88	Positive
Burkholderia cenocepacia	J2315/LMG16656^T	www.sanger.ac.uk	Positive	20	Positive
Burkholderia stabilis	LMG 14294^T	AF456031	Positive	63	Positive
Burkholderia pyrrocinia	LMG 14191^T	AF143794 BPP	Positive	75	Positive
Burkholderia ambifaria	LMG 19182^T	AF323985	Positive	48	Positive
Burkholderia dolosa	LMG 18943^T	AF323971	Positive	58	Positive
Burkholderia anthina	LMG 20980^T	AF456059	Positive	57	Positive
Burkholderia cenocepacia	CFLG	AF456021	Positive	25	Positive
Burkholderia ubonensis	LMG 20358^T	AY780511	—	—	—
PCR control species					
Bordetella parapertussis	LMG 14449^T	—	Positive	—	Negative
Rhizobium vitis	LMG 8750^T	—	Negative	—	Negative
Xanthomonas sacchari	LMG 471^T	—	Positive	—	Negative
Ralstonia metallidurans	LMG 1195^T	—	Negative	—	Negative
Ralstonia gilardii	LMG 5886^T	—	Positive	—	Negative
Ralstonia eutropha	JMP134/LMG1197	www.jgi.doe.gov	Positive	—	Negative
Neisseria elongata	LMG 5124^T	—	Negative	—	Negative
Mycobacterium smegmatis	MC^{2}155	—	Negative	—	Negative
Pseudomonas aeruginosa	C3719	—	Negative	—	Negative
Pseudomonas aeruginosa	PAO1	NC_002156	Negative	—	Negative

[a] —, *recA* sequence or RFLP not available or not determined.

[b] PCR amplification with BUR1 and BUR2 resulted in one product of the correct size, 869 bp, and another of approximately 400 bp which could not be resolved by optimization.

TABLE 2. Characteristics of selected *B. cepacia* complex strains

Species (*recA* phylogenetic cluster)	Strain name (other strain designations)	Source and relevant information	Reference(s)
B. cepacia (genomovar I type strain cluster)	ATCC 49709	Grass seed biological control strain	11
B. cepacia (genomovar I group K)	SAR-1	*Burkholderia* species metagenomic strain from Sargasso Sea	40
B. cepacia (genomovar I group K)	383 (ATCC 17660; LMG 6991)	Forest soil, Trinidad; genome draft available (http://genome.jgi-psf.org/draft_microbes/bur94/bur94.home.html)	32
B. cepacia (genomovar I type strain cluster)	ATCC 25416^T (LMG 1222)	Onion rot; *B. cepacia* type strain	32, 37
B. cepacia (genomovar I type strain cluster)	J1050	Clinical infection, USA	28
B. cepacia (genomovar I type strain cluster)	ATCC 17759 (LMG 2161)	Forest soil, Trinidad	32
B. cepacia (genomovar I type strain cluster)	LMG 14087	Wound swab, UK	37
B. cenocepacia (III-B)	M36	Corn rhizosphere, USA; registered biopesticide withdrawn from commercial use; type Wisconsin strain; encodes *B. cenocepacia* pathogenicity island	27
B. cenocepacia (III-B)	BC-1	Corn rhizosphere, USA; biological control strain; encodes *B. cenocepacia* pathogenicity island	24
B. cenocepacia (III-B)	BC-2	Corn rhizosphere, USA; biological control strain; encodes *B. cenocepacia* pathogenicity island	24
B. stabilis[a]	HI-2482	Veterinary shampoo contaminant, USA	This study
B. stabilis[a]	LMG 14294	Cystic fibrosis patient, Belgium	38
B. vietnamiensis[a]	G4 (ATCC 53617; R-1808)	Waste water, USA; capable of trichloroethylene degradation; derivative strain ENV435 effective in commercial field test on a contaminated aquifer; genome draft available (http://genome.jgi-psf.org/draft_microbes/bur08/bur08.home.html)	29, 33
B. ambifaria[a]	M54 (R-5142)	Corn rhizosphere, USA; registered biopesticide in commercial use; type Wisconsin strain with antifungal properties	27
B. ambifaria[a]	J82 (R-5140)	Corn rhizosphere, USA; registered biopesticide in commercial use; type Wisconsin strain with antinematodal properties	27
B. ambifaria[a]	BC-F	Corn rhizosphere, USA; U.S. Department of Agriculture biological control strain; production of antifungal agents	43
B. ambifaria[a]	AMMD^T (LMG 19182^T)	Pea rhizosphere, USA; biological control strain and species type strain	4, 27
B. ambifaria[a]	AU212 (LMG 19466)	Cystic fibrosis strain isolated from a patient in Wisconsin, USA	4
B. pyrrocinia[a]	BC11	Soil; antifungal-producing biological control strain	12
B. pyrrocinia[a]	ATCC 39277	Corn field soil; production of antifungal agents	25

[a] Species group and *recA* cluster are identical.

[all subsequent primer positions are given relative to this sequence]) and BUR2 (20-mer; bases 819 to 938).

Prior to laboratory testing, the OligoCheck software was used to examine the performance of primers in silico against all available *recA* sequences. A panel of nine non-*Burkholderia* control strains, each containing fewer than five putative mismatches to the PCR primers, were selected from the OligoCheck output for testing (Table 1). BUR1 and BUR2 produced an 869-bp amplicon from all *Burkholderia* species (Fig. 1A) except *B. andropogonis* (Table 1). Four of the control species, *Bordetella parapertussis*, *Xanthomonas sacchari*, *Ralstonia gilardii*, and *Ralstonia eutropha*, also produced the same-sized amplicon; the remaining control strains were negative (Table 1). Because of its broad specificity for *Burkholderia* species, the PCR employing BUR1 and BUR2 was subsequently used as a means to test a *recA*-based RFLP approach and to obtain further nucleotide sequence information for specific primer design.

RFLP analysis of *Burkholderia* species *recA* gene. BUR1 and BUR2 PCR amplification and digestion with HaeIII (21) were applied to analyze the *Burkholderia* genus. RFLP analysis with other enzymes was not performed since previous work had already shown HaeIII to be the most discriminatory enzyme for restriction of *B. cepacia* complex *recA* genes (21). A total of

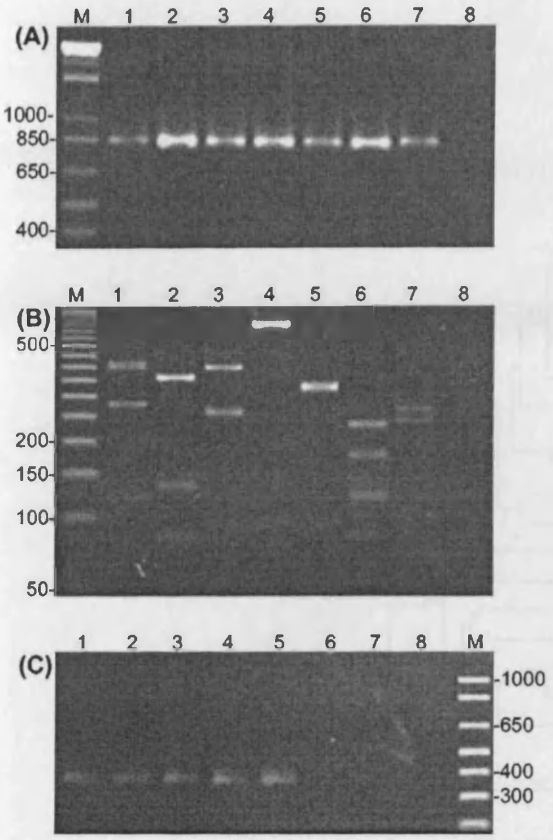


FIG. 1. *Burkholderia recA* gene PCR analysis. Panel A shows PCR products obtained with primers BUR1 and BUR2 from *Burkholderia* and control species as follows in each lane: 1, *B. stabilis* LMG14294; 2, *B. caryophylli* LMG2155^T; 3, *B. fungorum* LMG16225^T; 4, *B. graminis* LMG18924^T; 5, *B. plantarii* LMG9035^T; 6, *Bordetella parapertussis* LMG14449; 7, *Ralstonia eutropha* JMP134; 8, negative control. Panel B shows the RFLP analysis of the BUR1 and BUR2 *recA* amplicon performed by digestion of PCR products with the enzyme HaeIII. Samples in each lane are as follows: 1, *B. stabilis* LMG14294; 2, *B. caryophylli* LMG2155^T; 3, *B. fungorum* LMG16225^T; 4, *B. graminis* LMG18924^T; 5, *B. plantarii* LMG9035^T; 6, *Bordetella parapertussis* LMG14449; 7, *Ralstonia eutropha* JMP134; 8, negative control. Panel C shows the specific amplification of the *Burkholderia recA* gene with primers BUR3 and BUR4 with the following species run in each lane as follows: 1, *B. stabilis* LMG14294; 2, *B. caryophylli* LMG2155^T; 3, *B. fungorum* LMG16225^T; 4, *B. graminis* LMG18924^T; 5, *B. plantarii* LMG9035^T; 6, *Bordetella parapertussis* LMG14449; 7, *Ralstonia eutropha* JMP134; 8, H_2O negative control. Molecular size markers are shown in lane M for all panels, and the sizes of relevant bands are indicated in bp.

TABLE 3. *Burkholderia* and *Pandoraea* isolates found to possess matching *recA* RFLP types

Species	Strain name	HaeIII RFLP type
B. xenovorans	LB400^T	01
"*Burkholderia*" sp. nov.	R-15821	01
P. pulmonicola	BCC0150	15
P. pnomenusa	BCC0580	15
B. caribiensis	LMG 18531	18
B. cepacia	BCC0679	18
B. cepacia complex unknown	AU0553	38
B. cepacia complex unknown	BCC0095	38
B. pyrrocinia	LMG 21823	38
B. vietnamiensis	CLO	45
B. ubonensis	R-11767	45
B. ambifaria	MVP/C1 64	46
B. cenocepacia	MDII 367	46
B. cenocepacia	MVP/C1 73	62
B. cenocepacia	BELF 2	62
B. stabilis	LMG 14294	63
B. cenocepacia	PC184	63
B. gladioli	LMG 2216	70
B. gladioli	BCC0238	70
B. sacchari	LMG 19450	71
"*Burkholderia*" sp. nov.	R-20943	71
B. graminis	LMG 18924	77
"*Burkholderia*" sp. nov.	R-13392	77
"*Burkholderia*" sp. nov.	R-8349	84
"*Burkholderia*" sp. nov.	LMG 20580	84

103 isolates representative of both *Burkholderia* species (Table 1) and RFLP pattern diversity with the *B. cepacia* complex (21; E. Mahenthiralingam, unpublished data) were examined; 81 unique and 13 shared RFLP patterns were identified. Examples of unique RFLP types from *B. cepacia* complex members, *Burkholderia* species, and two non-*Burkholderia* species bacteria are shown in Fig. 1B. Shared RFLP types are described in Table 3.

