FINAL REPORT

Rhizosphere Bacterial Degradation of RDX, Understanding and Enhancement

SERDP Project ER-1504

February 2014

Stuart Strand
David Stahl
University of Washington

Neil Bruce
Elizabeth Rylott
University of York

Barbara MacGregor
University of North Carolina

Distribution Statement A
This document has been cleared for public release

SERDP
DOD ▪ EPA ▪ DOE

This report was prepared under contract to the Department of Defense Strategic Environmental Research and Development Program (SERDP). The publication of this report does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official policy or position of the Department of Defense. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the Department of Defense.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to the Department of Defense, Executive Services and Communications Directorate (0704-0188). Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
02-2014	Technical Report	2006-02-2014

4. TITLE AND SUBTITLE
Rhizosphere Bacterial Degradation of RDX, Understanding and Enhancement

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Stuart Strand
David Stahl
Neil Bruce
Elizabeth Rylott
Barbara MacGregor

5d. PROJECT NUMBER
ER-1504

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Washington
Dept. of Civil & Environmental Engineering - Box 355014
Seattle, WA 98105

8. PERFORMING ORGANIZATION REPORT NUMBER
ER-1504

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
SERDP/ESTCP
4800 Mark Center Drive, Suite 17D08
Alexandria, VA 22350-3605

10. SPONSOR/MONITOR'S ACRONYM(S)
SERDP/ESTCP

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Our overall goals were to identify RDX-degrading bacteria in plant rhizospheres, discover the factors that control their abundance and diversity, and develop probes that can be used in the field to detect them. We suggest that important controlling variables in determining RDX persistence in soil are carbon and nitrogen availability. This implies that soil bacteria do not effectively degrade energetic materials in situ unless they are associated with a carbon-rich environment that selects for populations active in either direct or cometabolic degradation of RDX. Further, we expected that activity of RDX-degrading bacteria would be favored by specific carbon sources in the root exudate spectrum, allowing the rhizosphere to be manipulated to enhance populations of RDX-degrading bacteria.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE		223	Stuart Strand
					19b. TELEPHONE NUMBER *(Include area code)* 970-491-8224

Reset

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

I. Table of Contents

I. Abstract 9
A. Objectives 9
B. Technical Approach 9
C. Results 9
D. Benefits 10
II. Objectives 11
III. Background 11
IV. Results and Discussion 14
A. Microbe plant interactions: Rhodococcus rhodochrous 11Y (Neil Bruce and Liz Rylott, and Astrid Lorenz, University of York, UK) 14
1. Background 14
2. Materials and Methods 15
3. Results 22
4. Discussion 49
B. Incorporation of plant-derived carbon into microbial rRNA: Isotopic probing of rhizosphere RNA/DNA (Barbara Macgregor, Univ. North Carolina, Chapel Hill) 52
1. Introduction 52
2. Methods 53
3. Results 56
4. Conclusions 63
C. Stable isotope probing of RDX degraders: Density separation of ^{15}N-DNA. (Stuart Strand, David Stahl, Peter Andeer, University of Washington, Seattle) 64
1. General Methods 64
2. Data analyses 80
D. The lateral transfer of genes for RDX degradation. (Stuart Strand, David Stahl, Peter Andeer, University of Washington, Seattle) 82
1. Introduction 82
2. Methods 83
a) Enrichment and isolation. 83
b) HPLC quantification of RDX. 84
c) Growth of Microbacterium sp. *MA1 using RDX as a sole nitrogen source.* 84
d) DNA extraction. 84
e) PCR amplification cloning, and sequencing. 84
f) Pulse field gel electrophoresis and Southern analysis. 85
g) Fosmid library construction and sequence analysis. 85
h) Phylogenetic analysis. 86
3. Results 86
4. Discussion 92
E. High sensitivity stable isotope probing by quantitative TRFLP. (Stuart Strand, David Stahl, Peter Andeer, University of Washington, Seattle) 93
1. Introduction 93
2. Materials and methods 95
a) Soil microcosms. 95
b) DNA extraction from soil microcosms. 96
c) Construction of 16S rRNA gene standards. 96
d) Preparation of pure culture genomic DNA. 96
e) DNA sequencing. 96
f) Isopycnic centrifugation and gradient fractionation. 97

g) Quantitative PCR. 97
h) TRFLP sample preparation and processing. 98
i) Data analysis. 98
3. Results 98
4. Discussion 105
F. Stable isotope probing of military training range soils degrading RDX. . (Stuart Strand, David Stahl, Peter Andeer, University of Washington, Seattle) 106
1. Introduction 106
2. Materials and methods 107
3. Results 110
4. Discussion 121
G. Cocultures of *Rhodococcus* sp. EG2B with military range soil isolates. (Stuart Strand, David Stahl, Peter Andeer, University of Washington, Seattle) 123
1. Introduction 123
2. Materials and methods 123
3. Results 124
4. Discussion 127
V. Conclusions and Implications for Future Research 128
VI. Appendices 132
A. Eglin Air Force Base sample information 132
B. SIP analysis workflow 141
C. Eglin SIP data 186
D. SigmaPlot stratistical reports 206
E. List of Technical Publications 215
F. References 216

List of Tables

Table 1. Arabidopsis with mutations in the phenylpropanoid pathway selected for root exudates analysis. 16
Table 2. Compounds detected in Arabidopsis root exudates..... 25
Table 3. Conditions used for growing Alfalfa in a gnotobiotic environment..... 40
Table 4. Mass balance (% RDX recovered). 48
Table 5. Commonly used primers..... 73
Table 6. Typical PCR compositions (20µl) and cycling conditions. All volumes are in microliters. 73
Table 7. The effect of fluorescein labeled primer [6'-FAM]-27F on qPCR standard curve values. 100
Table 8. Consistency between predetermined compositions of DNA mixtures and qPCR-TRFLP quantification of the compositions under varied amplication[a]. 101
Table 9. Calculated buoyant density shifts of peaks between the unlabeled and ^{15}N profiles. 104
Table 10. Overview of RDX degradation of extracted Eglin soil samples and *xplA* qPCR amounts. 111
Table 11. Soil microcosms for SIP analyses. 112
Table 12. DNA density gradients used for SIP analysis..... 113
Table 13. Example of buoyant density calculations and analysis for RF of 154 bps. 114
Table 14. The twelve most dominant RFs in the density gradients based on copy numbers... 115
Table 15. Comparisons of the average G+C contents and estimated BDs of the 16S rRNA genes and flanking regions to the genomic G+C contents[a]. 119
Table 16. Growth of strains used in cocultures on different nitrogen sources. 125
Table 17. Colony counts and the ratio of the colony counts of added strains to EG2B in the cultures at 24.5 hours[a]..... 127

List of Figures

Figure 1. Location of restriction sites in *xplA*. 19
Figure 2. Construction of pMEK-5X and pMEK-11X. The 5'-end of *xplA* is shown with the ribosome binding sites. 19
Figure 3. Set-up for gnotobiotic rhizosphere experiments. 21
Figure 4. Testing methods for collecting root exudates from Arabidopsis. 23
Figure 5. Mass spectroscopy analysis of root exudates. 25
Figure 6. HPLC traces of exudates obtained from Arabidopsis mutant tt9 grown in liquid culture. 26
Figure 7. HPLC traces of exudates obtained from Arabidopsis mutant tt8 grown in liquid culture. 27
Figure 8. HPLC traces of exudates obtained from Arabidopsis wild type Landsberg erectus (Ler) grown on sieves. 28
Figure 9. Exudate profiles obtained from wild type Ler Arabidopsis plants grown aeroponically for 17 days. 29
Figure 10. HPLC trace of exudates obtained from five seedlings *tt8* grown in 10 g vermiculite. 29
Figure 11. HPLC traces of exudate from Arabidopsis mutant line tt8 30
Figure 12. HPLC trace of root exudates from Arabidopsis (Ler) seedlings. 31
Figure 13. Colonization of Arabidopsis seedlings post-inoculation with Rhodococcus rhodochrous. 32
Figure 14. Removal of RDX in plant rhizospheres. 33

Figure 15. Colony forming units in growth media inoculated with *Pseudomonas fluorescens* and Arabidopsis. 34
Figure 16. Griess assays showing nitrite release from RDX by XplA-expressing bacteria. 35
Figure 17. RDX uptake by *P. fluorescens* transformed with two identical clones of pAX1-xplA (pAX1 and pAX2) and empty vector control (pJAK14). 35
Figure 18. Resting cell assays showing RDX uptake by *Pseudomonas* containing pMEK-5X and pMEK-11X. 36
Figure 19. Uptake of RDX by *P. fluorescens* containing pAX1 or pMETAX..... 37
Figure 20. Uptake of RDX by *P. fluorescens* WCS365 containing pME31X, pME30X or pAX1. 38
Figure 21. Uptake of RDX by P. fluorescens F113 containing pME6031-xplA..... 39
Figure 22. Percent RDX recovered from the rhizosphere 41
Figure 23. Percent RDX recovered from the rhizosphere 41
Figure 24. Percent RDX recovered from the bulk sand of pots containing alfalfa with and without inoculation with *Ps. fluorescens* with empty vector or containing xplA 42
Figure 25. Percent RDX recovered from the bulk sand with and without alfalfa and *Ps. fluorescens* 44
Figure 26. Percent RDX extracted from the rhizosphere with and without ALA amendment. . 45
Figure 27. Rhizosphere biomass of plants treated with and without ALA..... 45
Figure 28. RDX extraction from the plant tissue..... 46
Figure 29. Percent RDX recovered from the bulk soil and plant tissue four weeks after seedling transfer. Results are means ± standard error of 14 biological replicas. 47
Figure 30. ^{14}C-bicarbonate incorporation into bacterial and archaeal SSU rRNA..... 53
Figure 31. Incorporation of ^{14}C-bicarbonate into soil microbial RNA..... 57
Figure 32. In silico specificity of BG553 and ALF 963 probes. 58
Figure 33. Group-level capture probes with and without helper probes. 58
Figure 34. *In vitro* specificity of the ALF963 and BG553 capture probes..... 60
Figure 35. Optimization of rRNA-SSCP for pure-culture, Bact338-captured 16S rRNA. 61
Figure 36. Isolation and rRNA-SSCP separation of bacterial SSU rRNA from soil and rhizosphere..... 62
Figure 37. Large-scale RNA-SSCP separation of bacterial SSU-rRNA. 63
Figure 38. Absorbance spectrum of RDX. 68
Figure 39. Example of heat maps generated from qPCR-TRFLP analysis scripts..... 82
Figure 40 Phylogenetic tree of selected RDX-degrading bacteria inferred from 16S rRNA sequence relationships. Phylogenetic relationships of characterized RDX-degrading bacteria that carry xplA and Kimura two-parameter model (Kimura 1980, Felsenstein 1989). RDX degraders are shown in boldface. GenBank accession numbers are in parentheses..... 87
Figure 41 Growth of Microbacterium sp. MA1 on RDX. 88
Figure 42 Hybridization of *xplA* gene probe to *Microbacterium* sp. strain MA1 and *Rhodococcus rhodochrous* 11Y DNA resolved by PFGE..... 89
Figure 43 Distribution of transposases and IS elements in pMA1. 90
Figure 44. Functional annotation of 52.2 kbp of sequence from pMA1. 91
Figure 45. Tandem qPCR-TRFLP analysis of CsCl gradients. 95
Figure 46. qPCR-TRFLP analysis of 16S rRNA genes in SIP fractions..... 103
Figure 47. Stable isotope probing (SIP) profiles of soil DNA peak 144 and RFs of *Microbacterium* sp. MA1 in the buoyant density gradients determined using three methods. 105
Figure 48. RDX degradation in soil microcosms (ca. 12% w/v) incubated with RDX enrichment media. 112

Figure 49. Heat map showing the distribution of the RF copy numbers across four of the gradients..... 116
Figure 50. *xplA* gene copy number distribution in ^{14}N- and ^{15}N-density gradients..... 118
Figure 51. Nitrite production in cultures of *Rhodococcus* sp. EG2B degrading RDX..... 121
Figure 52. RDX nitrite concentrations in *Rhodococcus erythropolis* EG2B in monoculture and cocultures..... 126

List of Acronyms

cDNA Complimentary DNA
C+G Cytosine & Guanine
CYP177A1 Cytochrome P450 177A1, the classification of XplA
DGGE Denaturing gradient gel electrophoresis
DNA Deoxyribonucleic acid
FAD Flavin adenine dinucleotide
FMN Flavin mononucleotide
GTN Glycerol trinitrate
HMX High metling explosive (octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine)
HPLC High performance liquid chromatography
MEDINA Methylendinitramine
MG Mono Glucoside
MS Mass Spectrometry
NDAB 4-nitro-2,4-diazabutanal
NAD Nicotinamide adenine dinucleotide
NADP Nicotinamide adenine dinucleotide phosphate
PCR Polymerase chain reaction
pME6010, pME6030, pME6031, pBG1180
plasmid vectors used in genetic transformations
RDX Royal demolition explosive (hexahydro-1, 3, 5-trinitro-1,3,5-triazine hexahydro- 1,3,5-trinitro-1,3,5-triazine
RNA Ribonucleic acid
rRNA Ribosomal RNA
SSCP Single-stranded conformational polymorphism
TNT 2,4,6-Trinitrotoluene
T-RFLP or tRFLP
Terminal restrication fragment polymorphism, or DNA fingerprinting
UDP Uridine Diphosphate
UY University of York
UW University of Washington
XplA RDX-degrading enzyme from *Rhodococcus rhodochrous* 11Y
xplA Gene encoding RDX-degrading enzyme from *Rhodococcus rhodochrous* 11Y
XplB Reductase encoded adjacently to *xplA* in *Rhodococcus rhodochrous* 11Y genome
xplB Gene for reductase encoded adjacently to xplA in *Rhodococcus rhodochrous* 11Y genome

Keywords
RDX, Rhodococcus rhodochrous, XplA, XplB, rhizosphere, Pseudomonas fluorescens

Acknowledgements

Students:

Peter Andeer, University of Washington, Seattle

Astrid Lorenz, University of York, UK

The shuttle vectors pME6010, pME6030 and pME6031 were kindly donated by Dieter Haas Université de Lausanne, Switzerland.
The *Pseudomonas fluorescens* strain WCS 365 was kindly given by Dr G. Bloemberg (Leiden University, The Netherlands) and Dr F. O'Gara provided *P. f*luorescens F113 (University College Cork, Cork).
The nod-box system containing pBG1180 was kindly provided by Dr Rafael Rivella (University of Madrid, Spain).

We thank Guy Amplemann and the staff of the Biotechnology Research Institute, National Research Council Canada, for synthesis of the fully labeled ^{15}N-RDX.

I. Abstract

A. Objectives

Our overall goals were to identify RDX-degrading bacteria in plant rhizospheres, discover the factors that control their abundance and diversity, and develop probes that can be used in the field to detect them. We suggest that important controlling variables in determining RDX persistence in soil are carbon and nitrogen availability. This implies that soil bacteria do not effectively degrade energetic materials *in situ* unless they are associated with a carbon-rich environment that selects for populations active in either direct or cometabolic degradation of RDX. Further, we expected that activity of RDX-degrading bacteria would be favored by specific carbon sources in the root exudate spectrum, allowing the rhizosphere to be manipulated to enhance populations of RDX-degrading bacteria.

Our objectives were to test the following hypotheses:

Hypothesis 1: Soil bacteria that can degrade RDX are carbon limited. Therefore, bacterial RDX degradation is enhanced by carbon compounds exuded by roots in the rhizosphere.

Hypothesis 2: The rhizosphere bacteria community is nitrogen limited. RDX serves as a nitrogen source for rhizosphere bacteria.

Hypothesis 3: The type of carbon compound in root exudate influences RDX degradation.

B. Technical Approach

We characterized root exudates of Arabidopsis mutant lines inoculating the rhizospheres of Arabidopsis and slender wheatgrass with *R. rhodochrous* and *P. fluorescens* and determining the survival of the inoculant strains. The RDX degrading strain, The *xplA* gene was expressed in *P. fluorescens*. Degradation of RDX by defined cultures of sterile alfalfa and transformed *P. fluorescens* was assayed.

Wheatgrass grown in RDX degrading soil from training ranges was exposed to ^{14}C labeled CO_2 and soil samples analyzed for the identity of rhizosphere bacteria growing on plant exudates using rRNA separation techniques: RNA hybridization with appropriate probes and captured on beads and separation of ribosomal RNA by single stranded conformational polymorphisms (SSCP) on minigels. Group specific probes were developed for broad phylogenetic groupings of the Proteobacteria.

The xplA gene was localized on extrachromosomal elements in known RDX degrading bacteria. Stable isotope probing with ^{15}N-labeled RDX was used to identify bacteria in training range soils that assimilated nitrogen from RDX. A tandem qPCR-TRFLP protocol was developed that improved SIP resolution and allowed the degree of label incorporation to be determined for individual members of the bacterial population. This method was applied to soils obtained from the Eglin Air Force Base training range.

C. Results

HPLC profiles of hydroponic media of Arabidopsis mutants varied significantly between replicates, which complicated interpretation of differences between mutant lines. Treatment of the roots during extraction accounted for some of the variation, but not all. Culture conditions, especially the presence or absence of sucrose in the medium, or hydroponic, aeroponic, and vermiculite culture produced highly different root exudate profiles.

R. rhodochrous 11Y was inoculated into cultures of Arabidopsis, wheatgrass, and alfalfa, but it did not efficiently colonize the rhizospheres and RDX in these rhizospheres was not enhanced. Two efficient root colonizer strains of Pseudomonas fluorescens were transformed with xplA under inducible control. When these transformed strains were introduced into soils RDX removal increased compared to soils inoculated with *Ps. fluorescens* without xplA. Plants

grown in the soil inoculated with xplA-transformed *Ps. fluorescens* contained less RDX. Transformed *Ps. fluorescens* persisted in alfalfa rhizospheres even as other bacteria colonized the soil, but the presence of RDX did not enhance persistence of the transformed species. RDX removal was 30% higher in soils inoculated with transformed *Ps. fluorescens* compared with controls. These transformed strains lacked xplB.

Our efforts to label plant root exudates with ^{14}C and to use SSCP gel electrophoresis of RNA extracted from rhizosphere communities failed to detect labeled RNA. We developed group-specific capture probes for bacterial RNA and improved our SSCP and RNA extraction methods.

Pulse field gel analysis localized the RDX-degrading gene xplA to extrachromosomal elements in Rhodococcus and a distantly related Microbacterium. The *R. rhodochrous* 11Y and Microbacterium plasmid sequences in the vicinity of xplB and xplA were nearly identical and contained flanking insertion sequence (IS) elements, suggesting that xplA/B was transferred by horizontal gene transfer.

Because 15N SIP results in inadequate separation of labeled bands to clearly separate DNA variations due to C+G contents, we developed a tandem qPCR - TRFLP protocol that improves resolution by quantifying labeling of the different taxonomic groups independent of their C+G (Cytosine and Guanine) content. We verified separation of ^{15}N labeled DNA extracted from low and high G+C bacterial isolates and from soil microcosms amended with known amounts of genomic DNA from bacterial isolates.

We measured aerobic RDX degradation in surface soils extracted from a highly used target area of Eglin Air Force Base bombing range. RDX-degradation activity was spatially heterogeneous and dependent upon the addition of exogenous carbon sources to the soils. SIP analysis of soils exposed to fully-labeled (ring and nitro) ^{15}N-RDX in microcosms revealed several bacteria species that were fully labeled with ^{15}N-labeled DNA during and following RDX-degradation, including xplA-bearing organisms. A Rhodococcus species was the most prominent genus in the RDX-degrading microcosms and was completely labeled with ^{15}N-nitrogen from the RDX. Other highly-labeled species identified in the gradient included Mesorhizobium sp., Variovorax sp., Rhizobium sp. and unspecified Proteobacteria. A *Rhodococcus* sp.(EG2B) and a *Williamsia* sp. capable of degrading RDX were isolated from these soils and each possessed the genetic element encompassing the xplB and xplA genes identified in the xplA-bearing strains *R. rhodochrous* 11Y and *Microbacterium* sp. MA1. The presence of these genes indicate that xplA/B can persist in military range soils and would be a candidate genetic biomarker indicating the potential for RDX-degradation.

D. Benefits

This work advances fundamental understanding of the distribution of xplA/B in soil microbial communities. Our findings support the prevalence of Rhodococcus for RDX degradation in training range soils, while suggesting that RDX degradation may also occur as the result of Gram negative bacterial activity, resulting in assimilation of nitrogen derived from RDX.

An important observation was that RDX degradation potential and the occurrence of xplA was highly heterogeneous in samples taken from the target area at Eglin training range. More than half of the samples were unable to degrade RDX even with added carbon, and xplA was not detected in four of the soils. Munitions particulates deposited on soils that lack bacteria able express XplA are likely to leach RDX into the subsurface.

The second observation of importance is that soil samples incubated without added carbon were unable to degrade RDX. Thus the presence of both xplA or other RDX biological degradation mechanism and carbon substrates were necessary for RDX degradation in Eglin training range soil samples. This observation is consistent with the role of xplA as a nitrogen releasing mechanism in bacteria isolates growing on RDX.

Our results suggest that bioaugmentation (with xplA-bearing species) and biostimulation (with exogenous carbon sources) may be useful methods to increase RDX degradative potential in training range soils in target areas. Bioaugmentation and biostimulation could be accomplished using ground machinery, manned or remote controlled, by aerial application, or delivered on target ballistically.

II. Objectives

Our overall goals were to identify RDX-degrading bacteria in plant rhizospheres, discover the factors that control their abundance and diversity, and develop probes that can be used in the field to detect them. We suggest that an important controlling variable in determining RDX persistence in soil is nitrogen availability. This implies that soil bacteria do not effectively degrade energetic materials in situ unless they are associated with a carbon-rich environment that selects for populations active in either direct or cometabolic degradation of RDX. Further, we expect that activity of RDX-degrading bacteria will be favored by specific carbon sources in the root exudate spectrum, allowing the rhizosphere to be manipulated to enhance populations of RDX-degrading bacteria. Through a multifaceted experimental approach, we will define a suite of strategies that can be applied to enhance RDX degradation in the field.

Our objectives were to test the following general hypotheses:

Hypothesis 1: Soil bacteria that can degrade RDX are carbon limited. Therefore, bacterial RDX degradation is enhanced by carbon compounds exuded by roots in the rhizosphere. Incorporation of 14C labeled root exudate into different bacterial populations will be followed using a variety of molecular methods.

Hypothesis 2: The rhizosphere bacteria community is nitrogen limited. RDX serves as a nitrogen source for rhizosphere bacteria. We analyzed 16S rDNA sequences in the 15N enriched bacterial population labeled by using isotopically labeled RDX.

Hypothesis 3: The type of carbon compound in root exudate influences RDX degradation. Using mutant Arabidopsis, we manipulated the carbon composition of the root exudate and follow changes in RDX metabolism by the rhizosphere community

The expected results from this project were an understanding of (1) which bacteria are responsible for RDX degradation in plant rhizospheres, (2) which plant exudates stimulate RDX-degrading bacteria, and (3) which plant pathways or specific genes can be manipulated to favor growth of RDX-degrading bacteria in the rhizosphere. This understanding will allow us to define strategies for plant selection to enhance rhizosphere degradation of RDX. Our ultimate goal is to transition to field application by identifying strategies to select plants that favor RDX degradation in their rhizospheres. Additionally, these analyses will provide the foundation for developing molecular probes the screen plant systems for potential RDX degradation. In transitional work, quantitative PCR of both 16S rRNA and RDX-degrading enzymes will be applied to DNA extracts from field soils and greenhouse studies with plants from training range impact areas and surface runoff environments.

III. Background

Soil contamination by explosives on training ranges is a significant problem. Compounds used in explosives can be highly toxic, and their accumulation in the environment results in ecotoxicity and increased risk to human health. Containment or removal of energetic compounds from training ranges poses great challenges for existing technologies. Since very large land areas and groundwater volumes become contaminated during munitions training, soil or groundwater removal and/or treatment by conventional means is not feasible. Only remediation strategies that incorporate the biodegradative potential of microbes and plants offer the possibility of in-place destruction of contaminants. Plants are common interceptors of

pollutants in swales and wetlands, and bacterial degradation in runoff often occurs in communities associated with plant roots.

The use of microorganisms to remediate RDX contamination has been investigated, and several strains capable of degrading RDX have been isolated. Microorganisms have been shown to be able to attack RDX under aerobic conditions, generally when it is supplied as the sole nitrogen source. Bacteria identified include three strains of Corynebacterium (Yang, Wang et al. 1983, Fournier, Halasz et al. 2002) Stenotrophomonas maltophilia strain PB1(Binks, French et al. 1996), Rhodococcus sp., strain DN22 (Coleman, Spain et al. 2002) and Rhodococcus rhodochrous strain 11Y (Seth-Smith, Rosser et al. 2002). The metabolites produced from aerobic degradation of RDX do not include the toxic nitroso and hydrazine derivatives seen under anaerobic conditions, with end products identified including nitrite, ammonium, formate, formaldehyde, and 4-nitro-2,4-diazabutanal (NDAB)(Figure 1) (Fournier, Halasz et al. 2002, Seth-Smith, Rosser et al. 2002). NDAB degrades cometabolically by action of white rot fungi (Fournier, Halasz et al. 2004). Biodegradation of RDX generally occurs more quickly under aerobic rather than anaerobic conditions, and appears to form fewer toxic products. Thus aerobic degradation of RDX may represent a real possibility for the remediation of contaminated sites.

Initial biodegradation studies were performed under anaerobic conditions and several pure cultures of anaerobic RDX degraders have been isolated, including Clostridium bifermentans (Rosser, Basran et al. 2001). The products of RDX degradation by anaerobic bacteria have recently been elucidated (Hawari, Halasz et al. 2000). Biodegradation of RDX generally occurs more quickly under aerobic than anaerobic conditions, and appears to form fewer toxic products. Thus aerobic degradation of RDX in soils represents a promising possibility for the remediation of target areas on live fire training ranges contaminated with RDX.

Work on rhodococcal strains 11Y and DN22 has shown that a cytochrome P450 activity is responsible for the action on RDX. The genes which confer RDX degradative ability to strain 11Y have been cloned and identified as a novel P450 system (XplA), with a flavodoxin domain fused to the P450 domain (CYP177A1).(Seth-Smith, Rosser et al. 2002). Bacteria that degrade RDX in pure culture were all determined to be gram positive, and mostly identified as unique strains of Rhodococcus. Homologs of XplA have now been identified in rhodococci isolated from explosives-contaminated soil in the USA, UK, Belgium, Australia and Israel.

Compounds exuded from plant roots are known to stimulate bacterial activity in the rhizosphere, which in turn could enhance the degradation of munitions compounds such as RDX. Many rhizospheres are considered to be nitrogen limited (Jensen and Nybroe 1999, Koch, Worm et al. 2001, Ekblad and Nordgren 2002). Access to recalcitrant nutrients is also the basis of the “opine strategy,” in which desirable bacterial strains are selected in the rhizospheres of transgenic plants that have been engineered to produce the exotic opine amino acids that the only the engineered bacteria can degrade (Oger, Petit et al. 1997, Savka, Dessaux et al. 2002). Thus, available data strongly suggests that rhizosphere bacteria having the capacity to degrade RDX, and thereby recover nitrogen in nitrogen-limited environment, will have a competitive advantage in the rhizosphere of plants inhabiting RDX contaminated training ranges.

Alternatively, RDX may be degraded cometabolically by bacterial enzymes that attack root compounds structurally similar to RDX. For example, several Arabidopsis mutants have been

shown to have increased or decreased levels of root exudates that have some structural similarity to RDX and are catabolized by rhizosphere bacteria. Bacterial oxygenases are implicated in the cometabolic degradation of ring-structure plant exudates (e.g., indole derivatives that comprise up to 14% of root exudates(Narasimhan, Basheer et al. 2003)) and the degradation of aliphatic and aromatic pollutants (Donnelly, Hegde et al. 1994, Juhasz and Naidu 2000, Arp, Yeager et al. 2001, Joner, Corgie et al. 2002, Leigh, Fletcher et al. 2002, Van Eerd, Hoagland et al. 2003). While these studies suggest pathways of RDX degradation and the potential role of plant exudates in stimulating degradation, we need to identify the bacteria in the rhizosphere that degrade RDX and the environmental conditions conducive to enhanced degradation.

Molecular fingerprinting techniques have been used to follow changes in the plant rhizosphere community. However, while molecular approaches such as T-RFLP and the construction of clone libraries for rRNA or functional genes can produce lists of rhizosphere bacteria, they provide little information about the roles of these bacteria in the rhizosphere ecology. The flow of carbon from the roots and uptake by the microbial community has been investigated using isotope analysis; in general, 25% of the exudates are incorporated into microbial biomass and soil organic matter and 25% is found in the roots.

New techniques in molecular microbial ecology are emerging that allow links to be inferred between metabolism and bacterial phylogeny. These methods use isotopically-labeled compounds to trace which bacteria have primary access to particular substrates. For example, 13C-labeled substrates such as methane, naphthalene or phenol have been used to identify bacteria that preferentially degrade methane or naphthalene in soil microcosms(Radajewski, Ineson et al. 2000, Radajewski, Webster et al. 2002, Padmanabhan, Padmanabhan et al. 2003, Radajewski, McDonald et al. 2003) or phenol in bioreactors (Manefield, Whiteley et al. 2002). Density gradient centrifugation is used to separate DNA or RNA labeled with 13C, allowing identification of the 16S rRNA genes or rRNA and functional genes of bacteria that have first access to the labeled substrate. Functional phylogenetic analysis using radioactive substrates has also received limited use. For example, the combination of microautoradiography (MAR) with fluorescence in situ hybridization (FISH) has been used to identify single cells that have taken up radiolabel (MAR-FISH).(Nielsen, Andreasen et al. 1999). More recently, microarrays have been used to identify populations that incorporate specific radiolabeled substrates (radiomicroarray).(Adamczyk, Hesselsoe et al. 2003). Substrate-stimulated growth of populations can be monitored in a multiplexed parallel format using a microarray of DNA probes to capture rRNAs. Only growing populations will incorporate radiolabel in rRNA on the array.

We have also developed a variation of the SSCP (single-stranded conformational polymorphism) method, to separate mixtures of rRNAs by non-denaturing polyacrylamide gel electrophoresis (MacGregor and Amann, in prep.). Whereas most 16S rRNAs migrate together in denaturing gels, in which separation is based primarily on size, migration in SSCP gels is influenced by the sequence-dependent folded conformation of the molecules. Ideally, each different rRNA sequence will have a specific migration rate, although we do not yet know the limits of resolution. Individual rRNA bands can then be excised from stained gels for identification by RT-PCR and sequencing.

For the complex RNA mixtures expected in many environmental samples, banding patterns can be simplified by first capturing total 16S rRNA by magnetic bead capture hybridization

(MacGregor, Bruchert et al. 2002). Alternatively, we have found that oligonucleotide-directed ribonuclease H digestion of 16S rRNA can be controlled with single-mismatch discrimination (MacGregor et al., in prep); bands of interest (e.g. those derived from RDX degraders) can be identified by running RNaseH-digested and uncut samples in adjacent lanes.

IV. Results and Discussion

A. Microbe plant interactions: Rhodococcus rhodochrous 11Y (Neil Bruce and Liz Rylott, and Astrid Lorenz, University of York, UK)

1. Background

Rhodococcus rhodochrous strain 11Y was first isolated from explosives-contaminated soil using selective enrichments with RDX supplied as a sole source of nitrogen for growth (Seth-Smith, Rosser et al. 2002, Rylott, Jackson et al. 2006, Jackson, Rylott et al. 2007, Sabbadin, Jackson et al. 2009). The genes encoding the ability to metabolize RDX were isolated and the proteins, XplA and XplB characterized (Rylott, Jackson et al. 2006, Jackson, Rylott et al. 2007, Sabbadin, Jackson et al. 2009).

Microbial activity within the rhizosphere is relatively high, due to nutrient-rich root exudates including organic acids, sugars and amino acids (Lugtenberg and Kamilova 2009) thus rhizosphere soil can contain 10 to 200 times more bacteria than adjacent bulk soil (Miller and Wood 1996). As root exudates are known to stimulate bacterial activity in the rhizosphere, they could in turn enhance the degradation of munitions compounds such as RDX. Indole derivatives make up the majority of the nitrogenous root exudates, and 14% of the total root exudates in Arabidopsis. Two Arabidopsis mutant lines pap1 and fah1 have significantly altered root exudate profiles. The pap1 (Production of anthocyanin pigment 1) mutant has a disrupted MYB regulator of phenylpropanoid biosynthesis and over-accumulates anthocyanin pigments and other secondary metabolites, thus pap1 plants have purple pigmentation in vegetative organs throughout development. Conversely, fah1 contains a mutation in ferulic acid hydroxylase and exhibits reduced sinapoyl malate, syringyl lignin and sinapoyl choline. Sinapoyl esters comprise a significant portion of Arabidopsis root exudates and are easily identifiable using HPLC techniques. In addition, the exudate profiles of a range of transparent testa (tt) lines with mutations in the phenopropanoid pathway were tested.

The composition of the rhizosphere microorganism population is dependent on the root exudates, therefore, the rhizosphere can be seen as a variable system which can be manipulated by changing root exudation to create favorable conditions for RDX-degrading bacteria to thrive. In order to understand plant exudate-microbe interactions, knowledge of root exudate composition is essential. Thus root exudates profiling from wild type and mutant Arabidopsis plants was conducted. Following this, the ability of *R. rhodochrous* to colonize and grow on the roots of Arabidopsis was tested. Whereas the ability of *R. rhodochrous* to colonize roots is unknown, *Pseudomonas fluorescens* strains are known to colonize alfalfa roots efficiently (Villacieros, Whelan et al. 2005). By expressing XplA activity from *R. rhodochrous* 11Y in *P. fluorescens*, we can exploit this root-colonizing ability to investigate whether RDX degradation can be engineered in the rhizosphere.

2. Materials and Methods

a) Arabidopsis root exudate profiling

(1) Plant lines

Phenylpropanoid mutant Arabidopsis lines and wild type ecotypes were obtained from the Nottingham Arabidopsis stock Centre (Table 1).

Phenylpropanoid mutants and wild type were grown up to check the phenotype, fertility and plant vigor and to bulk up seed for further analysis. All plant lines except N83 (*tt2*) germinated and grew satisfactorily. The germination rate of N83 was low and seed set poor. This line was not analyzed further. The ferulic acid hydroxylase mutant, *fah1-7* (At4g36220) as published by Meyer *et al.* (1996) is in Ler background; however, NASC described N8604 to be in Col0 background. Both ecotypes were used in our studies.

Table 1. Arabidopsis with mutations in the phenylpropanoid pathway selected for root exudates analysis.

NASC stock no.	Background	allele	locus	Phenotype
N83	Ler	tt2-1	tt2	Yellow seeds due to absence of brown pigment in seed coat.
N84	Ler	tt3-1	tt3	Yellow seeds and anthocyaninless leaves and stems.
N85	Ler	tt4-1	tt4	Yellow seeds and anthocyaninless leaves and stems.
N86	Ler	tt5-1	tt5	Like tt4, perhaps brighter green.
N87	Ler	tt6-1	tt6	Brown-yellow seeds and reduced anthocyanin content in leaves.
N88	Ler	tt7-1	tt7	Like tt6. Controls flavonoid-3-' hydroxylation.
N89	Ler	ttg-1	ttg	Like *tt1*. Yellow seeds. No anthocyanin in leaves and stems. Trichomes absent. Aberrant seed coat structure.
N164	Ler		tt9	Pale brown seeds due to reduced levels of brown pigment in seed coat. Approximately wild-type anthocyanin levels observed in leaves, stems and other tissues.
N110	Ler	tt10-1	tt10	Brown-yellow seeds and reduced anthocyanin content in leaves.
N111	Ler/Enkheim	tt8-1	tt8	Yellow seeds due to absence of brown pigment in seed coat.
N3884	Col1	pap1-D		Purple pigmentation in many vegetative organs throughout development
N8604	Ler	fah1-7	fah1	No discernable phenotype
NW20	Ler	-	-	Characterized Ler characterized ecotype
N3176	Col1	-	-	Characterized Col1 background ecotype

(2) Plant growth conditions

Prior planting, seeds were sterilized using hydrochloric acid. Therefore approximately 300 seeds were transferred to eppendorf tubes which were placed with lids open into a plastic box. A beaker with 100 ml bleach was placed into the box and 3 ml concentrated hydrochloric acid was added to the bleach. The lid of the plastic box was immediately closed. After at least three hours lids of the eppendorf tubes were closed, seeds removed from the box and placed into the sterile flow hood. Lids were opened for 10 min to allow evaporation of remaining hydrochloric gases. Seeds were then transferred to plates.

For exudates collection different plant growth methods were developed and are described in the results. Unless stated otherwise, all plants were grown in rooms with a 16 h photoperiod (light 80 $\mu mol.m^{-2}.s^{-1}$) at 25 °C.

(3) Exudate extraction

Exudates were purified from the growth medium by passing the growth/extraction medium through a C_{18}-SPE column and eluting in 600 µl 100 % methanol containing 0.1 % acetic acid. Further concentration was achieved using centrifugation under vacuum. The volume was normalized to plant dry weight and analyzed by HPLC and/or LC-MS.

b) Colonization of the rhizosphere by R. rhodochrous

(1) *R. rhodochrous* root colonization assays

Root colonization was measured in vermiculite, perlite and mixtures of vermiculite: soil (2:1) and vermiculite:Terra-green (synthetic small-grain gravel):sand (1:1:1). The different growth media were chosen to optimize Arabidopsis growth in this gnotobiotic system. The mutant line *tt8* and *R. rhodochrous* was used for these assays.

Arabidopsis seeds were surface sterilized and sprinkled onto agar plates containing ½ MS with 0.68 % sucrose. Six-days-old-seedlings were inoculated with cultures of *R. rhodochrous* or *P. fluorescens*. Seedlings were incubated for 1-2 h, washed with sterile water then transferred to the pots containing sterile media. Inoculation of the no plant controls was performed using inoculated seedlings where, after one day, the above ground biomass was cut off, leaving just the roots in the medium. The pots were watered as necessary with nutrition solution. After 22 days, colony forming units (cfu) in the rhizosphere and control soils were determined.

(2) Determination of colony forming units and 16S analysis

To determine the cfu formed by *R. rhodochrous* 1.22 g of soil were diluted into 10 ml of 40 mM phosphate buffer (pH 7.2). Samples were incubated shaking (50 rpm) in the cold room overnight and vortexed for 5 s. Serial dilutions were spread onto LB agar plates to determine the colony forming units. Colony forming units of samples inoculated with *P. fluorescens* were determined by weighing 1g soil from rhizosphere and the bulk and diluting it into 1 ml phosphate buffer (40 mM, pH 7.2). Samples were shaken (50 rpm) for one hour at room temperature, further diluted and aliquots were plated onto agar containing the appropriate antibiotics. For 16S sequencing analysis the primers 16S-fd1 5' -AGAGTTTGATCCTGGCTCAG-3' and 16S-rD2 5'-AAGGAGGTGATCCAGCC-3' were used.

(3) Soil-based RDX degradation assays using *R. rhodochrous 11Y*

Five Arabidopsis or 20 slender wheatgrass seedlings were planted per pot and grown for six weeks. Plants were watered with 15 ml 180 µM RDX, which was watered into the plants 6-8 weeks after sowing. The pots were inoculated with cultures of *R .rhodochrous* 11Y three days prior and three days after dosing with RDX. Pots were placed onto saucers to enable watering of the plants over weeks without the RDX being lost. At the end of the experiment, RDX was watered out and RDX in the soil leachate determined as described (Jackson, Rylott et al. 2007).

c) Expressing xplA in Pseudomonas fluorescens

A number of constructs were created designed to test a range of promoters, ribosome binding sites and other transcription regulatory elements. The native *xplA* sequence, i.e. unmodified for the codon usage in *Pseudomonas*, was used, and the construction of all these plasmids is outlined below.

(1) Cloning of xplA into the shuttle vector pJAK14

The ITPG-inducible shuttle vector, pJAK14, was used to express *xplA* in *P. fluorescens* WCS365 and F113. Primers were used to introduce *HindIII* and *XbaI* restriction sites (Primer *HindIIIR* = 5'-aagcttcaggacaggacgatcgg-3', *XbaIF* = 5'-tctagatgaaccgacgtaactgtcctgttc-3'). The *xplA* gene was ligated into pJAK14 and the whole construct was named pAX1.

(2) Cloning of xplA into the constitutively-expressing shuttle vector pME6010

The shuttle vectors pME6010, pME6030 and pME6031 were kindly donated by Dieter Haas. Primers were designed to introduce *HindIII* and *NcoI* restriction sites: *HindIII* R = 5'-aagcttcaggacaggacgatcgg-3' and NcoI F = 5'-ccatggatgaccgacgtaactgtcct-3'-resulting in a construct named pMEX**.**

HindIIIR and NcofrXPLA = 5'-ccatggccatgaccgacgtaactgtcct-3', resulting in a construct named pMEXfr.

(3) Expressing xplA in pME6010 under the control of the TAC promoter.

The P_{TAC}-*xplA* cassette from pJAK14, was amplified without the repressor by engineering in an NcoI restriction site upstream of the TAC promoter, and a *HindIII* restriction site downstream of *xplA*. The TAC promoter sequence was used as published by (de Boer, Comstock et al. 1983). The cassette was ligated into pME6010 as described above. The construct was named pMETAX.

(4) Using XhoI and HindIII restriction sites to include Pseudomonas specific ribosome binding sites.

The following primer was designed to engineer in *HindIII* and *NcoI* restriction sites: *HindIII* R = 5'-aagcttcaggacaggacgatcgg-3'.

This primer was used in combination with either *XhoI* F = 5'-ctcgagatgaccgacgtaactgtcctg -3', or *XhoI* rbs5 F = 5'-ctcgagagaggagaatcatgaccgacgtaactgtcctg-3' or XhoI rbs11 F = 5'-ctcgagagaggatgtggagaatcatgaccgacgtaactgtcctg-3'. The gene *xplA* was amplified and ligated into the pCR2.1 TOPO vector. Due to the presence of an *XhoI* site in the *xplA* gene, x*plA* was digested from TOPO pCR2.1 in two independent reactions using firstly *XhoI* and *AflIII* and secondly *AflIII* and *HindIII* (Figure 1).

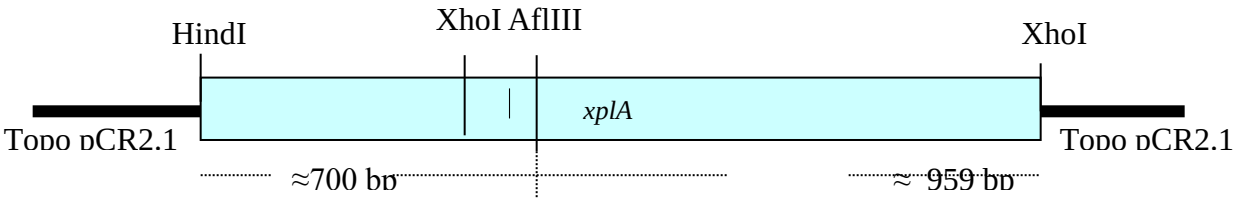


Figure 1. Location of restriction sites in *xplA*.

XplA fragments were separated using 1.5 % agarose gel, purified and ligated into pME6010 using methods described above. The resulting constructs were named pMEK-X, pMEK-5X and pMEK-11X (

Figure 2).

xhoI F 5'-**ctcgag**---------- **atg**accgacgtaactgtcctg -3',
xhoI **start**- *xplA*

xhoI rbs5 F 5'-**ctcgag** ---- -agagga------gaatc------**atg**accgacgtaactgtcctg-3'
xhoI RBS 5 bp **start**-*xplA*

xhoI rbs11 F 5'-**ctcgag**-----agagga---tgtggagaatc- **atg**accgacgtaactgtcctg-3'
xhoI RBS 11 bp **start**- *xplA*

Figure 2. Construction of pMEK-5X and pMEK-11X. The 5'-end of *xplA* is shown with the ribosome binding sites.

d) Cloning of xplA into the constitutively expressing shuttle vector pME6030 and pME6031.

pME6030 is
similar to pME6010, but does not contain the KanR promoter. pME6031 contains a TT_4 transcription terminator. The expression cassette P_{TAC}-*xplA* was digested from pMETAX using *NcoI* and *HindIII* and separated on a 1 % agarose gel. The ≈1760 bp band was excised and gel purified using the Promega Wizard gel purification kit. P_{TAC}-*xplA* was ligated into pME6030 and pME6031, previously digested with *NcoI* and *HindIII* and dephosphorylated. The constructs were called pME30TX and pME31TX**,** respectively.

Bacterial cells were transformed with the following plasmids, in the presence of the selective agents described below: for *E. coli* Rosetta gami B transformed with pET16b-*xplA*: carbenicillin (10 µg/ml), chloramphenicol (10 µg/ml), streptomycin (10 µg/ml) and tetracycline (10 µg/ml), pME6010, pME6030 and pME6031derivates: tetracycline (80 µg/ml). *P. fluorescens* F113: rifampicin (10 µg/ml) and pJAK14: kanamycin (100µg/ml).

e) Codon optimization of xplA.

The codon optimization for expression in *Pseudomonas* was carried out by GeneArt (Germany) and the sequence cloned into the vector pMA. To test if the optimization results in increased XplA activity, *xplA* was cloned into the *Pseudomonas*-specific vector pME6010. The *P. fluorescens* WCS365 was then transformed with pME6010 containing the optimized *xplA* (pMEOX10) and XplA activity measured using resting cell assays by resuspending 0.05 g cells, pre-grown in liquid LB medium, into 1 ml phosphate buffer, pH 7.2 containing 160 μM RDX. Reactions were stopped by taking an aliquot at different time points and adding the same volume of 10 % acetic acid. RDX concentration was measured by HPLC.

f) Whole cell assays

To measure the functional expression of XplA in *P. fluorescens*, Cultures were induced by the addition of IPTG and supplemented with 0.5 mM $FeCl_3$ and 1 mM α-amino levuleic acid (ALA) (unless stated otherwise) Activity towards RDX was followed using the Griess assay and by HPLC (Jackson, Rylott et al. 2007).

g) Griess assay

For the Griess assay cells were diluted 1, 1/2, 1/4 and 1/8 with LB, then RDX added to a final concentration of 100 μM. In a total reaction volume of 180 μl, microtitre plates were left shaking (160 rpm) overnight at 30 ºC and RDX breakdown determined.. Therefore, 50 μl sulphanilamide (10 mg/ml sulfanilamide in 0.68 M hydrochloric acid) was added to each well, followed 10 min later by 20 μl NED (10 mg/ml *N*-(1-naphthyl)-ethylenediamine dihydrochloride). Plates were incubated at room temperature. Color change was monitored visually for the next ten minutes.

h) HPLC assay

For HPLC analysis resting cell assays were performed as follows: overnight cultures of cells (1 g) were resuspended in 2 ml phosphate buffer (40 mM, pH 7.2) and 50 μl added to 450 μl of 162 μM RDX in phosphate buffer. Reactions were incubated shaking (160 rpm) at 30 ºC. Aliquots were taken at different time points and reactions were stopped by the addition of an equal volume of 10 % acetic acid and analyzed by HPLC.

i) Growth of XplA expressing P. fluorescens in a gnotobiotic system with alfalfa

The *P. fluorescens* cells were grown for 24h in liquid Luria-Bertani medium then in minimal medium, with 420 μM NH_4Cl (empty vector) or 140 μM RDX (*xplA*-expressing). Cells were washed and resuspended in phosphate buffer to OD_{600} of 0.3. Alfalfa seeds were sterilized and germinated and grown for 2 days under sterile conditions. Seedlings, or unplanted soil controls, were inoculated with 1 x 10^5 cfu /g of *P. fluorescens* containing either the empty vector or *xplA* and then two seedlings each were transferred into glass tubes containing 30 g sterile quartz sand and 4 ml of RDX medium (30 mM 2-(*N*-morpholino) ethanesulfonic acid, 4.24 mg Murashige and Skoog salts, 70 μM RDX, 2 mM α-aminolevuleic acid and 0.5 mM $FeCl_3$). The tubes were placed at 25 °C, with a 16 h light photoperiod (80 $\mu mol.m^{-2}.s^{-1}$). The experimental set-up is shown in Figure 3. One week after inoculation seedlings were treated with another 1 ml of Murashige and Skoog (1.06 g/l) dissolved in 30 mM 2-(*N*-morpholino)ethanesulfonic acid

(MES) containing 1 mM Ala and 0.5 mM $FeCl_3$. Four weeks after inoculation, RDX was extracted from the bulk and rhizosphere soil.

Figure 3. Set-up for gnotobiotic rhizosphere experiments.

j) Extraction of RDX from the aerial plant tissue

Plant tissue was freeze-dried, ground and RDX extracted as described in the EPA Method 8330. Samples spiked with TNT (0.05 mmol) estimated the recovery of RDX was 80 %. Samples were concentrated under vacuum and then diluted 1:1 with water prior to HPLC analysis.

k) Growth of XplA expressing P. fluorescens in rhizosphere of Arabidopsis phenylpropanoid mutants

Wild type and mutant Arabidopsis lines were grown under non-sterile conditions in a sand and Terra-Green mixture with a weekly nitrogen supplement. Five weeks after sowing, plants were inoculated with 5 ml of 180 µM RDX, 1 mM aminolevulinic acid, 30 mM MES and 384 µl of bacterial suspension (resulting in approx $5x10^7$ colony forming units/ml) containing *P. fluorescens* F113 with either the XplA-expressing vector or the empty vector. Five days after inoculation the plants were given a second inoculation of 3 ml RDX (180 µM), aminolevulinic acid (1 mM), MES (30 mM) and 403 µl of bacterial suspension (resulting in approx $5x10^7$ colony forming units/ml). The plants were grown for a further 12 days prior to flushing the pots with water and determination of RDX in soil leachate using HPLC.

l) Growth of XplA expressing P. fluorescens *in rhizosphere of alfalfa*

40 ml of sand were filled into plant pots, alfalfa seeds were germinated and plants grown for three weeks. Alfalfa plants were dosed twice with 5 ml 180 µM RDX and inoculated with *P. fluorescens* expressing either XplA or containing the empty vector control. Nine days later RDX was watered out using 5 ml H_2O, the leachate collected and separated using the HPLC. Alfalfa seedlings were also grown in 52 g of a sand : F2 compost mix (1:1) and inoculated with either *P. fluorescens* containing the empty vector or *xplA*-expressing *P. fluorescens* three weeks after planting (eight pots per treatment). After four days 10 mls of 160 µM RDX were applied to the top of each pot and 10 days later the pots were flushed with 15 ml of uncontaminated

water and the levels of RDX measured in the soil leachate. The volume of bacteria used relates to an estimated $5x10^7$ colony forming units/ml mixture.

3. Results

a) Liquid culture grown plants

Seeds were surface sterilized then sprinkled onto agar plates containing ½ MS-agar. For stratification, seeds of Arabidopsis wild type and phenylpropanoid mutant lines were imbibed for four nights. Seedlings were placed into the growth room and after 24 h approximately 150 seedlings were transferred into flasks containing 50 ml ½ MS media with 0.68 % sucrose. Flasks were grown shaking (130 rpm), under constant light (80 $\mu mol.m^{-2}.s^{-1}$), with 16 h photoperiod at 25 °C as shown in Figure 4A. After 11 days the medium was replaced with fresh ½ MS. Six days later the ½ MS medium was replaced with 40 ml sterile water for exudate collection. The water containing the exudates was collected after three days and kept frozen until sample preparation. The plant material was freeze dried and weighed.

b) Sieve grown plants

To avoid both the submergence of the leaves in the medium, which might alter the exudate profile (Cuyckens and Claeys 2004), and to reduce oxygen stress, plants were grown on metal sieves with roots in liquid medium. Stainless steel sieves, 100 mm in diameter, on legs 30 mm high (as shown in Figure 4C) were dipped into ½ MS and sucrose agar to provide a surface onto which to grow the Arabidopsis seeds. Seeds were surface sterilized, and sprinkled onto the sieves. The seedlings were grown for four weeks, as shown in Figure 4B, and then incubated overnight in 150 ml of water to collect the exudates for profiling. This method produced large volumes of liquid which required long freeze-drying times to concentrate the exudates for HPLC analysis. To optimize experimental design, exudate collection was performed using a smaller volume, by placing a small Petri dish (diameter 6 cm, Greiner bio-one, Stonehouse, UK) filled with 15 ml of water directly under the roots (as shown in Figure 4C). Secondly, for comparison, roots were dried overnight prior the exudates collection, as drought stress is known to enhance exudation.

c) Aeroponically grown plants

Arabidopsis seeds were sprinkled onto Petri dishes containing 1/2 MS medium ± 0.68 % sucrose or water agar and plates were placed upside down to allow seedlings to grow aeroponically, as shown in Figure 4D. After two weeks, exudates were extracted by removing the seedlings from the agar plate and incubating the roots for 1.75 h in 25 ml of sterile water, as described by (Narasimhan, Basheer et al. 2003). Exudates were concentrated by freeze drying and analyzed by HPLC.

d) Plants grown in vermiculite

Five seedlings were grown in 10 g vermiculite for 3.5 weeks and exudates were extracted by washing the vermiculite twice with 40 ml water. The collected washings were centrifuged and filtered through a 20 μm filter. Exudates were prepared by freeze drying and analyzed by HPLC.

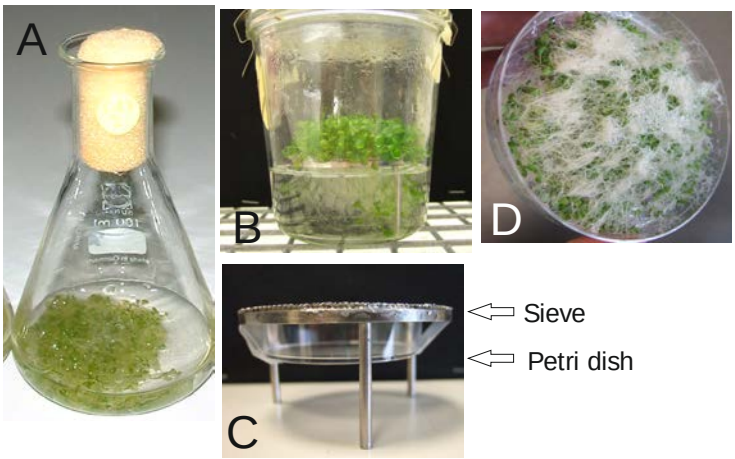


Figure 4. Testing methods for collecting root exudates from Arabidopsis.
(A) Arabidopsis seedlings grown in liquid culture and (B) grown on a stainless steel sieve, (C) improved sieve design. (D) Aeroponically grown Arabidopsis. Seeds were sprinkled onto the agar surface and plates were placed upside down to achieve aeroponical root-growth.

e) HPLC and LC-MS method development

Syringic acid, p-coumaric acid, ferulic acid, quercetin-3-rhamnoside and methyl 4 hydroxybenzoate were used as standards for exudates analysis. Standards (50 µl) were injected onto a TechSphere ODS 80A 5 µ (250 mm x 4.6 mm) column, at concentrations of 0.1 mg/ml, except Quercetin-3-rhamnoside, which had a concentration of 0.03 mg/ml.

Two HPLC methods were used to separate standards and exudates samples.

HPLC method 1:

0-3 min	10 % Acetonitrile (MeCN) + 0.1 % Acetic acid
3-28 min	Gradient to 40 % MeCN + 0.1 % Acetic acid
28-32 min	Gradient to 10 % MeCN + 0.1 % Acetic acid
32-39 min	10 % MeCN + 0.1 % Acetic acid

The following retention times were obtained:

syringic acid (13.11 min), p-coumaric acid (16.89 min), ferulic acid (18.12 min), quercetin-3-rhamnoside (20.32) and methyl-4-hydroxybenzoate (21.46 min).

HPLC method 2:

0-3 min	10 % MeCN + 0.1 % acetic acid
3-18 min	Gradient to 40 % MeCN + 0.1 % acetic acid
18-22 min	Gradient to 10 % MeCN + 0.1 % acetic acid
25 min	10 % MeCN + 0.1 % acetic acid

Following retention times were obtained:

syringic acid (11.33 min), p-coumaric acid (14.27 min), quercetin-3-glycoside (14.68) ferulic acid (14.99 min), quercetin-3-rhamnoside (16.00) and methyl-4-hydroxybenzoate (17.85 min).

For LC-MS analysis 5 µl standard mix were injected onto a TechSphere ODS 80A 5 µ (250 mm x 4.6 mm) column and eluted at 1 ml/min with the same protocol as described above. Peaks were collected and then infused at 5 µl/min into the ESI source. The source voltage was 4.02 kV, sheath gas 60 units, aux gas 10 units, capillary voltage was -28.33 V and capillary temperature 200 ºC. The data was collected with automated gain control on, max ion time 200 ms and with 5 micro scans per scan. MS2 data was collected with isolation widths of 3 m/z units and with collision energy of 35 %. Total ion spectrum was taken between 50-500 mass units and scanned in three segments (Figure 5):

Segment 1: 12 min with 2 scan events. Parent ion 197 g/mol (syringic acid)

Segment 2: 3 min with 3 scan events. Parent ion 163 g/mol (p-coumaric acid), 193 g/mol (ferulic acid) and quercetin-3-b-D-glucoside (464 g/mol)

Segment 3: 16-25 min with 3 scan events. Parent ion 447 g/mol (quercetin-3- rhamnoside) and 151 g/mol (methyl-4-hydroxybenzoate).

MS2 spectra for each parent ion were taken in the mass range specified above. The injection of 5 µl of quercetin-3-rhamnoside at a concentration of 0.03 mg/ml represented the detection limit. Using this method separation and quantification of exudates can be achieved in the event of co-elution.

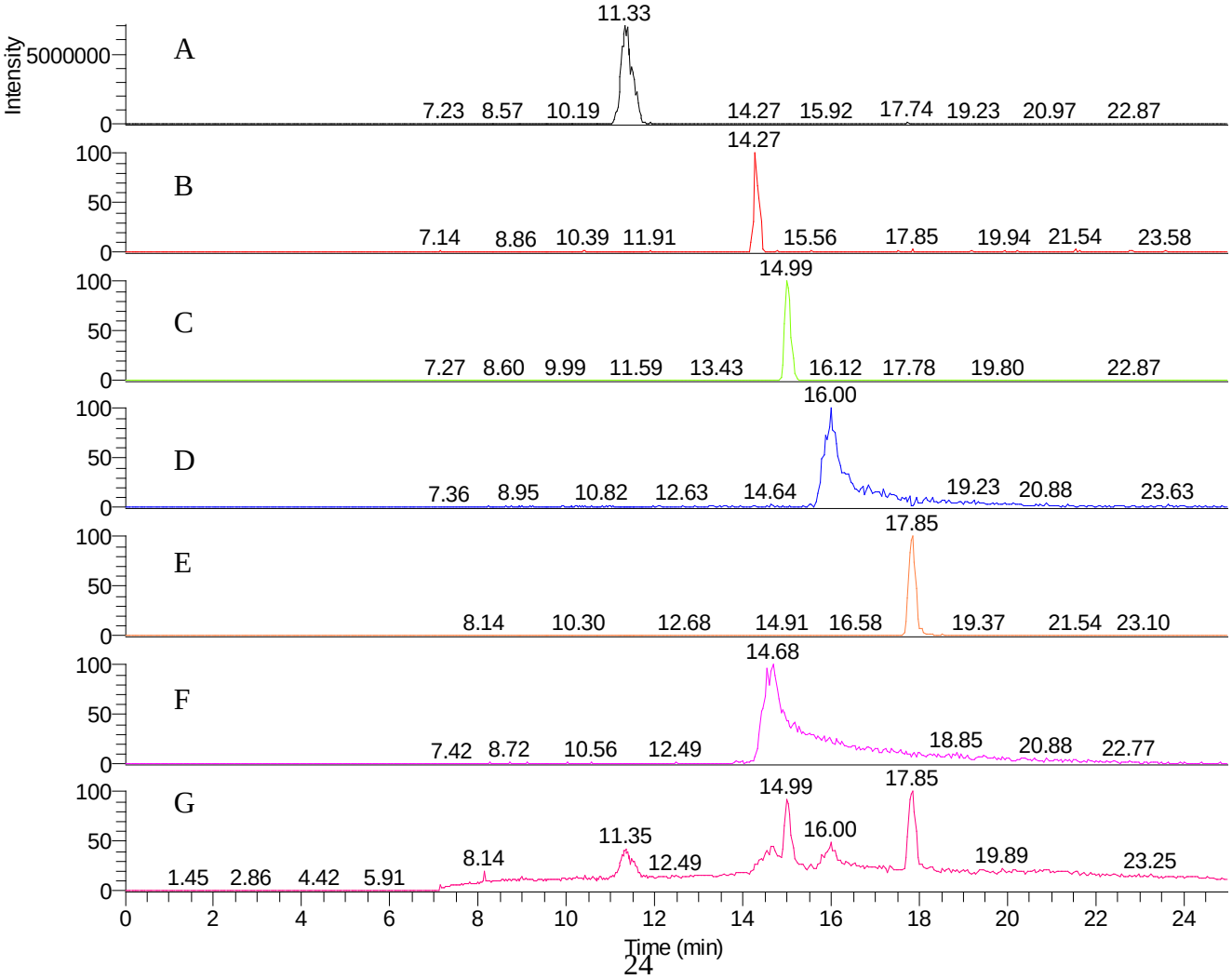


Figure 5. Mass spectroscopy analysis of root exudates.
(A)-(E) Detection of single compounds (A) syringic acid, (B) p-coumaric acid, (C) ferulic acid, (D) quercetin-3-b-D-glucoside, (E) quercetin-3-rhamnoside and (F) methyl-4-hydroxybenzoate. MS2 spectra for each parent ion were taken in the mass range specified above. (G) Full MS of standard mix (all compounds 0.1 mg/ml).

f) Arabidopsis root exudate profiling

Root exudates were collected from plants grown under different conditions. All the most commonly found Arabidopsis root exudates compounds; syringic acid, p-coumaric acid, ferulic acid and methyl-4-hydroxybenzoate, but not quercetin-3-rhamnoside, were detected in the exudates (Table 2).

Table 2. Compounds detected in Arabidopsis root exudates.
Exudates obtained from 150 or 300 Arabidopsis seedlings grown for 2 or 3 weeks in liquid culture. Values represent peak intensity.

Sample	2 weeks, 150 seedlings	2 weeks, 300 seedlings	3 weeks, 150 seedlings	3 weeks, 300 seedlings	Standard (0.1 mg/ml)
Syringic acid (198.17 g/mol)	$2*10^4$	$2.5*10^4$	$3*10^4$	$3*10^4$	$1*10^7$
p-coumaric acid (164.16 g/mol)	$3.6*10^4$	$5.3*10^4$	$6.2*10^5$	$7.8*10^5$	$8*10^5$
Ferulic acid (194.14 g/mol)	$9*10^4$	$4*10^4$	$4*10^5$	$2*10^5$	$6*10^6$
Quercetin-3-rhamnoside (448.4 g/mol)	Not detectable				$8*10^5$
Methyl-4-hydroxybenzoate (152.15 g/mol)	$2*10^3$	$3.6*10^3$	$5*10^3$	$3.9*10^3$	$2.7*10^6$

(1) Liquid culture grown plants

The HPLC profiles of liquid culture-grown exudates from phenylpropanoid mutants showed that there was significant variation in peak intensity and profile even after normalizing for plant dry mass. Examples showing the exudate profiles obtained from the Arabidopsis phenylpropanoid mutant lines *tt9* and *tt8* are shown in Figure 6 and Figure 7, respectively. As all plants were grown at the same time and treated equally, a large portion of this variation was considered to be due to extraction and processing of samples. Despite this variation, differences in the exudate profiles between lines were observed.

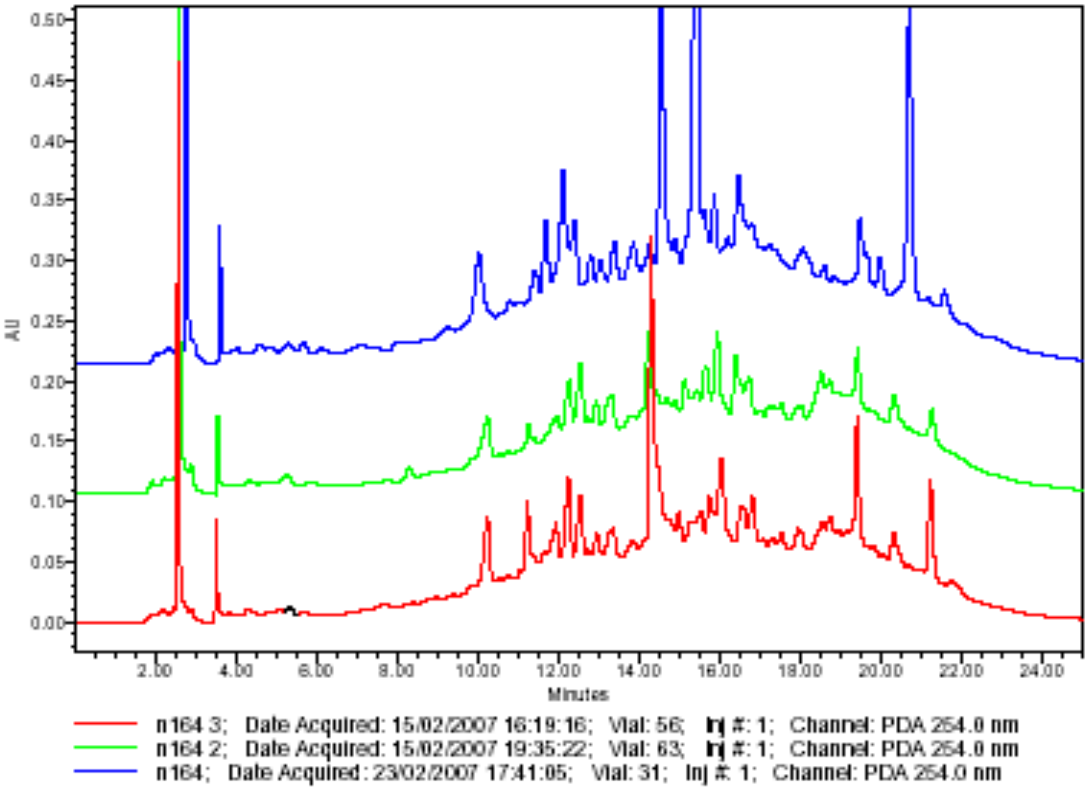


Figure 6. HPLC traces of exudates obtained from Arabidopsis mutant tt9 grown in liquid culture.
Graph shows variation in the triplicate exudates traces. Samples (50 µl) were injected onto a TechSphere ODS 80A 5 µ (250 mm x 4.6 mm) column and separated using HPLC method 2.

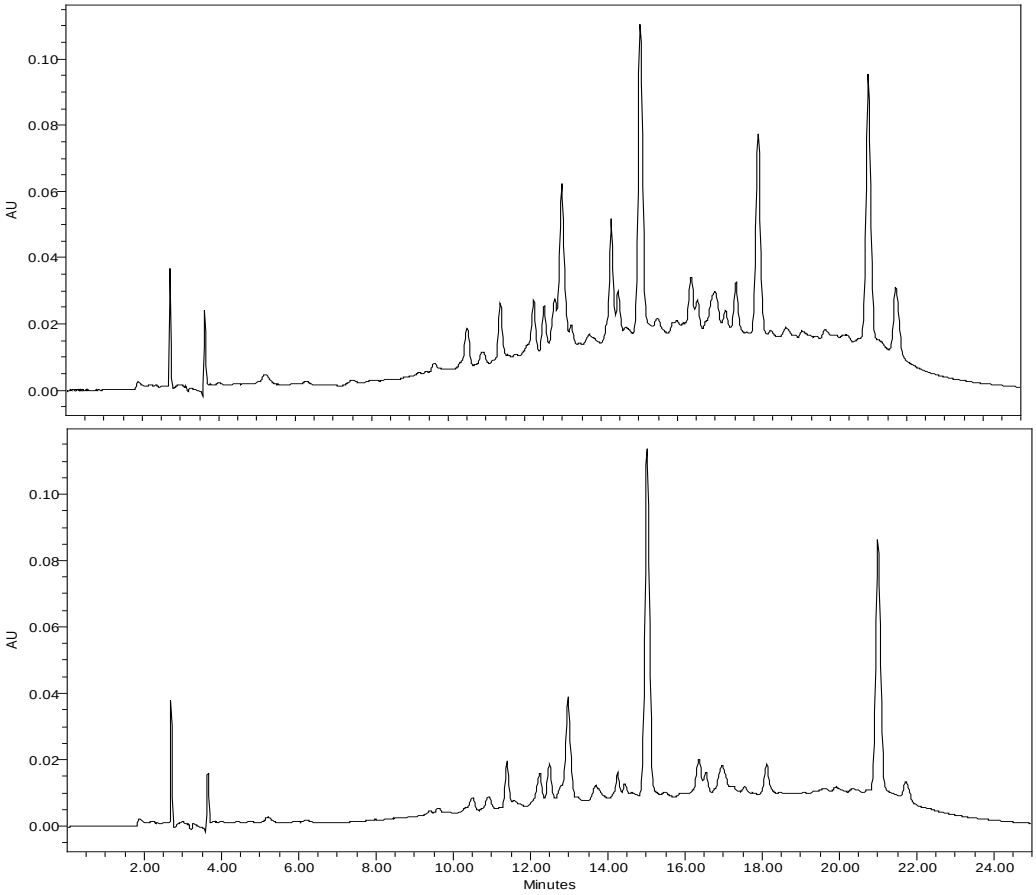


Figure 7. HPLC traces of exudates obtained from Arabidopsis mutant tt8 grown in liquid culture.
Graph shows duplicate exudates traces. Samples (50 µl) were injected onto a TechSphere ODS 80A 5 µ (250 mm x 4.6 mm) column and separated using HPLC method 2.

(2) Sieve grown plants

HPLC analysis of root exudates collected from seedlings grown on sieves showed that the profiles of exudates from air-dried and non air-dried roots were significantly different (Figure 8).

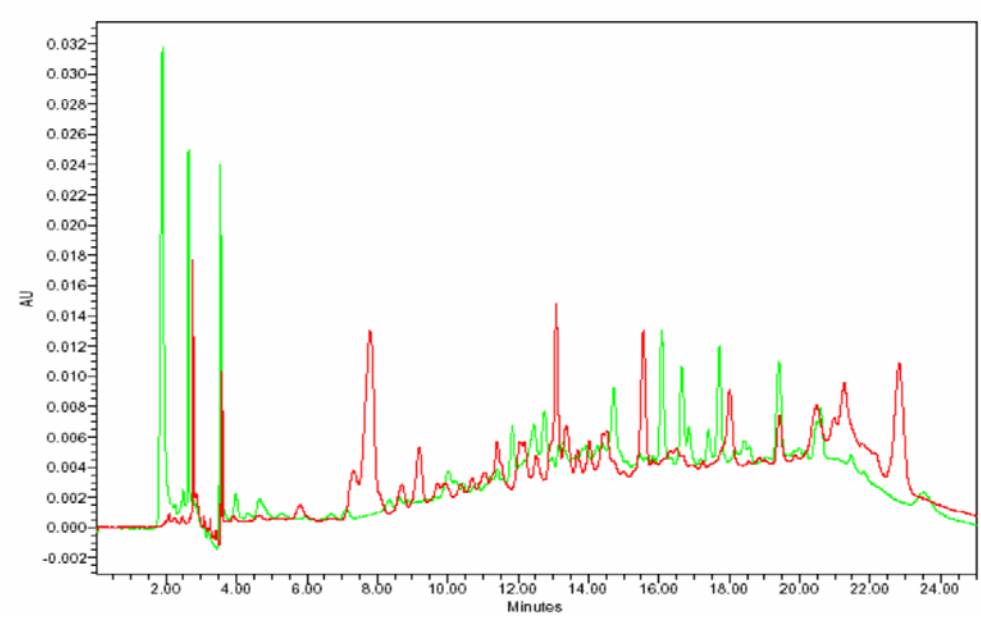


Figure 8. HPLC traces of exudates obtained from Arabidopsis wild type Landsberg erectus (Ler) grown on sieves.
Traces of air dried plants are shown in red and non air-dried plants in green. Samples (50 µl) were injected onto a TechSphere ODS 80A 5 µ (250 mm x 4.6 mm) column and separated using HPLC method 2.

(3) Aeroponically grown plants

Phenylpropanoid mutant lines and wild type ecotypes were grown aeroponically on agar media containing ½ MS ± sucrose to investigate if the addition of a carbon source alters the exudate profile; however, the mutant lines did not grow well in the absence of sucrose, therefore exudates were just collected from plants grown on agar containing sucrose. Exudate samples collected from aeroponically grown plants were normalized to plant fresh weight, and are shown in Figure 9.

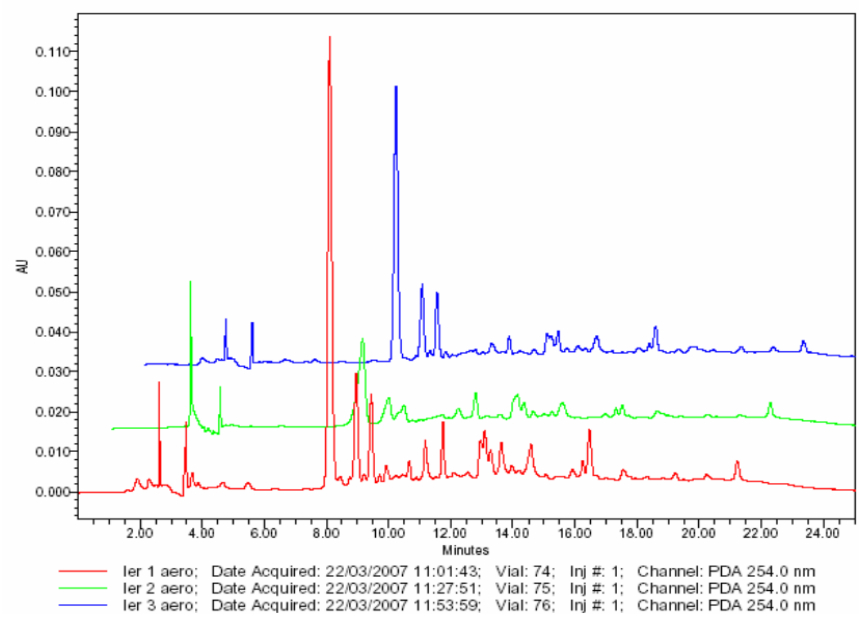


Figure 9. Exudate profiles obtained from wild type Ler Arabidopsis plants grown aeroponically for 17 days.
Data are shown are from triplicate repeats. Exudate samples were separated using the HPLC method 2.

(4) Vermiculite grown plants

Exudates extraction of plants grown in vermiculite was performed as described by (Kamilova, Kravchenko et al. 2006). Exudates could be detected when five Arabidopsis seedlings were grown in 10 g vermiculite for 3.5 weeks (Figure 10).

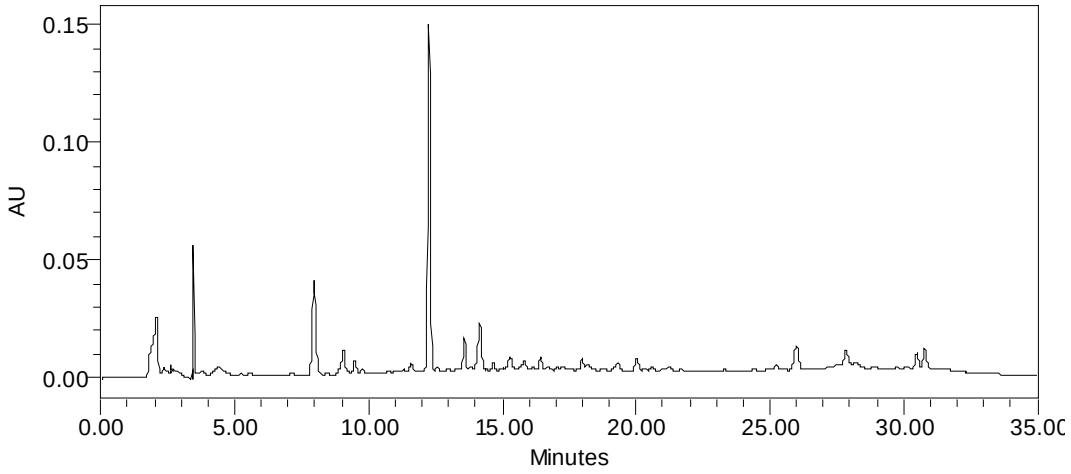


Figure 10. HPLC trace of exudates obtained from five seedlings *tt8* grown in 10 g vermiculite. Exudates were extracted after 3.5 weeks. HPLC method 1 was used for peak separation.

Exudate profiles of four methods were compared. Figure 11 shows profiles obtained from vermiculite exudate extractions (A) and liquid culture exudate extractions (B). Figure 12 shows profiles obtained from liquid culture and sieve grown plants. The profiles of both comparisons look significantly different. Exudates from plants grown in vermiculite were extracted 3.5 weeks after sowing; exudates of liquid culture grown plants 2.5 weeks after sowing. The difference in harvesting could have influenced the composition of the root exudates; however, the change was necessary to obtain enough root biomass for sufficient exudates extraction.

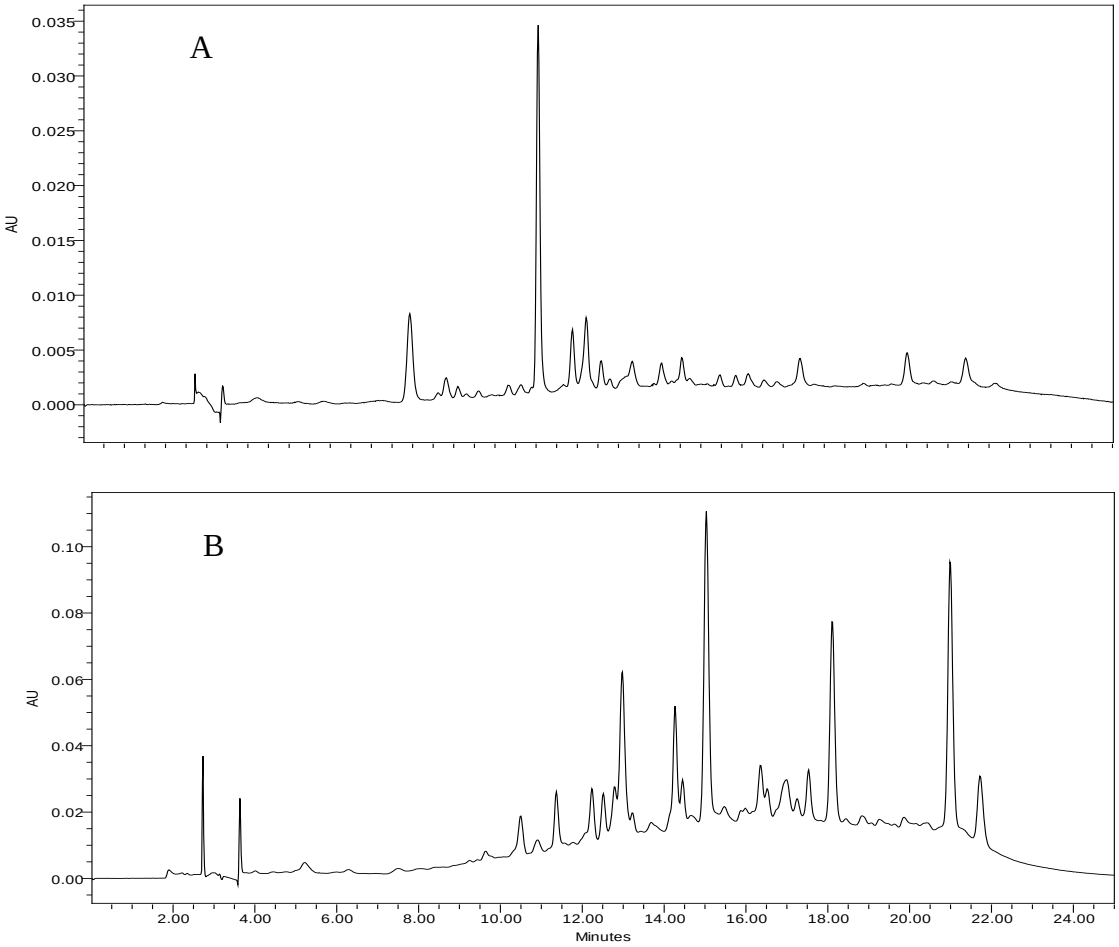


Figure 11. HPLC traces of exudate from Arabidopsis mutant line tt8
Grown in vermiculite (A) or in liquid culture (B). Exudate samples were separated using the HPLC method 2.

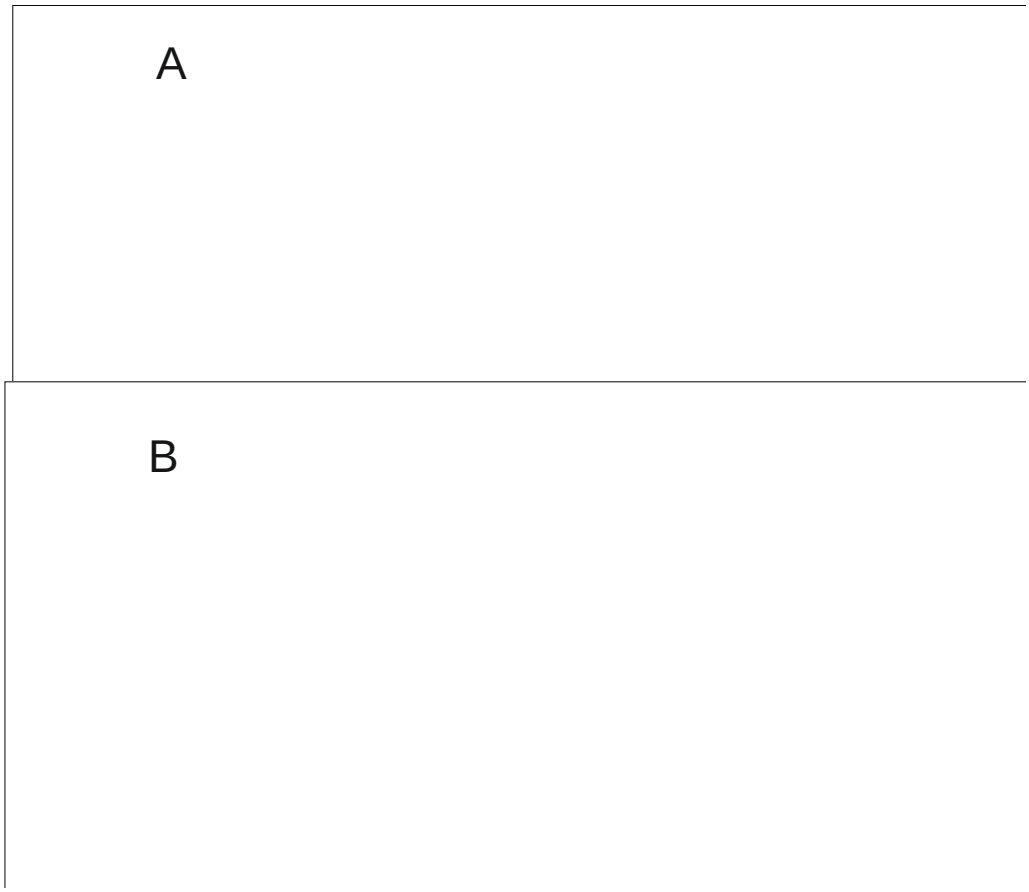


Figure 12. HPLC trace of root exudates from Arabidopsis (Ler) seedlings.
(A) sieve and (B) liquid culture grown plants. Exudate samples were separated using the HPLC method 1

g) Colonization of the rhizosphere by R. rhodochrous

(1) Colonization assays in a gnotobiotic system

As shown in Figure 13, after 22 days, bacterial counts in the vermiculite/perlite/soil samples containing plants were higher than the no plant controls (just vermiculite/perlite/soil and bacteria). Whilst the number of colony forming units in the *R. rhodochrous* samples were lower compared to the *P. fluorescens* samples Figure 15; it was still higher than at t = 0 (data not shown), indicating that the *R. rhodochrous* had grown.

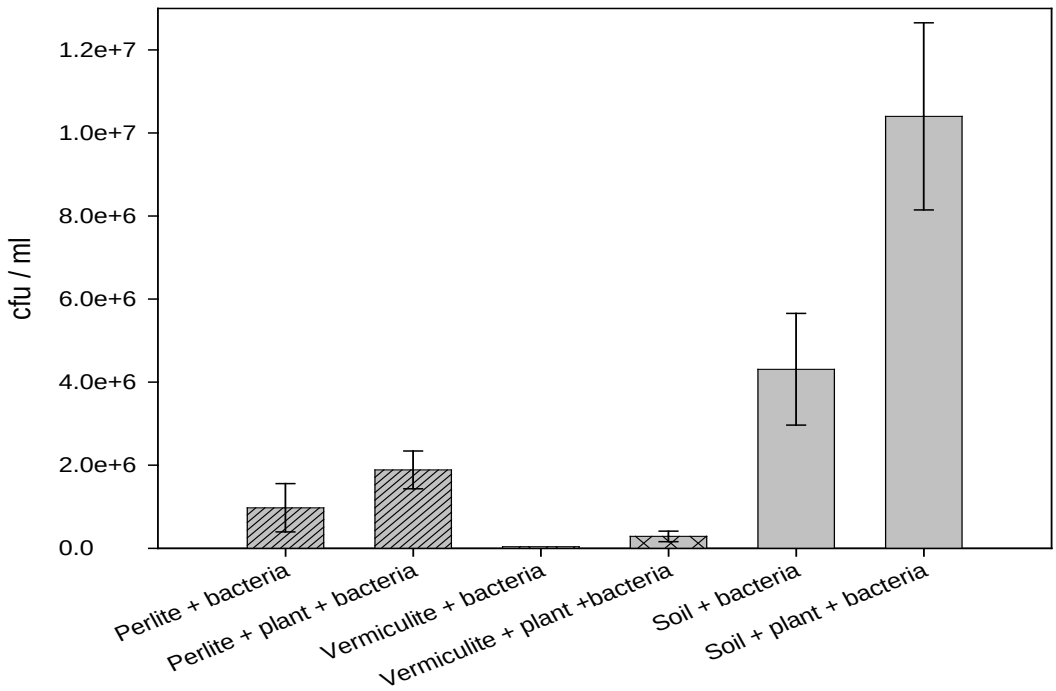


Figure 13. Colonization of Arabidopsis seedlings post-inoculation with Rhodococcus rhodochrous.
Results are means ± standard error of four replicas.

(2) Colonization of soil-grown Arabidopsis and wheatgrass by *R. rhodochrous*

The results in Figure 14 show that the presence of Arabidopsis or *R. rhodochrous* alone enhanced the removal of RDX form the soil. When Arabidopsis was inoculated with *R. rhodochrous* an additive effect in RDX uptake was seen. A similar pattern was seen for slender wheatgrass (Figure 14) indicating that *R. rhodochrous* did not successfully colonize the rhizosphere of Arabidopsis or Wheatgrass.

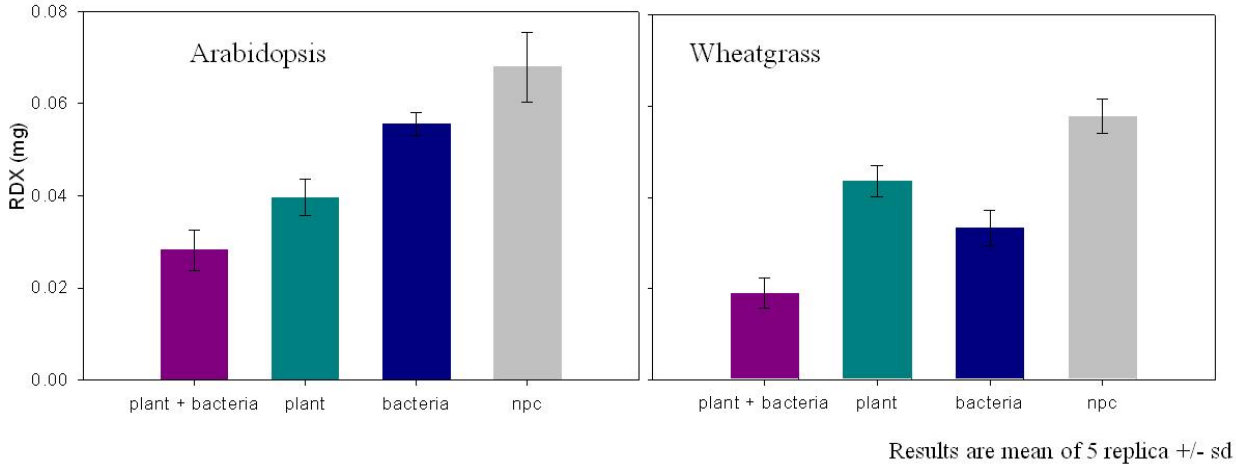


Figure 14. Removal of RDX in plant rhizospheres.
(A) Arabidopsis rhizosphere and (B) Wheatgrass rhizosphere. Levels of RDX in contaminated soil six weeks after addition of Arabidopsis and/or *R. rhodochrous* 11Y. Npc = no plants, no bacteria sample.

h) *Testing root colonization ability of* Pseudomonas fluorescens

As results indicated that *R. rhodochrous* did not colonize the roots of the plant species tested, the ability of the root colonizing bacterium *Pseudomonas fluorescens* to grow on Arabidopsis roots was tested. Figure 15 shows the number of colony forming units in three different types of growth media, Perlite, vermiculite and soil, 22 days following inoculation with *Pseudomonas fluorescens*. In the artificial Perlite and vermiculite media, the cfu counts were significantly higher in the presence of Arabidopsis, however, in soil, the presence of Arabidopsis did not have a significant enhancing effect on the cfu count. Arabidopsis may not be an optimal host for *Pseudomonas fluorescens*, and subsequent experiments using alfalfa, which is a known host species for *Pseudomonas fluorescens* were performed.