Five known species and several strains representing putative novel species possessed overlapping RFLP patterns (Table 3). Therefore, although the majority of strains and species analyzed (86%) possessed unique patterns, the discriminatory power of the *recA*-RFLP analysis was limited. However, it could serve as a useful primary screen in a *Burkholderia* species identification scheme in an analogous fashion to its use for determining species within the *B. cepacia* complex (21, 42). Further resolution was achieved with nucleotide sequence analysis of *Burkholderia recA*.

Development of *Burkholderia*-specific *recA* PCR. To confirm the sequence variations detected by RFLP analysis of the amplified *recA* fragment and facilitate the design of *Burkholderia*-specific PCR primers, the 869-bp BUR1 and BUR2 *recA* amplicons of 16 *Burkholderia* strains, nine indeterminate *Burkholderia* species strains, the *B. ubonensis* type strain, and four *Pandoraea* strains (Table 1) were sequenced. These novel *recA* sequences were analyzed by alignment with published *B. cepacia* complex sequences and genomic *recA* sequences.

A single *Burkholderia* genus-specific adenosine was identified at base 445 in the *recA* gene and used to design the downstream primer BUR4 (17-mer; bases 445 to 461). The upstream primer, BUR3 (17-mer; bases 76 to 92), was designed to prime off the same base as BUR1 but was shorter to facilitate potential future use in a nested PCR priming off a primary amplicon resulting from BUR1 and BUR2 PCR. BUR3 and BUR4 produced a 385-bp amplicon for all *Burkholderia* species examined (examples shown in Fig. 1C) except *Burkholderia andropogonis* and novel *Burkholderia* sp. strain R-15821. The *recA* sequence of the last strain was subsequently shown to be more closely related to that of *Bordetella* (see below; Table 1 and Fig. 2). BUR3 and BUR4 PCR did not produce amplicons with non- closely *Burkholderia*-related species, including those that had originally amplified with BUR1 and BUR2 (Table 1).

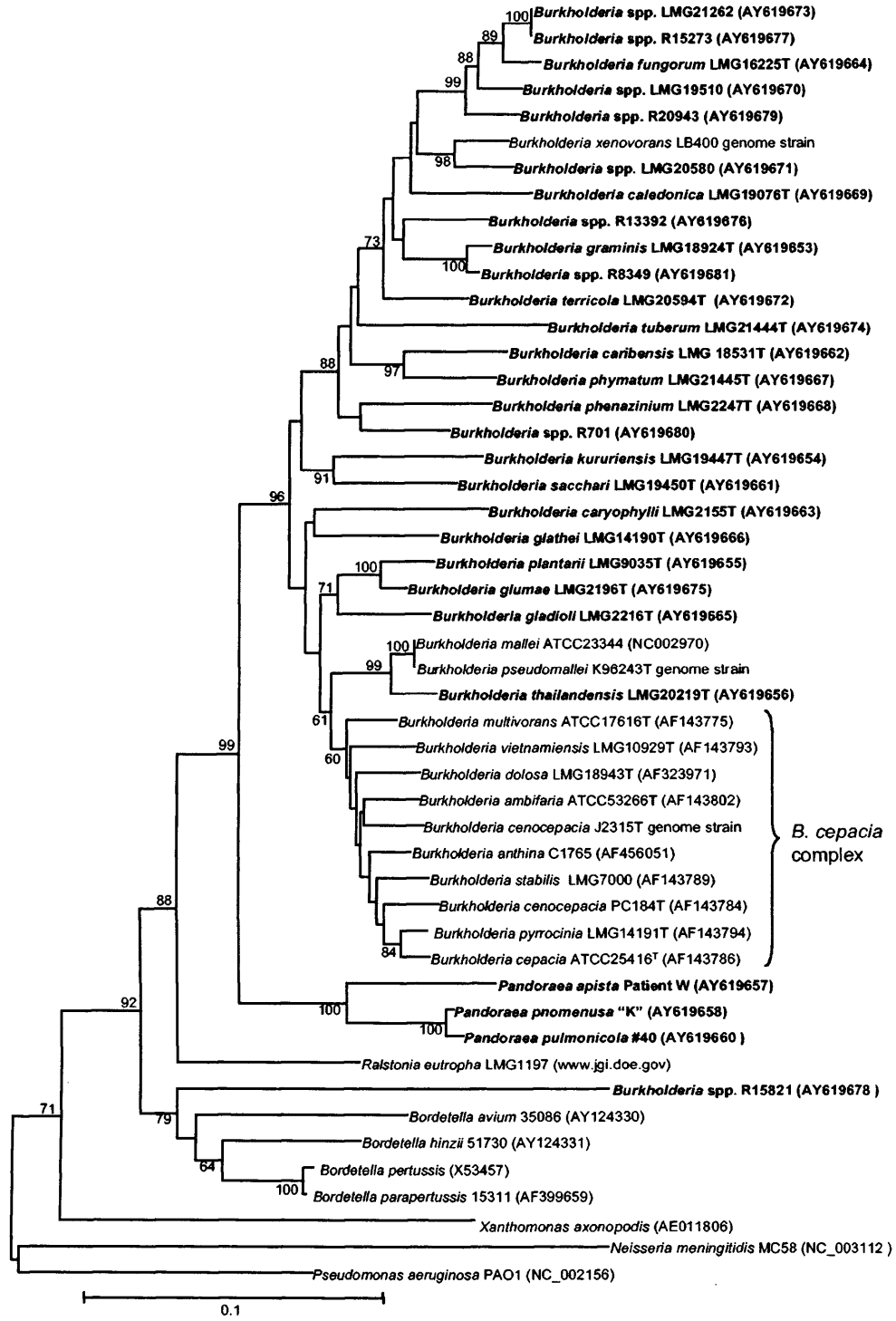


FIG. 2. Phylogenetic analysis of the *Burkholderia* genus using novel *Burkholderia recA* sequences. A phylogenetic tree comparing representatives from each species is shown; 27 novel *Burkholderia recA* genes and four novel *Pandoraea recA* genes are highlighted. Bootstrap values over 70% and genetic distance scale are indicated. The species name and strain number are shown, with the accession number in brackets. The tree was rooted with the *P. aeruginosa* PAO1 *recA* gene taken from the complete genome sequence as a representative member of a species outside the betaproteobacteria group. The following *recA* sequences were also included in the phylogenetic analysis as indicators of the ability of the analysis to differentiate among genera closely related to *Burkholderia*: *Ralstonia eutropha* JMP135, *Bordetella avium* AY124330, *Bordetella hinzii* AY124331, *Bordetella parapertussis* AF399659, *Bordetella pertussis* X53457, *Xanthomonas axonopodis* AE011806, and *Neisseria meningitidis* NC_0031112. Sequences from the species comprising the *B. cepacia* complex were included.

Phylogenetic analysis of *Burkholderia* species *recA* gene. A phylogenetic tree constructed with novel *recA* sequences from *Burkholderia* species and *Pandoraea* species (Table 1) is shown in Fig. 2 to illustrate the diversity of the *Burkholderia* genus. The *B. cepacia* complex formed a distinct phylogenetic cluster. *Burkholderia thailandensis*, *B. pseudomallei*, and *B. mallei* (the *B. pseudomallei* group) formed a cluster, which, although distinct, was more closely related to the *B. cepacia* complex than other *Burkholderia* species. The *Pandoraea* genus also formed a distinct cluster consistent with their recent separation from the *Burkholderia* genus. All the indeterminate *Burkholderia* species (Table 1) strains clustered within the *Burkholderia* genus cluster except strain R-15821, which was more closely related to *Bordetella* species.

To test if the same cluster assignments were made using sequence obtained from the smaller BUR3 and BUR4 amplicon, the sequences were trimmed to the 300 bases within the 385-bp amplicon produced by these primers and subjected to an identical analysis. The trimmed sequences produced a tree with the same topology and clusters that had been observed with the 800-base sequences derived from the 865-bp BUR1 and BUR2 amplicon (data not shown).

Phylogenetic diversity of the *B. cepacia* complex. Although the *recA* gene has proven to be a valuable tool in determining the taxonomy of the *B. cepacia* complex, species such as *B. cenocepacia* and *B. cepacia* are split into distinct phylogenetic clusters when analyzed in this way. Two *recA* lineages were originally observed in *B. cepacia* genomovars III, III-A, and III-B (21), and subsequently clusters III-C and III-D were reported when the formal name *B. cenocepacia* was proposed for this genomovar (36). Similarly, *B. cepacia* strains also divide into two lineages by phylogenetic polymorphism in the *recA* gene; one cluster includes the type strain (ATCC 25416^T) (21), while the second group was named group K based on the most common *recA* RFLP found within that cluster (42). Further novel *recA* phylogenetic groups have also been observed for other strains. However, the significance of these clusters was difficult to interpret within a phylogeny based solely on the *B. cepacia* complex (E.M., unpublished data), and so further analysis with the new data obtained in this study was carried out to resolve phylogenetic relationships on a broader scale.

A phylogenetic tree comprising a subset of 101 *B. cepacia* complex strain *recA* genes and including 28 novel *Burkholderia* and *Pandoraea* sequences (Table 1) is shown in Fig. 3. Seventeen phylogenetically distinct clusters were observed for *B. cepacia* complex species. As observed with the *Burkholderia* genus phylogeny (Fig. 2), the *B. pseudomallei* group clustered adjacent to the *B. cepacia* complex. *B. gladioli* and *B. plantarii*, and *B. glathei* and *B. carophylli* also formed distinct clusters. All the remaining *Burkholderia* species formed a diverse separate group.

Within the *B. cepacia* complex, all *B. stabilis*, *B. pyrrocinia*, *B. anthina*, *B. vietnamiensis*, *B. dolosa*, *B. ambifaria*, and *B. multivorans* strains and the *B. ubonensis* type strain formed single discrete phylogenetic clusters. As shown previously (21), *B. cepacia* strains were split into sublineages (Fig. 3), while *B. cenocepacia* III-A, III-B, III-C, and III-D were all distinct. Three *B. cepacia* clusters were observed, one containing the type strain, ATCC 25416^T, one composed of RFLP type K strains (42), and another composed of RFLP type AW strains that previously belonged to *B. cepacia* (9) (Fig. 3). However, it must be noted that *B. cepacia* RFLP group K cluster was the least robust group within the phylogenetic analysis (bootstrap value 30) (Fig. 3), suggesting that further species diversity may be present. Several strains which previously were difficult to assign (E.M., unpublished work), fell into distinct branches of the new tree. *B.* cepacia complex group 1 strains were closely related to *B. cenocepacia* III-A, while *B. cepacia* complex group 2 strains clustered between *B. cenocepacia* III-D and *B. anthina* (Fig. 3). Again all the clusters remained intact when the *recA* sequences were trimmed to 300 bp within the BUR3 and BUR4 amplicon (data not shown).