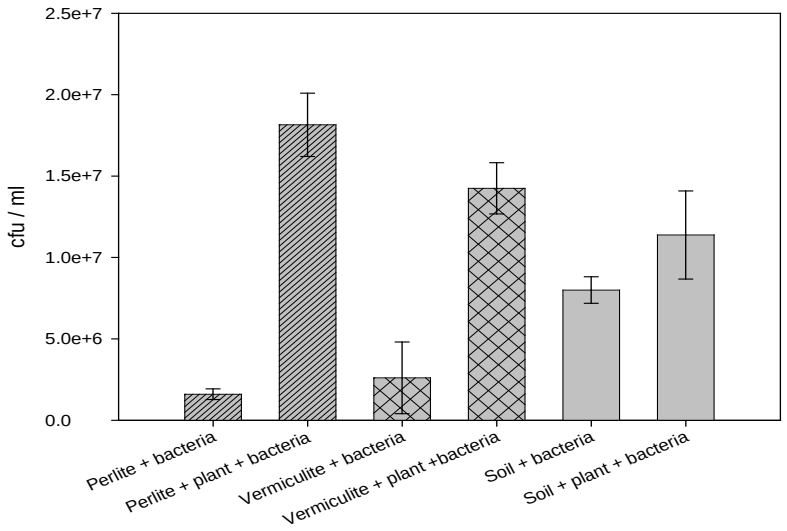


Figure 15. Colony forming units in growth media inoculated with *Pseudomonas fluorescens* and Arabidopsis.
Colony forming units (cfu) were counted 22 days post inoculation. Results are means ± standard error of four replicas.

i) Expressing XplA activity in Pseudomonas fluorescens

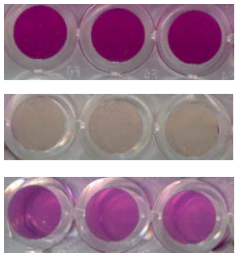
(1) Cloning of xplA into the shuttle vectors

The *xplA* gene was cloned into the IPTG-inducible shuttle vectors pJAK14, pME6010 and pME6010 derivatives pME6030 and pME6031. pME6010, under the control of the constitutive kanamycin promoter (KanR), is a useful vector for rhizosphere studies because of its high stability and constitutive-expression promoter in the root colonizing bacterium *P. fluorescens* ((Heeb, Itoh et al. 2000). Consequently, pME6010-*xplA* constructs do not require selective pressure (antibiotics) to maintain the plasmid nor IPTG to induce XplA expression.

(2) pJAK14-xplA

RDX transformation assays were used to study RDX degradation. To visualize XplA expression, samples of the whole cell transformation before and after IPTG induction were analyzed via SDS page; however results were inconclusive. Soluble expression of XplA in *P. fluorescens* WCS 365 was also tested using Western blot analysis and XplA antibodies. Samples were taken prior and after induction with IPTG, blotted onto nitrocellulose membrane and probed with XplA antibodies. Expression was observed; however, due to unspecific binding of the antibody, Western blot analysis was difficult to interpret and not routinely used to determine expression.

To test the ability of the *P. fluorescens* expressing XplA to degrade RDX, Griess assays were performed on whole cell transformations to detect levels of nitrite. The Griess assay showed pink coloration, an indication of nitrite, in the samples containing *P. fluorescens* transformed with pAX1-xplA and in *E. coli* Rosetta gami B transformed with pET16b-*xplA*, but no pink coloration was observed in the empty vector control (Figure 16).



XplA (pAX1) in *P.fluorescens* WCS 365
Empty vector control in *P.fluorescens* WCS 365
XplA (pET16b-*xplA*) in Rosetta gami B

Figure 16. Griess assays showing nitrite release from RDX by XplA-expressing bacteria.

Resting cell assays demonstrated that the *Pseudomonas* containing pAX1-*xplA* removed significantly more RDX than cells containing the empty vector, pJAX14 (Figure 17). Experiments performed with pET16b-*xplA* in *E. coli* Rosetta gami B, with the addition of $FeCl_3$ and ALA, showed complete degradation of RDX within 24 h (data not shown).

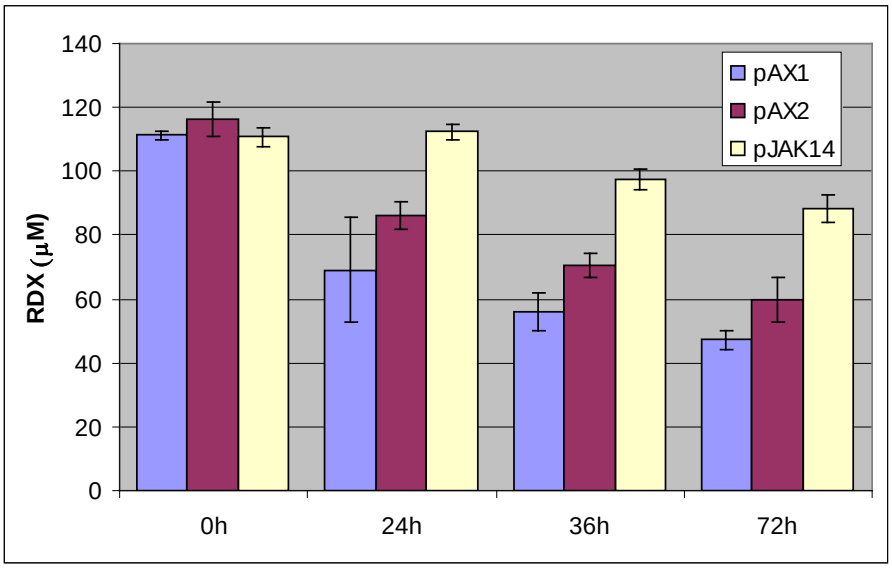


Figure 17. RDX uptake by *P. fluorescens* transformed with two identical clones of pAX1-xplA (pAX1 and pAX2) and empty vector control (pJAK14).
Cells used in this assay were grown at 37°C in LB to an OD_{600} = 0.8 and then induced with IPTG. No $FeCl_3$ and ALA were added. Cells were then grown at 20°C over night. Results are mean of 3 reactions ± one standard deviation of the mean. Resting cell assays were performed as described previously with 162 µM RDX starting concentration.

j) Testing XplA activity in a range of shuttle vectors

(1) Testing the P_{Kan} expression vectors

These pMEK-5X and pMEK-11X vectors were designed to express *xplA* under the P_{Kan} constitutive promoter, with a *Pseudomonas*-specific ribosome binding site (RBS) added either 5 or 11 base pairs upstream of the ATG start codon of *xplA*. A study of whole *Pseudomonas* cells transformed with these vectors showed uptake of RDX, with 46 % taken up after 24 h, compared to empty vector controls where there no significant uptake of RDX. RDX uptake was also seen without the addition of ALA and $FeCl_3$ (RDX loss up to 20 % after 24 h), but the efficiency was considerably lower; only 20 % of the RDX was removed after 24h (Figure 18).

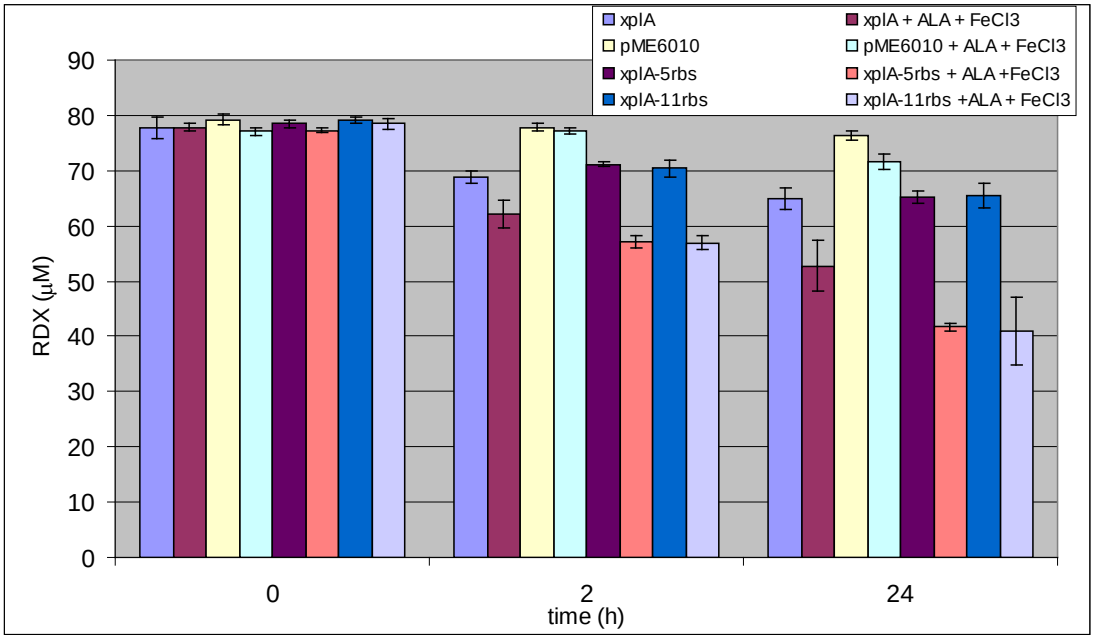


Figure 18. Resting cell assays showing RDX uptake by *Pseudomonas* containing pMEK-5X and pMEK-11X.
Cells were transformed and an aliquot was directly added to 25 ml LB culture containing 80 µg/ml tetracycline, ±1 mM ALA and ± 0.5 mM $FeCl_3$. Cells were grown for 24 h at 30°C and then a further 15 h at 20°C. Results are mean of 3 reactions ± standard deviation.

(2) Testing the expression vector pMETAX

To achieve constitutive expression, a pME6010 construct containing the TAC-*xplA* expression cassette without the lactose repressor was created and transformed into *P. fluorescens* WCS365 and RDX uptake assessed in whole cell assays (Figure 19). After 24h, approximately one third of the RDX had been taken up by the cells containing pMETAX. However, the uptake was slower than by cells containing the pAX1 vector.

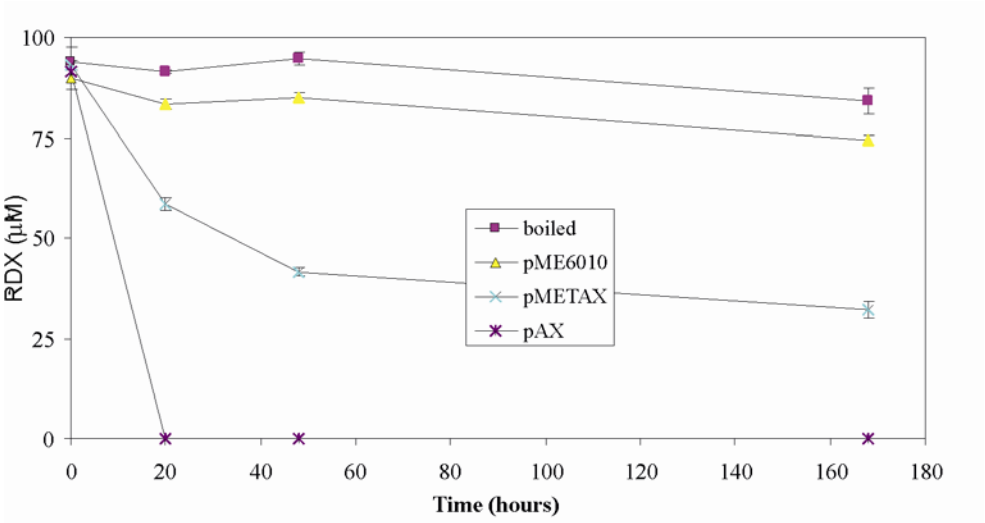


Figure 19. Uptake of RDX by *P. fluorescens* containing pAX1 or pMETAX.
One ml reactions contained 50 mg of *P. fluorescens* cells and 100 µM RDX in 40 mM potassium phosphate buffer at pH 7.2. Cells used in this assay were grown in overnight in LB the presence of $FeCl_3$ and ALA. Results are mean of three reactions ± one standard deviation of the mean.

(3) Testing the expression vectors pME6030 and pME6031

To test if the kanamycin promoter upstream of the TAC promoter in pMETAX reduces expression, pME6030 and pME6031, both promoter-less derivatives of pME6010, containing the TAC-*xplA* expression cassette from pJAK14 were transformed into *P. fluorescens* WCS365 and RDX uptake assessed in whole cell assays. After 24 h, approximately 50 % of the RDX was removed by cells containing pME6030 or pME6031 and supplemented with ALA and $FeCl_3$. Without supplements, only 35 % of the RDX was removed. Cells containing pAX with supplements performed better (approximately 90% removed after 24h) than cells containing pME6030 or pME6031 (Figure 20).

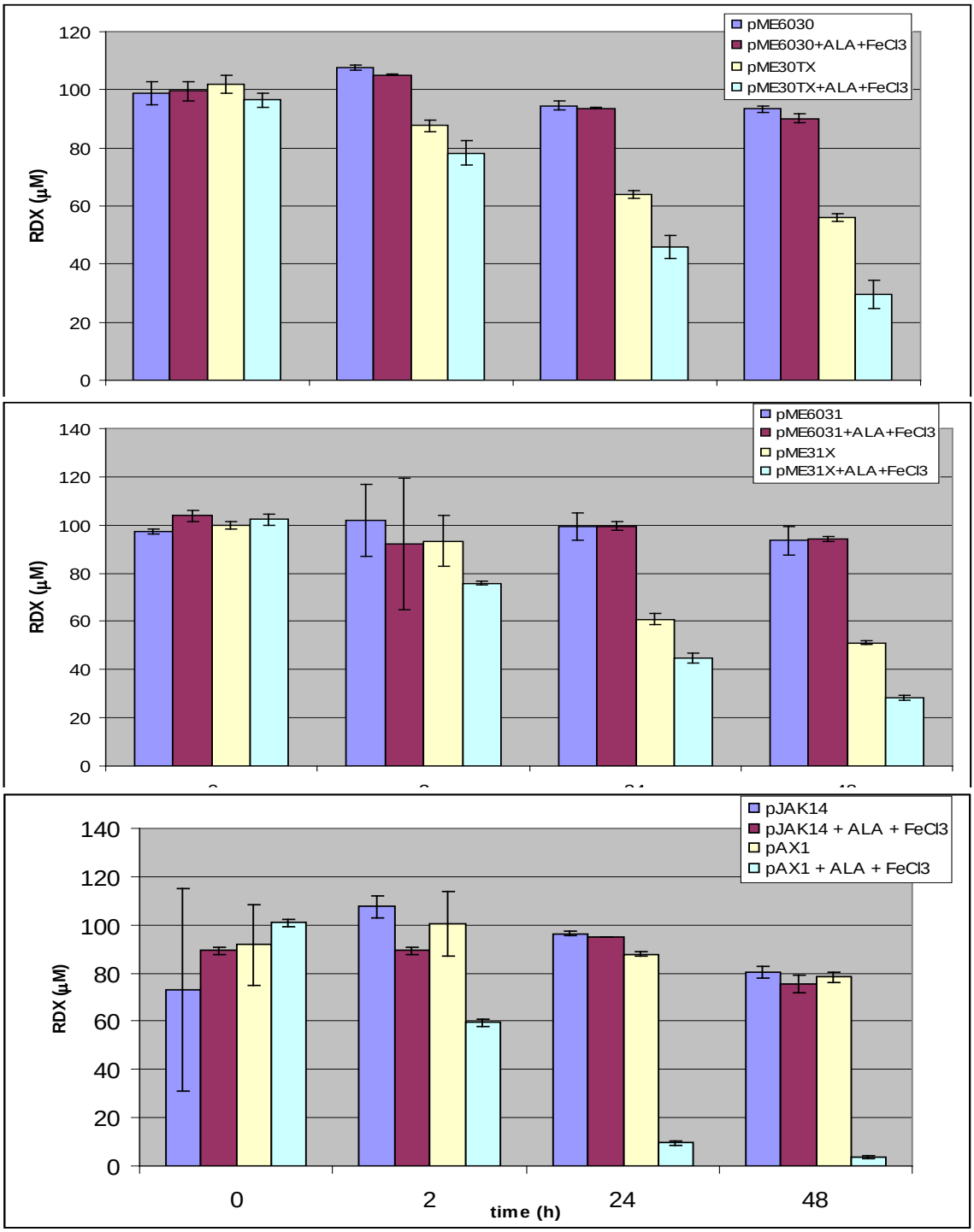


Figure 20. Uptake of RDX by *P. fluorescens* WCS365 containing pME31X, pME30X or pAX1.
500 µL reactions contained 25 mg of P. fluorescens cells and 162 µM RDX in 40 mM KH2PO4 buffer pH 7.2. pME31X and pME30X cells used in this assay were grown overnight in LB ± both FeCl3 and ALA. pAX1 cells were grown until OD600 = 0.7 and then induced with IPTG (100µM) and FeCl3 and ALA was added. All cells were grown over night at 30°C. Results are mean of three reactions ± standard deviation.

Resting cell assays were also carried out using a second strain; *P. fluorescens* F113 transformed with pME6031-*xplA* (Figure 21). The transformed cells removed 62 % of the RDX after 24

hours, compared to 6 % by the empty vector. There was no significant uptake of RDX by the untransformed strain. Based on the results above, the pME6031-*xplA* (Figure 20) vector was selected for subsequent rhizosphere studies.

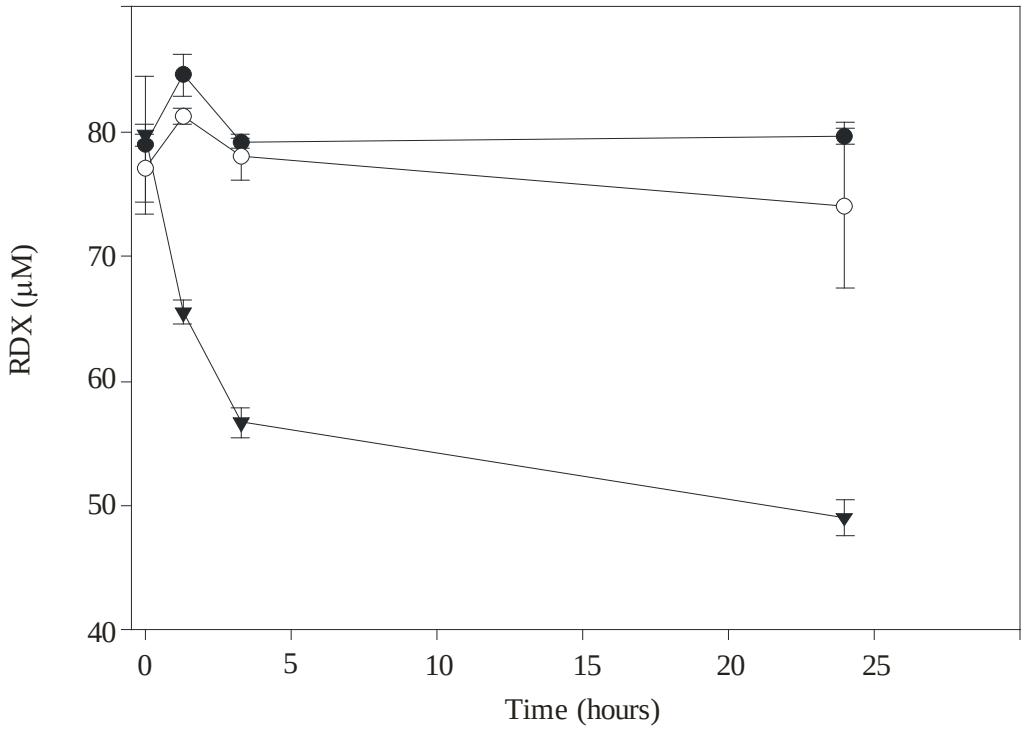


Figure 21. Uptake of RDX by P. fluorescens F113 containing pME6031-xplA
Cells were grown for 24 h at 30 °C and then transferred to minimal media containing RDX. Reactions were stopped by using the same volume of 10 % acetic acid (w/v). RDX-concentration was monitored by injecting an aliquot onto an 18 C-column and separating using an isocratic gradient of 50:50 Methanol/Water. (▼) *Pseudomonas fluorescens* F113 transformed with pME6031-*xplA* s, (○) *P. fluorescens* F113 containing the empty vector and (●) wild type *P. fluorescens* F113.

(4) Codon optimization of *xplA.*

To optimize gene expression the codon usage of *xplA* was modified to the preferable codon usage of *Pseudomonas* species. The *P. fluorescens* expressing the optimized XplA showed no RDX removal after 16 h, whereas approximately 23 % RDX was removed by *P. fluorescens* expressing XplA under the kanamycin promoter (*xplA* cloned into pME6010) and 53 % was removed by *P. fluorescens* expressing XplA under the lactose promoter (*xplA* cloned into pME6031). Thus, the non-optimized version of xplA was used for further studies.

k) Testing P. fluorescens expressing XplA in the rhizosphere

Research has shown that the *P. fluorescens* strains WCS 365 and F113 are well characterized as root colonizers, particularly of tomato and alfalfa roots, respectively {Lugtenberg, 1999, What makes Pseudomonas bacteria rhizosphere competent?;Lugtenberg, 2001, Molecular determinants of rhizosphere colonization by Pseudomonas.;Lugtenberg, 1999, Tomato seed and root exudate sugars: composition`, utilization by Pseudomonas biocontrol strains and role in

rhizosphere colonization. Sand has been used as the optimal growing medium for rhizosphere studies and was therefore used here. Seven experiments using Alfalfa were performed in series with progressive optimization steps.

(1) Gnotobiotic experiment using 1/2 MS and nitrogen-free nutrition medium

Experiments were performed using media containing 1 mM ALA and 0.5 mM $FeCl_3$ with nitrogen (1/2 MS) or without nitrogen. The eight treatments used are shown in Table 3.

Table 3. Conditions used for growing Alfalfa in a gnotobiotic environment

Treatment							
1	2	3	4	5	6	7	8
P. fluorescens WCS 365 + pME6031				*P. fluorescens* WCS 365 + pME31TX			
½ MS salts		Nitrogen free		½ MS salts		Nitrogen free	
2 ml RDX (140 µM)	3 ml RDX (140 µM)	2ml RDX (140 µM)	3 ml RDX (140 µM)	2 ml RDX (140 µM)	3 ml RDX (140 µM)	2 ml RDX (140 µM)	3 ml RDX (140 µM)

Plants were inoculated with approximately 4 x10^5 cfu/ml. Seven days after inoculation, four replica vials of each treatment were sacrificed to monitor the growth of *P. fluorescens*. The cfus observed when bacteria were grown on agar plates containing tetracycline were between 7 x 10^2 to 1 x 10^4 cfu/ml indicating that cells containing the vector were present twelve days after inoculation; however, counts were approximately 10-100 times lower than measured at the beginning of the experiment. Further monitoring of the bacterial growth was not performed. Four weeks after inoculation, RDX concentration in bulk sand and rhizosphere was measured. Both rhizosphere and bulk samples from plants inoculated with *P. fluorescens* WCS 365 expressing XplA had less RDX than samples obtained from plants inoculated with the empty vector control. When plants were treated with 2 ml of the 140 µM RDX solution (0.06 mg total RDX) and grown in ½ MS or in the nitrogen-free MS medium, both rhiozosphere and the bulk sand showed statistically significant differences in the RDX concentration between the XplA-expressing- and the empty vector control treatments (p-value < 0.05), detailed below:

In the bulk soil from plants inoculated with the empty vector control, 30 % of the RDX was recovered, whereas only 20 % was recovered from the bulk soil from plants inoculated with bacteria expressing XplA. In the rhizosphere soil from the empty vector control treatment, 80 % of the RDX was recovered compared to approximately 40 % when XplA-expressing bacteria were used (Figure 22).

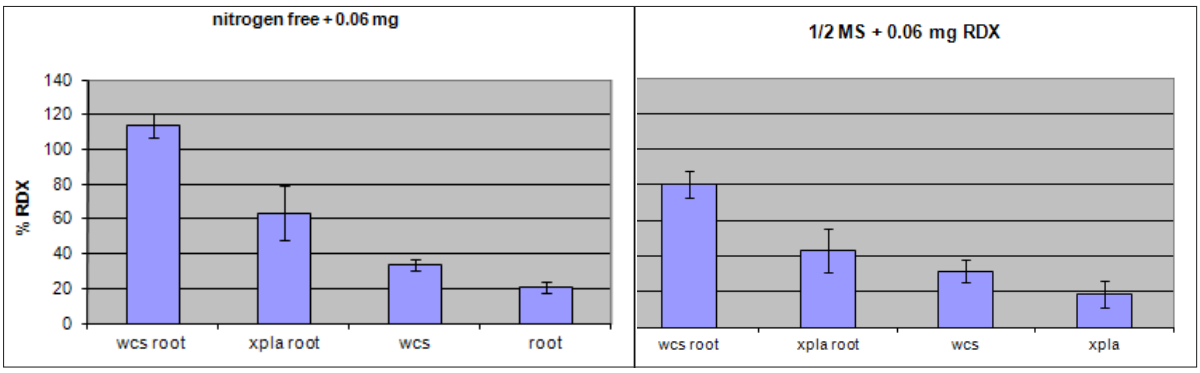


Figure 22. Percent RDX recovered from the rhizosphere
(wcs root = *P. fluorescens* containing the empty vector, xplA root = *P. fluorescens* expressing XplA) and the bulk sand (wcs = *P. fluorescens* containing the empty vector, xplA = *P. fluorescens* expressing XplA). Plants were grown in ½ MS salts or nitrogen free media dosed with 0.06 mg RDX

Bulk soil samples of plants dosed with 50 % more RDX (0.09 mg) and grown in ½ MS medium did not show any statistically significant difference in RDX removal (p-value = 0.09). However, the rhizosphere samples from the XplA-expressing bacteria/plant combinations contained significantly less RDX (50 %) than that recovered from the empty vector controls (70 %) (p = 0.008). In the nitrogen-free MS media the opposite was found: Significantly less RDX was recovered from the bulk soil in which plants inoculated with the XplA-expressing bacteria were grown (27 %) (p = 0.033) whereas 29 % of the RDX was recovered from the empty vector control treatment. There was no significant difference in the amount of RDX recovered from the rhizosphere samples (p = 0.249) (Figure 23).

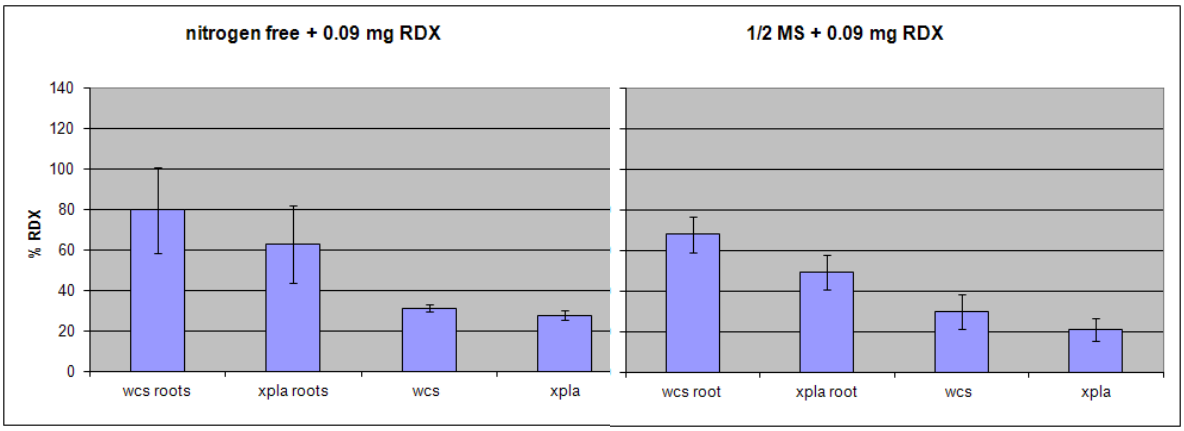


Figure 23. Percent RDX recovered from the rhizosphere
(wcs root = *P. fluorescens* containing the empty vector, xpla root = *P. fluorescens* expressing XplA) and the bulk sand (wcs = *P. fluorescens* containing the empty vector, xpla = *P. fluorescens* expressing XplA). Plants were grown in ½ MS salts or nitrogen free media dosed with 0.09 mg RDX

That significantly less RDX is recovered from the soil containing alfalfa inoculated with a root-colonizing strain of *Pseudomonas* bacterium expressing XplA, than from plants inoculated with empty vector-containing bacteria indicates that XplA is degrading the RDX. The greatest differences were found when 0.06 mg RDX was added and plants were grown in ½ MS medium. Therefore, repeat experiments were performed using these conditions. To dissect the

individual contributions of the plant and bacteria alone, additional controls were set up: plant only, bacteria only and no plant, no bacteria control.

(2) Gnotobiotic experiment using 1/2 MS medium and plant and bacteria only controls

As shown in Figure 24, bulk sand samples obtained from plants inoculated with *P. fluorescens* WCS365 expressing XplA contained less RDX (10 ± 1 %) than samples inoculated with the empty vector control (44 ± 7 %), or than all the other control treatments: bulk sand samples inoculated with bacteria expressing XplA only (65 ± 6 %), bulk sand samples inoculated with bacteria containing the empty vector control only (111 ± 19 %), plant bulk sand sample only (45 ± 11) or no plant no bacteria control bulk sand sample (100 ± 12 %) (Figure 24). A similar result was seen when RDX in the rhizosphere was measured; the quantity of RDX in the rhizosphere of samples inoculated with bacteria expressing XplA was significantly lower than of samples containing the plant inoculated with the empty vector control (Data not shown). Measurements were performed with five replicas for each treatment.

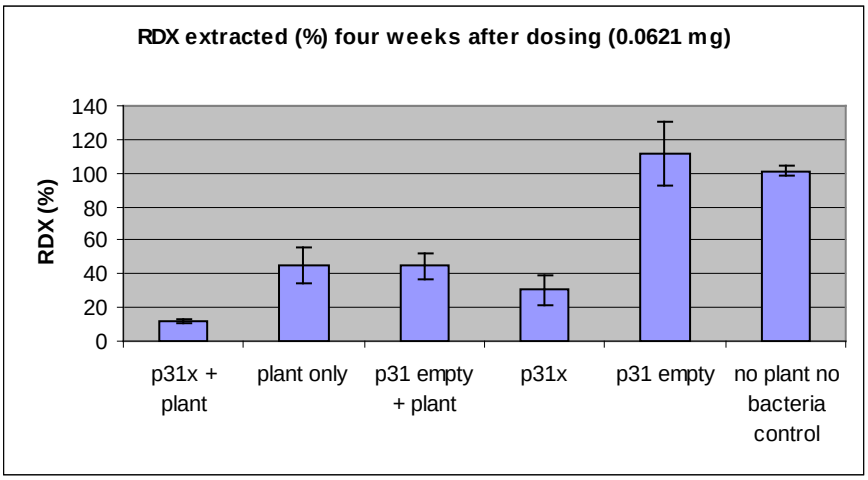


Figure 24. Percent RDX recovered from the bulk sand of pots containing alfalfa with and without inoculation with *Ps. fluorescens* with empty vector or containing xplA (p31x + plant = plants inoculated with *P.f.* expressing XplA in pME6031, p31 empty = plants inoculated with *P.f.* containing the empty vector pME6031, p31 = bacteria only containing pME6031 expressing XplA, p31 = bacteria only containing the empty vector pME6031). Plants were grown in ½ MS salts dosed with 0.06 mg RDX.

Bacterial growth on the roots was measured after four weeks by spreading a known volume of a serial dilution of the rhizosphere sample onto LB agar plates. To measure the stability of the plasmids, serial dilutions were also spread onto agar plates containing tetracycline, resistance to which is encoded by the plasmids. No bacterial growth was observed on plates containing LB agar with or without tetracycline when samples were taken from the bulk sand. The number of bacteria from the rhizosphere samples on LB agar only plates was uncountable as the number of cfus was too high to count at the dilutions used. The same sample spread onto plates containing LB and tetracycline contained approximately 800 cfu/g rhizosphere, which was a thousand times lower than the bacterial number immediately after inoculation.

(3) Testing the requirement of ALA amendment for RDX removal

To test if the addition of ALA, a heme precursor, is necessary for the activity of XplA, experiments were performed with and without the addition of ALA. In the samples amended with ALA, the results were in agreement with previous results. Significantly less RDX was measured (40 ± 2 %) in the soil from plants inoculated with XplA-expressing bacteria than in the control treatments: plant only, 71 ± 5 %; bacteria only, 70 ± 2 %; no plant no bacteria control, 100 ± 1 % RDX. In the samples without ALA, the level of RDX recovered from soil containing plants inoculated with XplA-expressing bacteria was not significantly different (34 ± 2 %) from that recovered from the empty vector treatment (36 ± 5 % RDX) when no ALA was added (Figure 25). There were also no differences in the rhizosphere samples, although here, the RDX concentration was significantly reduced with or without the addition of ALA (Figure 26). Interestingly, the rhizosphere biomass of plants treated with ALA was reduced by 60 % compared to the biomass of non-ALA treated plants (Figure 27). ALA is known to have plant growth promoting effects but can be phytotoxic at high concentrations ({Hotta, 1997, Effects of 5-aminolevulinic acid on growth of plant seedlings;Hotta, 1997, Promotive effects of 5-aminolevulinic acid on the yield of several crops} and this reduction in biomass also affected the RDX concentration.

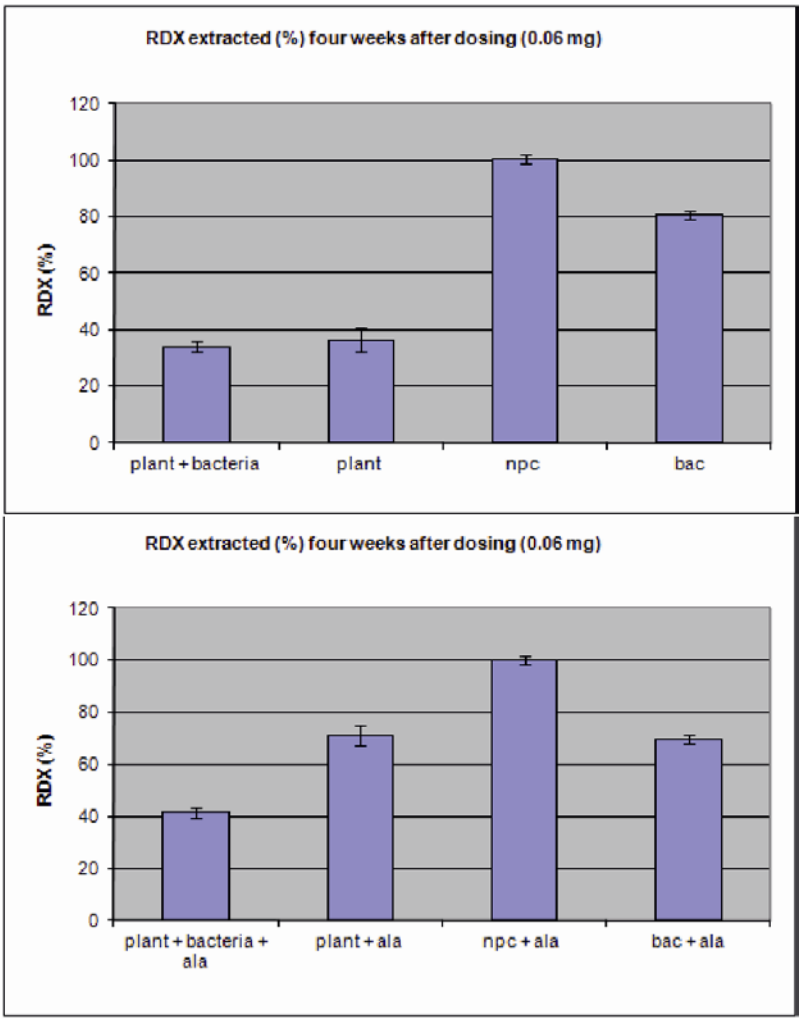


Figure 25. Percent RDX recovered from the bulk sand with and without alfalfa and *Ps. fluorescens*
(plant and bacteria = alfalfa inoculated with bacteria expressing XplA, plant = alfalfa inoculated with bacteria containing the empty vector, npc = no plant no bacteria control, bac = bacteria expressing XplA only). Plants were grown in ½ MS salts dosed with 0.06 mg RDX with and without amendment of ALA.

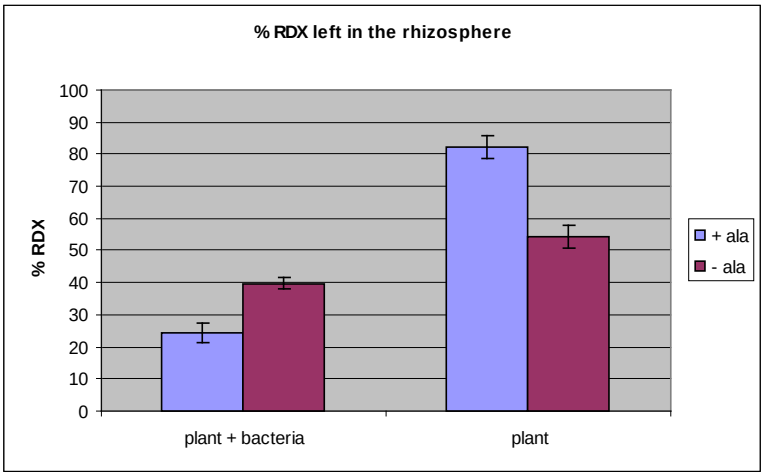


Figure 26. Percent RDX extracted from the rhizosphere with and without ALA amendment. Plant and bacteria = alfalfa inoculated with bacteria expressing XplA, plant = alfalfa inoculated with bacteria containing the empty vector.

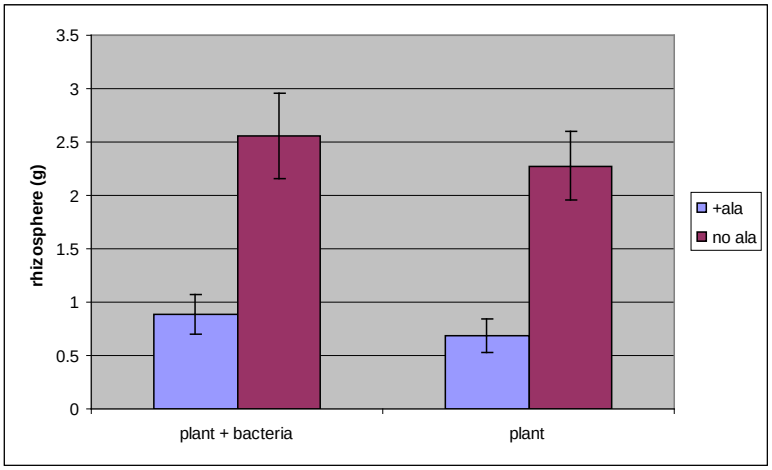


Figure 27. Rhizosphere biomass of plants treated with and without ALA.
Plant and bacteria = alfalfa inoculated with bacteria expressing XplA, plant = alfalfa inoculated with bacteria containing the empty vector.

(4) Testing the colonization abilities of Pseudomonas strain F113.

All experiments described so far were performed with *P. fluorescens* WCS 365. The root colonizing abilities of this strain have been characterized for a number of plant species, but not alfalfa; however, *P. fluorescens* F113 is known to colonize alfalfa roots, efficiently (Villacieros, Whelan et al. 2005). If the colonization of *P. fluorescens* F113 on alfalfa roots is better than that of *P. fluorescens* WCS365, then RDX removal might be increased. Therefore the experiment described above was repeated using *P. fluorescens* F113. Each treatment was performed with eight replicas. Plants were inoculated with 3.5×10^7 cfu bacteria per tube and the bacteria only control 5.2×10^7 cfu per tube. Plants were harvested after four weeks.

The results (not shown) for RDX removal were found to be similar to those measured in the previous experiment using WCS 365. This indicates that the two strains express XplA to similar levels and most likely colonize alfalfa roots to comparable extents.

(5) Use of sand pre-contaminated with RDX

In previous experiments, initial plant growth was established over one week prior to the addition of RDX; however, RDX removal might be enhanced if plants were transferred into RDX-containing media immediately after inoculation, as RDX could provide selective pressure increasing the stability of the plasmid containing *xplA* in the bacteria. Therefore the experiment was conducted as described above; with the alteration that RDX was added to the growth medium before the plants. Plants and growth medium were inoculated with approximately $1x10^5$ cfu. A similar amount of RDX removal was seen as measured in previous experiments. Therefore it was concluded that the time point of dosing did not affect the RDX removal.

In addition to the RDX extractions from the growth medium, RDX was also extracted from the aerial plant tissue. The RDX concentration found in the plant samples inoculated with bacteria expressing XplA was almost 50 % lower than that extracted from plants inoculated with the empty vector control (Figure 28).

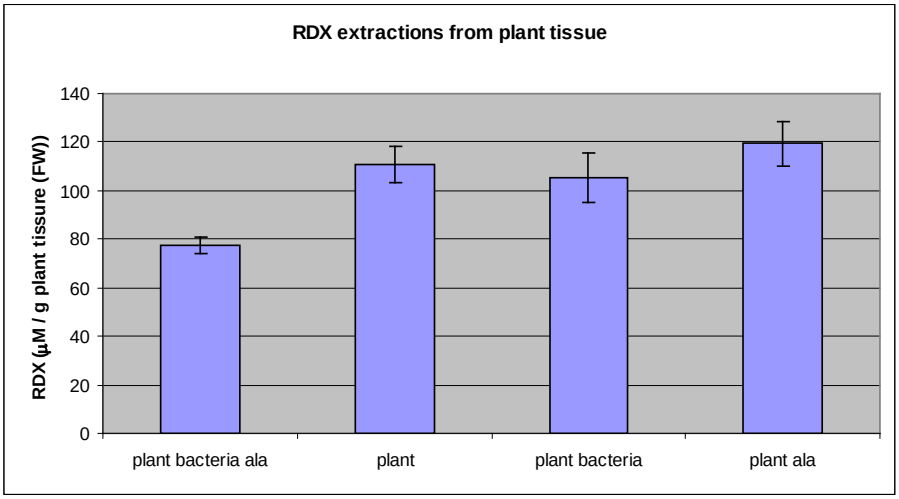


Figure 28. RDX extraction from the plant tissue.
Plants were treated with or without ALA (ala). Plant bacteria = alfalfa inoculated with bacteria expressing XplA , plant alfalfa inoculated with bacteria containing the empty vector. FW = fresh weight.

The total number of cfu from the rhizosphere and bulk soil from plants inoculated with *P. fluorescens* expressing XplA or containing the empty vector increased 100-fold after four weeks from $1x10^5$ cfu to $1.5x10^7$ cfu. However, when the same amount of bacterial suspension was grown on LB agar containing tetracycline, only $5x10^3$ cfu were detected, representing a 100-fold decrease in plasmid numbers from the day of inoculation. The presence of *xplA* in the tetracycline cfu was also confirmed by PCR using internal primers amplifying *xplA*. This indicates that the plasmid was not stable over the experimental period of four weeks.

(6) Replicates of optimized assays

To test the reproducibility of our optimized gnotobiotic assay system, replicate experiments were set up using *P. fluorescens* F113 containing pME31TX, or the empty vector pME6031 and inoculated Alfalfa seedlings. Control treatments included plants only, plants inoculated with the empty vector, bacteria only containing empty vector, bacteria only containing vector expressing *xplA*, no plant and no bacteria. The growth of bacteria was monitored by determining the colony forming units (cfu) on the plant roots, or in the bulk sand, at the beginning, and at the end, of the experiment. Plants and bulk soil were inoculated with $1x10^5$ cfu. Additionally, due to the reduction in root biomass caused by the toxicity of ALA seen in previous experiment, the amount of ALA was reduced from 1.4 to 1.2 mg.

Four weeks after inoculation, sand and plants were transferred onto sterile culture plates. The rhizosphere was separated from the bulk sand and aerial parts of the plants removed. Rhizosphere and bulk sand samples were dried, weighed and resuspended in water or buffer to determine the RDX concentration and the cfu count. As in previous experiments, both assays showed that bulk sand samples obtained from plants inoculated with *P. fluorescens* F113 expressing XplA contained less RDX than all the controls. As shown in Figure 29, just 21 ± 1.5 % of the RDX could be extracted from bulk sand from plants inoculated with *P. fluorescens* F113 expressing XplA, compared to 29.7 ± 1.3 % from plants inoculated empty vector. From bulk sand samples inoculated with *P. fluorescens* F113 expressing XplA, 51.2 ± 0.8 % of the RDX was extracted, 69.6 ± 0.8 % from the bulk sand samples inoculated with *P. fluorescens* F113 transformed with empty vector and 70.7 ± 0.9 from the no plant no bacteria, bulk sand samples. RDX levels from the aerial plant tissue were significantly lower in the plants inoculated with XplA- expressing bacteria compared to those from plants inoculated with the empty vector control (Figure 29).

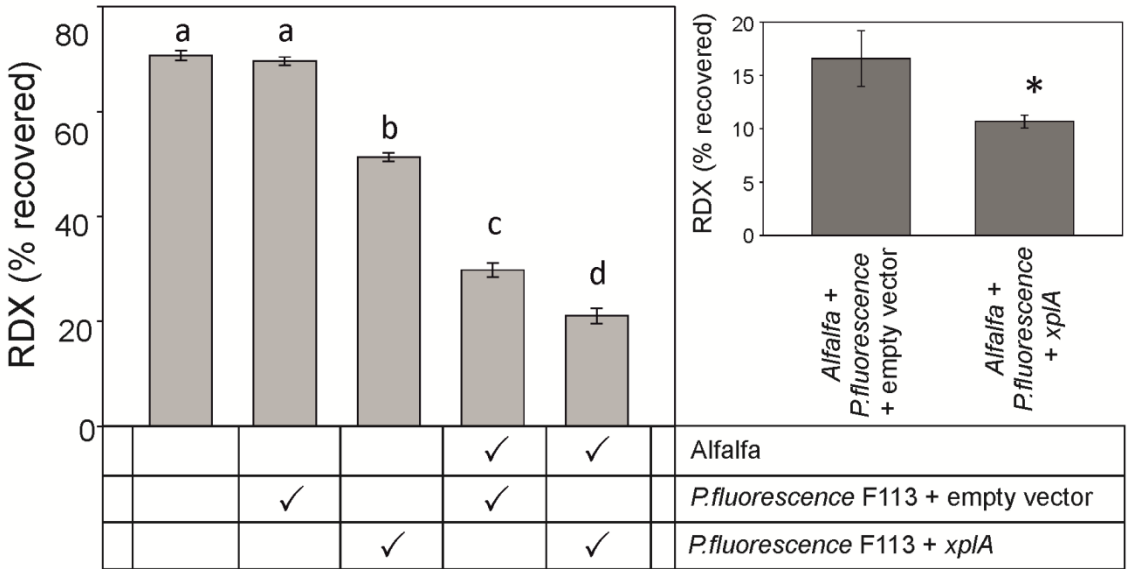


Figure 29. Percent RDX recovered from the bulk soil and plant tissue four weeks after seedling transfer. Results are means ± standard error of 14 biological replicas.

To determine the total loss of RDX, a mass balance was performed. As shown in Table 4, the RDX concentration of both the bulk sand and the rhizosphere of samples inoculated with XplA-expressing bacteria were lower than of samples representing the bulk

sand and rhizosphere inoculated with the empty vector control. *P. fluorescens* containing the empty vector had no effect on RDX removal. Interestingly, alfalfa itself might affect RDX removal, as not all of the extractable RDX could be recovered in the mass balance. Thus alfalfa may not just store RDX, but also break RDX down; however, no analysis of break down products was performed in this study

Mass balance (% RDX recovered)				
	Bulk	Rhizosphere	Aerial plant tissue	Total
No plant no bacteria control	70.7 ± 0.9			70.7 ± 0.9
P.f. F113 + empty vector	69.6 ± 0.8			69.6 ± 0.8
P.f. F113 + xplA	51.2 ± 0.8			51.2 ± 0.8
P.f. F113 + empty vector + alfalfa	29.7 ± 1.3	7.6 ± 3.0	16.6 ± 2.6	53.9 ± 6.9
P.f. F113 + *xplA* + alfalfa	21.0 ± 1.5	6.4 ± 0.4	10.7 ± 0.6	38.1 ± 2.5

Table 4. Mass balance (% RDX recovered).
The RDX concentration was measured in bulk sand, rhizosphere and aerial parts of the plant.

At the end of the experiment the cfu of all rhizosphere and bulk sand samples were determined. On non-selective medium, the rhizosphere sand samples were found to contain 5 x 10^6 cfu/g, but when grown on tetracycline, the bacterial number decreased to 5 x 10^3 cfu/g. A similar result was seen in the second replicate experiment (number 7). Bacteria grown on non-selective plates from the bulk sand were approximately 100 times less abundant compared to the rhizosphere samples. To determine if the reduced amount of bacteria obtained when agar containing tetracycline was used is due to plasmid instability, PCR using primers for *xplA* was performed on colonies grown on agar without the addition of tetracycline. Moreover, 16S sequencing on the tested colonies was performed to confirm that the bacterial species used for determining the cfus was *P. fluorescens* F113.

The analysis confirmed that all the bacterial colonies showing tetracycline resistance that were tested, were *P. fluorescens* containing *xplA*; however, when bacteria were grown on agar without tetracycline, two contaminating species were identified, counting for approximately 90 % of the total cfus and reducing the estimated amount of *Pseudomonas* to 5 x 10^5 cfu/g. Additionally, only 35 % of the *Pseudomonas* grown on LB agar without the addition of tetracycline still contained *xplA*.

l) Alfalfa inoculated with Rhodococcus rhodochrous *in a gnotobiotic system*

Previous experiments using *R. rhodochrous* 11Y in combination with plants with the aim of increasing degradation of RDX in the rhizosphere were unsuccessful. The development of a positive test system enabled us to re-test the root colonizing ability of *R. rhodochrous* and subsequent levels of RDX uptake. As a negative control, the non RDX-degrader CW25 was used to inoculate alfalfa. All steps were performed as described for previous gnotobiotic experiments; however, no enhanced degradation was seen when *R. rhodochrous* was used to inoculate the plants. Colonization levels of plant roots by bacteria were determined and it was

found that the amount of bacteria colonizing the plant roots significantly declined over the experimental period. The overall amount of bacteria per gram sand was calculated to be 270 cfu compared to $6x10^4$ cfu in the previous experiment. The decrease in bacterial numbers after inoculation indicates that *R. rhodochrous* 11Y is not an efficient root colonizer for alfalfa. In conclusion, *R. rhodochrous* 11Y seems not to be capable of colonizing roots of alfalfa efficiently under the conditions tested.

m) RDX degradation assays using Arabidopsis wild-type and mutant plants inoculated with P. fluorescens F113 expressing XplA

The construction of a RDX-degrading, efficient root colonizer, enabled us to test if different exudate profiles can influence the root colonization abilities of RDX degrading bacteria and alter levels of RDX degradation. We were unable to successfully grow Arabidopsis in the gnotobiotic system described for Alfalfa, therefore, the wild type (ecotypes Ler, Col0 and Col1) and mutant lines (*tt4-1, tt5-1, tt6-, tt7-1, ttg-1, tt10-1, tt8-1, tt9, pap1-D, fah1-7*) were grown under non-sterile conditions in a sand and Terra-Green (synthetic small grain gravel) mixture with a weekly nitrogen supplement and inoculated as described in the material and methods. RDX concentration in the soil leachate was measured and amounts of RDX expressed per plant fresh weight. There was no significant difference in all the treatments.

n) Alfalfa inoculated with P. fluorescens F113 expressing xplA-long term study

Experiments set up to study if the effect observed under sterile conditions was also seen when alfalfa are grown under non-sterile conditions showed no enhanced RDX degradation.

4. Discussion

a) Root exudate profiling

Root exudate profiling was set as the first milestone in this project and methods have been developed to extract exudates from Arabidopsis plants grown in liquid culture, in vermiculite or grown aeroponically. Methods were developed to analyze exudate profiles by HPLC and LC-MS using compounds found in root exudates. Exudate profiling was initially performed on plants grown in liquid culture. Extracts of exudates were obtained by concentrating large volumes of the liquid growth medium and then analyzing by HPLC. Surprisingly, replicate samples were shown to have great variations in their exudates profiles. This inconsistency in sample replication was thought to be due to the lengthly concentration process of the plant growth medium. Attempts to extract exudates from aeroponically grown plants have also shown differences in exudation, and moreover, mutant plants were found to grow poorly under these conditions and required the addition of sucrose, which may alter exudation. In summary, exudate profiling has proven to be complex as the comparison of exudate profiles has shown that there is significant variation in the profiles between the replicas, possible due to degradation, and variation between the same lines grown using different growth methods. Our findings show that the exudation of plants is very dependent on the growth medium. Therefore, plants grown in liquid culture may show a strong variation of exudation to plants grown in soil. Moreover exudate collection using Arabidopsis proved to be challenging as Arabidopsis is a small plant with a delicate root system and a low amount of exudation.

b) Colonization of the rhizosphere by R. rhodochrous

Rhocococcus rhodochrous 11Y, a gram positive soil bacterium, was found to efficiently degrade RDX and use it as a sole nitrogen source; however, it is not known if it is able to colonise roots. Therefore, assays were performed testing the ability of *R. rhodochrous* 11Y to colonize roots of alfalfa and Arabidopsis, which demonstrated that *R. rhodochrous* is not an efficient root colonizer. Furthermore, experiments using wheatgrass and Arabidopsis inoculated with *R. rhodochrous* 11Y showed no enhanced RDX removal when plants and XplA expressing *R. rhodochrous* 11Y bacteria were combined in the same pot. Thus our subsequent studies used *P. fluorescens*, a characterized root colonizing species.

c) XplA expression vectors

Pseudomonas fluorescens strains are known to colonize roots very efficiently (Tokala, Strap et al. 2002). However, no *Pseudomonas* strain is known to degrade RDX. *R. rhodochrous* 11Y is known to degrade RDX but not known to colonize roots. Therefore, *xplA* was cloned into *P. fluorescens* F113 and WCS365 two efficient root colonizer strains that have been well characterized.

Expression studies showed that XplA could be constitutively expressed in *P. fluorescens* WCS365 and F113 containing pMETAX, pMEK-X, pMEK-5X, pMEK-10X, pME30TX or pME31TX. The plasmid pME6010 and its derivates were found to be useful as expression vectors when used in soil studies as prior published work had showed them to be stabile without the need for antibiotics. Expression of soluble protein could be achieved without the addition of the heme precursor ALA, or $FeCl_3$; however, regardless of the presence of these additives, degradation of RDX was lower using *Pseudomonas* cells constitutively expressing XplA, than using either *E. coli* Rosetta gami cells inducing XplA expression with IPTG, or *P. fluorescens* cells containing the IPTG-inducible vector pAX1. The best expression temperature of all constructs was determined to be 30°C.

The pME6010 derivates pME6030 and pME6031 were constructed to avoid the potential hindrance of the kan promoter, but were not found to significantly increase RDX degradation. The incorporation of a ribosome binding site upstream of the XplA gene did not enhance RDX degradation in whole cell assays. The low expression of XplA in *Pseudomonas* containing pMETAX, pMEK-X, pMEK-5X, pMEK-10X, pME30TX or pME31TX, compared to XplA expression levels seen in *E. coli* rosetta gami B cells with p226 could be due to the different promoter systems. The pET-16b system used for expression in *E. coli* Rosetta gami B cells utilizes the T7 promoter system, with a highly efficient viral RNA polymerase that should drive strong expression of the target gene, *xplA*.

da pET-16b–based plasmid (in *E. coli*) or the pJAK14 plasmid in *Pseudomonas*; however, the use of IPTG in soil is not well studied. Villacieros *et al.*, 2005, overcame low expression by using the nod promoter system from Rhizobia species in *Pseudomonas*. The use of this promoter system resulted in a four-fold increase in levels of protein. A plasmid containing the nod system and primer sequences to amplify this have been obtained from Dr Rafael Rivella (University of Madrid, Spain), and was used to introduce this promoter system into pME6010; however, the insertion of the nod-promoter system into the plasmid pME6030/31 resulted not in an enhanced RDX removal in *P. fluorescens* F113 or WCS365. The lack of expression was almost certainly due to coding errors in the constructs. No further cloning attempts have been undertaken.

Western blot analysis was also performed using antibodies raised against XplA; however, no clear signal was seen (results not shown), indicating in agreement with previous results that expression of XplA in *Pseudomonas* is lower than expected from the initial *E .coli* studies as described above.

No RDX removal was observed in cells containing the codon optimized *xplA*. No further analysis were performed to investigate why there was not XplA-expression

d) Colonization of the rhizosphere by XplA-expressing P. fluorescens

In summary, experiments performed under gnotobiotic conditions showed that RDX removal from contaminated soil by plants increased when the plants were inoculated with *P. fluorescens* F113 or WCS 365 expressing XplA. This increase was from both the bulk and rhizosphere samples. The levels of RDX in the aerial parts of the plants were found to be lower in those plants inoculated with XplA-expressing bacteria than those inoculated with bacteria containing the empty vector.

The addition of ALA, a heme precursor was essential for the activity of XplA. ALA was shown to inhibit plant growth therefore reducing the root biomass at the concentrations initially used; however, the reduction of ALA in the later experiments did not result in increased RDX removal, most likely due to the limiting effect of ALA on the function of XplA.

Both, *P. fluorescens* WCS 365 and F113 could be used for the RDX degradation studies in a gnotobiotic system as both strains performed equally well; though, due to the fact that *P. fluorescens* F113 has been described in literature as an efficient colonizer of alfalfa roots, this strain was used for subsequent experiments (Tokala, Strap et al. 2002). *P. fluorescens* F113 was able to colonize Alfalfa roots efficiently; however, as sequencing revealed, at the end of the experiments approximately 90 % of all counted bacteria were contaminating species. This contamination was presumed to have originated from the seed as tests on the bacterial stocks used for inoculation did not find contamination. The addition of RDX immediately post-inoculation with the aim of providing increased selective pressure did not increase the stability of the *xplA* containing plasmid in the bacteria. Despite the reduction in overall numbers of *P. fluorescens* F113, this strain was still highly abundant. Investigating the plasmid stability revealed that approximately 35 % of the total counted *P. fluorescens* F113 colonies when grown on LB agar without selection still contained *xplA*. This indicates that the plasmid has some stability.

Enhanced removal of RDX was not observed when plants were grown under the non-sterile conditions tested. This was presumably due to the lack of competence of *P. fluorescens* expressing XplA; however, cfu were not determined when non-sterile systems were used. Moreover, even though that the plasmid was shown to be reasonably stable, expression of XplA in *P. fluorescens* F113 and WCS365 was found to be low. Taking all these factors together may explain why the result seen under gnotobiotic conditions could not be repeated when plants were grown non-sterile.

As discussed above XplA expression has proven to be low. Removal rates of RDX in the bulk sand by plants inoculated with XplA-expressing bacteria were significant, but just 30% lower when compared to plants inoculated with bacteria containing an empty vector. In previous

studies (Jackson, Rylott et al. 2007) we have shown that when both, XplA and its partnering reductase XplB were expressed in plants, XplA activity was increased 30 fold. Co-expression of XplA and XplB in the root colonizer *P. fluorescens* F113 could result in a similar enhancement in activity.

However, the introduction of *xplA* into the efficient root colonizing strains *P. fluorescens* F113 and WCS 365 has enabled us to create a rhizosphere bacterium capable of degrading RDX. Moreover, when used in combination with alfalfa we have shown this system enhances the removal of RDX in the rhizosphere. This plant-bacterial system could now be used to investigate whether root exudate composition is linked to RDX degradation.

B. Incorporation of plant-derived carbon into microbial rRNA: Isotopic probing of rhizosphere RNA/DNA (Barbara Macgregor, Univ. North Carolina, Chapel Hill)

1. Introduction

One of our initial hypotheses was that some part of the microbial community supported by plant carbon exudates might in turn supply plants with RDX-derived nitrogen. The main goal of the UNC group was to follow incorporation of plant-fixed carbon compounds by the soil microbial community. In outline, the intended approach was to isolate total RNA from a given sample; capture a phylogenetically defined fraction of this by magnetic bead capture hybridization; separate the RNA by denaturing or non-denaturing (SSCP) gel electrophoresis; expose gels on phosphorimager screens; and cut out radiolabeled bands for identification by RT-PCR and sequencing. However, although carbon transfer from plants to the soil community is assumed to be happening to some extent, it is apparently insufficient under the conditions tested for detection in non-eukaryotic RNA. We hesitate to interpret this data until heavier labeling, longer time course, and larger samples have been tested. As positive results, we have developed group-specific capture probes for bacterial RNA and improved our SSCP and RNA extraction methods, which we hope will prove useful for related future projects.

The objective of our portion of the project was to identify rhizosphere bacteria incorporating radiolabeled plant root exudates, to investigate whether plant carbon directly supports the growth of known RDX degraders. Plants that could support such species would be good candidates for *in situ* remediation of contaminated soils.

For this portion of the project, we have applied techniques for phylogenetically-based RNA separation that have been developed over the past several years: magnetic bead capture of small-subunit ribosomal RNA (MacGregor et al., 2002; MacGregor, Boschker, et al., 2006; Miyatake et al., 2009) and single-stranded conformational polymorphism for rRNA (rRNA-SSCP; MacGregor and Amann, 2006). Phylogenetically-specific rRNA capture can be combined with stable or radioactive carbon isotope characterization, either with label addition or at natural abundance, to identify or constrain the carbon sources used by particular microbial groups in a mixed community; Miyatake et al. (2009) have performed the most detailed such study to date, following incorporation of carbon substrates by sulfate-reducing bacteria in an estuarine mudflat. As another example, we were able to detect incorporation of ^{14}C-bicarbonate into both bacterial and archaeal RNA from a sandflat in the German Wadden Sea (Figure 30).

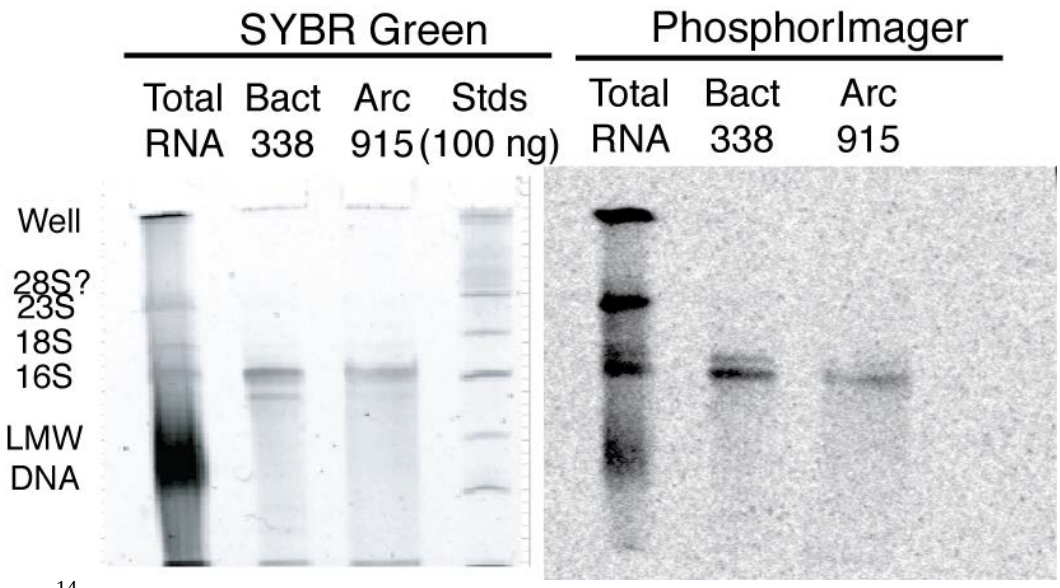


Figure 30. ^{14}C-bicarbonate incorporation into bacterial and archaeal SSU rRNA.
The total RNA shown was recovered from 0.2 g sediment (Janssand, Sylt; 2-3 cm depth); captured 16S rRNA bands represent 1.7 g sediment each. (N.Musat, B. MacGregor, and M. Kuypers, unpub.)

rRNA-SSCP can separate rRNAs of similar lengths but different sequence because gel conditions are such that they (apparently) maintain a stable secondary and tertiary structure. Our goal was to excise individual labeled bands for identification by sequencing (MacGregor and Amann, 2006), but with insufficient label for detection in even total bacterial small-subunit rRNA, this has not yet been attempted for the current sample set.

2. Methods

The RNA extraction, SSCP separation, and bead capture methods used in our laboratory have continued to evolve; these are the current protocols. Further optimization of the RNA method is needed for efficient extraction of plant (as opposed to microbial) RNA.

a) Large-scale TCA prep for RNA

After (McIlroy, Porter et al. 2008). This protocol is for use with a Braun Microdismembranator with 50 mL canisters. For preparation of 3M sodium TCA, see (McIlroy, Porter et al. 2008). DTT (dithiothreitol, Cleland's reagent); PVPP, polyvinylpolypyrrolidone.

TCA Lysis Buffer

Final concentration	Volume (for 100 mL buffer)
3M sodium TCA	67 mL of 4.5M
50 mM Tris-Cl (pH 8.0)	5 mL of 1M
15 mM NaEDTA (pH 8.0)	3 mL of 0.5 M
1% N-laurylsarcosine	10 mL of 10%
10 mM DTT	1 mL of 1M
DEPC'd water	to 100 mL

1. In baked canister, put:	0.1 mm beads	20 g
	0.4 mm beads	5 g

TCA lysis buffer	25 mL
PVPP	0.25 g
Antifoam B	418 µL
Sediment sample	5 - 15 g

2. Bead beat 2 x 40 sec, high speed.
3. Transfer to two Teflon Oak Ridge tubes. Important to get ~same amount of beads in each.
4. Centrifuge 10 min, 2500 rpm, 4˚C.
5. Transfer supernatant to Falcon tube.
6. Add 0.6 volume isopropanol, mix (can store overnight at –20˚C at this point).
7. Incubate ≥20 min on ice.
8. Centrifuge 30 min, 2500 rpm, 4˚C.
9. Wash pellets 2x with cold 70% ethanol, centrifuging 2 min each time.
10. Air dry with tube inverted on KimWipes, resuspend in DEPC'd water (volume depends on size of pellet).
11. Extract with phenol, phenol/chloroform/isoamyl alcohol, and chloroform/isoamyl alcohol as usual (details are in a separate protocol, not included here).

b) Magnetic Bead Capture of SSU rRNA

A magnetic bead capture method was used to concentrate single subunit rRNA (Mastrangeli, Micangeli et al. 1996, Bach, Hartmann et al. 1999, Bergin 2002, MacGregor, Bruchert et al. 2002, MacGregor, Boschker et al. 2006, Miyatake, MacGregor et al. 2009).

Hybridization buffer
This is designed to be used at a 9:1 ratio with samples. For Arc915, use 30%; Bact338 and Euk1379, 20% (30% to avoid all 23S); Uni1390, 10%.

This is for 30% formamide; adjust concentration by varying amount of water and formamide (water + formamide = 6.29 mL).
(From Roche, but without blocking reagent)

Final conc.	Add:
5X SSC	2.5 mL of 20X
0.1% N-laurylsarcosine	100 µL of 10%
0.1% NaCl	100 µL of 10%
0.02% SDS	10 µL of 20%
30% formamide	3 mL
RNase-free ddH_2O	3.29 mL

Maleic acid buffer

100 mM maleic acid	11.6 g/L
150 mM NaCl	8.8 g/L

Adjust to pH 7.5 with NaOH

Blocking reagent (Roche Molecular Biochemicals)
Prepare 10% stock solution in maleic acid buffer. Autoclave. Store refrigerated or frozen.

20X SSC

3M NaCl	175.3 g/L
0.3M Na citrate	88.2 g/L

c) RNA/probe hybridization

1. Mix RNA and hybridization buffer (in 100µL final volume). Incubate 70˚ C for 10 minutes, then RT for 30 minutes (Mastrangeli, Micangeli et al. 1996).
2. Add appropriate probe dilution (10-fold molar excess over estimated concentration of target sites seems to work well, but this should be tested). Incubate 22˚C on end-over-end mixer overnight.

d) Bead preparation (Dynal A.S. handout for Dynabeads M-280 Streptavidin and Bach et al.)

3. Pipette enough beads for all samples into an Eppendorf tube [50 µL/sample].
4. Using magnet, remove storage solution
5. Wash beads with 3 x [original volume] 0.5X SSC
6. Resuspend in [original volume + a little extra] 0.1% blocking solution (Roche Molecular Biochemicals)/0.5X SSC
7. Aliquot into reaction tubes. Incubate 1 hour RT on end-over-end mixer.

e) Capture and elute hybridized RNA

8. Remove blocking solution using magnet, add hybridization mix.
9. Incubate 2 hours RT on end-over-end mixer.
10. Remove [and save, if needed] hybridization mixture.
11. Wash beads with 3 x 100 µL of 7.5X SSC.
12. Resuspend in 100 µL RNase-free ddH_2O.
13. Elute at 90˚C for 3 minutes.
14. Separate supernatant from beads with magnet; repeat if necessary.

For RNA to be separated by SSCP, RNeasy (Qiagen Inc.) cleanup is suggested at this point, followed by isopropanol precipitation and resuspension in a smaller volume.

f) Precipitate RNA

15. Mix:

Sample	100µL
Isopropanol	100µL
7.5 M ammonium acetate	50µL

Incubate RT a few minutes

16. Centrifuge 15 min, 4˚C, 13,000 rpm.
17. Wash pellets once with 70% ethanol.
18. Resuspend in RNase-free ddH_2O.

g) SCP separation of 16S rRNA on minigels

The conditions here have worked best for the soil, sediment, and pure culture samples tested, but may need adjustments for other sample types. Urea, Tris, and borate concentrations are the easiest variables to manipulate; temperature changes are not recommended. Some reports on DNA-SSCP recommend adding 5 - 10% glycerol, but we have not found that to improve RNA separation.

10X 30 mM TBE (Liu et al., 2000)

For 100 mL:

	Final concentration	Amount
Tris base	0.3 M	3.6 g
Boric acid	0.3 M	1.8 g
Na_2EDTA	10 mM	0.37 g

Adjust to pH 8.3. Bring to 100 mL final volume.

5% Duracryl, 1.7M urea, 30 mM TBE

For one gel (7.5 mL):

30% Duracryl solution (30C, 2.6C)	1.25 mL
10X 30 mM TBE	1.0
10 M urea	1.25
dH_2O	4.0
	7.5 mL
TEMED	5 µL
10% APS	50 µL

Sample-loading dye premix (use 7.5 µL per 5 µL sample)

10 M urea	625 µL
1% bromophenol blue	62.5
RNase-free water	62.5

Prerun gel at 250 V in 4°C room for approx. 15 min. The initial current will be around 23 mA, and will decrease to about 7 mA during the prerun. Load samples. The blue dye will runs off in the first half hour or so. Gels are run for ~4 h and visualized by staining with SYBR Gold (Invitrogen Inc.).

3. Results

One of our initial hypotheses was that some part of the microbial community supported by plant carbon exudates might in turn supply plants with RDX-derived nitrogen. The main goal of the UNC group was to follow incorporation of plant-fixed carbon compounds by the soil microbial community, and in this we were unsuccessful. The approach taken was to grow wheatgrass plants on ^{14}C-bicarbonate in soil with a Milan, Eglin, or no inoculum; isolate total RNA from the root-associated soil by a combined TCA and phenol/chloroform procedure (see Appendix 1A); and capture bacterial small-subunit ribosomal RNA (16S rRNA) by magnetic bead hybridization (Appendix 1B), using a probe targeting all bacteria. Total and 16S rRNA fractions were separated by polyacrylamide gel electrophoresis, and the gels dried and exposed on phosphorimager screens for up to several weeks (Figure 31). 16S rRNA recovery was good, the bead capture hybridizations were efficient, and radiolabel could be detected in presumed eukaryotic rRNA in those incubations including plants. However, little or no label was detected in bacterial 16S rRNA, either in the presumed bacterial bands in the total RNA or in the captured 16S fraction. Assuming the microbial community is actively growing, this suggests it may derive the bulk of its cellular (or at least RNA) carbon from non-plant, non-bicarbonate sources. It is possible some subset of the community is specifically plant-dependent, which might be investigated with the more specific probes and SSCP separation methods developed early in the investigation (discussed below) but evidence suggests this would have to be a small

and/or slow-growing fraction. Larger samples might also yield a clearer signal, but these would be difficult to separate with the gel system currently employed.

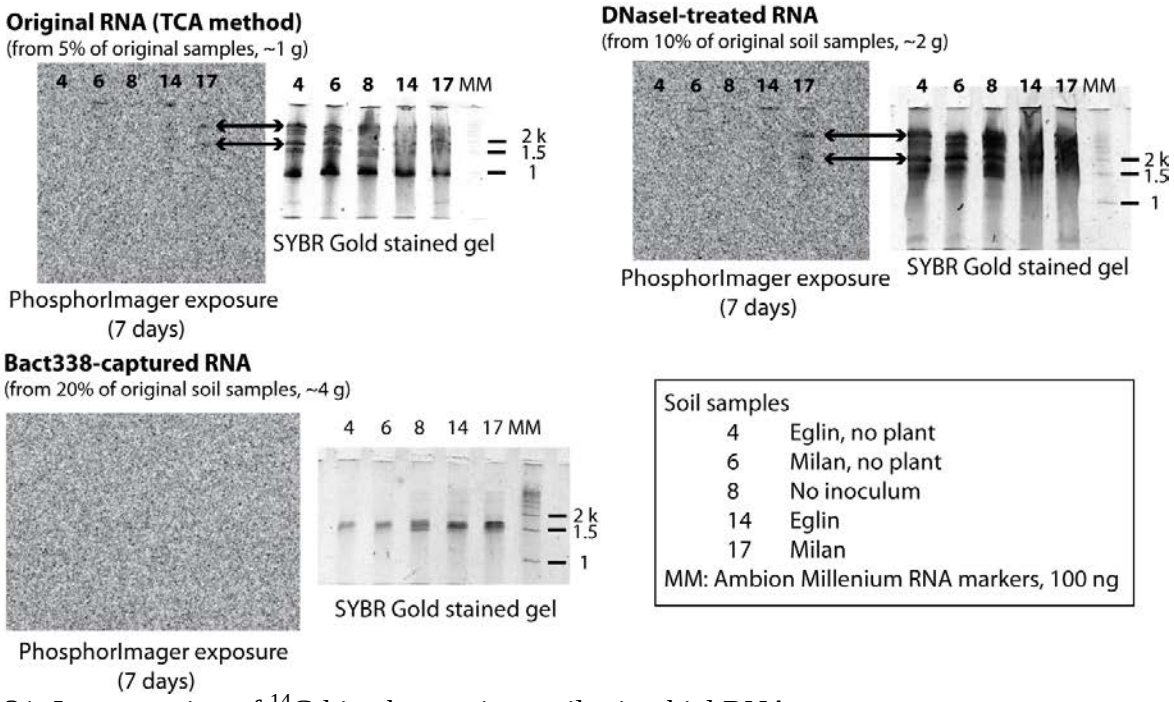


Figure 31. Incorporation of ^{14}C-bicarbonate into soil microbial RNA.
RNA was isolated from ~20 g of root-associated soil by beadbeating in a trichloroacetic acid/Sarkosyl solution (McIlroy, Porter et al. 2008) followed by phenol/chloroform cleanup (MacGregor, Moser et al. 1997). Bacterial RNA was captured using a general bacterial probe with helper probes as described (Miyatake, MacGregor et al. 2009). Total RNA, DNaseI-treated RNA, and bead captured RNA fractions were separated on denaturing polyacrylamide gels that were stained with SYBR Gold (Invitrogen Corp.), photographed, dried, exposed on phosphor screens, and imaged on a Storm PhosphorImager (GE Healthcare Inc.). Label incorporation into RNA was detected for bands of the size expected for eukaryotic large- and small-subunit rRNA (arrows), but not in probe-captured bacterial RNA.

a) Development of group-specific capture probes.

In order to move beyond measuring label incorporation into small subunit rRNA of the total bacterial community, group-specific probes of tested specificity are needed. The probes existing at the time we began our work had been designed against a much smaller 16S rRNA database than the current (and ever-expanding) one. We designed and tested new probe (Figure 32) and helper probe (Figure 33; (Fuchs, Glockner et al. 2000)) combinations targeting the Alpha and Beta/Gamma proteobacterial groups and identified optimal formamide concentrations for each by testing against pure-culture RNAs with zero, one, or two mismatches to the probe sequences (Figure 34). An unusually high formamide concentration of 70% was required to discriminate against *Photorhabdus luminescens* with the BG553 probe, although it has a central mismatch which would usually be considered a strong one. Hybridizations without helper probes showed less of an effect, but target RNA capture was also affected quite strongly (e.g. Figure 33). This difference between *in silico* and *in vitro* specificity highlights the need for empirical testing of probes, and ideally the sequencing of a representative sampling of captured RNA.

In addition to the two probes shown here, we have also contributed to an extensive study of new and existing probes targeting the Deltaproteobacteria (Lucker, Steger et al. 2007); but could not identify a probe targeting the entire group.

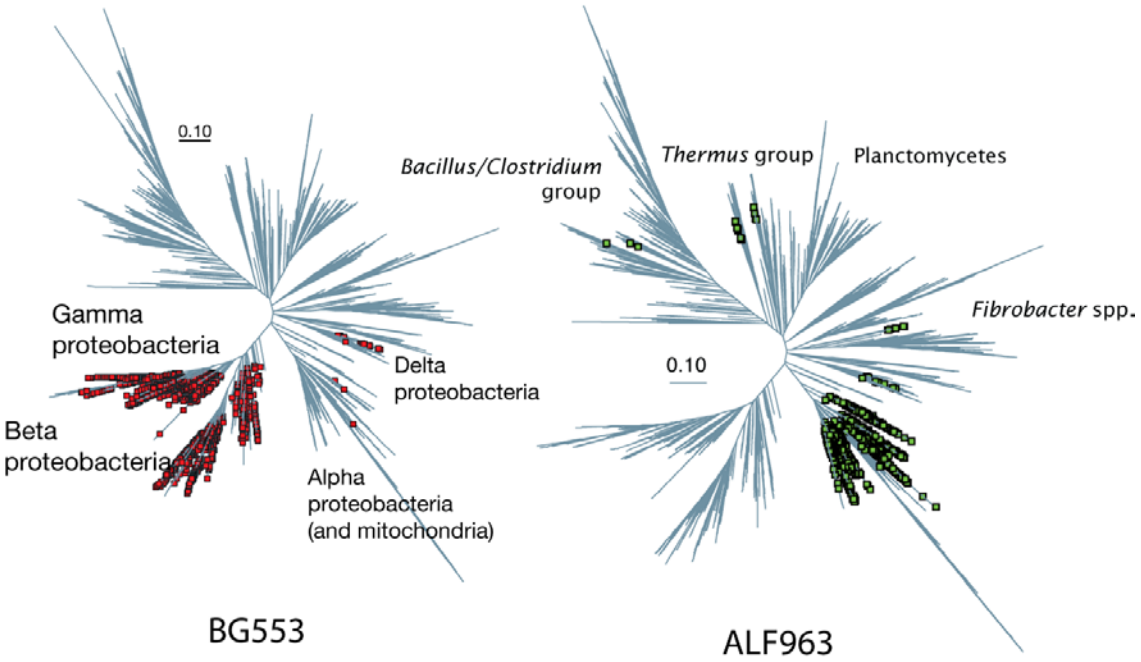


Figure 32. In silico specificity of BG553 and ALF 963 probes.
The probes were checked against the SILVA database (Pruesse, Quast et al. 2007) using the ARB sequence analysis package (Ludwig, Strunk et al. 2004).

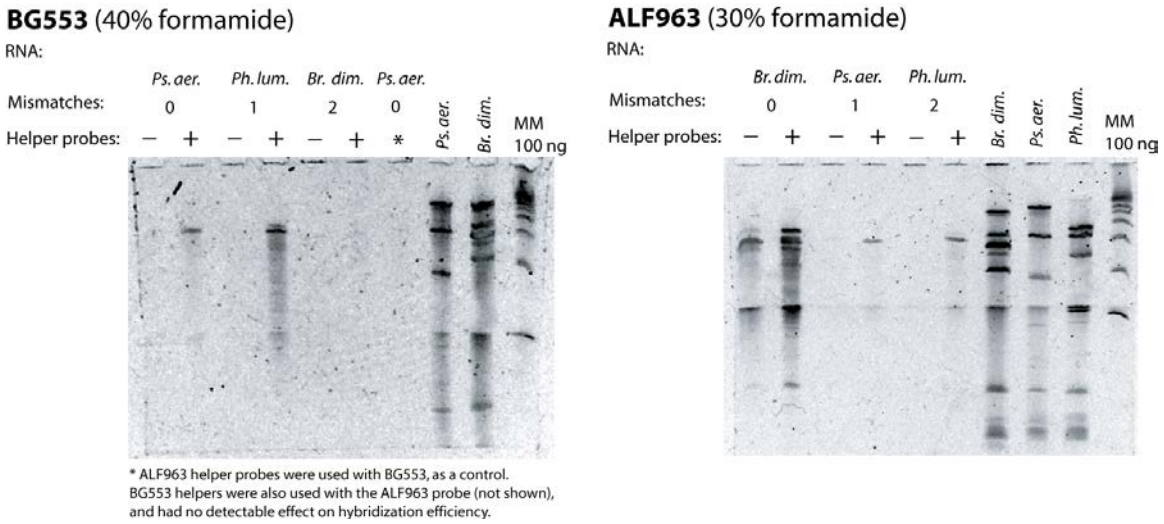
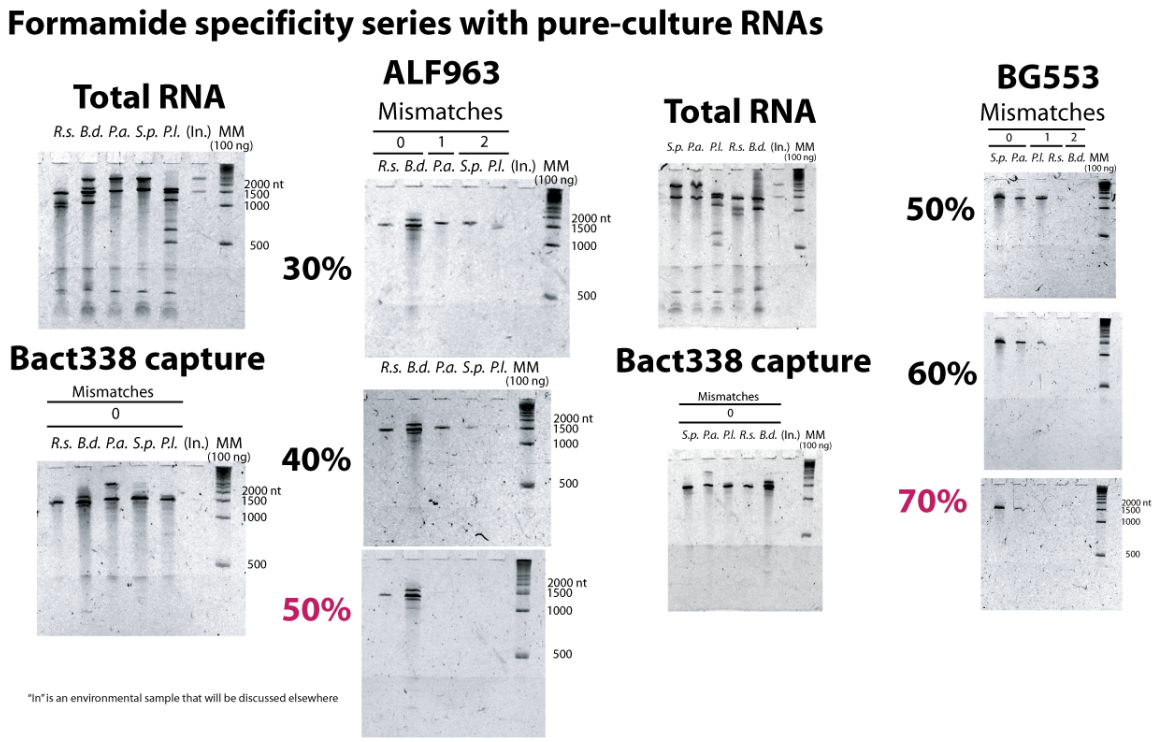


Figure 33. Group-level capture probes with and without helper probes.
Approximately equal amounts of 3 different pure-culture RNAs were hybridized with biotin-labeled oligonucleotides targeting the Beta/Gamma or Alpha Proteobacteria.

Helper probes are unlabeled oligonucleotides complementary to target-group consensus sequences upstream and downstream of the probe target site. When included, they were used at the same molar concentration as the capture probes. The pure-culture RNAs on the gels

represent 5 times the amount of RNA used in the hybridizations. Note that all three pure culture RNAs show evidence of rRNA processing and/or degradation. For example, the ALF963/B. diminuta hybridization yielded one long and several shorter species. These were also found in the starting RNA, so would not seem to result from degradation during the experiment; they may represent stages in rRNA maturation, or else result from degradation at specific sites during RNA isolation or handling. The bacterial strains used and the GenBank accession numbers for their SSU rRNA sequences are as follows: *Brevundimonas diminuta* (ATCC 11568; AJ227778), *Photorhabdus luminescens* (ATCC 29999; X82248), *Pseudomonas aeruginosa* (ATCC 10145; AF094713).



Probe and helper-probe target-region sequences

ALF963

Probe name:	**ALF963_help_up19**	**ALF963**	**ALF963_help_down18**
Probe sequence:	GAA TTA AAC CAC ATG CTC C	GGT TCT GCG CGT TGC TTC	SAT GTC AAR NGN TGG TAA
Target sequence:	GGA GCA TGT GGT TTA ATT C	GAA GCA ACG CGC AGA ACC	TTA CCA NCN YTT GAC ATS
R.s. Rhodobacter sphaeroides	GGA GCA TGT GGT TTA ATT C	GAA GCA ACG CGC AGA ACC	TTA CCA ACC CTT GAC ATG
B.d. Brevundimonas diminuta	GGA GCA TGT GGT TTA ATT C	GAA GCA ACG CGC AGA ACC	TTA CCA CCT TTT GAC ATG
P.a. Pseudomonas aeruginosa	GGA GCA TGT GGT TTA ATT C	GAA GCA ACG CG**A** AGA ACC	TTA CC**T** G**G**C CTT GAC ATG
S.p. Shewanella putrefaciens	GGA GCA TGT GGT TTA ATT C	GA**T** GCA ACG CG**A** AGA ACC	TTA CC**T** ACT CTT GAC ATC
P.l. Photorhabdus luminescens	GGA GCA TGT GGT TTA ATT C	GA**T** GCA ACG CG**A** AGA ACC	TTA CC**T** ACT CTT GAC ATC

BG553

Probe name:	**BG553_help_up21**	**BG553**	**BG553_help_down20**
Probe sequence:	AAC GCT YGC ACC CTM CGT ATT	CGC CCA GTA ATT CCG ATT	AAC CGC CTR CGN RCG CTT TA
Target sequence:	AAT ACG KAG GGT GCR AGC GTT	AAT CGG AAT TAC TGG GCG	TAA AGC GYN CGY AGG CGG TT
S.p. Shewanella putrefaciens	AAT ACG GAG GGT **C**CG AGC GTT	AAT CGG AAT TAC TGG GCG	TAA AGC GTG CGC AGG CGG TT
P.a. Pseudomonas aeruginosa	AAT ACG **A**AG GGT GCA AGC GTT	AAT CGG AAT TAC TGG GCG	TAA AGC GCG CGT AGG **T**GG TT
P.l. Photorhabdus luminescens	AAT ACG GAG GGT GCA AGC GTT	AAT CGG AAT **G**AC TGG GCG	TAA AGC GCA CGC AGG CGG T**C**
R.s. Rhodobacter sphaeroides	AAT ACG GAG GG**G** GC**T** AGC GTT	A**T**T CGG AAT TAC TGG GCG	TAA AGC GCA CGT AGG CGG **A**T
B.d. Brevundimonas diminuta	AAT ACG **A**AG GG**G** GC**T** AGC GTT	**GC**T CGG AAT TAC TGG GCG	TAA AG**G** GCG CGT AGG CGG **A**T

Figure 34. *In vitro* specificity of the ALF963 and BG553 capture probes.
For each probe, RNA from one or two perfect match, single mismatch, and double mismatch species were tested in hybridizations with increasing formamide concentrations. The formamide concentration yielding little or no detectable capture of non-target RNA is highlighted in each case. Helper probes were included in these hybridizations.

b) Optimization of SSCP separation conditions for pure culture and environmental samples.