Phylogenetic positions of clinical and environmental strains. Of the 101 *B. cepacia* strains analyzed phylogenetically (Fig. 3), 41 were from the environment and 60 were from human sources. No clinical strains were present in *B. cepacia* type AW, *B. pyrrocinia*, and *B. cenocepacia* III-C clusters, and no environmental strains were present in the *B. cenocepacia* III-D cluster. All the remaining *B. cepacia* complex phylogenetic clusters contained strains from both sources, with the exception of *B. ubonensis*, which comprised the single type strain (Fig. 3).

The phylogenetic positions of several environmentally, biotechnologically and clinically relevant strains were also resolved (Table 2). Strains with documented biological control activities were found to be present in the *B. cepacia* type strain cluster, *B. cenocepacia* III-B, *B. ambifaria*, and *B. pyrrocinia* (Table 2). Interestingly, all three *B. cenocepacia* III-B biological control strains contained the *B. cenocepacia* pathogenicity island (2) (Table 2), including strain M36, which was originally registered for commercial use in the United States and has now been withdrawn (23). One other registered biopesticide strain, M54 (27), was assigned to the *B. ambifaria* cluster, where the majority of strains with biological control activity were found (Table 2). The *recA* genes drawn from *Burkholderia* genomic resources clustered as follows (Fig. 1; Table 2): strain G4, a well-known bioremediation strain, clustered with *B. vietnamiensis*, as expected; and strain 383, one of the isolates from the pioneering study of Stanier et al. (32), clustered within the group K lineage of *B. cepacia*, as did the metagenomic Sargasso Sea strain SAR-1 (40).

Genetic identity of *B. cepacia* complex strains. Several strains were found to possess the same *recA* RFLP type and identical *recA* sequences, yet were of distinct environmental or clinical origin. To resolve their clonality to the strain type level, macrorestriction and PFGE fingerprinting was performed, resulting in the definition of four strain types, each comprising a strain from an environmental and a clinical source (Fig. 4). Each strain pair possessed almost identical genomic fingerprints (all Dice coefficients of similarity >0.93 for each pair) clearly defining each pair as a distinct genetic strain type. Pairs were found within the following species (see Table 2 and Fig. 4): *B. cepacia* strains, ATCC 17759 (environmental, isolated before 1966), and LMG 14087 (clinical, isolated in 1988 from a wound swab in the United Kingdom); ATCC 25416^T (environmental, isolated in the 1940s) and J1050 (clinical); *B. ambifaria* strains AMMD^T (environmental) and AU0212 (clinical); and *B. stabilis* strains HI-2482 (environmental) and LMG 14294 (clinical, isolated from sputum from a patient with cystic fibrosis in Belgium in 1993).

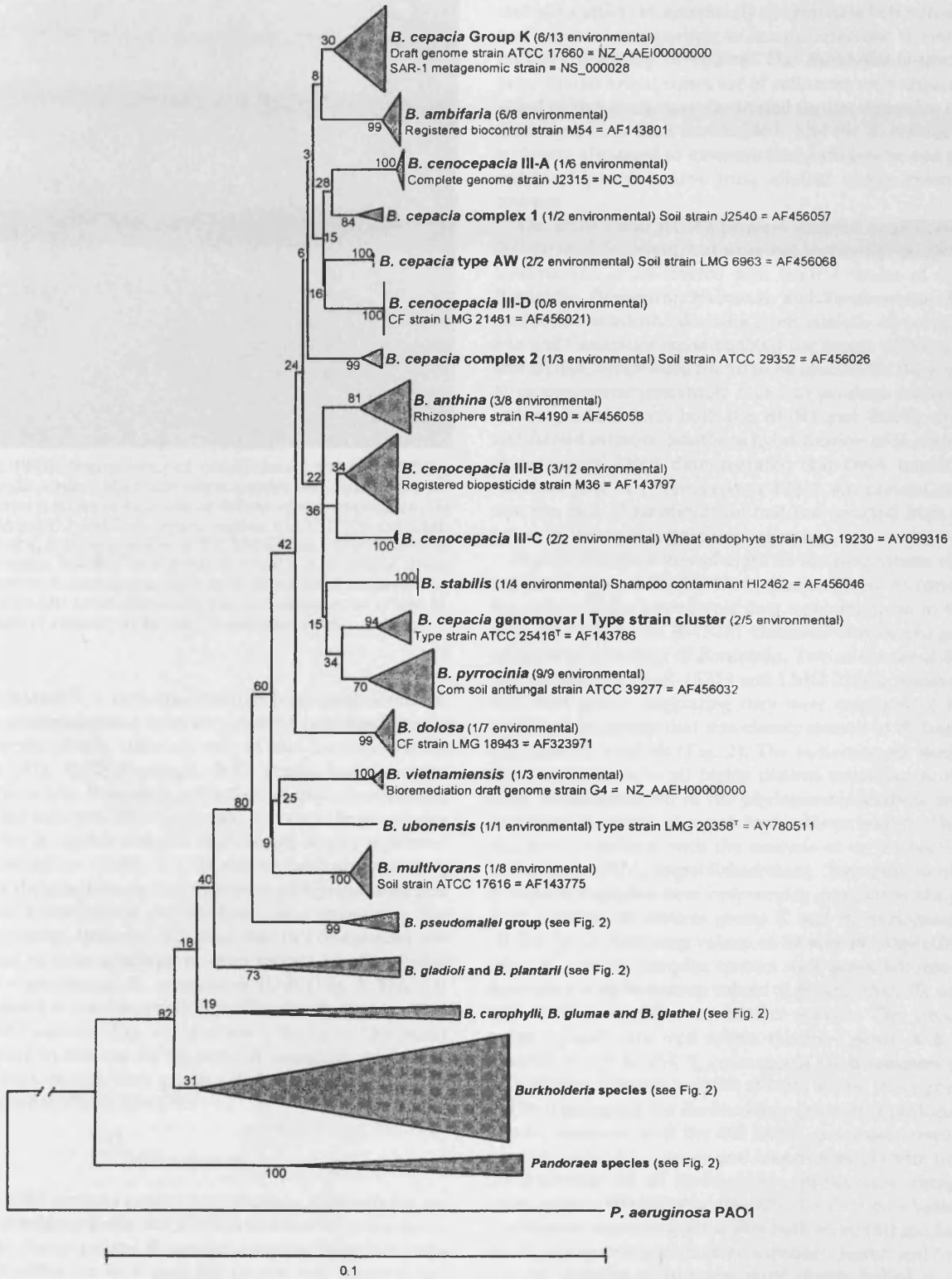


FIG. 3. Phylogenetic analysis of *B. cepacia* complex strains using novel *Burkholderia recA* sequences. A phylogenetic tree comparing 160 *B. cepacia* complex strains, 28 *Burkholderia*, and four *Pandoraea recA* sequences is shown as a composite of phylogenetically distinct sequence clusters. Bootstrap values and genetic distance scale are indicated. The species or group name for each cluster is shown in bold. Additional information for one or more reference or interesting strains within the group is also shown. The number of sequences deriving from environmental strains versus the total number of sequences in the group is shown in brackets.

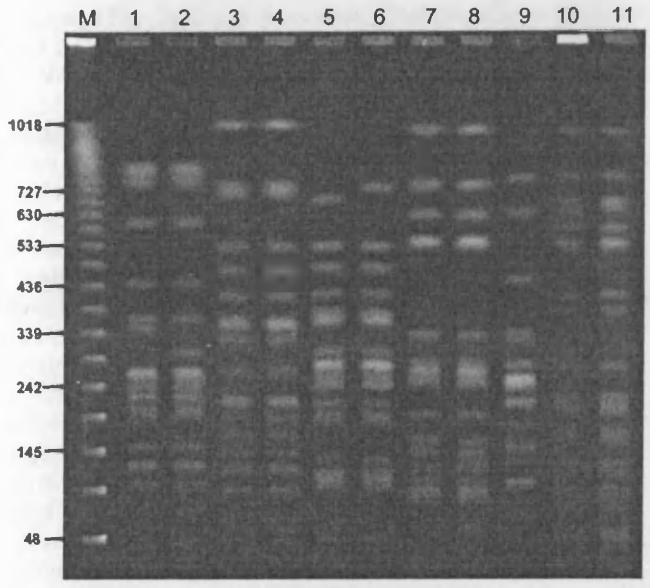


FIG. 4. PFGE fingerprinting of environmental and clinical *B. cepacia* complex strains. Macrorestriction analysis with SpeI of the following strains is shown in each lane as follows (given, respectively, for each strain pair): 1 and 2, *B. cepacia* strains ATCC 17759 and LMG 14087; 3 and 4, *B. cepacia* strains ATCC 25416^T and J1050; 5 and 6, *B. ambifaria* strains AMMD^T and AU0212; 7 and 8, *B. ambifaria* strains M54 and J82; 9, *B. cenocepacia* strain M36; 10 and 11, *B. stabilis* strains HI-2482 and LMG 14294. Molecular size markers were run in lane M, and the sizes of relevant fragments are indicated in kb.

Strain AMMD^T, a well-characterized biocontrol isolate (4, 27), was almost identical to strain AU0212 recovered from a cystic fibrosis patient, differing only in one macrorestriction fragment (Fig. 4). Interestingly, both strains had the same geographic origin, Wisconsin, even though their sources were distinct, soil and cystic fibrosis sputum. Genomic fingerprinting of the three *B. cepacia* complex biopesticide strains registered for commercial use (Table 2) (27) also revealed an interesting feature of these isolates in that they were all registered as type Wisconsin, a designation derived from their phenotypic features and source. However, it is clear that this designation was not related to their genotype or even species identity. Strain M36 was a member of *B. cenocepacia* III-B (Fig. 3; Table 3) and possessed a macrorestriction profile clearly distinct from strains M54 and J82 (Fig. 4). Moreover, the latter biocontrol isolates were in fact exactly the same *B. ambifaria* strain type (Fig. 4), even though each possess slightly different biopesticidal properties (Table 2) (27).

DISCUSSION

We have developed a genetic identification approach for the entire *Burkholderia* genus that can discriminate between members of the closely related *B. cepacia* complex. Sequence polymorphism within the *recA* gene has proven very useful in defining the taxonomy of the *B. cepacia* complex (6, 21), but the original approach could not be applied to *Burkholderia* species outside the complex. By designing new *Burkholderia* species *recA* PCRs, we have demonstrated that RFLP analysis of the gene can be used to differentiate the majority of *Burkholderia* strains. However, this pattern-matching technique was still limited in its ability to completely discriminate between all species and required comparison to known reference strains for preliminary grouping of strains. The *Burkholderia*-specific *recA* primers and an extensive set of reference *recA* sequences provided by this study have facilitated further definition of species diversity within both *Burkholderia* and the *B. cepacia* complex and were also used to examine the phylogenetic and genotypic relationships of strains from clinical versus environmental sources.