Separation of rRNA by single-strand conformational polymorphism (SSCP) (MacGregor and Amann 2006) allows molecules of similar size to be separated because of the different

conformations they assume. It depends sensitively on the details of the polyacrylamide gel electrophoresis conditions: temperature, buffer composition, acrylamide:bisacrylamide type and concentration, and urea concentration. We optimized these for minigels using a collection of pure-species SSU rRNAs and found that the clearest separation was obtained in 5% Duracryl (30T, 2.6C; Proteomic Solutions, Saint Marcel, France), 1.7 M urea, and "30 mM" TBE (Liu et al., 2000) (Figure 35). A detailed protocol for minigels is attached below (Appendix 1C). This can be scaled up for a DGGE apparatus; in that case the best separation was obtained by placing the apparatus in ice water in a 4˚C cold room and running gels at a constant 19˚C, the lowest temperature that could be reliably maintained for the ~17h run time required. An example of minigel separation applied to rhizosphere and soil samples is shown in Figure 36; multiple bands could be detected in both total RNA and captured 16S RNA.

rRNA-SSCP was also scaled up for a DGGE apparatus; in that case the best separation was obtained by placing the apparatus in ice water in a 4˚C cold room and running gels at a constant 19˚C, the lowest temperature that could be reliably maintained by a constant-power power supply for the ~17h run time required. Separation distances between pure-culture bands are larger with this method, but the bands to date have also been more diffuse (Figure 37). Given the difficulty of maintaining constant temperatures in large gels, minigels may be the best choice for most applications.

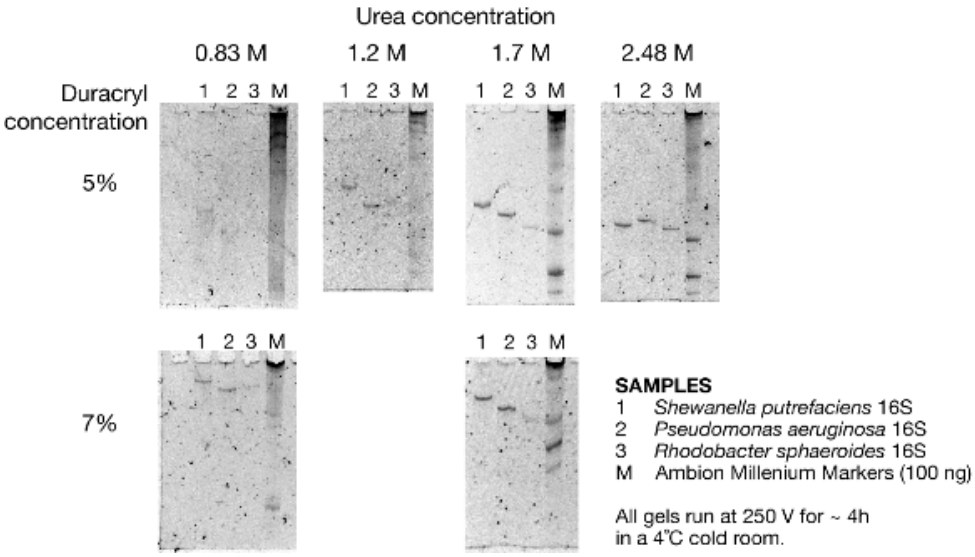


Figure 35. Optimization of rRNA-SSCP for pure-culture, Bact338-captured 16S rRNA. The best combination of sharp bands and separation was achieved with 5% Duracryl, 1.7 M urea.

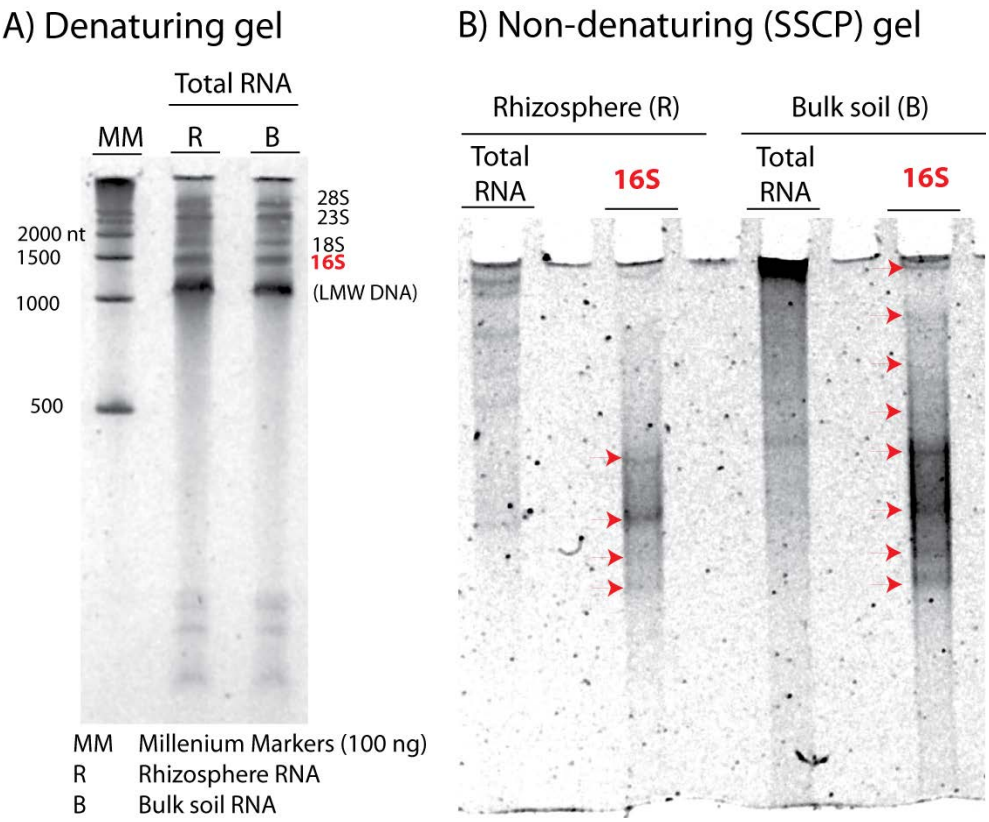


Figure 36. Isolation and rRNA-SSCP separation of bacterial SSU rRNA from soil and rhizosphere.
A) Total RNA was isolated from the rhizosphere or bulk soil of Western wheatgrass grown in a perlite/sand/peatgrass mixture, essentially as previously described (MacGregor, Boschker et al. 2006). B) Bacterial SSU rRNA (16S) was captured with Bact338 and separated by rRNA-SSCP (MacGregor and Amann 2006), using the conditions selected from the experiments in Figure 35. Arrows denote individual RNA bands separated by SSCP.

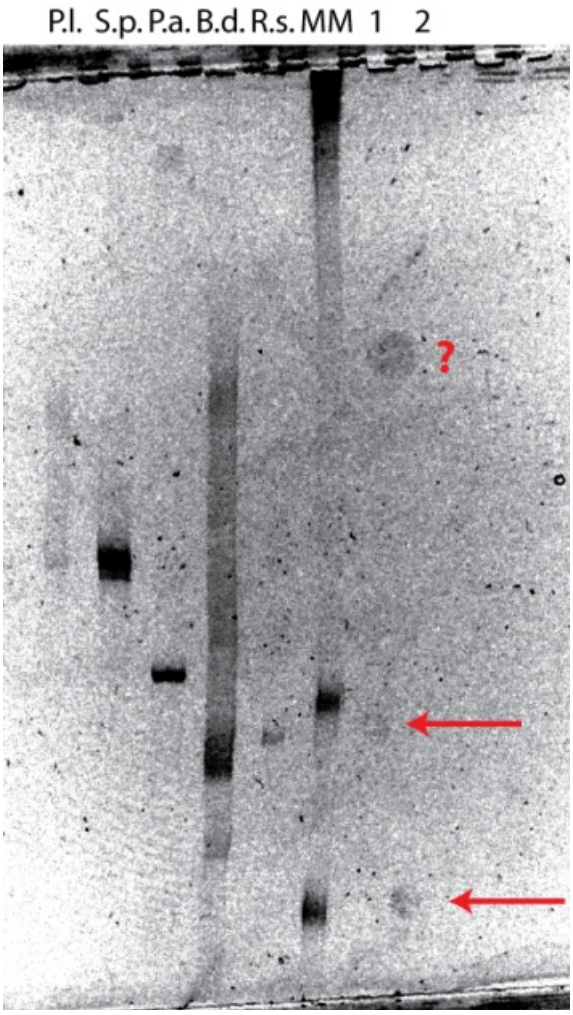


Figure 37. Large-scale RNA-SSCP separation of bacterial SSU-rRNA.
Probe Bact338 was used to capture 16S rRNA from total RNA of the species or samples shown. Soil-sample lanes represent about half of each sample (~1.3 g for rhizosphere, 6 g for bulk soil). P.l., *Photorhabdus luminescens*; S.p., *Shewanella putrefaciens*; P.a., *Pseudomonas aeruginosa*; B.d., *Brevundimonas diminuta*; R.s., *Rhodobacter sphaeroides*; MM, Ambion Millenium Markers (~100 ng); 1, Slender wheat #1 - Rhizosphere; 2, Slender wheat #1 - Bulk soil.

4. Conclusions

Our objective of identifying bacterial species incorporating plant-derived carbon was not met, but we believe it could be accomplished by some combination of heavier labeling, longer incubations, and larger sample size. In retrospect, it might have been best to start with a system where extensive transfer has already been demonstrated. We have been applying the methods developed and lessons learned here to several ongoing projects; in particular, working out the conditions for phosphorimager detection of ^{14}C-bicarbonate incorporation has encouraged us to begin looking for petroleum carbon incorporation by algae (Gutierrez and MacGregor, unpublished observations). The group-specific probes should be useful for the microbial ecology community in general - the use of nested sets of capture and hybridization probes is the ideal, but often falls apart in the middle range of specificity.

C. Stable isotope probing of RDX degraders: Density separation of ^{15}N-DNA. (Stuart Strand, David Stahl, Peter Andeer, University of Washington, Seattle)

1. General Methods

a) Growth media

Unless otherwise indicated chemicals (>98% pure) and media components were supplied by Sigma Aldrich (St. Louis, MO) Fisher Chemical (Fairlawn, NJ), or Difco (Becton Dickinson; Franklin, NJ). Water used in all media preparations, (ddH_2O) was deionized water purified with a Synergy 185 purification system (Millipore; Bedford, MA) that includes UV (185 nm), ion exchange, activated carbon and physical filtration (0.22 μm). Solutions were sterilized either by filter sterilization through: a 0.22 μm syringe filter (Pall; Port Washington, NY; or Millipore) or bottle top/prepackaged filtration units (Millipore), or through autoclaved (120°C @ 15 psi for 15 min unless otherwise stated). pH was adjusted with concentrated (12 N) HCl or 5N NaOH.

Potassium Phosphate Buffer, pH 7.2 (200 mM, 5x) was made by dissolving 6.72 g anhydrous potassium phosphate monobasic and 25.1 g of anhydrous potassium phosphate dibasic in 950 ml of water followed by pH and volume (1 L) adjustments.

Pfennig trace elements with $CaCl_2$ 100x solution:
This is the primary trace elements solution used in minimal media (Pfennig and Lippert 1966, Binks, Nicklin et al. 1995). Another recipe was tested (Rosenberger and Elsden 1960) but abandoned due to problems with precipitation. The following four stock solutions were made individually in 1 ml of water in sterile 1.5 ml microcentrifuge tubes:

15 mg $MnCl_2 \cdot 4\ H_2O$
5 mg $CuCl_2 \cdot 2\ H_2O$
10 mg $NiCl_2 \cdot 6\ H_2O$
15 mg $Na_2MoO_4 \cdot 2\ H_2O$

To 500 ml water, constantly stirring, the following components were added in order, dissolved and filter sterilized:

250 mg EDTA disodium salt
100 mg $FeSO_4 \cdot 7\ H_2O$
5.0 mg $ZnSO_4 \cdot 7\ H_2O$
100 μl of manganese stock solution
15 mg H_3BO_3
10 mg $CoCl_2 \cdot 6\ H_2O$
100 μl of copper stock solution
100 μl of nickel stock solution
100 μl of sodium molybdate stock solution
50 mg of $CaCl_2 \cdot 2\ H_2O$

$MgSO_4$ stock solution (50 mM, 200x) was made by dissolving 1.23 g of $MgSO_4 \cdot 7\ H_2O$ in 100 ml of H_2O and autoclaved.

Stock carbon sources used in most RDX cultures were: 10 % (w/v) glucose, 10 % (w/v) glycerol and 500 mM sodium succinate.

Glucose stock solution (10 % w/v; 105x) was made by dissolving 10 g of glucose in 90 ml of water, glycerol stock solution (10 % w/v; 105x) was made by mixing 8.0 ml of 99 % glycerol with 92 ml of ddH_2O and sodium succinate stock solution (500 mM; 100x) was made by dissolving 13.5 g of sodium succinate dibasic hexahydrate in 90 ml of water, and pH adjusted to 7.0 (+/- 0.2). All three solutions were adjusted to 100 ml and autoclaved.

Thauer vitamin solution: The following vitamin stock solution was incorporated into culture media where indicated (Brandis and Thauer 1981). 0.4 ml to 1 ml was used per liter of media.

To 1 L of ddH2O add:

- 20 mg biotin
- 20 mg folic acid
- 100 mg pyridoxine HCl
- 50 mg thiamine HCl
- 50 mg riboflavin
- 50 mg nicotinic acid
- 50 mg DL-pantothenic acid calcium salt
- 50 mg *p*-aminobenzoic acid
- 2 g choline chloride
- 10 mg vitamin B_{12}

The solution was sealed, autoclaved and stored in the dark.

b) Buffered mineral solution

The buffered mineral solution used in most of the culturing work (screening, enrichment, isolation and growth experiments) was based on previously reported RDX enrichment media (Binks, Nicklin et al. 1995, Seth-Smith, Rosser et al. 2002) composed of potassium phosphate buffer (40 mM, pH 7.2) supplemented with modified Pfennig trace elements (1 x), and $MgSO_4$ (0.5 mM).

c) Enrichment media

In most cases RDX enrichment cultures were conducted in buffered mineral solution amended with sodium succinate (5 mM), glucose (5 mM) and glycerol (10 mM) as carbon sources. RDX was added to media in one of two ways: 1) solvent from 10 mg/ml of RDX in acetone (Accustandard) or 99% ^{15}N- RDX (9.1 mg/ml; Defence Research and Development Canada; Valcartier, QC) in acetonitrile (ACN) was evaporated in autoclaved amber jars covered with AirPore tape sheets (Qiagen; Valencia, CA). Buffered mineral solution components were added and the solution was sonicated in a Branson 1200 sonicating water bath (Branson Ultrasonic Corporation; Danbury, CT) overnight to dissolve RDX. Carbon sources were then added and the media was filter sterilized. 2) Concentrated RDX-DMSO solution (>50 mg/ml) was made by evaporating solvent in sterile amber vials (20 ml) using N_2 supplied through a 0.22 mm filter and the RDX precipitant dissolved into 100% filter sterilized DMSO. RDX-DMSO stock solution was added in aliquots of 100 μl over several hours (to ~40 mg/L, 175 μM) to buffered mineral solution, shaken and allowed to sit in the dark at room temperature for 1 – 2 days. Carbon sources were then added as required and filter sterilized.

The first method provided media almost completely free of carbon sources (other than RDX) if required, but led to nitrite in the media due to decomposition of RDX (or possibly decomposition of contaminating HMX). Autoclaving RDX media led to RDX loss.

d) Alternative enrichment and culture media

Alternative carbon or nitrogen sources were used in a variety of screens and growth experiments. Alternative carbon sources used in RDX degradation experiments included yeast extract (5 mg/L), soil extract (Hurst and Knudsen 1997), sodium glycolate (5 mM), sucrose (5 mM) and the above carbon sources (glucose, glycerol, and succinate) at dilute (67 μM, 140 μM and 70 μM , respectively) concentrations. Carbon free incubations were also attempted. Alternative nitrogen sources employed included sodium nitrite (1 mM), ammonium nitrate (0.5 – 10 mM) and ammonium chloride (1 – 50 mM).

e) Solidified minimal media

For isolation of organisms on minimal agar plates, enrichment media (without RDX) with 1.5 % Noble agar was autoclaved and cooled to 50°C. 1 mg/ml of RDX (Accustandard) diluted in acetone and filter sterilized with PTFE filters (Millipore) was added to the agar at concentrations of 20 – 40 mg/L (85 – 175 μM) and swirled prior to pouring. Sterilized carbon and/or alternative nitrogen sources (e.g., sodium nitrite) were added following autoclaving as needed. When screening for RDX degraders, replicate plates without nitrogen were used for growth comparisons.

f) RDX overlay plates

RDX overlay plates (Seth-Smith, Rosser et al. 2002) were created by first pouring (ca. 15 ml) minimal media with carbon and without RDX solidified with 1.5% agarose as a shallow base. RDX in acetone was added to 50°C minimal media (concentration ~5 mM) with 1 % agarose and a thin layer (~ 5 ml) was poured on top of the solidified base.

g) Additional media employed

A variety of complex media were also used in this research for biomass production, culture purification, culture maintenance and general molecular methods (e.g. clone selection or screening). Miller LB media and agar, trypticase soy media, nutrient broth and agar, 1/10 nutrient agar, R2A agar, 1/10 R2A agar, plate count agar and Actinomycetes agar were all prepared following supplier instructions with noble agar supplemented to bring agar concentrations to 1.5% when 10x diluted solutions were used. SOB (super optimal broth) and SOC (super optimal broth with catabolite repression) were either provided in cloning kits (Invitrogen; Carlsbad, CA), or made from sterile Miller LB media amended with sterile $MgCl_2$ (10 mM) or sterile $MgCl_2$ (10 mM) and sterile glucose (20 mM), respectively prior to use (Sambrook and Russel 2001). Peptone yeast with brain heart infusion (PY-BHI) media was prepared with 1% peptone, 0.2% yeast extract, 0.2% brain heart infusion, 0.2% NaCl and 0.2% D-glucose, pH 7.2 (Yokota, Takeuchi et al. 1993) and solidified with 1.5% noble agar if required.

h) Sample collection and storage

Soil samples were obtained or collected from munitions contaminated sites to screen for RDX-degraders. Dr. Thomas Jenkins (USACE Research Engineer, January 2007) sent soils from several military bases screened positive for RDX and other explosives collected, and Paul Higgs (Environmental Coordinator) supplied contaminated soils from Milan Army Ammunition Plant (MAAP; Milan, TN; May 2007). MAAP samples were collected from two sources: soils excavated from munitions contaminated areas that had not yet been composted and soils from the same location composted with a mixture of chicken manure and potato scraps. Three 50 ml conical tubes taken directly from a new package were filled with each sample type.

Soils were collected from two sites: 'C52N Cat's Eye' an extensively used area of the Eglin Air Force Base (Eglin AFB; Eglin, FL; June 2007) training range and an unlined pit at Umatilla Chemical Depot (Umatilla, OR; April 2007) where runoff from munitions disposal/reclamation had once been channeled. Soils from 23 discreet locations were collected from Eglin AFB with the assistance of local EOD (explosive ordnance detection) personnel on a trip arranged by Edward O'Connell (Eglin AFB Environmental Scientist) and Michael Hunt. Some features of sample areas were: the interior and exterior of two former impact zones (craters), several areas lacking vegetation with high amounts of debris in the vicinity, areas with differing plant species, areas with differing soil appearances (e.g. brown sand, red sand); further information on sampling locations are located in Appendix A. Samples were excavated with trowels and placed into resealable bags (all purchased at a local Walmart) or into sterile 50 ml conical tubes. Trowels were wiped down with rubbing alcohol before each use to avoid cross contamination of samples. Samples were photographed and labeled according to the site description; unfortunately GPS coordinates were not obtained making replicate sampling at future dates difficult. Samples were then shipped overnight to the laboratory via Federal Express. Bulk soil samples from six locations believed to be in the same area at Eglin AFB were obtained by Dr. Lorraine Lillis (November, 2009). Soils from MAAP and Eglin AFB were both stored in the dark at 4°C.

i) Cultures

The following methods were used for the routine screening, enrichment and cultivation of isolated organisms for growth on RDX or other substrates. Collected samples were screened for aerobic RDX degradation activity either by direct addition of soils (1 – 10% w/v) to RDX enrichment media or inoculated (10 x – 100x dilution) with soil suspensions in 0.1% sodium pyrophosphate (1:10 w/v, soil:solution) (Hurst and Knudsen 1997). Cultures were grown in either baffled Erlenmeyer flasks (40 -80 mls in 250 ml flasks), or culture tubes (glass or polystyrene; 3 – 5 ml media) and shaken (100 – 200 rpm) at 28°C or 30°C. Soil slurries and cultures were screened for RDX degradation using HPLC analysis of aqueous phase samples.

j) HPLC methods

(1) Reagents

HPLC (High performance liquid chromatography) sample preparation and mobile phase reagents used were HPLC grade methanol (MeOH) and acetonitrile (ACN). Water was either ddH_2O or purchased 'HPLC grade' water. HPLC water provided less background absorbance, but both were acceptable. Reagents were supplied from: Sigma Aldrich, Mallinckrodt/ Baker

(Avantor; Phillipsburg, NJ) EMD Chemicals (Gibbstown, NJ) or Fisher Scientific (Pittsburg, PA). Mobile phases were degassed in a sonicating water bath (Branson 1200) for 20 minutes.

(2) Sample preparation: soil slurries and cultures

Plastic sterile serological pipets were used to remove samples for RDX concentration analysis. For other compounds, (e.g. TNT) sterile, borosilicate serological pipets were used to prevent sorption. Soil slurries or cultures sampled for HPLC analysis and DNA/RNA extractions, were first pelleted by centrifugation in a 1.5 or 2.0 ml microcentrifugetube at 16,000 – 20,000 x g for at least 10 minutes. The supernatant was transferred to a new tube and the soil/cell pellets were stored at -80°C. The supernatant was mixed with an equal volume of ACN or MeOH, mixed and centrifuged at 16,000 x g for 10 minutes. The added solvents arrested microbial activity, dissolved precipitated compounds (e.g. TNT metabolites), and promoted the pelleting of cells, salts and soil in samples. Following centrifugation, the supernatant (>500 µl) was transferred to a 1 ml amber HPLC sampling vial (Waters or equivalent).

(3) Sample preparation: soil samples

RDX concentrations in dry soil samples (environmental or dosed) were measured using a modification of EPA method 8330 (United States Environmental Protection Agency (USEPA) 1994). Samples were homogenized, and 2 – 3 grams were weighed in 15 ml conical tubes. 5 ml of acetonitrile was added, vortexed and stored in the dark until the day before analysis. Tube(s) were then sonicated overnight in a sonicating bath covered from light. Samples were then vortexed, and allowed to settle for 30 min. 300 – 750 µl of supernatant was removed and mixed with an equal volume of $CaCl_2$ solution (5 g/L), centrifuged for 10 min at >16,000 x g and transfered (>500 µl) to HPLC vials.

(4) HPLC run conditions

HPLC analyses were conducted on a modular Waters system outfitted with a C18 reverse-phase column (250 x 4.6 mm; Hypersil Gold, ThermoFisher or Waters) with spectrum analyses from (200 nm or 210 nm) – 400 nm using a photodiode array detector (PDA; Waters 2996). Mobile phase consisted of ddH_2O or HPLC water, and organic phase of either ACN or 5:1 MeOH:ACN. Initial separations were performed using 50/50 water/ACN at 0.8 or 1 ml/ min however this was changed to the water/MeOH-ACN mixture at 1 ml/ min isocratic flow ranging from 60/40 to 40/60 to improve separations between RDX and HMX as well as between TNT and metabolites. Most RDX samples (50 µl injection) were run isocratically with water/MeOH-ACN mobile phase (50/50) at 1 ml/min. Under these conditions, RDX eluted at about 6 min and was quantitated at 230 nm or 254 nm. These wavelengths were chosen based on the spectral analysis (Figure 2.1). The spectrum of the RDX peak was consulted frequently to verify peak purity. For plant extracts, the mobile phase ratio was changed to 40/60 (water/MeOH-ACN) to separate RDX from plant compounds, with the RDX peak eluting at approximately 12 minutes.

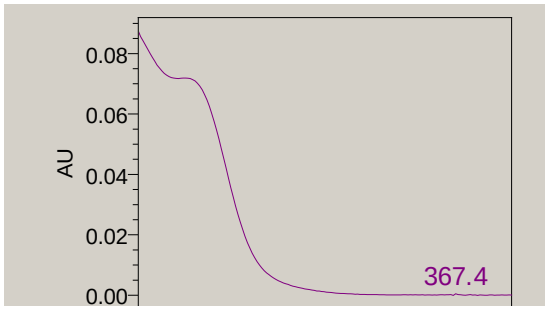


Figure 38. Absorbance spectrum of RDX.
Spectral scan (210 nm – 400 nm) of RDX using the Waters 2996 PDA.

k) Reagents for nucleic acid extractions

Three solutions were commonly used for molecular methods: 'DNA suspension buffer' which refers to

filtered, sterile, certified RNase and DNase free, Tris – EDTA (10 mM and 0.1 mM respectively) (Teknova; Hollister, CA), 'PCR water' which refers to filtered, sterile, certified for PCR and RNase/ DNase free (Teknova) and 'DEPC water' which refers to RNase/ DNase free filter-sterilized water treated with diethylpyrocarbonate (DEPC) (Applied Biosystems/Ambion; Austin, TX).

Sodium phosphate stock solutions (1 M, pH 6.6 and 8.0) were made to 50 ml in conical tubes with DEPC water. For 50 ml of buffers, 4.45 g of NaH_2PO_4 and 1.82 g of Na_2HPO_4 or 0.6 g of NaH_2PO_4 and 6.39 g of Na_2HPO_4 were dissolved into 40 ml of DEPC water for pH 6.6, and pH 8.0 buffers, respectively. 1 L of Tris stock solution (1M, pH 8.0) was made by dissolving 120 g of Trizma base into 900 ml of ddH_2O; pH and volumes were adjusted and autoclaved.
5 M NaCl stock solution (100 ml) was made by dissolving 29.2 g of NaCl in autoclaved ddH_2O.

Unless otherwise stated, the following solutions for DNA/RNA extractions were made using DEPC water in sterile 50 ml conical tubes and filter sterilized:

500 mM EDTA stock solution (500 ml) was made by adding 93 g of EDTA disodium salt to DEPC or PCR water (400 ml). pH was adjusted to 8.0 with 5 M NaOH and then autoclaved.
Aluminum sulfate solution (200 mM) was made by dissolving 6.66 g of $Al_2(SO_4)_3 \cdot 18\ H_2O$ into 40 ml of water and then pH was adjusted (c.a. 3.0) with 0.1 M NaOH and volume to 50 ml.
Dong salts solution (Dong, Yan et al. 2006) was made by combining 10.88 ml of sodium phosphate stock (pH 8.0), 45.4 ml of Tris stock, 4 ml of 5 M NaCl and 39.7 ml of DEPC water in an autoclaved jar and then filter sterilized.

Lysis buffer (50 ml) was made by combining: 5 ml Tris Stock, 10 ml EDTA stock, 3 ml NaCl stock, 5 ml 20% SDS solution and 27 ml DEPC water.

Sodium dodecyl sulfate (SDS, 20%) was made by adding 10 g of SDS to 35 ml of water (65^oC), and adjusting volume to 50 ml (no sterilization needed).

5 M NaOH stock made by dissolving 2g of NaOH pellets in 8 ml water and adjusting volume to 10 ml (1 M NaOH was made by diluting 5 M stock).

Sodium perchlorate solution (5 M, pH ~9.0 or pH 5.5) was made by dissolving 15.3 g of $NaClO_4$ in 18 ml of water (55^oC), pH was adjusted with 0.1 M NaOH and volume to 25 ml.

Phenol: chloroform: isoamyl alcohol (PCIA) was made by mixing chloroform, phenol and isoamyl alcohol at a ratio of 50:49:1. 100 mM Tris stock solution was then added (1/10 volume) to saturate phenol and maintain pH.

Chloroform: isoamyl alcohol solution (CIA) was made by mixing chloroform and isoamyl alcohol at a ratio of 49:1.

l) DNA/RNA extraction from cultures

Three primary methods were used for DNA extraction from cultures (isolates and enrichments):

1) Epicentre's Masterpure Gram Positive DNA/RNA extraction kit (Madison, WI), 2) a traditional method that combined several techniques (freeze thaw, enzymatic incubation, SDS and phenol-chloroform extraction) and 3) a modification of the soil extraction method. The first two methods provided high molecular weight DNA, but with many populations (e.g., Actinobacteria), the yields were much lower than achieved with the method three. Method three yields more DNA, but the DNA is more fragmented.

The Epicentre kit was used following the manufacturer's instructions with the following modifications made for extraction from Actinobacteria. Mutanolysin (150 U; Sigma-Aldrich), achromopeptidase (10 U; Sigma-Aldrich) and lysozyme solution (provided in the kit and used at twice the recommended dosage) were added to the suspended cell pellet and incubated for at least 45 minutes at 37°C. Proteinase K solution (provided in the kit or made fresh at 20 mg/ml) was added at 2 to 3 times the recommended concentration and incubations were conducted first at 37°C for 1 hr before moving to 65°C for 30 min.

2) Cultures were pelleted, resuspended in Tris-EDTA (10 mM/ 10 mM; 300 μl) with 1 mg/ml lysozyme (Sigma Aldrich) addition of mutanolysin and achromopeptidase optional) and incubated at 37°C for 15 min to 1 hr. Samples were then subjected to a series (2 or 3 times) of freeze thaw cycles moving directly between -80°C and 65°C for about 5 min at each temperature. Proteinase K was added (200 μg/ml to 1 mg/ml) and incubated at 55°C (1 hr to overnight) with periodic mixing. Sodium perchlorate (0.25 volumes, 55°C) was added and incubation continued for 10 min at 55°C. Samples were extracted twice with PCIA and once with CIA. Rnase A (5 μg) digestion was performed if required and then precipitated overnight (-20°C) with 0.6 to 1 volume of isopropanol. Nucleic acids were rinsed with 80% EtOH and suspended in DNA suspension buffer.

3) DNA and RNA extracted from cultures using the soils method, modified as follows: $AlSO_4$ and NaOH solutions were replaced with DEPC water and sodium phosphate buffer pH 8.0 was used instead of the pH 6.6 buffer.

m) DNA and RNA extraction from soils

A particular challenge in working with soils is to obtain DNA and RNA that is representative of the community present, and of sufficient quality, to allow for proper analyses (e.g., PCR, qPCR, and sequencing reactions). Because the centrifugation time for SIP is dependent upon the DNA fragment sizes, it is desirable to minimize DNA shearing. After testing several methods the following modification of Dong (Dong, Yan et al. 2006), was chosen for most soil extractions and for some pure cultures and enrichment cultures:

Soil slurry samples were transferred directly to lysis tubes. For fresh samples, or those frozen in bulk (e.g. > 5 g of soil slurry or mixed soils stored at -80°C), samples were weighed with a target of 300 – 400 mg/ lysis tube (Lysing Matrix E; MP Biomedicals; Solon, OH). Several tubes were used per sample when needed. Spatulas used for sample handling were stored in ethanol and wiped with 95% ethanol, flamed and cooled between samples.

Soil samples were incubated with aluminum sulfate ($AlSO_4$) at low pH (<7) to bind and precipitate humic acids prior to cell lysis to improve the purity of the extracted nucleic acids. $AlSO_4$ (100 μl) was added to samples, vortexed lightly and incubated on ice (10 s to 1 min depending on soil type). Potassium phosphate buffer, (pH 6.6; 200 μl) was then added and

mixed, causing brown precipitant to form. Dong salt solution (560 μl), 1 M NaOH (70 μl), 20% SDS (166 μl) and a drop (c.a. 50 μl) of chloroform were added to the tubes. 10 μl of selected samples were used to verify that pH was between 8 and 9 prior to SDS addition. Samples were left at room temperature until disruption to prevent SDS precipitation.

Sample disruption was performed in a FastPrep FP120 Cell Disrupter (MP Biomedicals). Initial disruption was conducted at speed '4.0' for 25 s and then centrifuged at 16,000 x g for 10 min at room temperature to pellet soil and debris. The supernatant was transferred to a 2.0 ml Safe-lock tube (Eppendorf; Hauppauge, NY); supernatants of larger scale extractions were pooled in 15 ml conical tubes (Becton Dickinson) with approximately 100 μl (2.0 ml tubes) or 500 μl (15 ml tubes) of chloroform and placed on ice.

Cell disruption was repeated twice to lyse more recalcitrant populations. For the second extraction, samples were incubated as above, but with the following solution volumes: 70 μl of $AlSO_4$ and 2 μl of concentrated HCl (~12 N; to counter sodium hydroxide carryover), 100 μl of sodium phosphate buffer, pH 6.6, 280 μl of Dong salts solution, 60 μl of 1 M NaOH and 80 μl of 20% SDS. Cells were then disrupted at setting '4.5' for 30 s, and centrifuged again. Samples were pooled with initial supernatants on ice. For the final extraction: 400 μl of Dong salts solution and 80 μl of 20% SDS were added and disrupted at '5.0' for 30 s. Samples were centrifuged and pooled with prior extractions.

Samples in chloroform were inverted and incubated on ice for 5 – 10 min and centrifuged at full speed (20,000 x g for 2.0 ml, 3220 x g for 15 ml tubes) for >10 min at 4^oC. The supernatant was transferred to clean tubes, no more than half full. $NaClO_4$ (0.25 volumes, 55^oC) was added, the sample inverted and incubated at 55^oC for 10 min. Samples were transferred to ice and 1 volume of ice cold CIA was added, gently mixed and incubated on ice or at -20^oC for at least 10 min. CIA extractions were centrifuged at maximum speed for 10 min at 4^oC and the upper phase was transferred to a fresh tube with blunt ended pipet tips (pipet tip ends cut with clean scissors wiped with 70% EtOH) avoiding the precipitated layer. CIA extraction was then repeated.

Sample were concentrated in 4 ml Amicon Ultra (30 kDa or 50 kDa) filtration devices (Millipore; 0.5 – 2 ml devices for smaller sample volumes) at half the recommended maximum speed until the volume was less than 500 μl (<60 μl for smaller devices). Dialysis against 8 volumes of DNA suspension buffer was performed twice in the same device and purified samples (100 – 200 μl) were transferred to clean microcentrifuge tubes. If needed, RNA was removed with a 30 minute incubation at 37^oC with 1 μl of RNase A (Ambion), followed by a final CIA extraction. Nucleic acids were precipitated overnight at -20^oC with 0.3 volumes of 10 M NH_4OAc and 1 volume of cold isopropanol and then centrifuged at 20,000 x g for 25 min at 4^oC. Supernatant was decanted and tubes inverted onto clean, lint free wipes. Pellets were rinsed twice with 500 μl of 70% EtOH and centrifuged 10 min at 16,000 x g before removing EtOH. A final centrifugation at full speed for 1 min to collect left over liquid, most of which was then carefully removed with a 20 μl pipet leaving < 10 μl of supernatant. Tubes were left open 5 to 10 min to evaporate remaining supernatant and DNA/RNA was suspended in 20 – 100 μl (dependent on sample size) of DNA suspension buffer (50^oC). Sample purity, homogeneity and concentrations were then estimated.

n) Secondary DNA purification

Further purification of DNA (up to 23 kb) was sometimes performed using the 'DNA Clean and Concentrator – 5' (Zymo Research; Irvine, CA). 10 – 100% of sample (up to 5 μg DNA per column) following the manufacturer's instructions with these specific methods: 5 volumes of binding buffer instead of 2 were used and DNA was eluted with 12.5 μl of heated DNA suspension buffer (or sequential 12.5 μl and 8 μl elutions) and volumes normalized to 20 μl. If 260/230 values were low (possibly due to EtOH carryover), a second purification using the kit was conducted.

o) DNA quality and quantification

Routine DNA quantity and quality estimates were performed on sample aliquots stored in TE or DNA suspension buffer (~1.2 μl) using the Nanodrop ND - 1000 (ThermoFisher Scientific; Wilmington, DE). At least one replicate was measured for concentration estimates and DNA purity was judged based on A260/A280 and A260/A230 ratios and visual inspection of the spectra (Sambrook and Russel 2001).

Quantification of dilute DNA samples (< 5 ng/ μl), was performed with SYBR® green I (Molecular Probes) in black polystyrene 96 well plates (Corning; Corning, NY) using a TECAN Infinite F500 plate reader (TECAN; Mannedorf, Switzerland). Either serial dilutions of *E. coli* genomic DNA or phage λ DNA of known concentration were used as standards. Samples (2 μl) were added to 80 or 100 μl of DNA suspension buffer with SYBR green I stock (10,000 x stock in DMSO; Invitrogen) diluted to working concentration (1x). Samples, standards and negative controls (no DNA) were all measured in replicate. Plates were incubated in the dark at room temperature for 10 minutes and then read on the TECAN plate reader using excitation and emission wavelengths of 485 nm and 535 nm, respectively. Initial reads were performed at machine calculated (optimized) gains; additional reads were performed on selected samples at user-defined gains as needed. Sample DNA concentrations were calculated from the serial dilution standard curves.

p) PCR methods

(1) Primers

The primers used for routine PCR, qPCR and TRFLP analyses are listed in Table 5. Primer stock solutions were stored at 1 mM or 100 μM. The working stock solutions used in reactions were diluted to 10 μM using PCR water.

Table 5. Commonly used primers.

Primer Name	Target	Sequence	Mod[a]	Use	Reference
27F	Bacterial 16S	5'-GTTTGAT CMTGGCTCAG-3'		Full 16S/ qPCR	(Lane 1991)
[FAM]-27F	Bacterial 16S	5'-GTTTGATC MTGGCTCAG-3'	[6'-FAM]	16S qPCR-TRFLP	(Lane 1991)
1492R	Bacterial 16S	5'-ACGGYTAC CTTGTTACGACTT-3'		Full 16S	(Lane 1991)
338R	Bacterial 16S	5'-GCTGCCTCCCG TAGGAGT-3'		16S qPCR	(Amann, Binder et
xplAF	*xplA* gene	5’-CCGACGTAA CTGTCCTGTTCGGAA-		*xplA* gene screen	(Rylott, Jackson et
xplAR	*xplA* gene	5’-CGGGTCCGTC CGCCGGCTGGAAGG-		*xplA* gene screen	(Rylott, Jackson et
xplAtaq-F	*xplA* gene	5'-GGAGGACAT GAGATGACCGCT-3'		*xplA* taqman assay	(Indest, Crocker et
xplAtaq-R	*xplA* gene	5'-CCTGTTGCAG TCGCCTATACC-3'		*xplA* taqman assay	(Indest, Crocker et
xplAtaq-Probe	*xplA* gene	5'-TCCCGAATTCAGG AACAACCCCTATCC-3'	[6'FAM] ,	*xplA* taqman assay	(Indest, Crocker et
dapBF	gene screen	5’-ATGACGAACA TCAGAGCTGTCGT-3			
dapBR	gene screen	5’-TTACAGTTCTTC GCGCACGATGTA-3’			

a. Primer modifications: FAM is 6-carboxyfluorescein and BHQ1a is ‘Black hole quencher 1’

b. A newly designed primer similar in location to the one reported in the reference but altered based on recent sequences submissions

c. Same sequence but different quencher used

(2) PCR reactions

Three primary PCR amplification solutions were used for routine screening, sequencing and clone insert generation. Two different solutions with Taq DNA polymerase (Fermentas and Lucigen; Middleton, WI) and Phusion enzyme (Thermo Scientific) were used. Table 6 lists the reaction composition and conditions typically used.

Table 6. Typical PCR compositions (20µl) and cycling conditions. All volumes are in microliters.

	Fermentas Taq	EconoTaq	Phusion
Buff	2	10*	4
MgCl2	2	X	0.15
Fwd Primer	0.4	0.4	0.6
Rev Primer	0.4	0.4	0.6
dNTPs	0.4	X	0.4
Water	11.7	7.2	11.55
DMSO	1	X	0.6
Enzyme	0.1	X	0.1
Template	2	2	2
Cycling (20 – 34 cycles)			
Initial Denature	95°C, 3 min	96°C - 98°C, 2 - 3 min	98°C, 30s - 2 min
Denature	95°C, 30 s	96°C, 30 s	98°C, 8s - 10s
Annealing	52°C - 62°C, 25s	52°C - 62°C, 25 s	58°C - 72°C, 15 - 25s
Extension	72°C, 1min/kb	72°C, 1 min/kb	72°C, 30s/1kb
Final Extension	72°C, 10 min	72°C, 10 min	72°C, 5 - 8 min

* - master mix with dNTPs and enzyme (Lucigen)

(3) Quantitative PCR

Quantitative PCR (qPCR) was usually performed in a MJ-Research PTC-200 gradient thermocycler with a Chromo 4 Real-Time PCR detector using the Opticon Monitor 3.1 software (BioRad; Hercules, CA) for quantification. Reactions were conducted in white, low-profile thermo strips with ultraclear flat caps (Thermo Scientific). Reactions (20 µl) consisted of 2 µl of template and 18 µl of master mix solution. For 16S rRNA gene amplification, master mix solutions contained (per reaction): 10 µl of SsoFast EvaGreen supermix (BioRad), 6.4 µl PCR water and 0.8 µl of each primer (10 µM stock solution, 8 pmoles/rxn each). *E. coli* K-12 MG1655 genomic DNA (Blattner, Plunkett et al. 1997) was used as a copy number standard for 16S rRNA gene amplification (1.38 x 10^6 copies/ng). Cycling conditions used were either:

1) Initial denaturation was performed at 98°C for 2 minutes, and each cycle was 98°C for 8s, 58°C for 12s and 72°C for 15s with a 5 minute final extension step added after cycling was finished; or

2) A 2 minute initial denaturation (98°C) and each cycle was 98°C for 8 s and 58°C for 28 s and fluorescence analysis.

For *xplA* Taqman assays reactions (20 µl) consisted of: Faststart Taqman Probe Master (10 µl, Roche), 6.1 µl PCR certified water, 0.7 µl of each primer (xplAtaqF, xplAtaqR; 7 pmol), 0.5 µl probe (xplAtaqprobe; 5 pmol) and 2 µl of sample. See Table 5 for primer sequences. The plasmid pHSX1, a vector with the *xplA* gene (Seth-Smith, Rosser et al. 2002) was used as a copy number standard for *xplA* quantification. An initial denaturation was performed at 95°C

for 10 min followed by cycling between 95°C for 15 s and 60°C for 40 s followed by fluorescence reading. Samples for digestion were stored at -20°C until purification.

(4) PCR purification

qPCR products (~300 bp) generated with the [6'-FAM –27F] and 338R bacterial primer set, were purified either using Princeton Separations HTS PCR purification kits or with Zymo's ZR-96 DNA Clean and Concentrator-5 kit to remove unincorporated primers. Both of these kits were used as directed and products eluted using an elution buffer (10 mM Tris buffer or DNA suspension buffer). PCR products for cloning or sequencing were either purified using a commercial PCR purification kit (Qiagen; Zymo) or size selected, excised from a 1% agarose gel and purified using the Montage DNA gel extraction kit (Millipore).

q) Cloning

(1) Clone library construction

PCR products, qPCR products and fragment DNA were ligated into the pCR4-TOPO, pCR2.1Blunt-TOPO or pCR4Blunt-TOPO vectors contained in Invitrogen's TA cloning and Shotgun Subcloning kits. PCR products generated using Taq DNA polymerase were purified and directly ligated into pCR4-TOPO vectors (Invitrogen) following manufacturer's instructions. Blunt ended PCR products (qPCR products and PCR products generated using Phusion DNA polymerase) were either: 1) purified and then incubated in 25μl of 1 x PCR buffer solution with Taq DNA polymerase (1U), $MgCl_2$ (1.2 mM) and dATP (0.2 mM) for 2 min at 72°C and ligated into the pCR4-TOPO vectors (Invitrogen) or ligated into the pCR4Blunt-TOPO vector (Invitrogen) following manufacturer's instructions.

(2) Electroporation

Electrocompetent *E. coli* DH5α cells were prepared following standard protocols (Sambrook and Russel 2001). 10 ml LB starter cultures were created from a single colony of DH5α on LB agar and grown to mid-late exponential phase at 37°C (14 to 16 hrs or an optical density of 0.8 – 1 at 600 nm (OD600)) in LB media. Culture was then transferred to 500 ml of LB media at 1:1000 dilutions in baffled flasks. After 2 – 5 hours, cultures (0.8 OD600) were transferred to an ice bath and divided into precooled 500 ml centrifuge bottles, and centrifuged at 4°C for 10 min at 5,000 x g. The supernatant was carefully decanted three times and cells were suspended and washed in 0.5 to 0.7 volumes of 10% ice-cold sterile glycerol and centrifuged at 5,000 x g. Cells were then suspended in a 1/400 of initial volume (1.25 ml for 500 ml culture) of 10% sterile glycerol, and divided into cryovials, (50 to 100 μl aliquots) flash frozen and stored at -80°C.