The BUR1 and BUR2 primers enabled amplification of an 869-bp *recA* fragment, but were not absolutely specific to *Burkholderia* and cross-reacted with control strains of the genera *Bordetella*, *Pandoraea*, *Ralstonia*, and *Xanthomonas*. However, nucleotide sequence deriving from analysis of novel *Burholderia* and *Pandoraea recA*s enabled the design of primers BUR3 and BUR4, which were found to be specific for the genus. Only *B. andropogonis* repeatedly failed to produce correct amplification products with both the BUR1 and BUR2, and BUR3 and BUR4 primers. Southern hybridization of *B. andropogonis* chromosomal DNA demonstrated that DNA homologous to the *recA* gene of *B. cenocepacia* J2315 was present, suggesting that the lack of amplification had not resulted from gene deletion in this isolate (data not shown).

Phylogenetic analysis of eight of the nine strains representing putative novel *Burkholderia* species (Table 1) corroborated the results of the phenotypic data assigning them to the genus (Fig. 2). Only strain R-15281 clustered outside the genus and adjacent to members of *Bordetella*. Two of the novel *Burkholderia* species strains, R-15273 and LMG 21262, possessed identical *recA* genes, suggesting they were members of the same novel species group that was closely related to *B. fungorum* by phylogenetic analysis (Fig. 2). The remaining six *Burkholderia* species strains were all highly distinct according to the bootstrap values observed in the phylogenetic analysis, suggesting they were members of novel *Burkholderia* species. These findings were consistent with the analysis of their whole-cell protein profiles (P.V., unpublished data). Two clusters within the *B.* cepacia complex were only weakly grouped by the phylogenetic analysis, *B. cepacia* group K and *B. cenocepacia* group III-B (Fig. 3; bootstrap values of 30 and 34, respectively). All other *B. cepacia* complex species *recA* genes fell into clusters associated with bootstrap values of greater than 70, consistent with their recognition as distinct species. The considerable polymorphism observed within the *recA* genes of *B. cepacia* complex group K and *B. cenocepacia* III-B indicates that further species diversity may be present within these groups.

The topology of the *Burkholderia* genus *recA* phylogeny (Fig. 2) was congruent with the 16S rRNA gene-based phylogenetic tree presented by Coenye and Vandamme (5) with respect to the following: (i) all *Burkholderia* species were distinct from other genera; (ii) all taxonomically classified *Burkholderia* species formed separate arms within both trees; (iii) species within the *B. cepacia* complex formed a distinct cluster; and (iv) the *B. cepacia* complex cluster was most closely linked to a group containing *B. mallei*, *B. pseudomallei*, and *B. thailandensis*. Clusters which were divergent in phylogenetic analysis based on the *recA* gene and 16S rRNA gene gene (5) were as follows: (i) *B. plantari* and *B. glumae* formed a distinct cluster in both trees, however, this group also contained *B. gladioli* in the *recA*

analysis (Fig. 2); (ii) *B. fungorum*, *B. caledonica*, *B. graminis*, and *B. caryophylli*, which formed a distinct cluster in the 16S rRNA gene tree, were not closely associated according to *recA*; and (iii) the positions of all the remaining *Burkholderia* species examined were not conserved between the *recA* and 16S rRNA gene analyses. The major advantage *recA* phylogenetic analysis provides over the 16S rRNA gene for classification of *Burkholderia* is the greater degree of resolution among closely related species within the genus, such as the *B. cepacia* complex.

Although RFLP analysis of the *recA* gene did not discriminate between all species of *Burkholderia*, sequence analysis of the BUR1 and BUR2 amplicon was sufficient to separate all strains, including their differentiation from members of the *B. cepacia* complex (Fig. 2 and 3). In particular, the degree of resolution of the *recA* phylogenetic trees for members of the *B. cepacia* complex was much greater than observed with 16S rRNA gene analysis (6). However, the presence of discrete *recA* lineages within species such as *B. cenocepacia* and *B. cepacia* adds an additional level of complexity. A promising finding of the study was that analysis of just a 300-bp region of *recA* sequence within the BUR3 and BUR4 amplicon produced phylogenetic trees with the same topology and discrimination as the nearly full-length trees derived from analysis of the BUR1 and BUR2 PCR. These data suggest that the BUR3 and BUR4 PCR could provide a useful and rapid non-culture-based approach to explore the *Burkholderia* diversity in various environments.

Genomic resources for *Burkholderia* have increased substantially in the last few years, and we have shown that by phylogenetic analysis of whole-genome sequence derived *recA*s the classification of both cultured and uncultured genomic strains can be clarified. The genome strains *B. cepacia* ATCC 17760 (strain 383) (32) and *Burkholderia* Sargasso Sea strain SAR-1 reconstructed from metagenomic data (38) were both found to be members of *B. cepacia* group K. The *recA* gene of SAR-1 was identical to that of strain R-12710 (data not shown), cultivated from sheep with mastitis (3), and another marine isolate (obtained from the Sea of Japan) also belongs to *B. cepacia* group K (P.V., unpublished data). While the presence of *B. cepacia* complex members in river water has been reported (15), their isolation from marine environments is thought to be rare. The novel primers developed in this study can be applied to investigate this area of scientific interest.

Phylogenetic analysis of *Burkholderia recA* genes also demonstrated that, like other taxonomic criteria, it could not be used as a means to distinguish environmental from clinical strains. Strains from both sources were found throughout the extensive phylogenetic tree derived from this study (Fig. 3). Although no clinical strains were found in four clusters and no environmental strains found in one other *B. cepacia* complex cluster (Fig. 3), these clusters contained small numbers of strains. It is likely that as more *B. cepacia* complex strains are examined, so that the number of strains in these clusters is increased, environmental and clinical stains will be identified for each cluster.

Analysis of the *recA* gene of *B. cepacia* complex strains in this study resulted in the identification of several clonal pairs of strains each from distinct environmental and clinical sources (Fig. 4). Such pairs were found for *B. cepacia*, *B. stabilis*, and *B. ambifaria* and clearly showed that genotypically identical *B. cepacia* complex strains can be isolated from both human infections and natural or environmental sources. This work extends the work of LiPuma and colleagues (17), who demonstrated that an epidemic cystic fibrosis strain of *B. cenocepacia* was identical to a strain found in the soil. *B. ambifaria* strain AMMD^T is the first strain with known biopesticidal properties (4, 27) to have an essentially clonal relative, strain AU0212, cultured from a patient with cystic fibrosis (Fig. 4). In conclusion, it appears that all *B. cepacia* complex bacteria are capable of colonizing a wide range of habitats, and in the case of infectious niches, this trait appears more dependent on the vulnerability of the host than the taxonomic or phylogenetic classification of the strain.

B. cepacia complex strains with useful biotechnological properties were also found throughout the *recA* phylogeny. *B. cepacia*, *B. cenocepacia*, *B. ambifaria*, and *B. pyrrocinia* strains with biopesticidal or antifungal properties were identified (Table 2). The ability to obtain accurate species information using *recA* sequence data may prove vital in the characterization of future biotechnologically useful strains, especially since detailed understanding of the strain taxonomy forms one of the major risk assessment criteria for commercial registration of bacteria (27). In addition, our study has shown that classification of biopesticidal strains based on phenotypic criteria alone is not adequate. The two type Wisconsin strains of *B. ambifaria*, M54 and J82, which remain registered for commercial use were found to be genotypically identical (Fig. 4) even though they possess different biopesticidal properties (Table 2) (27). This kind of phenotypic variability within a single clone has also be observed in isolates from cystic fibrosis patient infections (14). Although the *B. cepacia* complex has been highlighted as a group of bacterial species which require risk assessment as biological control agents (27), the case for this evaluation is not limited to these organisms.

In summary, the *recA*-based approach developed in this study provides molecular tools for the identification of *Burkholderia* species that will help to enable researchers to keep pace with the ever-increasing ecological, pathogenic, and genomic interest in the genus.

ACKNOWLEDGMENTS

G.W.P. is grateful for Ph.D. funding from the Natural Environment Research Council (NER/S/A/2002/10347A) and their Environmental Genomics Program (NER/T/S/2001/0299). Determination of the *B. cepacia* complex *recA* gene sequences was performed via funding from the Cystic Fibrosis Trust (project grant PJ472). T.C. and P.V. are indebted to the Fund for Scientific Research–Flanders (Belgium) for a position as postdoctoral fellow and research grants, respectively. J.J.L. is supported by a grant from the Cystic Fibrosis Foundation (U.S.).

We acknowledge K. E. Ashelford for assistance with in silico primer analysis.

REFERENCES

1. **Ashelford, K. E., N. Banning, J. C. Fry, and A. Weightman.** Assessing the effectiveness of commonly used 16S rRNA universal primer sets to target intended taxa: an *in silico* study. Manuscript in preparation.
2. **Baldwin, A., P. A. Sokol, J. Parkhill, and E. Mahenthiralingam.** 2004. The *Burkholderia cepacia* epidemic strain marker is Part of a novel genomic island encoding both virulence and metabolism-associated genes in *Burkholderia cenocepacia*. Infect. Immun. **72**:1537–1547.
3. **Berriatua, E., I. Ziluaga, C. Miguel-Virto, P. Uribarren, R. Juste, S. Laevens, P. Vandamme, and J. R. Govan.** 2001. Outbreak of Subclinical Mastitis in a Flock of Dairy Sheep Associated with *Burkholderia cepacia* Complex Infection. J. Clin. Microbiol. **39**:990–994.

4. **Coenye, T., E. Mahenthiralingam, D. Henry, J. J. LiPuma, S. Laevens, M. Gillis, D. P. Speert, and P. Vandamme.** 2001. *Burkholderia ambifaria* sp. nov., a novel member of the *Burkholderia cepacia* complex including biocontrol and cystic fibrosis-related isolates. Int. J. Syst. Evol. Microbiol. **51:**1481–1490.
5. **Coenye, T., and P. Vandamme.** 2003. Diversity and significance of *Burkholderia* species occupying diverse ecological niches. Environ. Microbiol. **5:**719–729.
6. **Coenye, T., P. Vandamme, J. R. Govan, and J. J. LiPuma.** 2001. Taxonomy and identification of the *Burkholderia cepacia* complex. J. Clin. Microbiol. **39:**3427–3436.
7. **Coenye, T., P. Vandamme, J. J. LiPuma, J. R. Govan, and E. Mahenthiralingam.** 2003. Updated version of the *Burkholderia cepacia* complex experimental strain panel. J. Clin. Microbiol. **41:**2797–2798.
8. **Engledow, A. S., E. G. Medrano, E. Mahenthiralingam, J. J. LiPuma, and C. F. Gonzalez.** 2004. Involvement of a plasmid-encoded type IV secretion system in the plant tissue watersoaking phenotype of *Burkholderia cenocepacia*. J. Bacteriol. **186:**6015–6024.
9. **Gillis, M., T. Vanvan, R. Bardin, M. Goor, P. Hebbar, A. Willems, P. Segers, K. Kersters, T. Heulin, and M. P. Fernandez.** 1995. Polyphasic taxonomy in the genus *Burkholderia* leading to an emended description of The genus and proposition of *Burkholderia vietnamiensis* sp. nov. for N_2-fixing isolates from rice in Vietnam. Int. J. Syst. Bacteriol. **45:**274–289.
10. **Hall, T. A.** 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. **41:**95–98.
11. **Jayaswal, R. K., M. Fernandez, R. S. Upadhyay, L. Visintin, M. Kurz, J. Webb, and K. Rinehart.** 1993. Antagonism of *Pseudomonas cepacia* against phytopathogenic fungi. Curr. Microbiol. **26:**17–22.
12. **Kang, Y., R. Carlson, W. Tharpe, and M. A. Schell.** 1998. Characterization of genes involved in biosynthesis of a novel antibiotic from *Burkholderia cepacia* BC11 and their role in biological control of *Rhizoctonia solani*. Appl. Environ. Microbiol. **64:**3939–3947.
13. **Karlin, S., G. M. Weinstock, and V. Brendel.** 1995. Bacterial classifications derived from *recA* protein sequence comparisons. J. Bacteriol. **177:**6881–6893.
14. **Larsen, G. Y., T. L. Stull, and J. L. Burns.** 1993. Marked phenotypic variability in *Pseudomonas cepacia* isolated from a patient with cystic fibrosis. J. Clin. Microbiol. **31:**788–792.
15. **Leff, L. G., R. M. Kernan, J. V. McArthur, and L. J. Shimkets.** 1995. Identification of aquatic *Burkholderia* (*Pseudomonas*) *cepacia* by hybridization with species-specific rRNA gene probes. Appl. Environ. Microbiol. **61:**1634–1636.
16. **LiPuma, J. J., B. J. Dulaney, J. D. McMenamin, P. W. Whitby, T. L. Stull, T. Coenye, and P. Vandamme.** 1999. Development of rRNA-based PCR assays for identification of *Burkholderia cepacia* complex isolates recovered from cystic fibrosis patients. J. Clin. Microbiol. **37:**3167–3170.
17. **LiPuma, J. J., T. Spilker, T. Coenye, and C. F. Gonzalez.** 2002. An epidemic *Burkholderia cepacia* complex strain identified in soil. Lancet **359:**2002–2003.
18. **LiPuma, J. J., T. Spilker, L. H. Gill, P. W. Campbell 3rd, L. Liu, and E. Mahenthiralingam.** 2001. Disproportionate distribution of *Burkholderia cepacia* complex species and transmissibility markers in cystic fibrosis. Am. J. Respir. Crit. Care Med. **164:**92–96.
19. **Lowe, P., C. Engler, and R. Norton.** 2002. Comparison of automated and nonautomated systems for identification of *Burkholderia pseudomallei*. J. Clin. Microbiol. **40:**4625–4627.
20. **Mahenthiralingam, E., A. Baldwin, and P. Vandamme.** 2002. *Burkholderia cepacia* complex infection in patients with cystic fibrosis. J. Med. Microbiol. **51:**533–538.
21. **Mahenthiralingam, E., J. Bischof, S. K. Byrne, C. Radomski, J. E. Davies, Y. Av-Gay, and P. Vandamme.** 2000. DNA-Based diagnostic approaches for identification of *Burkholderia cepacia complex*, *Burkholderia vietnamiensis*, *Burkholderia multivorans*, *Burkholderia stabilis*, and *Burkholderia cepacia* genomovars I and III. J. Clin. Microbiol. **38:**3165–3173.
22. **Mahenthiralingam, E., T. Coenye, J. W. Chung, D. P. Speert, J. R. Govan, P. Taylor, and P. Vandamme.** 2000. Diagnostically and experimentally useful panel of strains from the *Burkholderia cepacia* complex. J. Clin. Microbiol. **38:**910–913.
23. **Mahenthiralingam, E., and P. Sayre.** 2003. *Burkholderia cepacia* complex, p. 41–44. *In* M. D. Licker (ed.), McGraw-Hill yearbook of science and technology. McGraw-Hill, New York, NY.
24. **Mao, W., R. D. Lumsden, J. A. Lewis, and P. K. Hebbar.** 1998. Seed treatment using pre-infiltration and biocontrol agents to reduce damping-off of corn caused by species of *Pythium* and *Fusarium*. Plant Dis. **82:**294–299.
25. **Meyers, E., G. S. Bisacchi, L. Dean, W. C. Liu, B. Minassian, D. S. Slusarchyk, R. B. Sykes, S. K. Tanaka, and W. Trejo.** 1987. *Xylocandin*: a new complex of antifungal peptides. I. Taxonomy, isolation and biological activity. J. Antibiot. (Tokyo) **40:**1515–1519.
26. **Mondello, F. J.** 1989. Cloning and expression in *Escherichia coli* of *Pseudomonas* strain LB400 genes encoding polychlorinated biphenyl degradation. J. Bacteriol. **171:**1725–1732.
27. **Parke, J. L., and D. Gurian-Sherman.** 2001. Diversity of the *Burkholderia cepacia* complex and implications for risk assessment of biological control strains. Annu. Rev. Phytopathol. **39:**225–258.
28. **Rabkin, C. S., W. R. Jarvis, R. L. Anderson, J. Govan, J. Klinger, J. LiPuma, W. J. Martone, H. Monteil, C. Richard, S. Shigeta, et al.** 1989. *Pseudomonas cepacia* typing systems: collaborative study to assess their potential in epidemiologic investigations. Rev. Infect. Dis. **11:**600–607.
29. **Shields, M. S., S. O. Montgomery, S. M. Cuskey, P. J. Chapman, and P. H. Pritchard.** 1991. Mutants of *Pseudomonas cepacia* G4 defective in catabolism of aromatic compounds and trichloroethylene. Appl. Environ. Microbiol. **57:**1935–1941.
30. **Shields, M. S., M. J. Reagin, R. R. Gerger, R. Campbell, and C. Somerville.** 1995. Tom, a new aromatic degradative plasmid from *Burkholderia* (*Pseudomonas*) *cepacia* G4. Appl. Environ. Microbiol. **61:**1352–1356.
31. **Speert, D. P., M. Bond, R. C. Woodman, and J. T. Curnutte.** 1994. Infection with *Pseudomonas cepacia* in chronic granulomatous disease: role of nonoxidative killing by neutrophils in host defense. J. Infect. Dis. **170:**1524–1531.
32. **Stanier, R. Y., N. J. Palleroni, and M. Doudoroff.** 1966. The aerobic *pseudomonads*: a taxonomic study. J. Gen. Microbiol. **43:**159–271.
33. **Steffan, R. J., K. L. Sperry, M. T. Walsh, S. Vainberg, and C. W. Condee.** 1999. Field-scale evaluation of in situ bioaugmentation for remediation of chlorinated solvents in groundwater. Environ. Sci. Technol. **33:**2771–2781.
34. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
35. **Vandamme, P., D. Henry, T. Coenye, S. Nzula, M. Vancanneyt, J. J. LiPuma, D. P. Speert, J. R. Govan, and E. Mahenthiralingam.** 2002. *Burkholderia anthina* sp. nov. and *Burkholderia pyrrocinia*, two additional *Burkholderia cepacia* complex bacteria, may confound results of new molecular diagnostic tools. FEMS Immunol. Med. Microbiol. **33:**143–149.
36. **Vandamme, P., B. Holmes, T. Coenye, J. Goris, E. Mahenthiralingam, J. J. LiPuma, and J. R. Govan.** 2003. *Burkholderia cenocepacia* sp. nov.-a new twist to an old story. Res. Microbiol **154:**91–96.
37. **Vandamme, P., B. Holmes, M. Vancanneyt, T. Coenye, B. Hoste, R. Coopman, H. Revets, S. Lauwers, M. Gillis, K. Kersters, and J. R. Govan.** 1997. Occurrence of multiple genomovars of *Burkholderia cepacia* in cystic fibrosis patients and proposal of *Burkholderia multivorans* sp. nov. Int. J. Syst. Bacteriol. **47:**1188–1200.
38. **Vandamme, P., E. Mahenthiralingam, B. Holmes, T. Coenye, B. Hoste, P. De Vos, D. Henry, and D. P. Speert.** 2000. Identification and population structure of *Burkholderia stabilis* sp. nov. (formerly *Burkholderia cepacia* genomovar IV). J. Clin. Microbiol. **38:**1042–1047.
39. **Vandamme, P., B. Pot, M. Gillis, P. de Vos, K. Kersters, and J. Swings.** 1996. Polyphasic taxonomy, a consensus approach to bacterial systematics. Microbiol. Rev. **60:**407–438.
40. **Venter, J. C., K. Remington, J. F. Heidelberg, A. L. Halpern, D. Rusch, J. A. Eisen, D. Wu, I. Paulsen, K. E. Nelson, W. Nelson, D. E. Fouts, S. Levy, A. H. Knap, M. W. Lomas, K. Nealson, O. White, J. Peterson, J. Hoffman, R. Parsons, H. Baden-Tillson, C. Pfannkoch, Y. H. Rogers, and H. O. Smith.** 2004. Environmental genome shotgun sequencing of the Sargasso Sea. Science **304:**66–74.
41. **Vermis, K., T. Coenye, J. J. LiPuma, E. Mahenthiralingam, H. J. Nelis, and P. Vandamme.** 2004. Proposal to accommodate *Burkholderia cepacia* genomovar VI as *Burkholderia dolosa* sp. nov. Int. J. Syst. Evol. Microbiol. **54:**689–691.
42. **Vermis, K., T. Coenye, E. Mahenthiralingam, H. J. Nelis, and P. Vandamme.** 2002. Evaluation of species-specific recA-based PCR tests for genomovar level identification within the *Burkholderia cepacia* complex. J. Med. Microbiol. **51:**937–940.
43. **Zhou, H., F. Yao, D. P. Roberts, and T. G. Lessie.** 2003. AHL-deficient mutants of *Burkholderia ambifaria* BC-F have decreased antifungal activity. Curr. Microbiol. **47:**174–179.