Ligation reactions were diluted 3 fold (1 to 1:9) in PCR water and 1 μl was added to 24 - 49 μl of cells in a 1 mm gap electroporation cuvette (BioRad). Electroporation was conducted in a Gene Pulser II Electroporation System (BioRad) set as follow: 1.7 kV, 25 μF, 200 Ω. 0.5 ml of SOB media was immediately added and used to transfer cells to a 2 ml tube and shaken for 1 hr at 37°C. Aliquots of 10 – 100 μl were then spread on LB plates with kanamycin (50 μg/L) or carbenicillin (50 – 100 μg/L).

r) Sequencing

DNA sequencing was either performed at the University of Washington High Throughput Sequencing (UW HTSEQ) facility from recombinant colonies supplied cryogenically in LB with 10% glycerol or sequencing reactions were conducted in the Stahl lab and analyzed at the DNA sequencing facility at UW Biochemistry Department or at the UW Comparative Genomics Center. ABI BigDye v3.1 (Applied Biosystems) protocols were followed for sequencing. Typical sequences reactions (10 µl) consisted of 2 µl of BigDye v3.1 cycle sequencing mix, 0.4 µl of (10 µM) sequencing primer, 1 µl of 5x BigDye v3.1 buffer and PCR with DNA template (50 – 500 ng for PCR product or plasmid templates) to 10 µl. 20 µl reactions were the same as above with 5x BigDye v3.1 increased to 3 µl and PCR water and template increased accordingly.
Cycling for standard sequencing reactions was performed following the BigDye v3.1 protocols (Applied Biosystems) for standard clone sequencing and direct fosmid sequencing.

s) TRFLP

(1) Digestion

10 µl of purified qPCR products, two negative controls and two qPCR standards were digested for TRFLP analysis. 30 µl of MnlI digestion mixture (prepared as a master mix): 4 µl of 10x NEB (New England Biolabs; Ipswich, MA) buffer 4, 0.4 µl of 100x BSA (NEB), 25.2 µl PCR water and 0.4 µl of MnlI, was added per sample. Samples were incubated in a thermocycler for 3.5 hrs at 37°C followed by twenty minute incubation at 65°C (to deactivate and remove bound enzyme from the fragments) and then cooled to room temperature. 10 µl of MspI mixture consisting of 1 µl NEB buffer 4, 8.2 µl of PCR water and 0.8 µl of MspI were then added to the samples, mixed and centrifuged. Digestion with MspI proceeded for 2.5 hrs at 37°C.

(2) TRFLP purification and analysis

Unincorporated primers were removed from TRFLP reactions either by extraction with phenol-chloroform followed by purification with Centri-Sep size exclusion plates (Princeton Separations) or purified directly using the ZR-96 DNA Sequencing Clean-up kit (Zymo Research) eluted with 16 µl of PCR water. 4 µl of sample, or 5x dilution with PCR water to a volume of 6 µl was added to 96 well plates (MicroAmp; ABI or equivalent) and 12.5 µl of size standard master mixture was added. The master mix contained 12.25 µl of Hi-Di formamide (Applied Biosystems) and 0.25 µl of the size standard (30, 50, 100, 125, 150, 175, 200, 210, 250, 300, 400, 550 single stranded bps labeled with carboxy-X-rhodamine (5-ROX); Bioventures; Murfreesboro, TN) per reaction. Just prior to analysis, samples were centrifuged and denatured in a thermocycler for 3 minutes at 97°C and immediately transferred on ice and analyzed using an ABI 3730 capillary sequencer (48 samples, 1.25 hrs/ run) at the UW Comparative Genomics Center.

t) Fosmid cloning

For large molecular weight DNA, extraction was performed using the digestion method and DNA was sheared to approximately 40 kb by pipetting through a 20 µl pipet tip. Sheared DNA was end repaired using the reagents supplied in the CopyControl^{TM} Fosmid Library Production Kit (Epicentre), separated on a large (25 cm) 1% low-melt agarose gel (run conditions) with

SYBR green I (1x; Invitrogen) along with provided 40 kb fosmid control DNA size standard (Epicentre). Lanes with the sample for Fosmid cloning were carefully removed from the gel prior to visualization. The remaining gel with size standards was imaged and printed at actual size. The excised gel portion with sample was then placed on the printout and the gel section with the appropriate sized fragment was excised. The following steps were all performed with the Epicentre kit (listed above) following manufacturer's instructions. The fragment from the low melt gel was digested using GELase, end repaired and ligated into the pCC1FOS vector. The ligated construct was then packaged into MaxPlax™ Lambda packaging extracts, infected and plated using EPI300-TI® (Epicentre) phage resistant strains. Recombinant colonies were selected on LB agar with chloramphenicol (34 μg/ml).

u) Pulse field gel electrophoresis

Cell plugs were made of late exponential phase cells grown in LB media. Cells were harvested, and suspended at several dilutions into 50 μl of 10mM Tris and heated to 50°C in a waterbath and mixed with an equal volume of 50°C 2% SeaKem Gold agarose (Lonza; Rockland, ME) in 0.5x TBE (Tris-Borate EDTA) and quickly added to cell plug molds (BioRad; Hercules, CA). After solidifying at 4°C for 20 min, plugs were transferred to microcentrifuge tubes to lyse cells following the modified BioRad protocol listed below supplied by Dr. Nicolás Pinel. 500 μl of lysozyme solution (10 mM Tris, 50 mM NaCl, 50 mM EDTA, 0.2% SLS (sodium lauryl sarcosine) and 1 mg/ ml lysozyme) was added and incubated for 24 hrs at 37°C. The lysozyme was solution was removed, the plug rinsed with 1 ml of wash buffer (20 mM Tris and 50 mM EDTA) and 500 ml of proteinase K solution was added (100 mM EDTA, 0.2% SLS and 1 mg/ml proteinase K) and incubated for 24 hrs at 55°C. The plug was then washed four times with PFGE wash buffer and stored at 4°C.

Cell plugs were embedded into 1% agarose (SeaKem Gold agarose; Lonza) in 0.5x TBE by casting the agarose gel around the comb with the cell plugs adhered. Pulse field gel electrophoresis (PFGE) was performed in a Chef DR II system (BioRad) for 24 hours at 6 V/cm with a 10 to 100 second switch time ramp at a 120° angle with 0.5x TBE buffer circulating at 14°C. The *S. cerevisiae* YNN295 and Lambda Ladder markers (BioRad) were used as size standards. The gel was stained with 1x SYBR green I (Invitrogen) dissolved in 0.5x TBE and rocked for 30 minutes at room temperature.

v) Southern analysis

DNA from the pulse-field gel was transferred to a Magnacharge membrane (Micron Separations Inc; Westborough, MA) by overnight capillary transfer using the alkaline transfer method (Sambrook and Russel 2001). Alkaline phosphatase-labeled *xplA* DNA probe (c.a. 400 bp) was generated using the Alkphos Direct™ kit (GE Healthcare; Piscataway, NJ) of a DNA fragment amplified from *Microbacterium* sp. MA1 using the xplAF – xplAR primer set (Table 2.1) following manufacturer's instructions. The Gene Images Alkphos Direct Labeling and Detection System kit (GE Healthcare) with the CDP-*Star* chemiluminescent detection reagent (GE Healthcare) was used for probe hybridization and alkaline phosphatase labeling. Probe-labeled membrane was exposed to Hyperfilm ECL (GE Healthcare).

w) SIP methods

(1) Materials and reagents

The OptimaMax (bench-top) Ultracentrifuge, a TLA-110 fixed angle rotor and OptiSeal 4.7 ml pollyallomer plug sealed tubes (Beckman Coulter, Brea, CA) were used for ultracentrifugation. Stock cesium chloride (CsCl) solution was made with Ultrapure cesium chloride (Invitrogen) dissolved in DNA suspension buffer to a buoyant density (BD) of approximately 1.71 g/ml with EDTA added (0.5 M stock solution) to a final concentration of 0.1 M EDTA and then filter sterilized. Gradient buffer (100 mM Tris (pH 8.0), 100 mM KCl and 10 mM EDTA) was made with DEPC water and stock solutions and filter sterilized.

(2) Sample and run conditions

For effective separation of DNA in cesium chloride gradients, it is important to optimize initial solution density, centrifugation speed and time. Initial solution density and speed were determined from the rotor operating manual and separation goals. Degree of separation is primarily dependent upon three parameters: centrifugation speed, the tube and rotor dimensions (minimum and maximum radii) and initial solution density. Centrifugal forces vary across the length of the centrifuge tube in a nonlinear manner. Therefore, density gradients are not linear across the length of the centrifuge tube with steeper slopes towards the bottom of the tubes. For this reason, it is desirable to use a starting concentration at or as near to the particle density (1.69 – 1.73 g/ml for unlabeled DNA) as possible while avoiding precipitation of CsCl (Osterman 1984).

Slower centrifugation speeds lead to shallower gradients that increase the separation between particles. Thus, ideal separations would be achieved at the slowest possible speed and highest possible initial density. Conversely, the centrifugation times need to be long enough to achieve a stable gradient and allow particles to reach its equilibrium position within the gradient. The former rate is relatively fast and dependent upon equipment parameters (28.4 hrs for the rotor used). The second rate is inversely influenced by centrifugation speed and particle size, so slower speeds and smaller DNA fragments lead to longer centrifugation requirements (Osterman 1984).

For these reasons, the chosen run parameters were: initial densities of 1.66 – 1.68 g/ml (maximum permissible value is ~1.71 g/ml), centrifugation speed of 55,000 rpm (88,100 x g – 164,000 x g relative centrifugal field range), which were in agreement with published values (Buckley, Huangyutitham et al. 2007). To minimize the number of DNA fragments that had not reached equilibrium, centrifugation times of 93 – 96 hrs were used, suitable for the theoretical equilibration of DNA fragments of as small as 1100 bps (Osterman 1984).

(3) Sample preparation and setup

Samples were prepared directly in ultracentrifuge tubes to prevent sample loss. Initial volumes used were 4.44 mls of CsCl stock solution, 168 μl of gradient buffer and 96 μl of purified sample in DNA suspension buffer. DNA masses per gradient ranged from 1.0 μg to 7.0 μg. Tubes were topped off when needed with the above CsCl sample solution prepared with DNA suspension buffer instead of sample. Tubes were then lightly plugged and mixed through inversion. 100 μl were taken from each tube and carefully weighed to estimate densities (135AB – S/FACT balance, readable to 0.01 mg, certified to 1 mg; Mettler Toledo; Columbus,

OH) and replaced. Tubes were weighed and masses were adjusted with the DNA free CsCl solution and the balanced tubes were loaded into the rotor.

(4) SIP fractionation and measurements

Following centrifugation, the centrifuge was slowed at the lowest braking level (one above coasting) to minimize disruption of the gradient. Samples were then fractionated. The tube was carefully immobilized on a stand with a clamp. An IM1 (Becton Dickinson) needle was inserted approximately 1 cm above the bottom of the tube with a slightly upward angle with the needle's bevel horizontal in the center (radially) of the tube. A 1/16" barb was inserted in the needle housing to reduce drop size. Light mineral oil was then pumped either by hand or with a syringe pump through the top of the tube through a syringe outfitted with tubing and an IM1 needle inserted into the tube at the 'bell' at the top. 8 drops/ fraction (~85 – 100 µl/ fraction) were collected in PCR tube strips which were immediately sealed, and any alterations (missed or extra drops) in a given fraction were noted. A Foxy Jr® fraction collector (Teledyne ISCO; Lincoln, NE) outfitted for a 96 well plate was used to collect fractions. Fraction collection based on time provided more consistent fractionation; however, reliability of the syringe pump available excluded this option. Additionally, the extra tubing from the needle to the fraction collector may lead to DNA losses due to sorption. Therefore, the drop method described was used.
After all tubes were fractionated, the refractive index (n) of 6 µl aliquots were measured in an AR200 hand held digital refractometer (Reichart, Ithica, NY) modified with electrical tape to reduce the required sample volume (Buckley, Huangyutitham et al. 2007). The refractometer was wiped with wet and dry tissues in between each sample and the tape was frequently changed. Refractive index values were converted to buoyant density (BD) values (g/ml) using the following equation:

$BD = 10.928 \times n_{20} - 13.593$ (Osterman 1984)

While temperature can have a slight influence on measurements, 20°C was used unless noted. Buffer can also affect refractive index values (Osterman 1984), however differences in refractive index readings between CsCl solutions prepared in ddH_2O and those prepared as used in density gradients were negligible. Calculated BD values were then plotted against fraction number and were fitted with a polynomial trendline to account for small variations. Fractions were stored at -20°C until purification.

(5) Fraction purification

Fraction volumes were recorded as they were transferred to microcentrifuge tubes. Several fraction purification methods were tested: ethanol precipitation, dialysis with microcon centrifugation filter units (Millipore), and column purification (Zymo/ Princeton Separations) before selecting precipitation with polyethylene glycol (PEG) (Neufeld, Vohra et al. 2007). 40 - 60 µl of DNA suspension buffer with 5 µg of the co-precipitant linear acrylamide (Ambion) was added to each fraction followed with 300 µl (2 volumes) of PEG precipitation solution. Following overnight incubation at room temperature, samples were centrifuged at 16,000 x g at room temperature for 30 minutes. Samples were inverted on tissue and 500 µl of 75% EtOH was added to samples and centrifuged at 16,000 x g for 30 minutes three times. After the second wash, the cap and lip of the microcentrifuge tubes were wiped with a tissue with 70% EtOH and/or compressed air to remove dried salts. Samples were air dried for 5 – 10 minutes,

and suspended in 40 to 60 µl of 65°C DNA suspension buffer. Samples were then stored at -20°C until analyses. DNA concentrations in the fractions were measured using the SYBR green method.

(6) SIP gradient analyses

Density gradient fractions were selected for analysis based on the DNA concentrations and BD measurements. Fractions of the ^{14}N gradients were analyzed corresponding to the ^{15}N fractions selected regardless of DNA concentrations. Samples were thawed the day as needed at 50°C and *xplA* and 16S rRNA gene qPCR analyses were performed. For TRFLP amplification, fractions were pooled after initial amplification based on the number of 16S rRNA gene copies detected in order to stop amplification when reactions were in exponential phase as best as possible. Following the separation of digested products TRFLP and qPCR-TRFLP analyses were performed and used for comparative analysis between labeled (^{15}N) and unlabeled samples.

2. Data analyses

The following computational methods were performed throughout this research.

(1) Sequence analysis

Sequence data was downloaded and imported into Sequencher software (v4.6 or v4.9; Gene Codes Corp.; Ann Arbor, MI) for initial trimming and contig formations. BLAST (Basic Local Alignment Search Tool; (Altschul, Gish et al. 1990)) was used for initial sequence identification and preliminary phylogenetic characterization.

(2) Phylogenetic tree construction

After primers sequences were trimmed, sequenced 16S rRNA genes were imported into the ARB software package and added to the current SSU Reference database (Small Subunit rRNA Database) (Ludwig, Strunk et al. 2004) downloaded from the SILVA website (http://www.arb-silva.de/documentation/background/release-108/), which is composed of 16S rRNA sequences that have been trimmed of lower quality sequences, and sequences shorter than 1,200 bps.

(3) Fosmid annotation

The fosmid sequence was submitted to the JCVI Annotation Service for automated annotation using the JCVI prokaryotic annotation pipeline. This service includes gene finding using Glimmer, Blast-extend-repraze (BER) searches, HMM (hidden Markov model) searches, TMHMM (transmembrane hidden Markov model) searches, SignalP predictions and AutoAnnotate. All of this information was stored in a MySQL database and associated files which were downloaded for review and manual annotation using the Manatee manual annotation tool downloaded from SourceForge (manatee.sourceforge.net). Gene predictions were verified using GeneMark.hmm for Prokaryotes (v2.4) using *Mycobacterium avium paratuberculosis* as a model organism (Lukashin and Borodovsky 1998). Coding sequence start sites were subsequently changed as needed. Inverted repeats were queried using the *Palindrome* software (Institut Pasteur and Ressource Parisienne en Bioinformatique Structurale) distributed by Pasteur's Mobyle portal.

(4) TRFLP analysis

Initial analysis of TRFLP profiles was performed using DAx Data Acquisition and Analysis, v7.0 (Van Mierlo Software Consultancy; Eindhoven, The Netherlands). Peaks below an empirically set relative peak area threshold (0.5 – 1.0%, defined as the area of a given peak divided by the total peak areas of that chromatogram) were removed from calculations in the initial analysis. The minimum thresholds were decreased in subsequent analyses to as low as 0.1% after initial inspection of the gradient RF profiles. Spurious bp marker peaks were removed and bp sizes were entered for the ROX labeled markers under 'annotation'. Sample bps for each peak were then calculated by the software for the chromatogram using the markers for calibration.

(5) qPCR – TRFLP analysis of fractionated gradients

qPCR-TRFLP data from fractionated gradients (15 – 25 fractions, 2 – 3 chromatograms/ fraction) were modified from the DAx software output using Micosoft® Excel spreadsheets programmed with a series of Macros written in Visual Basic. Computing code and screen images of the different worksheets illustrating the workflow of processing the qPCR-TRFLP data can be found in Appendix B. Three spreadsheets containing four scripts were used to organize, process and analyze the data. BD data, qPCR data (copies/ μl) and DAx output (Peak #, Peak Height/Voltage (V), Peak Area (V * min), Relative Peak Area (%), and Base Pairs) were entered into the first spreadsheet. The script calculated relative peak heights (peak height divided by total peak heights) and then determined the number of unique bps among all of the fractions and calculated the peak copy numbers by multiplying the qPCR copy numbers by both relative peak areas and heights. The output of the script contained the copy numbers for each restriction fragment, RF (based on area and height), in each qPCR organized by BD with replicate samples color coded. The data was then manually binned by comparing replicate samples and by comparing RFs across the gradient.

The binned peak copy numbers were transferred to the second Excel program with script that calculated and outputted average values, standard deviations and the number of samples for each peak at each BD value. The third program scripts 1) used the output of the second script to report (ascending by RF bps and descending by maximum copy number) the three highest copy number values for each peak, the corresponding buoyant densities for each value and calculated p-values between each of the three copy numbers, 2) used linear interpolation to calculate copy numbers at BD values at 0.0015 g/ml intervals, and 3) used these values to generate two heat maps. The first heat map colored the cells based on percentiles calculated at each position (bp vs BD) with respect to all values in the gradient. The second heat map colored the cells based on percentiles calculated at each position with respect to all values for each bp. The heat maps are for visual comparisons of RF copy number profiles between gradients. An example heat map is shown below in Figure 39.

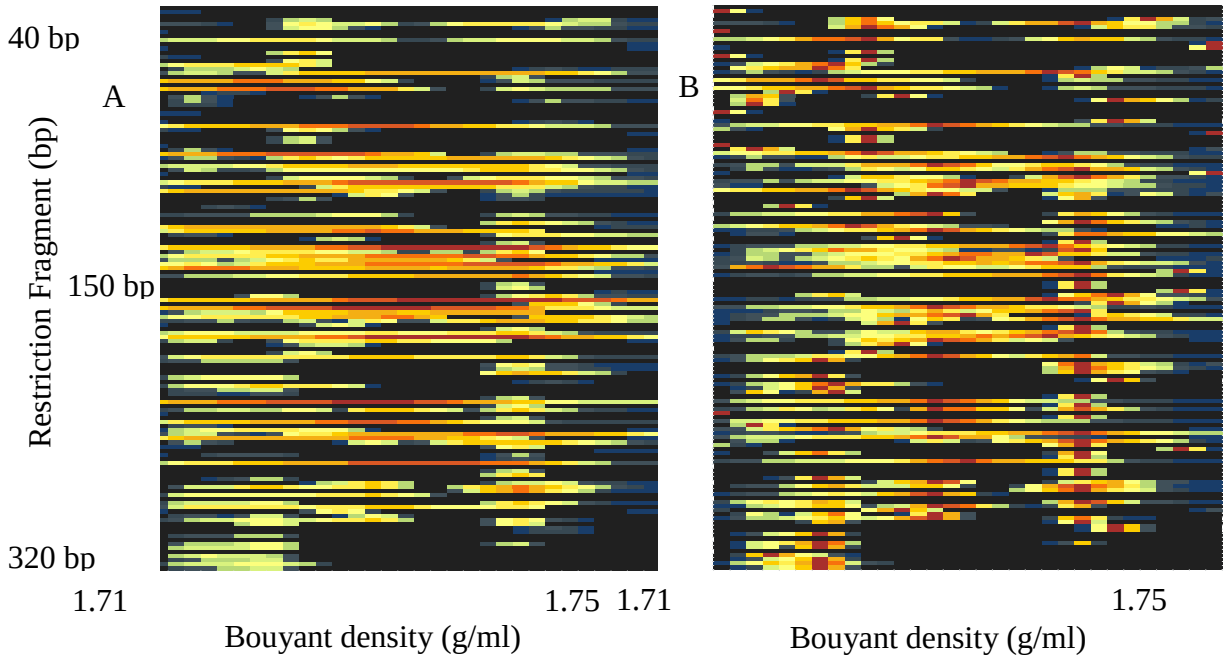


Figure 39. Example of heat maps generated from qPCR-TRFLP analysis scripts.
Two heat maps are created for each gradient based on the RF copy numbers (orange and red represent high copy numbers, blue are low copy numbers and black are not detected). The left map (A) creates heat maps based on percentiles calculated for each copy number based on the total number of copy numbers on the gradient so abundant RFs can be identified by quick visual inspection. The percentiles calculated in the right map (B) are among each RF so BD values corresponding to the highest copy numbers can be identified visually.

(6) Statistical analyses of microcosms

Statistical analyses were performed using SigmaPlot (Systat Software, Inc.; Chicago, IL). Where possible, ANOVA (analysis of variance) with the Shapiro-Wilk normality test was performed. The Holm Sidak method was used for multiple pairwise comparisons between groups in one way ANOVAs. When an ANOVA was not valid, either Kruskal-Wallis one way analysis of variance on ranks (Breslow 1970) or Mann-Whitney rank sum test was used for comparisons.

D. The lateral transfer of genes for RDX degradation. (Stuart Strand, David Stahl, Peter Andeer, University of Washington, Seattle)

1. Introduction

Past practices of production, application, and disposal of RDX have resulted in widespread contamination. Environmental contamination is aggravated by its high mobility, contributing to more widespread contamination of groundwater than by other commonly used explosives (Pennington and Brannon 2002). Ingestion or inhalation of RDX is associated with neurological disorders and organ failure (U.S. Environmental Protection Agency 2005), and exposed wildlife show behavioral changes and suffer liver and reproductive damage (U.S. Environmental Protection Agency 2005). The U.S. Environmental Protection Agency (EPA) has classified RDX as a possible human carcinogen (U.S. Environmental Protection Agency

2006). These adverse effects have provided motivation to better understand the microbiology and biochemistry of RDX degradation.

As yet there is relatively limited information concerning natural rates or mechanisms of microbial RDX degradation that are needed to predict or control rates of degradation in the environment. Of the three general pathways for RDX degradation or transformation based on metabolite analysis outlined in the review by Crocker and associates (Crocker, Indest et al. 2006), aerobic degradation initiated by XplA is among the better-characterized systems. This enzyme, a novel cytochrome P450, with a fused flavodoxin reductive domain (Seth-Smith, Rosser et al. 2002, Jackson, Rylott et al. 2007), was first identified by Seth-Smith et al. in *Rhodococcus rhodochrous* 11Y to be encoded by *xplA* (Seth-Smith, Rosser et al. 2002). This gene has been identified in 24 bacterial isolates of the Corynebacterineae capable of utilizing RDX as a sole nitrogen source (Coleman, Nelson et al. 1998, Seth-Smith, Rosser et al. 2002, Thompson, Crocker et al. 2005, Nejidat, Kafka et al. 2008, Seth-Smith, Edwards et al. 2008). While mammalian nitric oxide synthase family enzymes are known to be P450-like enzymes with fused flavodoxin domains, there are very few identified examples of this type of protein fusion among characterized microbial species (Munro, Lindsay et al. 1996, Cao, Bulow et al. 2000, Hunter, Roberts et al. 2005, Jackson, Rylott et al. 2007). Subsequent studies by Jackson and associates (Jackson, Rylott et al. 2007) demonstrated that XplA, in association with an electron transferring flavodoxin reductase (XplB), functions to efficiently denitrate RDX aerobically to the aliphatic 0) (Jackson, Rylott et al. 2007). NDAB has been shown to serve as a viable nitrogen source for *Methylobacterium* sp. strain JS178 (Fournier, Trott et al. 2005) and degraded by *Phanerochaete chrysosporium* (Fournier, Halasz et al. 2004). Thus, complete mineralization often appears to be mediated by multiple microbial populations.

The capacity for microbial degradation of recalcitrant organics, many of which are apparently new to the biosphere as a result of chemical manufacture, is often determined by plasmids and associated mobile genetic elements (van der Meer, de Vos et al. 1992, Trefault, de la Iglesia et al. 2004). Plasmids serve both as a reservoir of genetic information and to promote metabolic innovation, since their replication is independent of the chromosome and they do not generally encode essential functions. Although it was earlier suggested that genes in *Rhodococcus sp.* Strain DN22 associated with initial steps of RDX degradation are plasmid encoded (Coleman, Spain et al. 2002), no direct evidence for an extrachromosomal location was provided. We now show that near-identical genes for XplA and XplB are encoded on plasmids in two phylogentically and geographically distinct bacterial isolates - *Microbacterium sp*. MA1 isolated from North America (Milan, Tennessee, USA) and *Rhodococcus rhodochrous* 11Y isolated from England (United Kingdom) (Seth-Smith, Rosser et al. 2002). Thus, these genes are more broadly distributed within the Actinomycetales than previously recognized and the near-identity of gene sequence (6710 of 6721 bp) in these divergent genera is indicative of recent plasmid-mediated transfer. Analysis of approximately 52 kbp of sequence near *xplA* and *xplB* in strain MA1 revealed closely linked genes for transport and degradation that are flanked by transposable elements, suggesting that plasmid encoded *xplA/xplB* are part of a larger class I transposable element encoding for both transport and degradation of RDX.

2. Methods

a) Enrichment and isolation.

Medium described by Binks et al. (Binks, Nicklin et al. 1995), with RDX (Accustandard, New Haven, CT) as a sole nitrogen source (110 μM - 250 μM of RDX), was inoculated with soil

suspensions from RDX contaminated soil from the Milan Army Ammunition Plant. Excavated soil was added to a 0.1% sodium pyrophosphate solution (1:10 w/v) (Hurst and Knudsen 1997), suspended by vortexing briefly, and shaken for at least 1 hour at ≥200 rpm (28°C) before adding to the growth medium (1:100 v/v). RDX degradation was monitored using HPLC and RDX-degrading bacteria were recovered by repeated colony isolation on 1.5% agar plates containing either the enrichment medium or the complex media R2A (Reasoner and Geldreich 1985). RDX degradation of individual colonies was confirmed by clearing of RDX overlay plates (Seth-Smith et al. (Seth-Smith, Rosser et al. 2002)) and by monitoring RDX loss from broth cultures.

b) *HPLC quantification of RDX.*

RDX concentrations in cultures were analyzed using a modular Waters HPLC system consisting of a Waters 717+ autosampler, two Waters 515 HPLC pumps and a Waters 9926 photodiode array detector. A 4.6 x 250 mm, Waters C18 column was used for separation using run conditions similar to those outlined previously (Seth-Smith, Rosser et al. 2002) with concentration determined based on absorbance at 240 nm. Peak integration and analysis was conducted using the Millennium[32] software (Waters, Milford, MA).

c) *Growth of* Microbacterium sp. *MA1 using RDX as a sole nitrogen source.*

Growth studies using RDX (approximately 190 μM) as a sole nitrogen source were conducted in triplicate along with a control flask that was not inoculated under conditions described previously (Seth-Smith, Rosser et al. 2002). Cultures were regularly sampled to monitor turbidity (600 nm) and RDX concentration. Samples taken for RDX determination (800 μl) were processed by first removing cells by centrifugation (20,000 x g for 15 minutes in microcentrifuge) and amending 250 μl of supernatant with 10% w/v sodium azide to a final concentration of 0.1% w/v and stored at 4°C until analyzed with HPLC

d) *DNA extraction.*

For genomic DNA extractions, cultures were grown to late exponential phase before harvest. Cells were recovered by centrifugation (10 min at 10,000 x g) and resuspended to approximately 20 mg per ml of sucrose lysis solution (400 mM sucrose, 100 mM EDTA, 100 mM Tris pH 8.0, 1 mg/ ml lysozyme, 120 U/ml Mutanolysin). Following an overnight incubation at 37°C with gentle shaking (100 rpm), cells were lysed using an SDS-proteinase K lysis solution following established protocols (Gerhardt, Murray et al. 1994) followed with RNaseA (0.5μg/ml) incubation, phenol chloroform extraction and DNA precipitation using standard protocols (Sambrook and Russel 2001). DNA was suspended in Tris-EDTA buffer and concentration estimated by measuring A_{260} using a NanoDrop ND-1000 spectrophotometer (ThermoFisher Scientific, Wilmington, DE).

e) *PCR amplification cloning, and sequencing.*

Sequences for *xplA* were amplified using xplAF (5'-CCGACGTAACTGTCCTGTTCGGAA-3') and xplAR (5'-CGGGTCCGTCCGCCGGCTGGAAGG-3') as PCR primers as previously described (Rylott, Jackson et al. 2006). A region of sequence for the *R. rhodochrous* 11Y FAD/NADH binding domain protein was amplified using dapBF (5'-ATGACGAACATCAGAGCTGTCGT-3') and dapBR (5'-TTACAGTTCTTCGCGCACGATGTA-3') primers designed for this study. Well-characterized

primers for the bacterial 16S rRNA genes (27F and 1492R) were used to recover sequences for phylogenetic analysis (Lane 1991). Correct sized amplification products were ligated into the pCR4 vector (Invitrogen, Carlsbad, CA) and transformed using the TOPO-TA cloning kit (Invitrogen). Vector priming sites were used to determine 400 - 1100 bps of sequence from each end of an insert using two University of Washington sequencing services. Recombinant colonies were either submitted directly to "High-Throughput Sequencing Solutions" (www.htseq.org) or, alternatively, the BigDye v3.1 kit (Applied Biosystems, Foster City, CA) was first used to generate product from recombinant plasmid DNA for submission to the sequencing facility maintained by the Department of Biochemistry, University of Washington, Seattle, WA.

f) *Pulse field gel electrophoresis and Southern analysis.*

The Bio-Rad CHEF DRII system was used for pulse field gel electrophoresis (PFGE). Cultures of *Microbacterium* sp. MA1 and *Rhodococcus rhodochrous* 11Y were grown and harvested from late exponential phase growth LB broth. Cell plugs were molded according to manufacturer's instructions and embedded in 1% SeaKem Gold agarose dissolved in 0.5x TBE and run for 24 hours at 6 V/cm with a 10 to 100 second switch time ramp at a 120° angle with buffer recirculating at 14°C. The *S. cerevisiae* YNN295 and Lambda Ladder markers (Bio-rad, Hercules, CA) were used as size standards. SybrGreen was used to stain the gel for visualization.

DNA from the pulse-field gel was transferred to a Magnacharge membrane (Micron Separations Inc, Westborough, MA) by overnight capillary transfer using the alkaline transfer method (Sambrook and Russel 2001). PCR amplified DNA probe hybridization and detection was done using the Gene Images Alkphos Direct Labeling and Detection System kit (GE Healthcare, Piscataway, NJ) using the CDP-Star chemiluminescent detection reagent (GE Healthcare) by exposing it to Hyperfilm ECL (GE Healthcare).

g) *Fosmid library construction and sequence analysis.*

A fosmid library of the *Microbacterium sp.* MA1 DNA was constructed using the pCC1FOS vector from the CopyControl Fosmid Library Production Kit and Phage T1-Resistant EPI300-T1 *E. coli* Plating Strain (Epicentre Biotechnologies, Madison, WI) following the instructions provided. Approximately 400 fosmid clones were screened for the *xplA* gene by PCR amplification using the previously described xplAF/xplAR primer set. A subset of the positive clones were selected for shotgun sequence analysis using the TOPO-TA shotgun sequencing kit with pCR4 vector (Invitrogen) following manufacturer's instructions.
Vector priming sites were used for initial end-sequencing of the shotgun library (as previously described) and for subsequent sequencing of subclones. The Sequencher 4.6 software (Gene Codes Corp., Ann Arbor, MI) was used for initial assembly. Restriction mapping (NotI, KpnI, PvuII, MscI, BamHI, SacI, EcoRI, EcoRV, BsmI, MluI, HindIII, AscI and DraI) was then used to order contigs and to direct subcloning (data not shown) into the TOPO-pCR4 Zero Blunt vector (Invtirogen) and subsequent sequencing. The GenBank accession numbers for the MA1 16S rDNA sequence and partial plasmid (pMA1) sequence are FJ357539 and FJ577793 respectively.

The fosmid sequence was submitted to the JCVI Annotation Service for automated annotation using the JCVI prokaryotic annotation pipeline. This service includes gene finding using Glimmer, Blast-extend-repraze (BER) searches, HMM searches, TMHMM searches, SignalP

predictions and AutoAnnotate. All of this information was stored in a MySQL database and associated files which were downloaded for review and manual annotation using the Manatee manual annotation tool downloaded from SourceForge (manatee.sourceforge.net). Gene predictions were verified using GeneMark.hmm for Prokaryotes (v2.4) using *Mycobacterium avium paratuberculosis* as a model organism (Lukashin and Borodovsky 1998). Coding sequence start sites were subsequently changed as needed. Inverted repeats were queried using the *Palindrome* software (Institut Pasteur and Ressource Parisienne en Bioinformatique Structurale) distributed by Mobyle.

h) *Phylogenetic analysis.*

The ARB software package was used for 16S rRNA gene sequence alignment and tree construction (Ludwig, Strunk et al. 2004). 16S rRNA gene sequences for other RDX degrading bacteria were downloaded from the NCBI database and other bacteria used in the alignment and analysis were imported from the Silva database (Pruesse, Quast et al. 2007). The PHYLIP software package was used to determine bootstrap values using the Neighbor-Joining method using the Kimura 2-parameter model (Kimura 1980, Felsenstein 1989)**.**

3. Results

Microbacterium sp. MA1 (Figure 40) was isolated from contaminated soil from the Milan Army Ammunitions Plant (Milan, TN) based on its capacity to use RDX as a sole nitrogen source. Growth of MA1 was directly correlated with loss of RDX, with nearly complete degradation (190 - 195 μM initial RDX concentration) after 48 hours (Figure 41). PCR analysis of MA1 with primers for *xplA* produced the predicted 403 bp product.

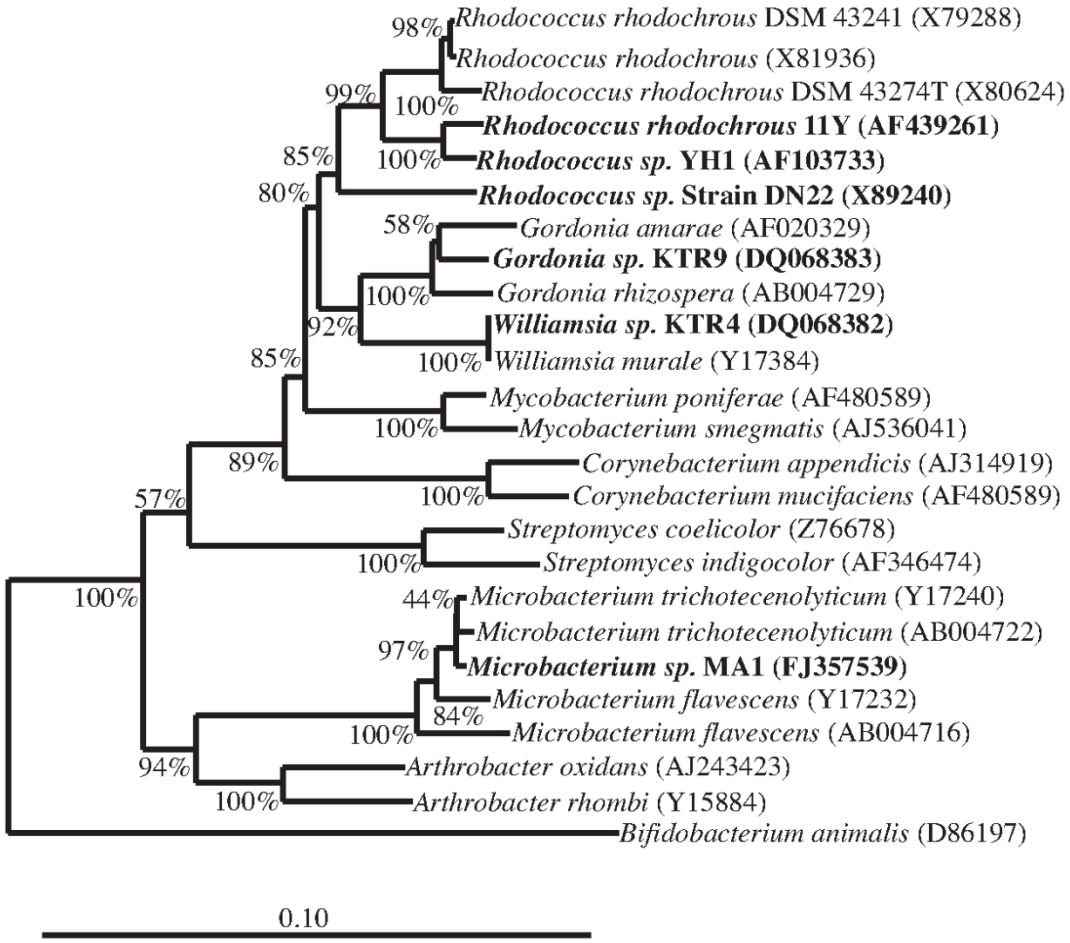


Figure 40 Phylogenetic tree of selected RDX-degrading bacteria inferred from 16S rRNA sequence relationships. Phylogenetic relationships of characterized RDX-degrading bacteria that carry xplA and Kimura two-parameter model (Kimura 1980, Felsenstein 1989). RDX degraders are shown in boldface. GenBank accession numbers are in parentheses.

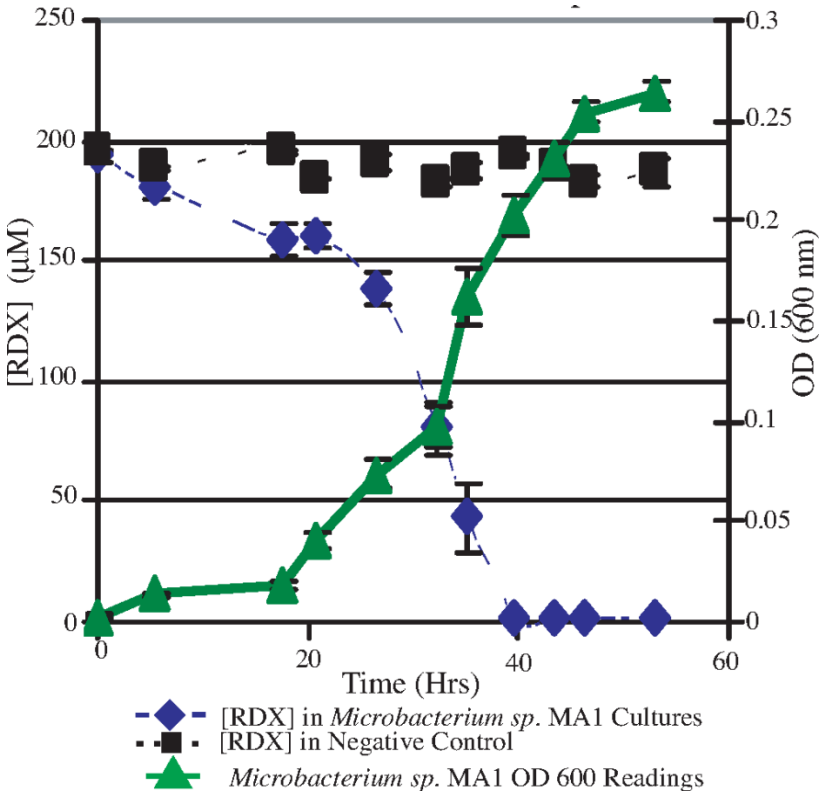


Figure 41 Growth of Microbacterium sp. MA1 on RDX.
Microbacterium sp. MA1 was grown in 80 ml cultures using RDX as a sole source of nitrogen. Aliquots were routinely collected for optical density measurements at 600 nm and RDX concentration determined by HPLC.

Initial characterization of *Microbacterium sp.* MA1 and *R. rhodochrous* 11Y DNA by pulse field electrophoresis revealed extrachromosomal elements (putative plasmids) in each, migrating near the 145.5 kb Lambda marker in MA1 (pMA1) and between the 225 and 245 kb markers in 11Y (p11Y) (Figure 42A). Both species contained nearly identical *xplA* sequences that were shown to be localized to the extrachromosomal element by hybridization with a 403 bp *xplA*-specific gene probe (Figure 42B). This is the first description of *xplA* outside the Corynebacterineae (*Rhodococcus*, *Gordonia* and *Williamsia*) (Coleman, Nelson et al. 1998, Seth-Smith, Rosser et al. 2002, Thompson, Crocker et al. 2005). The near identity of *xplA* sequences in *Microbacterium sp.* MA1 and *Rhodococcus rhodochrous* 11Y, despite their very different phylogenetic affiliations (Figure 40), is most consistent with recent lateral transfer of *xplA*.

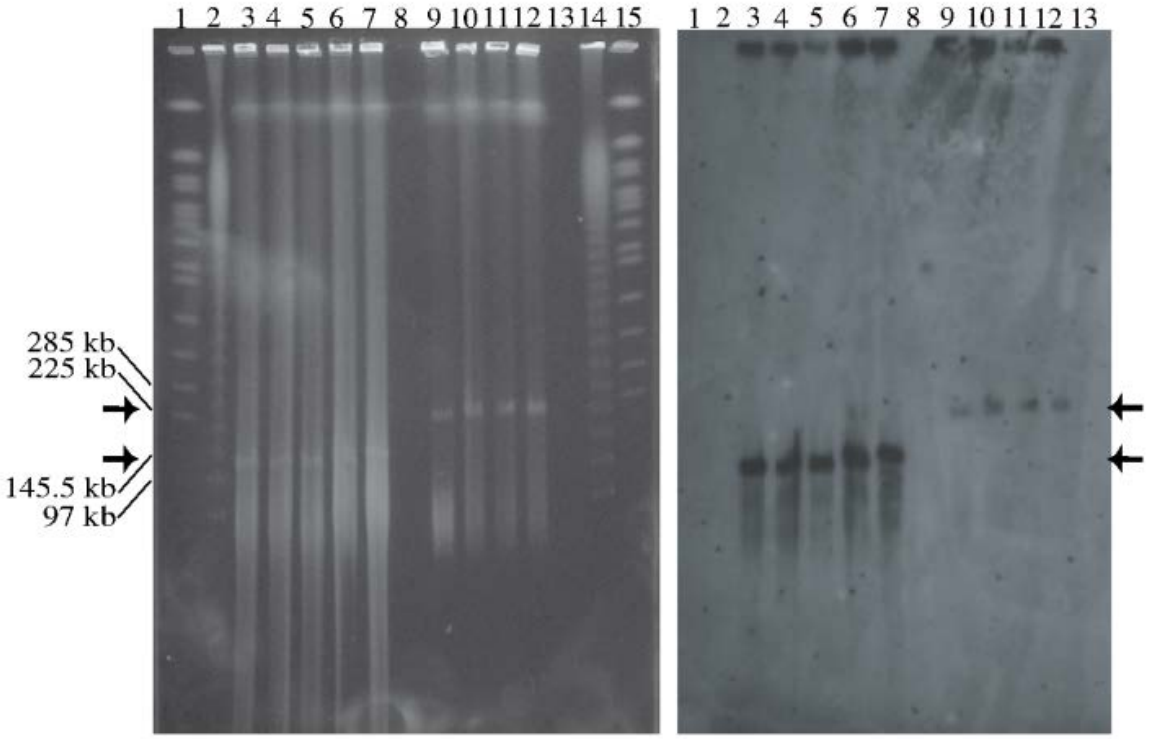


Figure 42 Hybridization of *xplA* gene probe to *Microbacterium* sp. strain MA1 and *Rhodococcus rhodochrous* 11Y DNA resolved by PFGE.
(A) SYBR green I-stained gel. Lanes 3 to 7, *Microbacterium* sp. strain MA1; lanes 9 to 12, *R. rhodochrous* 11Y; lanes 1 and 15, *S. cerevisiae* YNN295 marker; lanes 2 and 14, Lambda ladder. (B) Hybridization with a 403-bp fragment of the *xplA* gene.