Application of a *recA* gene-based identification approach to the maize rhizosphere reveals novel diversity in *Burkholderia* species

George W. Payne[1], Alban Ramette[2], Helen L. Rose[1], Andrew J. Weightman[1], T. Hefin Jones[1], James M. Tiedje[2] & Eshwar Mahenthiralingam[1]

[1]Cardiff School of Biosciences, Cardiff University, Cardiff, Wales, UK; and [2]Centre for Microbial Ecology, Michigan State University, East Lansing, MI, USA

Correspondence: Eshwar Mahenthiralingam, Cardiff School of Biosciences, Main Building, Museum Avenue, Cardiff University, Cardiff CF10 3TL, Wales, UK. Tel.: +44 (0)29 20875875; fax: +44 (0)29 20874305; e-mail: mahenthiralingame@cardiff.ac.uk

Received 15 December 2005; revised 17 February 2006; accepted 26 March 2006. First published online May 2006.

doi:10.1111/j.1574-6968.2006.00257.x

Editor: Sergio Casella

Keywords
Burkholderia; rhizosphere diversity; cultivation-independent analysis; PCR

Abstract

Burkholderia species are widely distributed in the natural environment. We evaluated the use of the *recA* gene in a cultivation-independent approach to examine the *Burkholderia* diversity associated with the maize rhizosphere. Two types of *recA* gene library were constructed, one with broad-specificity *recA* primers (BUR1 and BUR2) and a second from the products of nested PCRs using *Burkholderia*-specific primers (BUR3 and BUR4). The broad-specificity primer set provided near full-length *recA* sequences (869 bp) suitable for the creation of robust environmental sequence data sets; however, the nested PCR approach demonstrated the greatest specificity (84%) for detection of *Burkholderia* species *recA* genes. In addition, the screening approach was able to identify *recA* phylotypes matching *Burkholderia cepacia* complex species previously cultivated from the maize samples and discriminate these from other *Burkholderia*. The ecological benefit of *Burkholderia* species cultivated from maize rhizosphere is well documented, however, the fact that the majority of *Burkholderia recA* genes detected in this study (90%) were suggestive of novel taxa indicates that a wealth of potentially important interactions with uncultivated *Burkholderia* species remain unstudied in this habitat.

Introduction

Bacteria from the genus *Burkholderia* occupy multiple niches of major ecological importance. Many *Burkholderia* species have been isolated from soil and have been reported to be closely associated with the plant rhizosphere (Parke & Gurian-Sherman, 2001; Coenye & Vandamme, 2003). Beneficial environmental interactions of *Burkholderia* species include their ability to facilitate both plant protection and growth promotion as biopesticidal agents (Parke & Gurian-Sherman, 2001), the capacity of certain species to fix nitrogen (Minerdi *et al.*, 2001; Caballero-Mellado *et al.*, 2004; Reis *et al.*, 2004), and versatile catabolic capacity that allows degradation of numerous major pollutants such as trichloroethylene and polychlorinated biphenyls (Shields *et al.*, 1991; Goris *et al.*, 2004). In contrast, several *Burkholderia* species may also cause disease in vulnerable humans (Mahenthiralingam *et al.*, 2005), animals and plants (Coenye & Vandamme, 2003). The *Burkholderia* genus currently comprises 34 formally described species (Coenye & Vandamme, 2003), with nine species forming the *Burkholderia cepacia* complex (Mahenthiralingam *et al.*, 2005). *Burkholderia* taxonomy is complex and recent studies have undertaken a polyphasic approach to characterize new species (Coenye & Vandamme, 2003). Studies on environmental *B. cepacia* complex bacteria have involved cultivation of organisms on selective media followed by further identification including molecular analysis of 16S rRNA gene (Salles *et al.*, 2002) or *recA* (Fiore *et al.*, 2001), or both (Miller *et al.*, 2002; Ramette *et al.*, 2005). The *recA* gene has recently been used as a cultivation-independent approach, however, this study was limited to the detection of only *B. cepacia* complex species occurring in the maize rhizosphere because of the specificity of the PCR used (Pirone *et al.*, 2005). Exploration of the diversity of the entire *Burkholderia* genus in environmental samples has, however, been largely limited to 16S rRNA gene analyses (Salles *et al.*, 2002; Coenye & Vandamme, 2003), revealing *Burkholderia* as important endosymbiotic species such as those forming leaf galls (Van Oevelen *et al.*, 2004).

The *recA* gene has been established as a useful target for the identification of the *B. cepacia* complex species; it affords greater discriminatory power than the 16S rRNA gene for differentiation within this closely related group

Table 1. PCR primers used in the study

PCR name	Primer pair (forward and reverse, 5′–3′)	Product size (bp)	Annealing temperature (°C)	Specificity	Reference
BCR1.2	BCR1, TGACCGCCGAGAAGAGCAA BCR2, CTCTTCTTCGTCCATCGCCTC	1043	58	*Burkholderia cepacia* complex only	Mahenthiralingam *et al.* (2000)
BUR1.2	BUR1, GATCGARAAGCAGTTCGGCAA BUR2, TTGTCCTTGCCCTGRCCGAT	869	58	Broad specificity including *Burkholderia* species	Payne *et al.* (2005)
BUR3.4	BUR3, GARAAGCAGTTCGGCAA BUR4, GAGTCGATGACGATCAT	385	57	*Burkholderia* species only	Payne *et al.* (2005)

(Mahenthiralingam *et al.*, 2000). We recently expanded the *recA*-based approach to identify the entire *Burkholderia* genus by designing two useful primer sets, BUR1 and BUR2 (Payne *et al.*, 2005), which amplify an almost full-length *recA* gene (869 bp) from all *Burkholderia* species. These primers were found not to be genus-specific in that they amplified other betaproteobacterial species. However, they were useful for the generation of primary sequence data from which the specific primers BUR3 and BUR4 were designed and found to produce a 385 bp amplicon only from *Burkholderia* species (Payne *et al.*, 2005). Recently, Ramette *et al.* (2005) used a high-throughput cultivation-enrichment approach in conjunction with colony hybridization and PCR using 16S rRNA gene and *recA*-derived probes to examine the extent of cultivable *B. cepacia* complex species diversity present in maize-associated soil samples. In this study, we evaluated the *Burkholderia* species *recA* approach as a cultivation-independent assay performed on the same maize rhizosphere samples from which *B. cepacia* complex species have already been cultivated (Ramette *et al.*, 2005).

Materials and methods

Environmental sampling

Maize rhizosphere samples were obtained as described previously (Ramette *et al.*, 2005) from the Michigan State University, W. K. Kellogg Biological Station Long-Term Ecological Research (KBS) site (Hickory Corners, MI), and were stored at −20 °C. Total DNA was extracted from 0.5 g of thawed root system samples with the Bio 101® Systems FastDNA® spin kit for soil according to the manufacturer's instructions (Qbiogene, Cambridge, UK). Four replicate DNA extractions were performed on soil aliquots from the root systems of two maize plants (Plant 1 and Plant 3; Ramette *et al.*, 2005).

PCR detection of *Burkholderia* species *recA* genes

DNA extracted from the maize rhizosphere samples was subjected to PCR to detect *recA* genes from the *Burkholderia cepacia* complex and other *Burkholderia* species using the primers shown in Table 1. *Burkholderia cepacia* complex *recA* genes were amplified using a PCR with primers BCR1 and BCR 2 (BCR1.2 PCR; Table 1) exactly as described by Mahenthiralingam *et al.* (2000). The PCR methods described by Payne *et al.* (2005) were used to amplify *recA* genes from other *Burkholderia* species. PCRs with primers BUR1 and BUR2 (BUR1.2 PCR; Table1) were used to amplify near full-length *recA* gene products from a broad range of species including *Burkholderia* (Payne *et al.*, 2005). Specific PCR of *Burkholderia recA* genes was achieved using PCR with primers BUR3 and BUR4 (BUR3.4 PCR; Table 1) as described (Payne *et al.*, 2005). PCR was performed using Qiagen reagents (Qiagen Ltd., Crawley, UK) in 25 μL reactions. Each PCR contained the following: 1 U *Taq* polymerase, 250 μM of each dNTP, 1 × PCR buffer (containing 1.5 mM $MgCl_2$), 10 pmol of each appropriate oligonucleotide primer, and 20 ng of template DNA. Thermal cycling was carried out in a Flexigene thermal cycler (Techgene Ltd., Cambridge, UK) for 30 cycles of 30 s at 94 °C, annealing for 30 s at the appropriate annealing temperature (see Table 1) and extension at 72 °C for 45 s. The following modifications were included: (i) to amplify fragments suitable for cloning, a 20 min final extension time was included in the thermal cycle; (ii) a nested BUR3.4 PCR was performed using a 1000-fold dilution (*c.* 20 ng) of the BUR1.2 product as the template; (iii) a temperature gradient PCR with annealing temperatures between 57 and 64 °C was performed to optimize the BUR3.4 PCR for amplification of high copy number plasmid template DNA in a Dyad DNA Engine thermal cycler (Bio-Rad Laboratories Ltd., Hemel Hempstead, UK). Approximately 2 μL of each PCR product was visualized by agarose gel electrophoresis as described by Mahenthiralingam *et al.* (2000).

Construction and screening of *recA* gene libraries

PCR products from BUR1.2 and BUR3.4 PCRs were cloned with pGEM®-T Easy vector in competent *Escherichia coli* JM109 cells using the manufacturer's protocols (Promega, Southampton, UK). Recombinant clones were picked into

96-well plates containing Luria–Bertani broth (LB) and the antibiotic ampicillin (100 μg mL^{-1}), grown overnight and then stored at −80 °C after the addition of 8% dimethylsulfoxide to each well. Clones from the BUR1.2 library were screened individually for the presence of putative *Burkholderia*-specific *recA* genes using a high throughput 96-well PCR approach as follows: after revival and growth in 96-well plates containing 150 μL LB broth with ampicillin selection, a small amount of each resulting culture was transferred with a 96-point replicator to 5 μL of lysis solution (Ramette *et al.*, 2005) in a 96-well PCR plate and boiled at 99 °C for 10 min.A cocktail of complete BUR3.4 PCR reagents was added to the boiled lysate and subjected to thermal cycling and agarose gel electrophoresis as described above.

Nucleotide sequence determination

Plasmid DNA from selected clones was prepared using the Wizard® *Plus* SV Minipreps (Promega) as instructed by the manufacturer. Both strands of the *recA* clones were sequenced directly with the M13F and M13R primers as described (Mahenthiralingam *et al.*, 2000). Sequencing reactions were prepared using Applied Biosystems Big Dye Terminator ready reaction mix version 3.1 and analysed using an ABI-PRISM 3100 Genetic Analyser capillary electrophoresis system running Performance Optimized Polymer 6 (POP-6) in accordance with the manufacturer's instructions (Applied Biosystems, Foster City, CA). Raw sequences were aligned, and a consensus sequence was derived using the CAP contig assembly program within the BioEdit software (Hall, 1999). Putative sequence identity was determined using the Basic Local Alignment Search Tool (BLAST; www.ncbi.nlm.nih.gov). A total of 47 representative *recA* sequences were deposited in GenBank under accession numbers DQ076251 through DQ076297 and the aligned sequence set is available from ftp://cepacia.bios.cf.ac.uk/pub/.