Sequence analysis of approximately 52 kbp of DNA flanking the *xplA* gene, encoding a cytochrome P450 previously shown to be required for RDX degradation (Seth-Smith, Rosser et al. 2002, Jackson, Rylott et al. 2007), appears to be part of a larger metabolic module (underlined in Figure Figure 43A; pMA1.029 – pMA1.034, Figure 42) that shares high similarity with the 7.5 kbp of sequence available for the region near *xplA* in *R. rhodochrous* 11Y (Seth-Smith, Rosser et al. 2002).

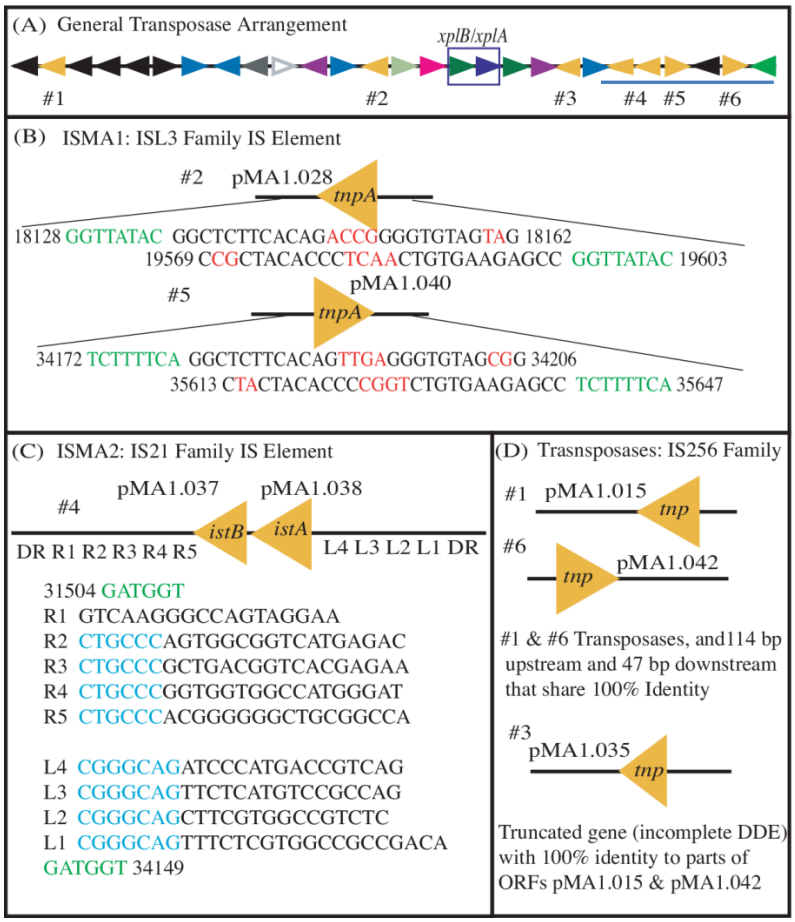


Figure 43 Distribution of transposases and IS elements in pMA1.
(A) The six ORFs associated with transposition in the 52 kbp sequence of pMA1 are shown in relation to *xplB/xplA*. The apparent metabolic module that xplB/xplA belongs to is underlined in green. (B) Two identical ISL3 family elements (ISMA1), each encoding a single transposase (ORFs pMA1.028, pMA1.040). Imperfect indirect repeat and direct repeat sequences characteristic of ISL3 elements are shown (Mahillon and Chandler 1998). (C) IS21 family element (ISMA2) encoding an ATP binding domain protein (pMA1.037) and an integrase (pMA1.038). Direct and indirect repeat sequences are displayed below. Repeat sequences found throughout the indirect repeats highlighted in blue. (D) Three IS256 family elements shown. pMA1.015 and pMA1.042 and 161 bps of sequence flanking each share 100% identity encoding a transposase. pMA1.035 is a truncated gene with incomplete DDE motif, but shares 100% nt identity with portions of pMA1.015 and pMA1.042.

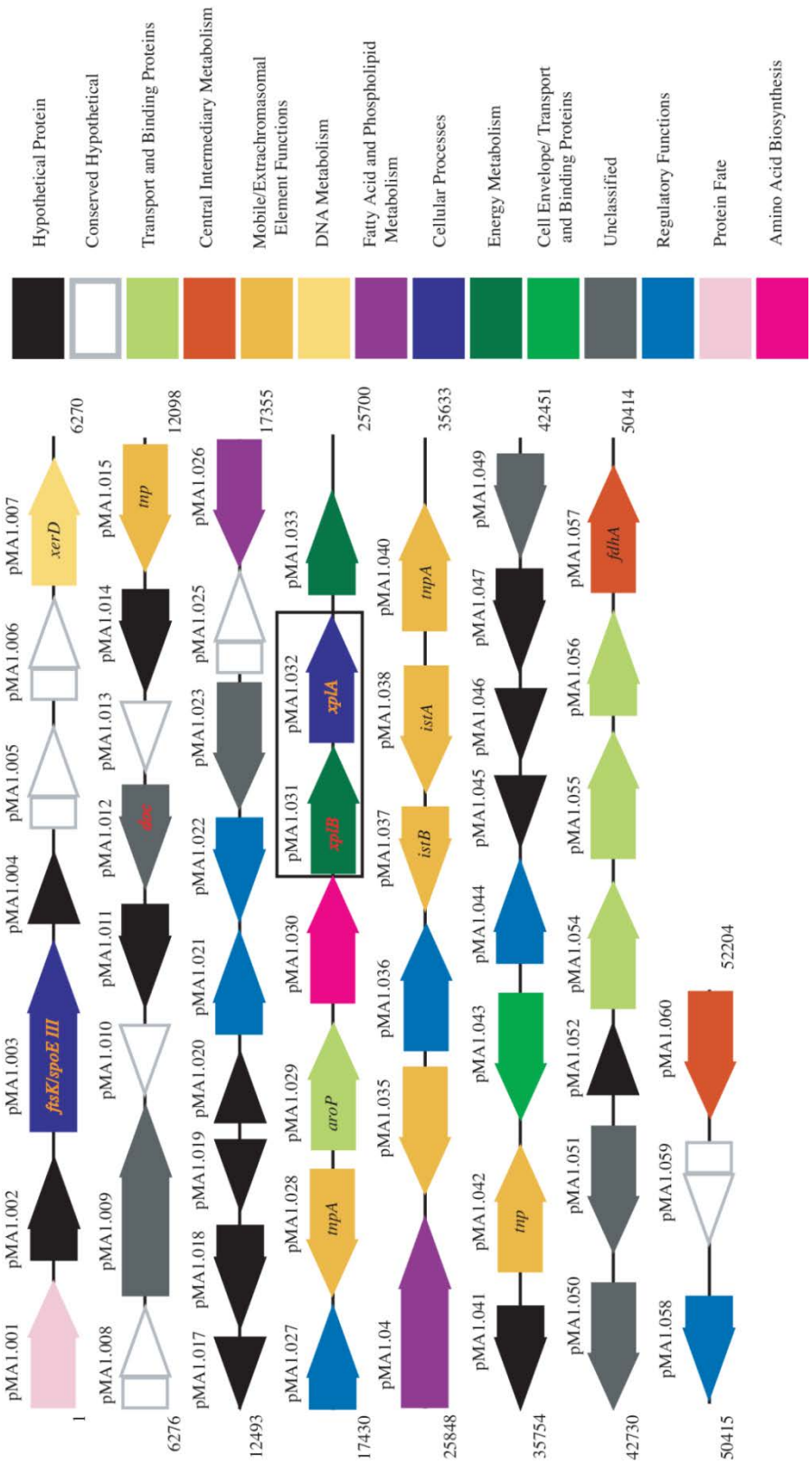


Figure 44. Functional annotation of 52.2 kbp of sequence from pMA1. ORFs of the two overlapping fosmid clones were numbered and assigned to the functional categories (Riley 1993) by the JCVI (J. Craig Venter Institute) automated annotation service (www.jcvi.org). Selected genes have been labeled and the *xplB* and *xplA* genes have been boxed to show relative positions.

A coding region (pMA1.057) annotated as a glutathione-independent formaldehyde dehydrogenase (*fdhA*) (Ito, Takahashi et al. 1994) found downstream from this region (Figure 3.5) may function in metabolism of formaldehyde, a previously identified product of aerobic

RDX metabolism (Fournier, Halasz et al. 2002, Seth-Smith, Rosser et al. 2002, Thompson, Crocker et al. 2005, Jackson, Rylott et al. 2007). Two closely linked coding regions, Ftsk/SpoIIIE (pMA1.003) and an integrase/ recombinase (pMA1.007), are associated with dimer resolution (Errington, Bath et al. 2001) and a FtsK/SpoIIIE homolog (TcpA) has shown to be essential for transfer of the conjugative plasmid pCW3 in *Clostridium perfringens* (Parsons, Bannam et al. 2007), are consistent with localization of *xplA* to a plasmid.

The genes associated with RDX degradation also appear to be associated with mobile elements. At least six transposases, encoded by three different types of insertion sequence (IS) elements, are present within the 52 kbp sequence (Figure 43A). Two identical copies (pMA1.028 and pMA1.040) of an ORF encoding a transposase related to TnpA is the only gene encoded by an ISL3 family IS element (designated ISMA1 – Figure 43B) (Cirillo, Barletta et al. 1991, Mahillon and Chandler 1998). An IS21 family-type IS element (designated ISMA2) carries an ATP binding domain protein (pMA1.037) and an integrase (pMA1.038 - Figure 3.4c) (Mahillon and Chandler 1998). The remaining three elements (pMA1.015, pMA1.035, and pMA1.042) are related to the IS256 family of transposable elements (Figure 43 D) (Mahillon and Chandler 1998), two of which (pMA1.015 and pMA1.042) share complete nucleotide identity including 114 bp upstream and 47 bp downstream of each. pMA1.035 is a truncated transposase that is not likely to be active because its DDE sequence motif, a highly conserved acidic amino acid triad found in the catalytic sites of many transposases including pMA1.015 and pMA1.042 (Figure 43D) (Mahillon and Chandler 1998), is incomplete.

4. Discussion

These data have established that genes required for RDX degradation are plasmid encoded and likely part of a class I transposable element as suggested by the presence of several flanking pairs of IS elements. Transposition between plasmids has likely promoted transfer of the capacity for RDX degradation among diverse species, as is now supported by the observation of near identical sequences in two suborders of Actinomycetales. While the flavodoxin domain of XplA has homology (>35% amino acid identity) to several amino acid sequences deposited in GenBank, the P450 domain of XplA protein has significant relationship with only one other deposited sequence (Seth-Smith, Edwards et al. 2008). The near identity of sequences for the *xplA* gene and flanking sequences from plasmids from phylogenetically distant members of the Actinomycetales, *R. rhodochrous* 11Y and *Microbacterium sp.* MA1, provide compelling evidence for recent lateral transfer. A contribution of functions encoded by plasmids and associated mobile elements to the degradation of xenobiotics is now well established (Top, Springael et al. 2002). For example, as for genes (atzA, *atzB*, and *atzC*) encoding enzymes that transform the herbicide atrazine, a xenobiotic with a triazine backbone, to cyanuric acid (de Souza, Seffernick et al. 1998). However, the discovery of nearly identical gene clusters on plasmids carried by phylogenetically divergent microorganisms, independently isolated from different continents, indicates a remarkably rapid dissemination of this novel catabolic activity – possibly within the 70 year period since first environmental contamination.

Our analysis of a 52 kbp region of the *Microbacterium* plasmid sequence also suggests that *xplA* and *xplB* may be part of a larger gene cluster (pMA1.029 – pMA1.034) associated with RDX degradation. In addition to *xplA* and *xplB*, the gene cluster includes a gene highly similar to an *E. coli* general aromatic amino acid permease (*aroP*), an FAD/NAD(P) binding domain protein, an aldehyde dehydrogenase domain protein, and an acetyl-CoA synthetase homolog. The proximity of the *aroP* gene to *xplA*, suggests a potential role in the cellular uptake of RDX.

The remaining genes in the cluster are less likely to be directly involved in RDX degradation as *in vitro* experiments have shown XplB and XplA are capable of breaking down RDX (Jackson, Rylott et al. 2007). However, a formaldehyde dehydrogenase (pMA1.057) located on pMA1 outside the described gene cluster could aid the cell through removal of the toxin formaldehyde, an identified degradation product along with nitrite and NDAB in the *xplA*-bearing isolates: *R. rhodochrous* 11Y, *Rhodococcus sp.* Strain DN22, *Williamsia sp.* KTR4, *Gordonia sp*. KTR9, as well as *in vitro* experiments using XplB and XplA (Fournier, Halasz et al. 2002, Seth-Smith, Rosser et al. 2002, Fournier, Halasz et al. 2004, Thompson, Crocker et al. 2005, Jackson, Rylott et al. 2007).

The *xplA* gene has been found in every bacterial isolate examined for the gene that aerobically uses RDX as a nitrogen source (Seth-Smith, Rosser et al. 2002, Thompson, Crocker et al. 2005, Jackson, Rylott et al. 2007, Nejidat, Kafka et al. 2008, Seth-Smith, Edwards et al. 2008), and has been recovered from RDX contaminated soils (Andeer et al., unpublished observations) suggesting that this gene should provide a useful monitoring tool in applications of bioremediation. Recognition that XlpA is plasmid encoded and likely part of a larger metabolic module carried on a transposable element could provide a foundation for better process control, for example, by promoting environmental conditions that foster its transfer among resident microbial populations. The presence of several IS elements in the vicinity of the *xplA* gene cluster also suggests that these genes could be readily integrated into different broad range plasmids for selective transfer to disparate microbial species (Davison 1999).

Among mechanisms for lateral gene transfer, conjugative transfer of plasmids and phage mediated transfer are important. Conjugative transfer of the xplA/B complex was shown by Jung et al (Jung, Crocker et al. 2011), but only after construction of plasmids with homologous flanking regions and antibiotic resistance. Transfer of the xplA/B region by phage infection has not been demonstrated, thus the mechanism of its horizontal transmission in natural environments remains speculative. Several bacteria, including gram negatives, and several plants have been transformed with xplA/B and shown to be functional, although modification of codons was necessary in some cases to promote efficient expression. The xplA gene alone provides for weak activity, which is increased more than ten fold by the presence of xplB.

E. High sensitivity stable isotope probing by quantitative TRFLP. (Stuart Strand, David Stahl, Peter Andeer, University of Washington, Seattle)

1. Introduction

Since its introduction, stable isotope probing (SIP) has served as a powerful technique for linking microbial function to community structure by identifying populations that metabolize selected substrates (Radajewski, Ineson et al. 2000, Lueders, Pommerenke et al. 2004, Buckley, Huangyutitham et al. 2007, Kalyuzhnaya, Lapidus et al. 2008). In SIP, the system of interest is first challenged with a substrate labeled with a stable isotope, usually a heavier isotope of carbon (^{13}C) or nitrogen (^{15}N). Then, DNA or RNA from organisms that have incorporated the isotope are separated from unlabeled nucleic acid by cesium chloride (CsCl) or cesium trifluoroacetate (CsTFA) density gradient centrifugation (Manefield, Whiteley et al. 2002, Neufeld, Vohra et al. 2007), respectively. Populations that have incorporated label are identified by comparative analysis of gradient fractions containing heavy and light DNA, generally by selective amplification and sequencing of specific genes (Radajewski, Webster et al. 2002), fingerprinting techniques such as terminal restriction fragment length polymorphism (TRFLP) (Lueders, Pommerenke et al. 2004, Buckley, Huangyutitham et al. 2007, Schwartz

2007) or denaturing gradient gel electrophoresis (DGGE) (Manefield, Whiteley et al. 2002, Griffiths, Whiteley et al. 2003, Rangel-Castro, Killham et al. 2005), or metagenomic analysis (Schwarz, Waschkowitz et al. 2006, Kalyuzhnaya, Lapidus et al. 2008, Neufeld, Chen et al. 2008).

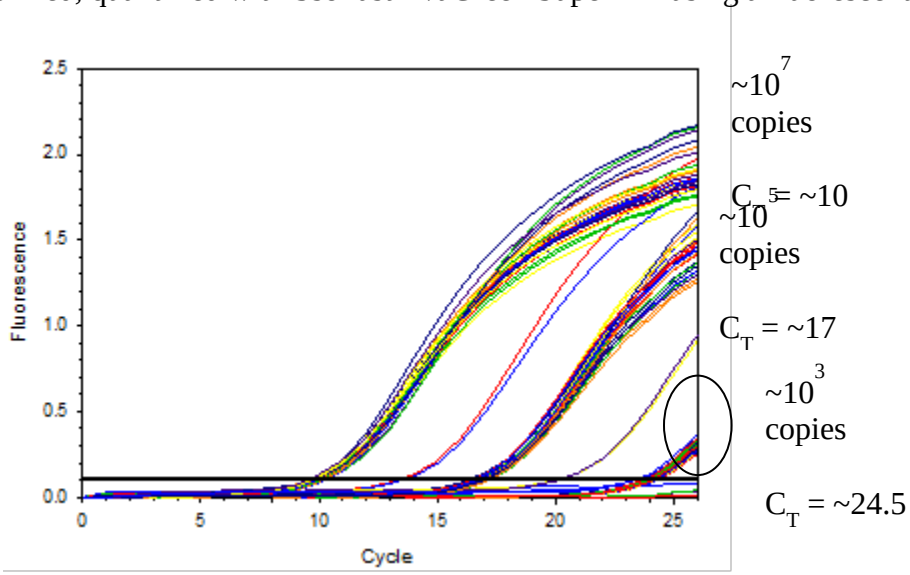
The most statistically robust conclusions are derived from simple presence or absence analysis, for example as determined by diagnostic TRFLP fragments (Culman, Gauch et al. 2008). However, presence-absence assessment is dependent upon complete separation of heavy and light DNA. For this reason, with some notable exceptions (Buckley, Huangyutitham et al. 2007, Buckley, Huangyutitham et al. 2007, Schwartz 2007, Wawrik, Callaghan et al. 2009), most analyses have focused on substrates labeled with ^{13}C-carbon, since the smaller increases in buoyant density (BD) from ^{15}N incorporation provides lower resolution of labeled and unlabeled nucleic acids (Cupples, Shaffer et al. 2007, Addison, McDonald et al.). In particular, small changes in density may not be sufficient to resolve labeled DNA from unlabeled DNA of high G+C content (Cadisch, Espana et al. 2005, Buckley, Huangyutitham et al. 2007, Cupples, Shaffer et al. 2007, Addison, McDonald et al.). Although resolution can be improved by using AT selective intercalating dyes which exaggerate G+C bias (Karlovsky and Decock 1991, Buckley, Huangyutitham et al. 2007, Buckley, Huangyutitham et al. 2007), because the distribution of nucleic acids is Gaussian (Osterman 1984), nucleic acids derived from an abundant population are often distributed throughout a gradient. Thus, TRFLP analysis alone may not be sufficient for accurate determination of peak position within a gradient or for comparisons of gradients containing variable amounts of DNA.

To address this limitation, we developed a general protocol for accurate determination of peak position by combining a fluorophore-labeled primer with an intercalating dye to quantify individual TRFLP restriction fragments (RFs) in individual fractions of the gradient. The intercalating dye is used for real time monitoring of amplification so reactions can be terminated before amplification efficiency is compromised by reaction component limitations (von Wintzingerode, Gobel et al. 1997). The fluorophore-label is subsequently employed to quantify individual RFs (Figure 45). Importantly, the method is not constrained by the requirement for an additional internal hybridization site, as required for the Taqman protocol (Yu, Ahuja et al. 2005), providing greater flexibility in the design of primers used to assess sequence diversity.

Tandem qPCR-TRFLP analysis of CsCl gradients

Step 1. DNA extraction, SIP separation, fractionation and purification

Step 2. Fractions amplified, quantified with SsoFast EvaGreen Supermix using a fluorescently-labeled



qPCR reactions terminated in exponential phase to minimize # of cycles and associated bias

Step 3. qPCR product purified, digested (MspI and MnlI), and separated on an ABI 3730 DNA Analyzer

Step 4. Individual RF abundance calculated from qPCR of total 16S rRNA gene copy numbers.
RF copy number = total copy numbers * RF fraction

Step 5. Copy numbers of each restriction fragment plotted against buoyant densities in each gradient.
^{15}N-SIP profiles are compared to unlabeled controls to identify assimilating populations.

Figure 45. Tandem qPCR-TRFLP analysis of CsCl gradients.
Flow diagram showing the steps used to assign quantities for individual RFs (Steps 2 – 4) and how these calculations are used in SIP analysis.

2. Materials and methods

a) Soil microcosms.

Soil slurries containing 13% (w/v) soil were incubated for 3.5 days with an enrichment media for RDX (1,3,5-trinitro-1,3,5-triazine) degraders. RDX degradation was monitored using the previously described HPLC method.

b) DNA extraction from soil microcosms.

DNA was extracted from soil slurries using a modification of a previously published protocol by incorporating a pre-incubation step with aluminum sulfate to remove humic acids (Dong, Yan et al. 2006). Approximately 300 mg of soil was loaded into a lysing matrix E tube (MP Biomedicals; Solon, OH) with 130 µl of 100 mM $AlSO_4$ solution (pH 3), 200 µl of 100 mM sodium phosphate buffer (pH 6.6), 60 µl of 1 M NaOH (to pH ~9.0), 560 µl of extraction buffer (55 mM $NaPO_4$, 225 mM Tris, 100 mM NaCl, pH ~8.5) and 160 µl of 20% SDS solution. The samples were disrupted in a FastPrep 1200 bead beater (MP Biomedicals) for 30s at a machine setting of 4.0, centrifuged for 10 min at 16,000x g, and the supernatant transferred to a clean microcentrifuge tube on ice. Concentrated HCl (2 µl) was then added to the soil pellet, and extraction repeated using a 200 mM $AlSO_4$ solution and all reagent volumes 50% of the initial extraction. After centrifugation, the supernatants were combined, incubated on ice or at -20° C for 20 min, and centrifuged for 10 min at 20,000 x g at 4°C to remove excess SDS. A 0.2 volume of 5 M $NaClO_4$ (pH 9) was added to the combined supernatants and the mixture incubated for 10 min at 55°C. Following two extractions with cold CIA (chloroform:isoamyl alcohol, 24:1), nucleic acids were recovered by the addition of 1 volume isopropanol and washed using standard protocols (Sambrook and Russel 2001).

c) *Construction of 16S rRNA gene standards.*

Near-complete 16S rRNA gene sequences from *Flavobacterium* sp., *Acidovorax* sp., *Arthrobacter* sp., and *Microbacterium* sp. were amplified using bacterial primers 27F and 1492R (Lane 1991) and ligated into the Invitrogen pCR4 vector (Life Technologies; Carlsbad, CA). Plasmids containing the cloned fragments were extracted using the Qiagen plasmid miniprep kits (Qiagen; Valencia, CA) following manufacturer's instructions, and quantified by measuring A_{260} using a NanoDrop ND-1000 spectrophotometer (ThermoFisher Scientific; Wilmington, DE). Purified plasmid DNA was diluted to qPCR stock solutions of 25 ng μl^{-1} and concentrations were verified with triplicate measurements. Mixes of the four species 16S rRNA clones were prepared at various gene copy ratios (from equimolar to 10:1:1:1).

d) *Preparation of pure culture genomic DNA.*

E. coli K12 MG1655, *Rhodococcus rhodochrous* 11Y, *Microbacterium* sp. MA1, and *Variovorax* sp. were grown on the previously described minimal media used for enrichment of RDX degrading organisms with either RDX, 99% [^{15}N]RDX (Defence Research and Development Canada; Valcartier, QC), ammonium nitrate or 98% [^{15}N]ammonium [^{15}N]nitrate (Sigma-Aldrich; St. Louis, MO) supplied as a nitrogen source. DNA for genomic DNA mixtures was prepared from cells harvested by centrifugation, washed, and DNA isolated following protocols described previously with the exception that sucrose was not used in the lysis solution. For the SIP experiment, DNA was prepared using the above soil extraction protocol excluding aluminum sulfate incubation.

e) *DNA sequencing.*

16S rRNA gene sequences originating from this work were generated using the bacterial primers 27F and 1492R (Lane 1991), ligated into the pCR4 vector (Invitrogen), transformed using the TOPO-TA cloning kit (Invitrogen) and recombinant colonies were submitted directly to "High-Throughput Sequencing Solutions" (www.htseq.org). Genbank accession numbers will be obtained for the sequences of all relevant clones and microbial species generated for this

study following review. The Genbank accession numbers for the partial 16S rRNA gene sequences of *E. coli* K-12 MG1655 (Blattner, Plunkett et al. 1997), *R. rhodochrous* 11Y (Seth-Smith, Rosser et al. 2002) and *Microbacterium* sp. MA1 are NC_000913, AF439261, and FJ357539 respectively.

f) *Isopycnic centrifugation and gradient fractionation.*

DNA (~4.9 μg) recovered from a soil microcosm using the described protocol was added to a mixture of unlabeled or ^{15}N-labeled genomic DNA (~620 to 670 ng) from *E. coli* K12 MG1655, *Rhodococcus rhodochrous* 11Y, and *Microbacterium* sp. MA1. The combined DNA was then added to a TE/CsCl solution (10mM Tris, 10mM EDTA pH 8.0, CsCl to BD (buoyant density) of 1.71 g ml^{-1}) with a volume of 4.8 ml in OptiSeal polyallomer tubes (Beckman Coulter; Brea, CA) and a final BD of ~1.67 g ml^{-1} based on density measurements. Gradients were established in a TLA110 rotor run for 96 hrs at 55,000 rpm in an OptimaMax Ultracentrifuge (Beckman Coulter) (Buckley, Huangyutitham et al. 2007). Gradients were then displaced with light mineral oil pumped into the top of the tube, collecting 90 to 105 μl fractions drop-wise from a point near the bottom of the tube. The BD of each fraction was determined using a modified AR200 digital refractometer (Reichart; Ithaca, NY) as described by Buckley *et al*. (Buckley, Huangyutitham et al. 2007). To account for possible measurement error, fraction number was plotted against buoyant density and fit with a linear curve for analysis. Fractions were combined, concentrated and dialyzed against DNA Suspension Buffer (10 mM Tris, 0.1 mM EDTA, pH 8.0) using 30 kDa Microcon membrane spin filters (Millipore; Billerica, MA) centrifuged at 6,000 x g. Eluates were all adjusted to 60 μl by addition of the same buffer.

g) *Quantitative PCR.*

qPCR (20 μl) consisted of 10 μl of SsoFast EvaGreen supermix (Biorad; Hercules, CA), 6.4 μl of PCR certified water (Teknova; Hollister, CA), 8 pmoles each of the 6'-carboxyfluorescein labeled, [6-FAM]-27F (Eurofins MWG Operon; Huntsville, AL) and unmodified 338R (Amann, Binder et al. 1990), and 2 μl of template DNA or water. PCR and amplification monitoring was run in duplicate for each dilution using a MJ-Research PTC-200 gradient thermocycler with a Chromo 4 real-time PCR detector with the Opticon Monitor 3.1 software (BioRad). Initial denaturation was performed at 98°C for 2 minutes, and each cycle was 98°C for 8s, 58°C for 12s and 72°C for 15s with a 5 minute final extension step added after cycling was finished. A dilution series from 250 pg μl^{-1} to 2.5 fg μl^{-1} of *Flavobacterium* sp. 16S rRNA gene clones (4.3 x 10^{7} to 4.3 x 10^{2} copies μl^{-1}) was used to generate a standard curve for quantification of 16S rRNA gene clones. For genomic DNA and spiked soil samples, *E coli* K12 MG1655 from 25 ng μl^{-1} to 250 fg μl^{-1} (3.44 x 10^{7} to 3.44 x 10^{2} copies μl^{-1}) was used for standard curve generation to avoid overestimation of the copy numbers that can occur when using plasmids for standards (Hou, Zhang et al. 2010). An initial amplification was performed to estimate concentration in each fraction. Fractions were then amplified in duplicate or triplicate in batches based on threshold cycle (C_T) values so amplification could be terminated in exponential phase. The Opticon software was used for analysis without curve smoothing and subtracting global minimum baselines. Threshold fluorescence levels were set to the lowest levels that minimize error in standard curves, typically at values between 0.1 and 0.15. Slopes and R^2 values of semi-log regression curves of the standards were routinely -3.3 +/- 0.3 and > 0.99, respectively.

h) *TRFLP sample preparation and processing.*

qPCR products were purified using the PSIclone HTS PCR 96 purification kit (Princeton Separations; Freehold, NJ) following manufacturer's instructions. DNA was digested with the 0.5 μl each of FastDigest MnlI and MspI enzymes (Fermentas; Glen Burnie, MD) for an hour, followed by a second addition of enzyme and incubation from 2 hours to overnight. Digests were purified by addition of phenol chloroform followed by centrifugation, and 20 μl of supernatant applied to the Centrisep 96 plates (Princeton Separations). Aliquots (4 μl) of properly diluted reactions were mixed with 12.5 μl of HiDi formamide and 0.2 μl of ROX-labeled custom mapmarker (30, 50, 100, 150, 200, 250, 300, 400, 500, 510 and 550 bps; Bioventures; Murfreesboro, TN). Samples were processed on an Applied Biosystems 3730 DNA Analyzer (Life Technologies).

i) *Data analysis.*

Analysis of TRFLP profiles was performed using DAx Data Acquisition and Analysis, v7.0 (Van Mierlo Software Consultancy; Eindhoven, The Netherlands). Peaks were binned manually and quantification of peaks was performed based on both relative area and relative height of the peaks when normalized by total profile area or total profile height, respectively.

3. Results

Since a linear relationship between fluorescence intensity and target sequence abundance could be achieved by early termination of the amplification reaction (Figure 45), the relative copy number of each RF should be equal to the relative fluorescence intensity. Several initial tests were performed to validate this equivalency. The effects of replacing a typical primer in the qPCR reaction with a fluorescein-labeled primer (Figure 45, Step 2) were first analyzed. Then, a series of experiments were conducted to develop and evaluate the method of calculating the copy numbers associated to individual RF values (Figure 45, Step 4). Finally, a SIP experiment was conducted to demonstrate improved resolution of density-shifted populations using this method versus the use of a TRFLP analysis alone.

The influence of the fluorescein-labeled primer on qPCR quantification was investigated by comparing standard curves developed for DNA from *E. coli* (3.44×10^7 to 3.44×10^3 16S rRNA gene copies μl^{-1}) and *Microbacterium* sp. MA1 (170 pg ml^{-1}, 17pg ml^{-1} and 1.7 pg ml^{-1}). Following amplification with either 27F or [6-FAM]-27F paired with 338R the threshold was set empirically to maximize R^2 for the standard curve, subtracting the global minimum baseline.

Table 7 shows a comparison of the standard curves (log [gene copy #] vs C_T) parameters, as well as the average calculated copy numbers, for the MA1 DNA. The major difference between the two datasets is the increased threshold used to calculate C_T values (0.154 vs. 0.016) for the FAM-labeled primer amplification due to the FAM background fluorescence. Nonetheless, the parameters and calculations generated at the respective threshold values with the fluorescent primer were very close to those obtained using the unlabeled primer.

Table 7. The effect of fluorescein labeled primer [6'-FAM]-27F on qPCR standard curve values.
Data was generated from *E. coli* and copy number calculations for *Microbacterium* sp. MA1 using either 27F or [6'-FAM]-27F with 338R[a].

	Threshold[b]	Slope[c]	Intercept[c]	R^{2c}	MA1- (Copy #)	MA1- (Copy #)	Coefficient (%)
27F	0.016	-3.51	34.0	1.00	6.48E+04	8.19E+03	12.64
[6'-FAM]-	0.154	-3.57	37.4	1.00	5.50E+04	7.42E+03	13.51

[a] *E. coli* genomic DNA was diluted from 10^7 to 10^3 copies of the 16S rRNA gene per µl and MA1 copy number statistics generated using 3 fold dilutions of the DNA within the standard curve's range. Amplifications were performed in duplicate under the same cycling conditions.

[b] Threshold value used for C_T calculation in respective quantifications

[c] Standard curve [log (gene copy #) vs C_T]

The effects of DNA concentration and degree of amplification on quantification were initially examined using mixtures of four pCR4 vectors containing *Flavobacterium* sp., *Arthrobacter* sp., *Microbacterium* sp. and *Acidovorax* sp. 16S rRNA gene sequences with restriction fragments of 44, 150, 156 and 280 bps respectively when amplified, digested and analyzed as described. Ten mixtures of the four sequences were prepared in ratios ranging from 1:1:1:1 to 10:1:1:1, diluted to concentrations between 250 pg μl^{-1} and 25 fg μl^{-1}, and amplified using either 16 or 26 temperature cycles. Gene ratios were determined from relative peak heights and relative peak areas for each fragment in the TRFLP profiles. Data from each of the amplification conditions (Table 4.2) were consolidated into linear regression plots of mixture composition versus detected composition (ratios of RFs in TRFLP profiles). The results are summarized in Table 8.

Table 8. Consistency between predetermined compositions of DNA mixtures and qPCR-TRFLP quantification of the compositions under varied amplication[a].

Template DNA[b]	16S rRNA gene copy #s	Avg. C_T	# of qPCR cycles	Linear regression values: Area[c] % add vs. $\frac{\text{RF area}}{\text{Chromatogram Area}} x\ 100$			Linear regression values: Height[d] % add vs. $\frac{\text{RF height}}{\text{Chromatogram height}} x\ 100$		
				Slope	Intercept	*R*2	Slope	Intercept	*R*2
Plamid mixtures	10^6	13	16	1.03	-0.75	0.97	1.02	-0.56	0.99
	10^7	10	16	0.93	1.74	0.93	0.94	1.46	0.95
	10^3	24.5	26	0.81	4.8	0.84	0.84	4.12	0.89
	10^5	17	26	0.80	4.97	0.90	0.80	5.04	0.97
	10^7	10	26	0.65	8.86	0.84	0.71	7.39	0.93
Genomic DNA mixtures	10^6	13	19	0.92	2.02	0.98	0.91	2.74	0.98
	10^5	16.5	19	1.00	-0.96	0.99	1.01	-0.30	0.99

a Concentrations and cycle amplifications were varied from 10^3 to 10^7 gene copies/µl and 16 to 26 cycles for gene standards and 10^5 – 10^6 gene copies/µl and 19 cycles for genomic DNA respectively (Supplementary Data). The results of each of the four gene standards were plotted together at each condition (60 to 104 data points per plot). Similarly, the data points for each genomic DNA dilution (45 and 51) were plotted together.
b. Plasmid mixtures (upper half) and genomic DNA mixtures (lower) were quantified with 16S rRNA gene standards and *E. coli* genomic DNA, respectively.
c. Linear regression values: percentage added vs TRFLP percentage calculated by RF area divided by total chromatogram area. Intercept units are detected-copies μl^{-1} and slope units are detected-copies calculated-copies^{-1}.
d. Same as (c) with calculations based on RF height divided by total chromatogram height

While 16S rRNA gene ratio calculations were generally more reproducible and accurate using relative peak heights, as opposed to relative peak areas, both calculation methods were reliable. The best correlations between the amount of DNA added and relative amounts inferred from the TRFLP profiles were observed in Sample Group A which was amplified for only 16 cycles and terminated in mid-exponential phase (3 cycles beyond its threshold cycle). Sample Group B, which had 10-fold higher starting concentrations than sample Group A and also terminated in late exponential phase, were also well correlated even though the regression curve parameters began to deviate from the ideal. When amplification was extended from 16 to 26 cycles, the sequence ratio correlations deviated further. Slope values fell to about 0.8 and intercept values increased to 4.5 +/- 0.5 in Groups C and D which were in early exponential and linear amplification phases, respectively, but which were much better than the parameters of Group E which was well into plateau phase. R^2 values for Group C were lower than the other samples, likely due to the lower concentrations of template. These findings illustrate the importance of stopping the amplification in exponential phase for accurate quantification of RFs.

Genomic DNA from the low G+C *E. coli* and the high G+C organisms, *Variovorax* sp. and *Rhodococcus rhodochrous* 11Y, were amplified alone and in mixtures ranging from 10% to 80% each of 16S rRNA gene copy numbers for each species. Because *R. rhodochrous* 11Y and

the *Variovorax* sp. have not been sequenced, their gene copies ng^{-1} was estimated using the *E. coli* DNA as a qPCR standard. 16S rRNA gene copies ng^{-1} of DNA for *Variovorax* sp. and *Rhodococcus rhodochrous* were estimated empirically over several orders of magnitude from multiple qPCR to be 4.04×10^5 (S.D. 2.4×10^4) and 2.75×10^5 (S.D. 2×10^4), respectively. Estimates of 16S rRNA gene copies were derived using three *Rhodococcus* strains and six Comamonadaceae family members presently included in the Microbesonline.org database (Dehal, Joachimiak et al.). The 16S rRNA gene copy numbers of these relatives of 11Y and *Variovorax* sp. ranged from 6.58×10^5 to 1.44×10^6 copies (ng of DNA)$^{-1}$ for the *Rhodococcus* and 9.46×10^5 and 1.68×10^6 copies (ng of DNA)$^{-1}$ for the Comamonadaceae, This result suggests that our empirical data for these groups may be underestimations. TRFLP profiles from all three species generated primary and secondary RF peaks that were verified *in silico* and summed for abundance calculations. *E. coli* had primary and secondary peaks at 122 bps and 172 bps, *Variovorax* sp. at 231 bps and 273 bps and *R. rhodochrous* at 158 bps and 168 bps, respectively. Table 8contains the parameters found by linear regressions of these data.

When amplification was terminated closer to the C_T values, better correlations were observed. Nonetheless, both dilutions correlated well with R^2 values greater than 0.98 regardless of dilution or calculation metric, along with slopes above 0.9 and intercepts less than 3 detected-copies μl^{-1} in all cases.

The qPCR-TRFLP method was then tested using DNA extracted from a soil microcosm spiked with DNA from pure cultures of *E. coli*, *Variovorax* sp. and *R. rhodochrous* 11Y. After first determining the 16S rRNA gene copy number in the soil extract (5.08×10^5 copies ng^{-1}, S.D. 1.3×10^5), three mixtures of DNA from each pure culture and the soil DNA were prepared (with DNA from each pure culture varying from 5% to 80% of total DNA; Table S4), and TRFLP profiles generated at three concentrations (25 ng μl^{-1} to 25 fg μl^{-1}) of each mixture. As inferred from the slopes of the soil DNA and diluted genomic DNA amplification efficiency was between 87% and 101% (ca. -3.3 to -3.7 detected-copies calculated-copies^{-1}). The slopes of the linear regressions between the copies of the pure culture DNA added versus copy number determined by qPCR-TRFLP were all above 0.69 with R^2 values above 0.96 in most cases. Instances of lower correlation may be attributable to variation in amplification efficiency between the various populations due to differences in GC content and sequence variability around priming sites (Polz and Cavanaugh 1998). For example, quantification of *E. coli* DNA was better at low numbers of amplification cycles, whereas quantification of high G+C bacteria was better at higher cycle numbers. Additional analyses of these mixtures at different concentrations and amplification cycles were comparable. TRFLP analysis of the soil DNA (data not shown) revealed small RFs corresponding to the genomic DNA RFs accounting for the increase in intercept values from zero.

Finally, the optimized qPCR-TRFLP method was used to determine shifts in buoyant density (BD) of individual operational taxonomic units (OTUs). DNA isolated from a soil microcosm established using RDX as sole nitrogen source was combined with a known mixture of unlabeled genomic DNA from *Microbacterium* sp. MA1, *Rhodococcus rhodochrous* 11Y (high G+C) and *E. coli* (low G+C). A reference gradient contained the soil DNA and ^{15}N-labeled genomic DNA from the same three isolates. Relative percentages of the RF copy number in each fraction were determined by dividing individual RF peak heights by the sum of the peaks heights. Appendix shows an example calculation, from one of the fractions in the ^{15}N gradient demonstrating how the qPCR values shown in Figure 46A were distributed among the various populations (Figure 46: B- F). RF quantities for each fraction were determined by multiplying

the relative numbers by the total number of bacterial 16S rRNA genes. If a species had multiple peaks, the peaks were summed. Figure 46 shows the abundance of 16S rRNA amplicons as a function of buoyant density: total 16S rRNA genes (A), restriction fragments for *E. coli* 16S rRNA genes (B), *Rhodococcus rhodochrous* 11Y 16S rRNA genes (C), *Microbacterium* sp. MA1 16S rRNA genes (D) and for two restriction fragments from the soil (E, F).

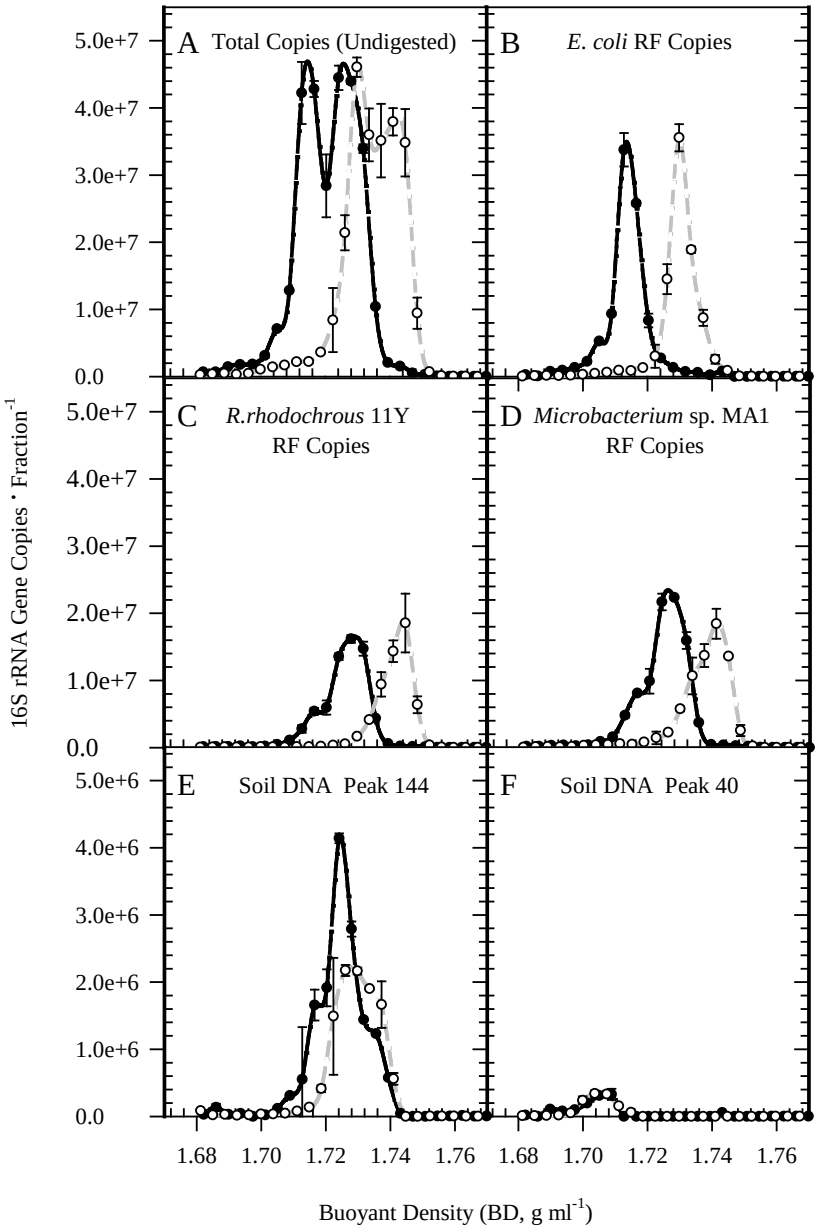


Figure 46. qPCR-TRFLP analysis of 16S rRNA genes in SIP fractions.
(A) 16S rRNA gene copies ([FAM]-27F and 338R) per fraction of the CsCl gradients with soil DNA amended with unlabeled genomic DNA (black) and ^{15}N-genomic DNA (gray dashed). (B - D) Profiles of unlabeled (black) and ^{15}N - labeled (gray dashed) 16S rRNA gene fragment copies of *E. coli*, *R. rhodochrous* 11Y and *Microbacterium* sp. MA1. (E-F) 16S rRNA gene copy profiles of two RF fragments from the unlabeled soil DNA in the two gradients. Panels B - F were generated from TRFLP analysis of digested (MnlI and MspI) qPCR products (A); RF copies per fraction were calculated from relative peak heights multiplied by full copy number. Error bars are standard deviations of duplicate qPCR (A) and calculated RF quantities (B - F).

Panels B-D in Figure 46 show clear shifts in the RF profiles of the added ^{15}N-DNA (black lines) relative to the RF profiles of the unlabeled DNA (gray lines). In contrast, the two RF profiles of indigenous bacteria in the unlabeled soil DNA (Panels E, F) were not shifted in the gradient with ^{15}N-labeled DNA from the isolates. The shifts of the diagnostic 16S rRNA gene RFs presented in Figure 46: B – F were quantified by calculating differences in buoyant density values between gradients at the RF profile maximums (Table 9). The shifts observed in the three genomic DNA samples are similar to the predicted shift of 0.016 g ml^{-1} (Buckley, Huangyutitham et al. 2007).