Phylogenetic analysis

A number of different phylogenetic methods had been evaluated in previous studies including neighbour-joining, maximum-likelihood and maximum-parsimony analyses (Mahenthiralingam *et al.*, 2000; Payne *et al.*, 2005). The following phylogenetic scheme was found to be straightforward to apply and provided excellent correlation to the current species distribution in the *Burkholderia* genus (Mahenthiralingam *et al.*, 2000; Vermis *et al.*, 2002; Baldwin *et al.*, 2005; Payne *et al.*, 2005). Multiple nucleotide sequence alignments spanning 385 nucleotides of the *recA* gene were constructed using CLUSTAL W (Thompson *et al.*, 1994). Phylogenetic and molecular evolutionary analyses were conducted using genetic distance-based neighbour-joining algorithms within MEGA version 2.1 (http://www.megasoftware.net/). Gaps and missing data were completely deleted by MEGA2.1 before trees were constructed using the Jukes and Cantor (1969) substitution model with random sequence input order and 1000 data sets examined by bootstrapping. Trees were rooted with the *Bordetella pertussis recA* gene as an outgroup (accession number X53457). To check for the presence of chimeric *recA* genes, sequences were divided into two fragments of equal length and phylogenetic analysis performed with each half. The novel alignment of the majority of CL1.2 and CL3.4 clones was not due to the presence of chimeric *recA* sequences as clustering remained consistent in trees constructed from each half of the sequence.

Results and discussion

The *recA*-based PCR approaches developed in a previous study (Payne *et al.*, 2005) enabled the identification and taxonomic assignment of *Burkholderia* species using DNA extracted from cultivated isolates. To test their efficacy as primers for the investigation of *Burkholderia* diversity in the natural environment, we applied them as a cultivation-independent approach to soil from the root systems of maize plants known to be positive for cultivated members of the *Burkholderia cepacia* complex (Ramette *et al.*, 2005). Amplification of putative *recA* genes using the broad-range BUR1.2 PCR demonstrated that three of four DNA extractions from Plant 1, and 4 of four extractions from Plant 3 were positive for the expected 869 bp PCR product. Testing of the same samples with a BCR1.2 PCR (Table 1; Mahenthiralingam *et al.*, 2000) demonstrated that only two of four extractions from Plant 3 were positive for *B. cepacia* complex-specific *recA* genes. None of the maize rhizosphere samples were positive for the *Burkholderia*-specific BUR3.4 PCR when it was applied directly to the extracted DNA. Given the success of the BUR1.2 and BCR1.2 PCRs, the failure of the BUR3.4 PCR to amplify the soil-extracted DNA is difficult to explain; lack of amplification may have occurred for a number of reasons. The BUR3.4 PCR used 17-mer primers that were shorter than the 19–21 base primers used in the successful *recA*-specific PCRs (Table 1). In combination with the annealing temperature of 57 °C required to obtain the correct 385 bp product from the BUR3.4 PCR, this may have produced conditions which were too stringent to facilitate amplification of limited number of *Burkholderia recA* genes present in the soil samples. In addition, the target site of primer BUR4 was within a central region of the gene that may have a reduced capacity for amplification in comparison with the other *recA*-specific oligonucleotides which were designed to prime at each end of the 1 kb gene. The samples from Plant 3 were selected for further study because: (i) the positive results of the BCR1.2 PCR indicated that they contained DNA from

B. cepacia complex species and (ii) previous cultivation-based analysis had demonstrated that moderate numbers of these bacteria (6.5 $\log_{10}$ CFU per g soil) were recovered from the rhizosphere of this plant (Ramette *et al.*, 2005). Also, by fulfilling the latter criteria, the sample provided a suitable test of discriminatory power of the *Burkholderia* genus *recA* identification approach (Payne *et al.*, 2005) and its ability to distinguish members of the *B. cepacia* complex from other members of this genus.

To estimate *Burkholderia* diversity in the DNA extracted from Plant 3, a two-stage *recA* gene library screening strategy was applied as follows. In the first stage, products from a BUR1.2 PCR (Table 1) were cloned to construct the *recA* library, CL1.2, and clones from this library were screened for the presence of putative *Burkholderia*-specific sequences using a high throughput 96-well BUR3.4 PCR (Table 1). A total of 47 putative *Burkholderia recA*-positive clones were identified from 384 clones screened and 42 of these were successfully sequenced. Eleven (26%) of the sequenced 869 bp *recA* clones were found to be representative of *Burkholderia* species *recA* genes using BLAST analysis. In addition, 28 clones from the CL1.2 library which were negative for amplification with the BUR3.4 PCR were selected at random and analysed to determine the background of non-*Burkholderia* species associated with the sample. None of these were *Burkholderia recA* sequences, but were most closely related to *Betaproteobacteria* such as *Herbaspirillum* spp., *Rubrivivax* spp., and *Dechloromonas* spp., and *Gammaproteobacteria* such as *Legionella* spp., *Xylella* spp., and *Methylococcus* spp. (data not shown). In the second stage screen, a nested PCR with the *Burkholderia*-specific BUR3.4 primers (Table 1) was performed on the BUR1.2 PCR products described above. The resulting 385 bp products were subcloned to create the library CL3.4 and 60 clones were subjected to nucleotide sequence analysis. Forty-four of the CL3.4 clones produced good sequence data and 37 (84%) sequences were found to be representative of *Burkholderia* species. In total, 114 *recA* sequences were determined from the maize rhizosphere sample.

To determine the prevalence of *Burkholderia* DNA associated with maize rhizosphere, the *recA* clone library sequences were subjected to phylogenetic analysis. The tree topology observed correlated well with previous studies (Mahenthiralingam *et al.*, 2000; Payne *et al.*, 2005; Ramette *et al.*, 2005) separating *Burkholderia* species from other genera, clearly defining the *B. cepacia* complex and discriminating the species within this group (Fig. 1). Screening *recA* gene library CL1.2 was the least sensitive means of detecting *Burkholderia* species, resulting in a final detection specificity of 26%. Eleven of the *Burkholderia* species *recA* sequences present in library CL1.2 clustered within the genus and were associated with three novel groups, only distantly related to known cultivated species, (Groups 1, 2, and 4, Fig. 1; bootstrap values > 50%). Only the *Burkholderia phenazinium recA* sequence clustered with these novel *Burkholderia recA* sequences and all of the other cultivated species within the genus fell outside the cluster. Of the remaining CL1.2 *recA* sequences, one was distantly related to *Burkholderia glathei* (clone 4c11; Fig. 1) and three identical CL1.2 *recA* sequences clustered closely with *Burkholderia gladioli* (Group 6; Fig. 1).

The *recA* gene library CL3.4 derived from the nested BUR3.4 PCR (Table 1) proved a more effective means of detecting *Burkholderia* within the maize soil sample, with 84% of the determined *recA* sequences falling within the genus. The majority of CL3.4 clones (35) clustered within novel phylogenetic groups, including those already defined by CL1.2 (Groups 1, 2, and 4; Fig. 1) as well as two further unique clusters (Groups 3 and 5; Fig. 1). Of the 11 BUR3.4 clones within Group 1, three were identical and eight were very closely related (the sequence difference across the entire group was 0.7%). The sequence variation across the four Group 2 BUR3.4 clones was 1%. Group 3 sequences were composed of six closely related clones (2.4% variation across the group) and were part of a novel cluster that included two discrete CL3.4-derived clones (1b1 and 2a7; Fig. 1). Groups 4 and 5 both contained a total of four closely related clones and the sequence variation within each of the latter groups was identical at 1.4%. The Group 5 sequences were distantly related to *B. glathei*. Only two of the CL3.4 clones, 1c5 and 2a1, were assigned to *B. cepacia* complex species, *Burkholderia ambifaria* and *Burkholderia pyrrocinia*, respectively (Fig. 1).

Determination of *recA* sequences present in an environmental DNA sample using this PCR approach clearly demonstrated the utility of *recA* as a cultivation-independent approach to examine *Burkholderia* diversity. It is an effective means to identify *Burkholderia* species and members of the *B. cepacia* complex in a single screen that would not be possible using methods based on the 16S rRNA gene where limited sequence diversity precludes such discrimination within this group of bacteria. Although amplification with the BUR1.2 PCR was less successful at revealing the presence of *Burkholderia*, the 869 bp product constitutes > 80% of the full length of the average *recA* gene and hence has utility for the generation of sequence databases not only for *Burkholderia*, but also can target other closely related species. The high copy number of the plasmid template DNA from the CL1.2 library may have accounted for the reduced specificity of the BUR3.4 PCR screening of CL1.2 clones. A temperature gradient PCR (annealing temperatures 57–64 °C) was subsequently performed to optimize the screen, resulting in the positive amplification of *Burkholderia* CL1.2 clones up to 60.5 °C, while non-*Burkholderia* clones from CL1.2 failed above 59 °C (data not shown).

Overall, nested PCR with BUR3 and BUR4 showed much greater specificity for the detection of *Burkholderia* DNA

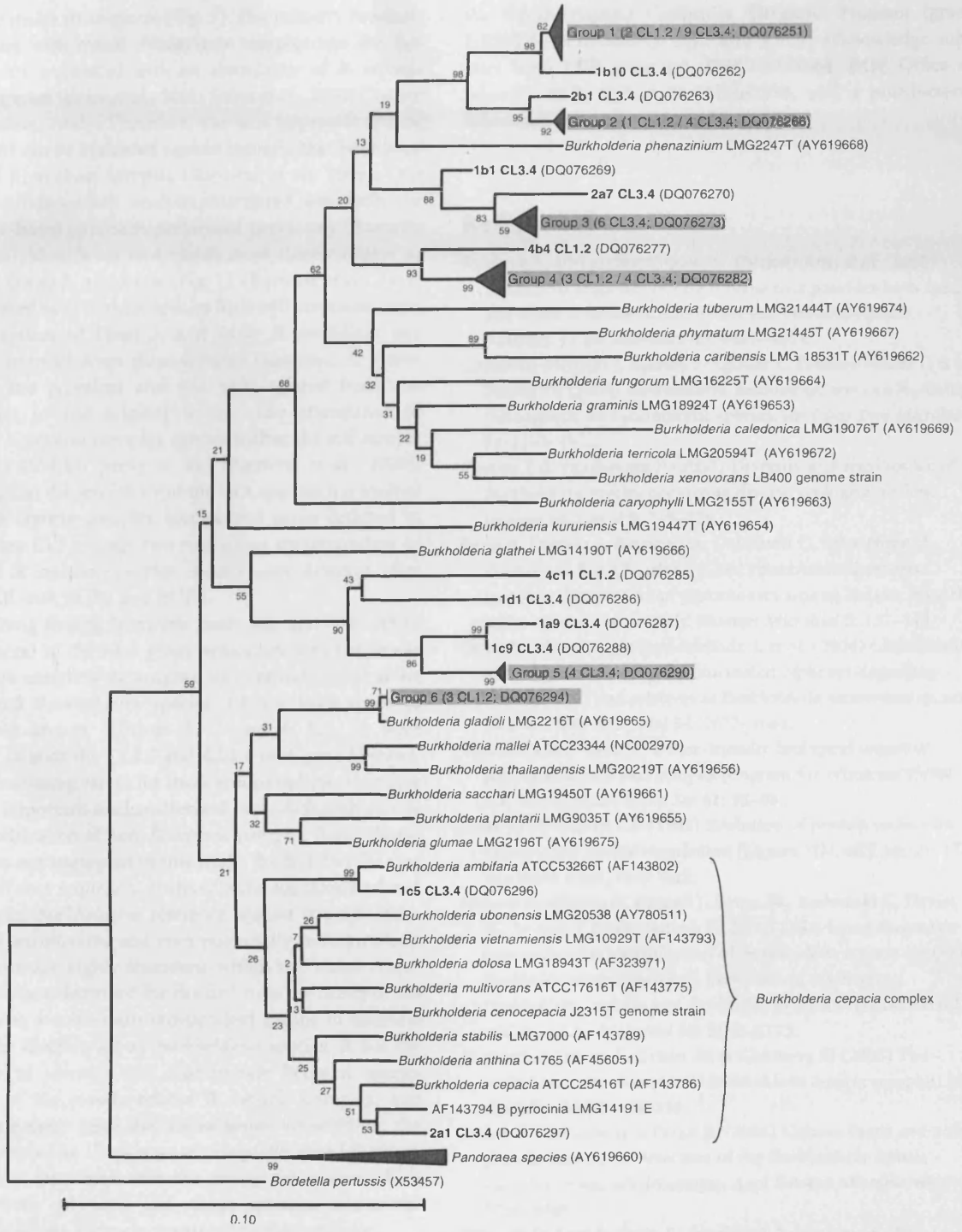


Fig. 1. Phylogenetic analysis of the *Burkholderia recA* sequences obtained from the maize soil environment. Nucleotide sequence alignments of *recA* (385 bases) were constructed and analysed phylogenetically using genetic-distance based neighbour-joining algorithms (Jukes–Cantor matrix model; bootstrapping with 1000 replications; Payne *et al.*, 2005). Six clusters (Groups 1–6) containing multiple novel *Burkholderia recA* genes are shown compressed. The clone library from which the sequence originated, the number of sequences in each cluster, and accession number for each representative *recA* gene are shown in parentheses. Bootstrap values and genetic distance scale (number of substitutions per site) are indicated. Species names and strain numbers for reference *Burkholderia recA* genes are shown with accession numbers in parentheses. Three *Pandoraea recA* sequences were included as a closely related genus and the *Burkholderia cepacia* complex group are also indicated.

within the maize rhizosphere (Fig. 1). The primary rationale for working with maize rhizosphere samples was the fact that they are associated with an abundance of *B. cepacia* complex species (Fiore *et al.*, 2001; Salles *et al.*, 2002; Coenye & Vandamme, 2003). Therefore, the *recA* approach (Payne *et al.*, 2005) can be evaluated against bacteria that have been cultivated from these samples (Ramette *et al.*, 2005). Our cultivation-independent analysis correlated well with the cultivation-based approach performed previously (Ramette *et al.*, 2005), identifying *recA* clones most closely related to *B. ambifaria* and *B. pyrrocinia* (Fig. 1). (Ramette *et al.*, 2005) had cultivated both of these species from soil associated with the root system of Plant 3, and while *B. ambifaria* was recovered from all seven plant samples examined, *B. pyrrocinia* was less prevalent and was only isolated from one other plant in the original survey. The abundance of these two *B. cepacia* complex species within the soil sample was $< 1 \times 10^6$ CFU per g of soil (Ramette *et al.*, 2005), indicating that the sensitivity of the *recA* approach is limited with no *B. cepacia* complex species *recA* genes detected in clone library CL1.2. Only two *recA* genes corresponding to cultivated *B. cepacia* complex species were detected after nested PCR with BUR3 and BUR4.

An exciting finding from this study was that 90% (43 of 48 sequences) of the *recA* genes associated with the maize rhizosphere sample were assigned to putatively novel as yet uncultivated *Burkholderia* species. Clones from three of these novel groups (Groups 1, 2 and 4; Fig. 1) were identified in both the CL1.2 and CL3.4 *recA* gene libraries. The high bootstrap values for these groups indicate that they represent important uncharacterized taxa. Although simultaneous cultivation of non-*B. cepacia* complex *Burkholderia* species was not attempted in this study, the fact that the vast majority of *recA* sequences analysed were not closely related to cultivated *Burkholderia* reference species strongly indicated that uncultivated and even potentially endosymbiotic *Burkholderia* are highly abundant within the maize rhizosphere. We have described for the first time the utility of the *recA* gene as a cultivation-independent means to examine directly the diversity of all *Burkholderia* species. It has the resolution to identify and discriminate between species members of the closely related *B. cepacia* complex, and differentiate them from the rest of genus. In addition, the analysis revealed an abundance of potentially novel *Burkholderia* species associated with the maize soil system which have not been cultivated and whose potential role in the maize rhizosphere is clearly worthy of further analysis.

Acknowledgements

G.W.P. is funded by a PhD studentship from the UK Natural Environment Research Council (grant NER/S/A/2002/10347A), and we also acknowledge additional funding from the Environmental Genomics Thematic Program (grant NER/T/S/2001/00299). A.R. and J.M.T. acknowledge support from NSF grant no. DEB-00755564, DOE Office of Science (BER) DE-FG02-98ER62535, and a postdoctoral fellowship from the Swiss NSF to A.R.

References

Baldwin A, Mahenthiralingam E, Thickett KM, *et al.* (2005) Multilocus sequence typing scheme that provides both species and strain differentiation for the *Burkholderia cepacia* complex. *J Clin Microbiol* **43**: 4665–4673.

Caballero-Mellado J, Martinez-Aguilar L, Paredes-Valdez G & Santos PE (2004) *Burkholderia unamae* sp. nov., an N_2-fixing rhizospheric and endophytic species. *Int J Syst Evol Microbiol* **54**: 1165–1172.

Coenye T & Vandamme P (2003) Diversity and significance of *Burkholderia* species occupying diverse ecological niches. *Environ Microbiol* **5**: 719–729.

Fiore A, Laevens S, Bevivino A, Dalmastri C, Tabacchioni S, Vandamme P & Chiarini L (2001) *Burkholderia cepacia* complex: distribution of genomovars among isolates from the maize rhizosphere in Italy. *Environ Microbiol* **3**: 137–143.

Goris J, DeVos P, Caballero-Mellado J, *et al.* (2004) Classification of the biphenyl- and polychlorinated biphenyl-degrading strain LB400T and relatives as *Burkholderia xenovorans* sp. nov. *Int J Syst Evol Microbiol* **54**: 1677–1681.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**: 95–98.

Jukes TH & Cantor CR (1969) Evolution of protein molecules. *Mammalian protein metabolism* (Munro HN, ed.), pp. 21–132. Academic Press, New York.

Mahenthiralingam E, Bischof J, Byrne SK, Radomski C, Davies JE, Av-Gay Y & Vandamme P (2000) DNA-based diagnostic approaches for identification of *Burkholderia cepacia complex, Burkholderia vietnamiensis, Burkholderia multivorans, Burkholderia stabilis*, and *Burkholderia cepacia* genomovars I and III. *J Clin Microbiol* **38**: 3165–3173.

Mahenthiralingam E, Urban TA & Goldberg JB (2005) The multifarious, multireplicon *Burkholderia cepacia* complex. *Nat Rev Microbiol* **3**: 144–156.

Miller SC, LiPuma JJ & Parke JL (2002) Culture-based and non-growth-dependent detection of the *Burkholderia cepacia* complex in soil environments. *Appl Environ Microbiol* **68**: 3750–3758.

Minerdi D, Fani R, Gallo R, Boarino A & Bonfante P (2001) Nitrogen fixation genes in an endosymbiotic *Burkholderia* strain. *Appl Environ Microbiol* **67**: 725–732.

Parke JL & Gurian-Sherman D (2001) Diversity of the *Burkholderia cepacia* complex and implications for risk assessment of biological control strains. *Annu Rev Phytopathol* **39**: 225–258.

Payne GW, Vandamme P, Morgan SH, *et al.* (2005) Development of a *recA* gene-based identification approach for the entire *Burkholderia* genus. *Appl Environ Microbiol* **71**: 3917–3927.

Pirone L, Chiarini L, Dalmastri C, Bevivino A & Tabacchioni S (2005) Detection of cultured and uncultured *Burkholderia cepacia* complex bacteria naturally occurring in the maize rhizosphere. *Environ Microbiol* **7**: 1734–1742.

Ramette A, LiPuma JJ & Tiedje JM (2005) Species abundance and diversity of *Burkholderia cepacia* complex in the environment. *Appl Environ Microbiol* **71**: 1193–1201.

Reis VM, Estrada-de los Santos P, Tenorio-Salgado S, *et al.* (2004) *Burkholderia tropica* sp. nov., a novel nitrogen-fixing, plant-associated bacterium. *Int J Syst Evol Microbiol* **54**: 2155–2162.

Salles JF, De Souza FA & van Elsas JD (2002) Molecular method to assess the diversity of *Burkholderia* species in environmental samples. *Appl Environ Microbiol* **68**: 1595–1603.

Shields MS, Montgomery SO, Cuskey SM, Chapman PJ & Pritchard PH (1991) Mutants of *Pseudomonas cepacia* G4 defective in catabolism of aromatic-compounds and trichloroethylene. *Appl Environ Microbiol* **57**: 1935–1941.

Thompson JD, Higgins DG & Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673–4680.

Van Oevelen S, De Wachter R, Vandamme P, Robbrecht E & Prinsen E (2004) '*Candidatus Burkholderia calva*' and '*Candidatus Burkholderia nigropunctata*' as leaf gall endosymbionts of African Psychotria. *Int J Syst Evol Microbiol* **54**: 2237–2239.

Vermis K, Coenye T, Mahenthiralingam E, Nelis HJ & Vandamme P (2002) Evaluation of species-specific recA-based PCR tests for genomovar level identification within the *Burkholderia cepacia* complex. *J Med Microbiol* **51**: 937–940.