Table 9. Calculated buoyant density shifts of peaks between the unlabeled and ^{15}N profiles.

	^{14}N-Gradient	^{15}N-Gradient	
	BD at Max (g ml^{-1})	BD at Max (g ml^{-1})	BD Shift (g ml^{-1})
E. Coli K-12 MG1655	1.713	1.730	0.017
R. rhodochrous 11Y	1.728	1.745	0.017
Microbacterium sp. MA1	1.724 – 1.728	1.741	0.013 – 0.017
Soil RF 40	1.704 – 1.707	1.705 – 1.709	0.001 – 0.005
Soil RF 144	1.724	1.726 – 1.730	0.002 – 0.006

A direct comparison of the qPCR-TRFLP method to commonly used methods for comparing RF distributions demonstrates its improvement over the conventional approaches (Figure 47). Panels A through C of Figure 47 compare alternative methods to determine the position of RF 144. Panel A presents RF 144 relative peak height values normalized by maximum height, as commonly used in published ^{15}N SIP analyses (Buckley, Huangyutitham et al. 2007). Panel B displays the relative peak heights normalized using the sum of all peak values from each chromatogram. Panel C is an analysis of peak position the qPCR-TRFLP protocol. The same comparison of analysis methods was made for the RFs specific for *Microbacterium* sp. MA1 (Figure 47, panels D – F). Using the RF abundance (Panels C, F), the distributions of these two populations along the gradients are clearly resolved and the buoyant density shift of MA1 indicative of ^{15}N assimilation is observed.

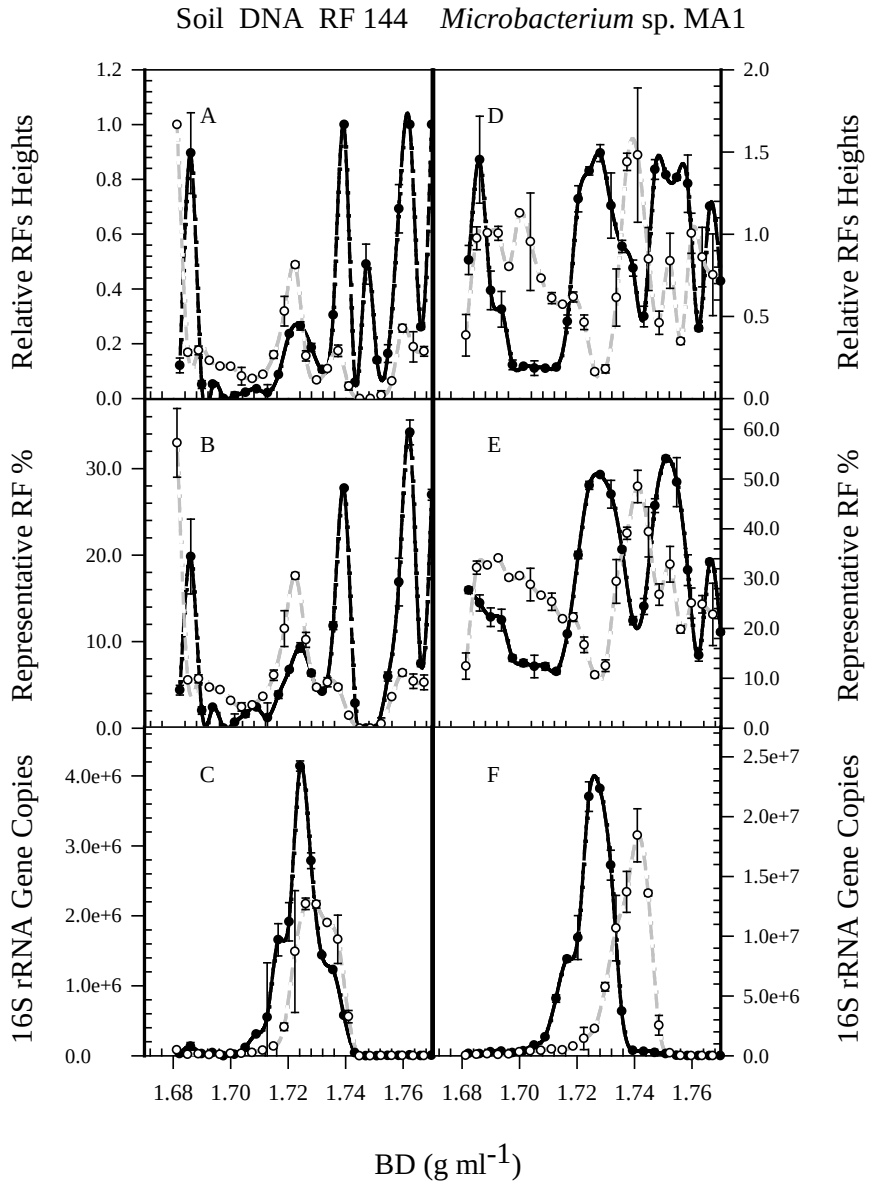


Figure 47. Stable isotope probing (SIP) profiles of soil DNA peak 144 and RFs of *Microbacterium* sp. MA1 in the buoyant density gradients determined using three methods. (A,D) SIP profiles of relative RF values (RF peak heights divided by the largest RF of each chromatogram; values in panel D rise above 1 due to multiple RF peaks of MA1 summed). (B, E) SIP profiles of relative peak heights (RFs % of total peak heights). (C, F) SIP profiles of RF copy numbers (Values in panels B and E multiplied by respective qPCR values of each reaction).

4. Discussion

The main objective of this study was to advance SIP analysis by developing a sensitive and robust method for measuring copy number of specific OTUs at different buoyant densities, a modification essential for assessing minor changes due to partial ^{13}C incorporation or any degree of ^{15}N labeling of DNA. In brief, the fluorescence intensity of EvaGreen provided for highly sensitive quantification not compromised by inclusion of a fluorescein-labeled primer ([6-FAM]-27F) in the reaction mix. Although higher threshold values were needed to calculate

C_T values with inclusion of the labeled primer, the calculations derived at the respective threshold levels were comparable (Table 9). Subsequent testing with GM3 (Muyzer, Teske et al. 1995) or a 5-carboxy- 2'4,4',5'7,7'-hexachlorofluorescein modified version, [5-HEX]-GM3 paired with 338R, displayed comparable baseline levels and no change in threshold was required for C_T calculation. However, FAM is one of the most commonly used labels for TRFLP, and recent tests comparing primers labeled with either FAM or HEX found that more OTUs could be identified using FAM (Pandey, Ganesan et al. 2007). Additionally, many qPCR quantification software packages contain methods such as the second derivative method (Rasmussen 2001) which calculate a Cp (crossing point) value independent of the baseline. Despite the difference in G+C content of the *E. coli* (50%) and *R. rhodochrous* (ca. 66%) in the genomic DNA mixes, TRFLP profiles were predictive of DNA mixture content (slopes within 0.1 of 1 and R^2 values >0.98, Table 3) when amplification was held below 20 cycles. Slightly better results were achieved when the reaction was stopped closer to the C_T values (Suzuki and Giovannoni 1996, Polz and Cavanaugh 1998, Egert and Friedrich 2003). Abundance calculations based on relative peak height or relative peak area were comparable, and would generally provide a useful cross-check in general sample analysis. Although greater deviation was observed between the amount of reference DNA added and the amount predicted from gradient profiles when it was mixed with soil DNA, the R^2 values remained consistently high and would have little influence on the detection of minor shifts in density (Figure 46).

The qPCR-TRFLP analysis applied to gradients with bisbenzamide, which increases the separation of DNA based on G+C content (Karlovsky and Decock 1991, Buckley, Huangyutitham et al. 2007), may possibly obviate the requirement for a second centrifugation step to identify minor populations in ^{15}N SIP (Karlovsky and Decock 1991). Unlike previous reports linking TRFLP with real-time PCR (Yu, Ahuja et al. 2005), quantification using Evagreen allows for a wider range of applications, such as targeting multiple regions of the 16S rRNA gene and potentially quantifying changes within a population's functional gene diversity (Horz, Yimga et al. 2001, Bernhard, Donn et al. 2005). In addition, since an intercalating dye is used for quantification, melt curve analysis can be used to verify products prior to TRFLP analysis. Finally, when used in conjunction with an unlabeled control gradient, the method can identify populations within a community having very small increases in buoyant density, as is associated with ^{15}N assimilation. Therefore, the method should have broad application to SIP analyses, allowing the identification of populations assimilating lower amounts of ^{13}C isotopes than in previously reported SIP experiments.

F. Stable isotope probing of military training range soils degrading RDX. . (Stuart Strand, David Stahl, Peter Andeer, University of Washington, Seattle)

1. Introduction

The leaching of toxic chemical from high explosives used on live-fire ranges into groundwater has become a great concern due to the detection of RDX (hexahydro-1,3,5-trinitro-1,3,5-triazine) in a 'sole-source drinking water aquifer' near the Army training range, Camp Edwards (Cape Cod, MA) (Clausen, Robb et al. 2004). RDX in groundwater originates from the dissolution of high explosives particulates (e.g. Composition B) dispersed from low-order (incomplete) ordnance detonations (U.S. Environmental Protection Agency 2005). To prevent RDX from leaching into groundwater, effective containment strategies are needed that are applicable to large areas contaminated on their surface with particulate high explosives. *In situ* microbial RDX degradation could potentially prevent migration of RDX offsite. RDX concentrations tend to be highest in surface soils of training ranges (Clausen, Robb et al. 2004,

Walsh, Taylor et al. 2010); thus, aerobic microbial RDX degradation in these areas may be the most effective way to prevent RDX intrusion deep into the soil column. However, most research on microbial RDX degradation has focused on anaerobic RDX-degraders (Ye, Singh et al. 2004, Crocker, Indest et al. 2006). While aerobic RDX-degraders have been isolated from contaminated environments (Coleman, Nelson et al. 1998, Seth-Smith, Rosser et al. 2002, Thompson, Crocker et al. 2005, Nejidat, Kafka et al. 2008, Seth-Smith, Edwards et al. 2008, Bernstein, Adar et al. 2011), aerobic RDX-degradation in military range soils has not been reported.

Several samples extracted from a long used bombing range were examined. The collected samples were screened for RDX-degradation activity with and without exogenous carbon supplied. Culture independent stable isotope probing (SIP) of soil microcosms supplied with fully- labeled ^{15}N-RDX and targeted isolation were performed to identify potential RDX degraders in the range soils. Quantitative PCR analysis of the *xplA* gene was performed in the density gradients to determine if *xplA*-bearing strains assimilated ^{15}N nitrogen from RDX degradation.

2. Materials and methods

a) Media

RDX enrichment media (soil microcosms) was made as described previously. For pure cultures, (0.4 ml/ 1 L of 1000x solution) vitamin solution ((Brandis and Thauer 1981), w/o lineolic acid) was added. Additional media with twice autoclaved soil extract (8% of a 33% w/v soil solution) (Hurst and Knudsen 1997) or yeast extract (5 mg/L) were also tested for stimulation of RDX-degradation. Nitrite minimal media was identical to RDX enrichment media except the RDX was replaced with sodium nitrite (500 μM and 5 mM). Solidified media used for isolation of microbial populations consisted of nitrite minimal media solidified with noble agar (1.5%), R2A (Reasoner and Geldreich 1985), nutrient agar, 10x diluted nutrient agar and LB agar. All agar was supplied by Difco (Becton Dickinson; Franklin, NJ).

b) Sample collection

Soils were collected into previously unopened resealable freezer bags or sterile 50 ml conical tubes from 23 discreet locations within ‘C52N Cat’s Eye’, an extensively used area of the Eglin Air Force Base (Eglin AFB; Eglin, FL) training range (June 2007) with the assistance of local EOD (explosive ordnance detection) personnel. Sampling locations were chosen based on area features, including: the interior and exterior of two former impact zones (craters), several areas lacking vegetation with high amounts of debris in the vicinity, areas with differing plant species, areas with differing soil appearances (e.g. brown sand, red sand). Additional samples, referred to as ‘bulk soil’ was obtained from six additional locations in the same area at Eglin AFB on a subsequent visit in November, 2009.

c) RDX degradation in soil microcosms

RDX degradation potential of soils collected from Eglin AFB was evaluated using mixtures of 5 or 6 soil samples (10 – 12% w/v soils), and individually using the above culture media. RDX concentrations were monitored over time in the aqueous phase by removing 500 μl samples, centrifuging at 16,000 x g for 10 minutes. The supernatant was mixed with an equal volume of acetonitrile, centrifuged and analyzed by HPLC.

For stable isotope probing experiments, soils (11.5% w/v) were incubated in 45 ml RDX enrichment media amended with ca. 40 ppm unlabeled or fully-labeled ^{15}N-RDX. Microcosms contained either soils showing no RDX-degradation in screening experiments or included soils that demonstrated RDX degradation. “Inactive soils” were composed entirely of soils obtained on the second sampling trip that did not show RDX degradation. “Active soils” had the same total mass, but 20% of the mass was replaced with a mixture of six soil samples that had screened positive for RDX degradation. Microcosms were shaken at 100 rpm at 30°C in baffled flasks. Samples were removed for RDX concentration measurements as above except the soil pellet was stored at -80°C for DNA extraction.

d) HPLC analysis

HPLC analysis was conducted using a modular Waters HPLC system consisting of a Waters 717+ autosampler, two Waters 515 HPLC pumps and a Waters 9926 photodiode array detector. A 4.6 x 250 mm Hypersil Gold reverse phase column (ThermoFisher) was used for separation with a 50:50 mobile phase (HPLC water: 83% MeOH/ 17% ACN) monitored at 254 nm. Peak integration and data analysis was conducted using the Millennium[32] software (Waters, Milford, MA).

e) DNA extraction

DNA was extracted from the soil samples and soil microcosms used for SIP analysis using the aluminum sulfate method described previously. Genomic DNA was extracted from pure cultures using the Gram Positve DNA Extraction Kit (Epicentre), modified by adding 1 μl of mutanolysin (150 U; Sigma-Aldrich; St. Louis, MO) and 1 μl of achromopeptidase (10 U; Sigma-Aldrich).

f) Density gradient centrifugation, fractionation and purification

Soil DNA (1.5 – 6 μg; 98 μl) in DNA suspension buffer (10 mM Tris- 1 mM EDTA; Teknova; Hollister, CA) was combined with 172 μl of gradient buffer (100 mM Tris, 100 mM KCl, 10 mM EDTA) and 4.53 ml of cesium chloride (CsCl) solution (ca. 1.71 g/ml buoyant density, by mass) in 10 mM Tris-10 mM EDTA). Centrifugation conditions, fractionation into PCR strip tubes (VWR, International; Radnor, PA) and buoyant density measurements were performed as described. The buoyant densities of the midpoint gradients’ fractions were all too high (i.e., the buoyant densities were greater than the other gradients, labeled and unlabeled, by values not attributable to ^{15}N assimilation), possibly due to evaporation prior to measurements. For this reason, buoyant densities of these gradient fractions were adjusted so the *xplA* profiles in these gradients would be aligned with the endpoint gradient profiles of this gene. For precipitation of DNA in gradient fractions, volumes were normalized to approximately 150 μl in microcentrifuge tubes with PCR certified water (Teknova) with 5 μg of linear acrylamide (Ambion) and mixed with 300 μl of PEG precipitation solution (30% PEG 6000, 1.6 M NaCl). DNA was precipitated as described (Neufeld, Vohra et al. 2007), rinsed three times with 70% EtOH and then resuspended in DNA suspension buffer (50 – 80 μl). Samples were stored at -20°C until analysis.

g) Quantitative PCR (qPCR)

All qPCR analyses were performed in duplicate (or greater replication) using a MJ-Research PTC-200 gradient thermocycler with a Chromo 4 real-time PCR detector with the Opticon Monitor 3.1 software (BioRad). Gene copies of *xplA* were measured in fractions using the primer-probe set described previously (Indest, Crocker et al. 2007) with a modified reverse primer sequence: xplAtaqF: GGAGGACATGAGATGACCGCT, xplAtaqR: CCTGTTGCAGTCGCCTATACC, xplAtaqprobe: [6'-FAM]-TCCCGAATTCAGGAACAACCCCTATCC-[BHQ1a] (6'-carboxyfluorescein, black hole quencher 1, Eurofins). Reactions (20 μl) consisted of Faststart Taqman Probe Master (10 μl, Roche), 6.1 μl PCR certified water, 0.7 μl of each primer (7 pmol), 0.5 μl probe (5 pmol) and 2 μl of sample. An initial denaturation was performed at 95°C for 10 min followed by cycling between 95°C for 15 s and 60°C for 40 s followed by fluorescence reading. The pHSXI plasmid (Seth-Smith, Rosser et al. 2002) containing the *xplA* gene was used as a quantification standard for *xplA*. For quantification of the *xplA* gene in soils, a total of four DNA samples were run for each sample, a 10 – fold and 20 – fold dilution of two separate DNA extractions.

qPCR of the 16S rRNA genes for TRFLP was performed as described previously, except temperature cycling was modified: 98°C for 2 min (initial denature), 98°C for 8 s and 58°C for 28 s and fluorescence analysis. For 16S rRNA gene quantification of soil DNA, the FAM-labeled primer was replaced with an unlabeled primer.

h) TRFLP analysis

Quantitative PCR products were purified using the ZR-96 DNA Clean & Concentrator-5 kit (Zymo Research; Irvine, CA) following manufacturer's instructions. Purified product (10 μl) was digested with MnlI (0.4 μl; New England Biolabs (NEB); Ipswich, MA) in a 30 μl reaction for 3.5 hours at 37°C, heat deactivated (65°C for 20 min), and digested with MspI (0.8 μl, NEB) for 3.5 hours at 37°C in a total reaction volume of 40 μl. Completed digests were purified using the ZR-96 DNA Sequencing Clean-up kit (Zymo Research). Digested reactions were separated on a 3730 DNA analyzer (Applied Biosystems; Austin, TX) and analyzed using DAx Data Acquisition and Analysis, v7.0 (Van Mierlo Software Consultancy; Eindhoven, The Netherlands).

i) Clone library generation

Clone libraries were generated from selected fractions for targeted sequencing. The 16S rRNA genes were amplified as above, except the forward primer was replaced with a non-labeled 27F primer. Following qPCR amplification, reactions were purified using the Zymo clean and concentrator-5 kit and incubated with Taq (Fermentas) and dATP (0.2 mM) for 5 min at 72°C. Reactions were again purified and cloned into the PCR4 vector using the TOPO TA-cloning kit (Invitrogen) following manufacturer's instructions. Transformed colonies were grown in 96 deep well plates (Nunc) and plasmids extracted from cultures that were pooled by row. TRFLP was performed on the extracted plasmids as described. Two 96 well plates of clones were sequenced based on the TRFLP profiles.

j) DNA sequencing

Recombinant colonies from clone librarie were submitted to "High-Throughput Sequencing Solutions" (www.htseq.org), and sequenced using the T7 promoter. 16S rRNA sequences of isolates in pure culture were PCR amplified from extracted DNA using the 27F, 1492R primer set (Lane 1991). The following primers were used to amplify the *xplB*, *xplA* and associated genes: ampermF ATGAGTCCGTGGGCGAACTT, dapBF: ATGACGAACATCAGAGCTGTCGT, dapBR TTACAGTTCTTCGCGCACGATGTA, xplBF GGCTCCTTCGCTACGG, xplBR TCAGCAGACCGATTCGGC , xplAF CCGACGTAACTGTCCTGTTCGGAA, xplAR CGGGTCCGTCCGCCGGCTGGAAGG (Rylott, Jackson et al. 2006), xplARend CGGCAGTTTTCGGTAGCG , dehR TCACCGCTGCACGTAGAAGAC. The Phusion High-Fidelity PCR kit (NEB) was used to amplify the genes following manufacturer's recommendations, purified and sequenced using the ABI BigDye v3.1 sequencing mix following manufacturer's recommendations, separated on a 3730xl DNA Analyzer. DNA sequences were trimmed and read using Sequencher 4.9 (Gene Codes Corp.; Ann Arbor, MI).

k) Nitrite assays

Nitrite concentrations were analyzed using the Griess assay (Green, Wagner et al. 1982). Samples (10 μl) or sodium nitrite (10 μl, 0.5 – 250 μM) were mixed with ddH_2O (70 μl) in clear 96 well polystyrene plates. Griess reagent (20 μl; (Green, Wagner et al. 1982)) was added and incubated at room temperature for 10 min and plates were read in a TECAN Infite500 (TECAN; Mannedorf, Switzerland) using an absorbance of 540 nm. Sample concentrations were determined based on sodium nitrite standard curves.

3. Results

Stable isotope probing (SIP) was conducted on military training range surface soils that degraded fully labeled ^{15}N-RDX aerobically. Surface soils collected from a heavily used military bombing range were screened for aerobic RDX degradation. The *xplA* gene was detected in DNA extracted from several of the soil samples, prior to enrichment, which degraded RDX. A mixture of soils that were able to degrade RDX was exposed to either unlabeled or ^{15}N-labeled RDX in aqueous slurries. These cultures were amended with exogenous carbon sources to facilitate degradation. DNA was extracted from samples for cesium chloride (CsCl) gradient separations. Putative RDX-degraders assimilating ^{15}N-nitrogen, including *xplA*-bearing species, were identified from density gradients using tandem qPCR-TRFLP of the 16S rRNA gene and Taqman qPCR of *xplA* (Indest, Crocker et al. 2007). Isolation and cultivation experiments succeeded in isolating *xplA*-bearing *Williamsia sp.* and *Rhodococcus* sp. The *Rhodococcus* sp. was prominent in the density gradients as well as a number of other species capable of using nitrite as a nitrogen source.

Twenty three surface soil samples were taken from 'C-52N Cat's Eye', a heavily used bombing range at Eglin Air Force Base (Eglin AFB; Valparaiso, FL) in May 2007 with the help of onsite explosive ordnance detection personnel. Soil samples were combined in mixtures (five or six soil samples each) and supplemented with: no exogenous carbon, yeast extract, soil extract or a mixture of glucose, glycerol and succinate. RDX degradation was observed only in samples amended with glucose, glycerol, and succinate (Table 10). Samples taken from the same area on a subsequent visit in November 2009 were also screened, but failed to degrade RDX. Six

samples amended with glucose, glycerol, and succinate positive for degradation were selected for further experimentation.

Table 10. Overview of RDX degradation of extracted Eglin soil samples and *xplA* qPCR amounts.

Sample	RDX Degradation[a]	*xplA* copies/ 16S rRNA gene copies (Avg, (Std Dev))[b]	Number of reactions with *xplA* detected[c]
1	x	5.08E-06	1
2**	+	ND	0
3	x	NT	
4**	+	5.75E-06 (4.77E-06)	4
5	x	3.31E-06 (1.75E-06)	2
6**	+	1.27E-05 (5.88E-06)	4
7**	+	ND	0*
8	x	1.32E-05 (7.57E-06)	3
9	x	4.09E-06 (1.64E-07)	2
10	x	2.31E-03 (1.48E-03)	4
11	+/-	NT	
12 - 13	x	ND	0
14**	+	5.78E-05 (1.76E-05)	4
15	x	6.82E-06	1
16	x	1.51E-05 (7.56E-06)	3
17	x	NT	
18**	+	1.55E-04 (7.64E-05)	4
19	x	NT	
20	x	3.33E-06 (2.12E-06)	2
21 - 23	x	NT	

a – Sample screening performed by Dr. Lorraine Lillis. ‘x’ – no degradation observed, ‘+’ – positive for degradation, ‘+/-‘ partial degradation after > 2 weeks

b – qPCR analysis of two DNA extractions from the soil samples, two dilutions (10x and 20x) were made of each sample for four reactions total per sample. Average value of: # of *xplA* copies/# 16S copies for the respective DNA sample. ND – no *xplA* copies were detected in any of the four samples, NT – soil was not analyzed with qPCR.

c – The number of the DNA samples(of 4) that amplified in the *xplA* qPCR reactions.

* - only two dilutions of 1 sample were performed

** - samples combined for SIP analysis

Quantitative PCR detected the *xplA* gene in four of the six samples that degraded RDX. These six samples (Table 10) positive for RDX-degradation were mixed with soils that did not show degradation in 1:5 ratios, termed ‘active soils’. Soils that did not show degradation were used as negative controls and are referred to as ‘inactive soils’. Three microcosms with active soils and two microcosms with inactive soils were each incubated with (ca. 40 ppm) unlabeled-RDX or ^{15}N-RDX (Table 11).

Table 11. Soil microcosms for SIP analyses.

Samples	Soil type	RDX
1 – 3*	Active soils	^{14}N-RDX
4 – 6	Active soils	^{15}N-RDX
7 – 8	Inactive soils	^{14}N-RDX
9 – 10	Inactive soils	^{15}N–RDX

* - the flask with the third unlabeled RDX sample broke

In active soil mixtures, degradation of RDX began after 120 hours and RDX was completely degraded in all active samples by 194 hrs. Soil samples were extracted at approximately 50% degradation, and the remaining soils were harvested following complete degradation for SIP analyses (Figure 48).

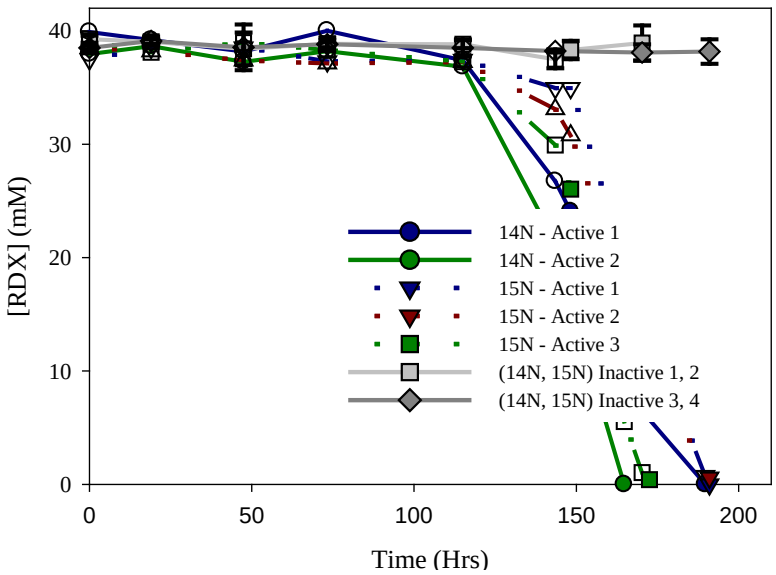


Figure 48. RDX degradation in soil microcosms (ca. 12% w/v) incubated with RDX enrichment media.
Samples with active soils (20% soils mass, colored lines) showed degradation in 49 to 80 hours following a delay of ca. 115 hours. Unlabeled RDX concentrations are plotted solid lines and ^{15}N-RDX concentrations with dashed lines. Gray lines show average concentrations of RDX in inactive soil microcosms (both unlabeled and ^{15}N) with error bars showing one standard deviation. Small soil samples (< 100 mg) were taken at all times except for filled marker points where samples were extracted for SIP analyses.

DNA yields from the soil samples degrading RDX: ^{15}N – active 1 and 2, ^{14}N – active 1 were c.a. 4.5 x higher (5.71 μg DNA/ mg soil, S.D. 0.56 μg DNA/ mg soil) than in the inactive sample (ca. 1.12 μg DNA/ mg soil). The higher yields indicate that much of the increase was due to RDX addition, and not just the addition of exogenous carbon.

Four DNA samples were separated in the first of two ultracentrifugation runs performed on DNA extracted from the soil samples as summarized in Figure 48 and Table 12. DNA density gradients used for SIP analysis..

Table 12. DNA density gradients used for SIP analysis.

Sample	Time Point	DNA Mass (μg)	Centrifugation Run	Sample	Time Point	DNA Mass (μg)	Centrifugation Run
^{14}N-Active Soils 1	Final	5.95	First	^{15}N – Active Soils 3	Final	1.50	Second
^{15}N – Active Soils 1	Final	5.90	First	^{14}N – Active Soils 2	Mid	4.80	Second
^{15}N – Active Soils 2	Final	5.50	First	^{15}N – Active Soils 1	Mid	1.90	Second
^{15}N – Inactive Soils 1	Final	2.10	First	^{15}N – Active Soils 2	Mid	2.20	Second

DNA profiles of the first four gradients (First column of Table 12) showed no clear differences in DNA buoyant density profiles between the ^{14}N control gradient and the ^{15}N inactive sample. However, clear increases in BD were observed in the ^{15}N profiles (data not shown).
The tandem qPCR-TRFLP method was used to quantify restriction fragment (RF) copy numbers in the fractions of each gradient. Over 190 distinct RFs between 40 and 332 bps were detected across the gradients. In each gradient, BD values for each RF were determined by comparing copy numbers between fractions. For each RF, t-tests were applied to the three fractions with the highest copy numbers. If one fraction had a significantly higher copy number (p-value > 0.05, Table 12) than the other two fractions, the BD of that fraction was used. If there was no significant difference between the top two or three fractions, the BDs for those fractions were averaged (Table 13).

Table 13. Example of buoyant density calculations and analysis for RF of 154 bps.

Sample	Fraction with the Maximum Average Copies of RF 154	Fraction with the Second Highest Average Copies of RF 154	Fraction with the Third Highest Average Copies of RF 154	Estimated BD for RF 154 (g/ml)	RF 154 BD Shift (g/ml)
Unlabeled:	**1.719**[a] 4.85E+07(1.06E+07; 0.133, 0.054)[b]	**1.715** 2.99E+07(1.66E+06; 0.034)[c]	**1.723** 1.60E+07(3.33E+06)[d]	1.717[e]	
15N (1):	**1.731** 5.23E+07(1.49E+07; 0.432, 0.143)	**1.735** 4.25E+07(1.00E+07; 0.136)	**1.739** 2.69E+07(3.60E+06)	1.735	0.018
15N (2):	**1.734** 5.67E+07(6.58E+05; 0.062, 0.001)	**1.730** 4.10E+07(7.21E+06; 0.02)	1.738 1.61E+07(1.89E+06)	1.732	0.015
15N (3):	**1.733** 1.14E+07(4.34E+05; 0.036, 0.023)	1.736 9.42E+06(3.24E+05; 0.959)	1.730 9.41E+06(3.84E+04)	1.733	0.016
				Average BD shift:	0.016[f] (g/ml)

a. Buoyant density of fraction with highest average copy numbers for the specific restriction fragment.

b. Number of copy numbers/ fraction. In parentheses: (standard deviation of copy numbers; p-value calculated from t-tests performed between highest and second highest average copy numbers, p-value calculated from t-tests performed between the highest and third highest average copy numbers).

c. Number of copy numbers/fraction. In parentheses: (standard deviation of copy numbers; p-value calculated from t-tests performed between the second and third highest copy numbers)

d. Third highest copy numbers with standard deviation in parentheses.

e. Average of BD values in bold (significantly greater; p value < 0.05) for that sample.

f. Average of the three BD shifts calculated between each ^{15}N sample and the unlabeled control.

The most abundant RF fragment in the gradients had a electrophoretic peak at 154 bps, with eleven others prevalent in all of the gradients (Table 14). Due to the high abundance of these 12 populations in all active samples, these operational taxonomic units (OTUs) were the focus of subsequent analyses. Using the clone libraries we identified the dominant peak (154) as belonging to one or more *Rhodococcus* species. Peak 154 was identified as a *Rhodococcus* sp. using eight clones from two of the active soil sample gradients. Comparisons of the eight clone sequences revealed four sequences with sequence identity between 96 and 99%; however, most of the variation may be due to differences in multiple gene copies in the same organism (further explanation below and in Appendix C). Several proteobacteria were identified among the other eleven. In some instances, multiple clones were isolated from different gradients with equivalent RFs. See Appendix C for a complete list of sequenced clones and predicted RFs.

Table 14. The twelve most dominant RFs in the density gradients based on copy numbers.

RF [a] (bp)	Relative abundance	Phylotype	^{15}N Avg BD (g/ml)	^{14}N Avg BD (g/ml)	BD Shift (g/ml)
154	1	Rhodococcus sp.[#]	1.733	1.717	0.016
273	5 – 19	Unidentified	1.734	1.719	0.015
231	4 – 13	Variovorax sp.[**]	1.734	1.719	0.015
113	11 – 16	Unidentified	1.733	1.719	0.014
119	2 – 3	Mesorhizobium sp.[**]	1.729	1.717	0.012
144	3 – 6	Alphaproteobacteria[**]	1.733	1.721	0.012
163	4 – 10	Gammaproteobacteria[**] Betaproteobacteria[**]	1.731	1.719	0.012
245/ 246	2 – 7	Rhizobium sp.[#]	1.729	1.719	0.010
156	9 – 14	Streptomyces sp.[**] Arthrobacter sp.[#]	1.724	1.721	0.003
139	2 – 6	Terrabacter sp.[**]	1.731	1.730	0.001
40	7 – 15	Chitinophaga sp.[**]	1.709	1.708	0.001
204	8 – 13	Betaproteobacteria[**]	1.717	1.719	-0.002

a – RFs are sorted by BD shift in descending order
** - identified in clone library
# - identified in clone library and isolated

Buoyant density calculations indicated that peak 154 was shifted in the ^{15}N- buoyant density gradients by the largest amount, 0.016 g/ml, but (OTUs) with seven other RFs also shifted by at least 0.01 g/ml in the ^{15}N-gradients (Table 14, Figure 49). Visualization as a heat map of the RF copy numbers in each fraction facilitated identification of the seven shifting OTUs (Figure 49).

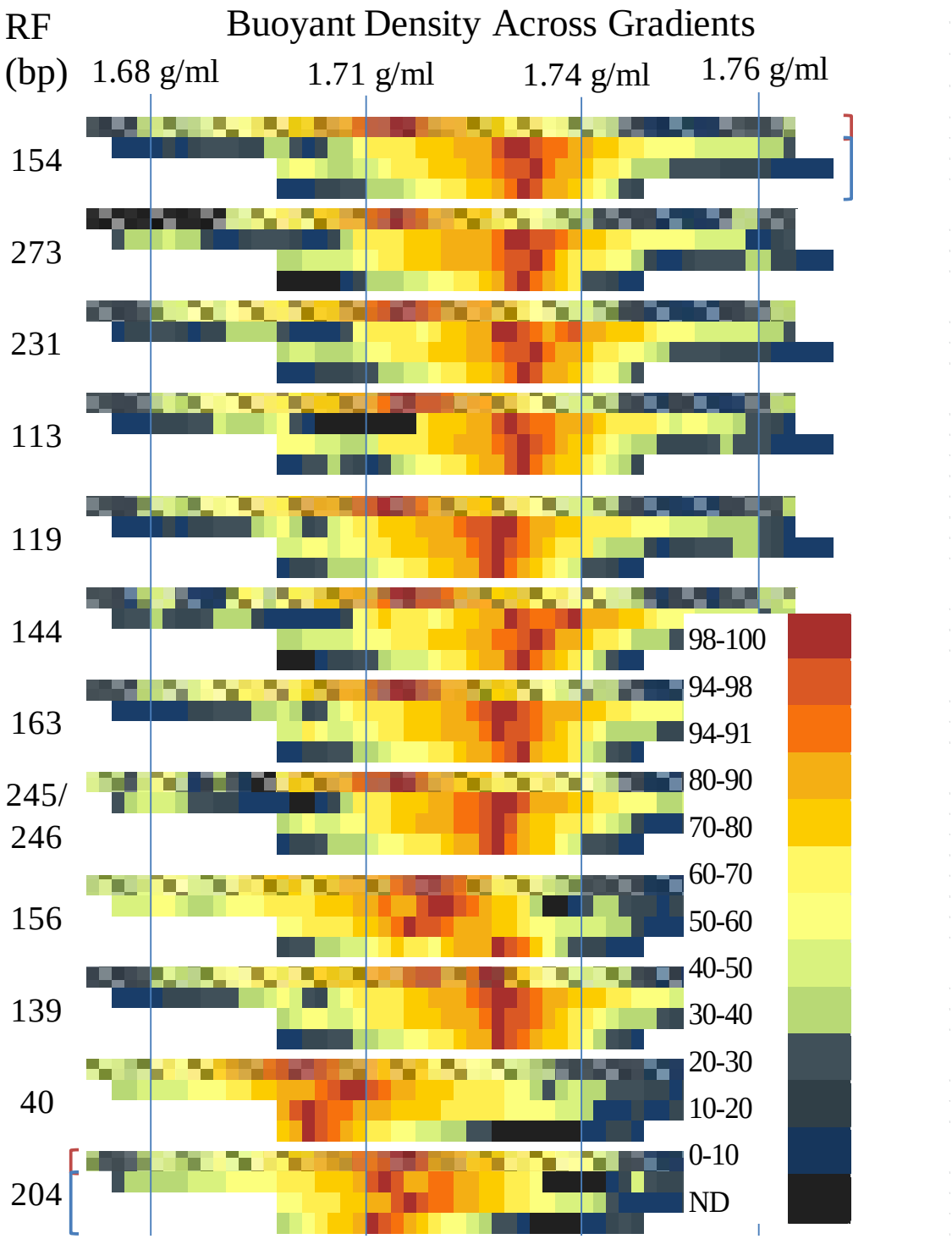


Figure 49. Heat map showing the distribution of the RF copy numbers across four of the gradients.

The top row for each fraction represents the copy number distribution in the ^{14}N-active soils 1 gradient (red bracket top right and bottom left) and ^{15}N-active soils 1 – 3 gradients (blue brackets top right and bottom left) at complete RDX degradation. Heat map colors indicate the percentile of the copy numbers of each fragment (bottom right overlay) in each fraction interpolated for every 0.0015 g/ml BD value. ND = 'Not detected'

The buoyant densities of these twelve RFs did not change much from the samples taken at midpoint and endpoint of RDX degradation in the respective microcosm (^{15}N active soils 1 and 2, Table 12). Any changes in RF buoyant densities between the two times were moderate (< 0.008) and inconsistent between microcosms. Relative abundance of individual RFs (Rank in Figure 49) varied between the midpoint and endpoint samples; however, peak 154 representative of *Rhodococcus* sp. was also the most abundant RF in the gradients of the midpoint samples. The most dramatic change in the relative abundance of these twelve samples occurred with the unidentified RF 113 which increased in relative abundance by thirteen and sixteen positions in its relative abundance between the midpoint and endpoint samples.

All twelve RFs were detected in the inactive sample gradient. Some of the RFs from the inactive gradient were identified through clone libraries. Peak 154 in the inactive gradient was not identified, but its relative abundance was much lower than in the active samples. Of the sequences listed in Table 14, the alphaproteobacteria (RF 144), gammaproteobacteria (RF 163), *Rhizobium* sp. (RF 245/246), *Arthrobacter* sp. (RF 156) and one of two closely related (98% identity) *Mesorhizobium* spp (RF 119) were identified in the inactive clone library. A more complete listing of fractions selected for clone library generation, sequenced clones and the gradients retrieved from are listed in Appendix C.

The finding that a *Rhodococcus* sp. dominated the ^{15}N-labeled populations was consistent with the isolation of several *Rhodococcus* sp. from munitions contaminated sites capable of using RDX as a sole nitrogen source (Coleman, Nelson et al. 1998, Seth-Smith, Rosser et al. 2002, Nejidat, Kafka et al. 2008, Seth-Smith, Edwards et al. 2008). Since the *xplA* gene has been identified in all of the strains examined (Seth-Smith, Edwards et al. 2008), evidence that *xplA*-bearing populations were involved in RDX-degradation was sought. Quantification of the *xplA* gene in these gradients indicated that an *xplA*-bearing population was prominent in these microcosms and assimilated nitrogen from the labeled RDX (Figure 50).

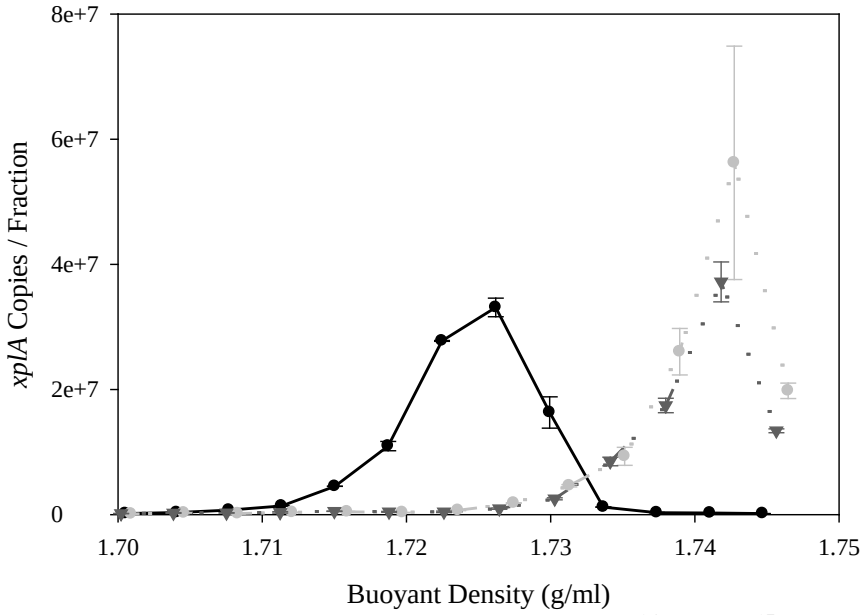


Figure 50. *xplA* gene copy number distribution in ^{14}N- and ^{15}N-density gradients.
xplA gene copy numbers were quantified in density gradient fractions. A buoyant density increase of 0.016 – 0.017 g/ml was found in replicate gradients with ^{15}N (dashed and grayscale) relative to the ^{14}N control (solid black).

The relative amount of the *xplA* gene in the midpoint gradients was higher than the endpoint gradients (ca. 40,000 copies/ng and 16,000 copies/ng respectively). It is important to note that a plasmid was used for *xplA* quantification. Research has demonstrated that using supercoiled plasmids as quantification standards leads to overestimation of the target copy numbers (Hou, Zhang et al. 2010); thus, the xplA copy number estimates reported here are likely inflated. The decrease in the abundance of the *xplA* gene between ca. 50% and 100% RDX degradation indicates that non-*xplA*-bearing strains became more abundant as RDX degradation progressed. The buoyant density of the *xplA* gene in the gradients was much higher in the gradients than the 16S rRNA profiles of any of the RFs. This discrepancy between the BDs of these genes was probably due to variation in G+C content within the genomes of the organisms. To investigate whether this was a plausible explanation for the variation in buoyant densities, the G+C content of the 16S rRNA genes and flanking regions of the closest sequenced population to the identified *Rhodococcus* sp., *Rhodococcus erythropolis* PR4, and three other Actinobacteria were examined (Table 15).

Table 15. Comparisons of the average G+C contents and estimated BDs of the 16S rRNA genes and flanking regions to the genomic G+C contents[a].

Species	No. of 16S rRNA genes/ genome	Genome G+C content (%)	Avg. G+C content of 16S rRNA genes (%)	Avg. G+C content of 16S rRNA genes and flanking sequences[b] (%)			
				500 Bps	1,000 bps	5,000 bps	10,000 bps
Rhodococcus Erythropolis PR4	5	63.04	57.76 (0.09)	54.99 (0.33)	55.34 (0.42)	58.25 (0.49)	60.20 (0.40)
		[1.723]	[1.719]	[1.715]	[1.715]	[1.718]	[1.720]
Mycobacterium sp. JLS	2	68.34	58.15 (0.07)	57.69 (0.07)	59.15 (0.41)	62.47 (0.51)	65.73 (1.18)
		[1.728]	[1.718]	[1.718]	[1.719]	[1.722]	[1.725]
Mycobacterium sp. KMS	2	68.44	58.25 (0.21)	57.69 (0.18)	59.20 (0.43)	62.53 (0.55)	65.80 (1.26)
		[1.728]	[1.718]	[1.718]	[1.719]	[1.723]	[1.726]
Arthrobacter aurescens TC1	6	62.34	57.14 (0.03)	55.84 (0.32)	56.43 (0.16)	58.47 (1.07)	60.00 (1.39)
		[1.722]	[1.717]	[1.716]	[1.717]	[1.719]	[1.720]

a- estimated buoyant densities are provided in brackets (g/ml) underneath each average G+C calculation

b - In each column, the G+C content was calculated for the 16S rRNA genes plus the fragment size listed on both sides of the genes. For example G+C contents listed under 500 bps represents average G+C content of the 16S genes + 1000 bps (500 bp before and 500 bps after the genes). Values in parentheses are one standard deviation of the average.

The G+C content of the *xplB* and *xplA* genes is ca. 67%. DNA fragment sizes generally varied between 1500 bps and 20,000 bps. Table 15 shows that the G+C contents of DNA fragments encompassing the 16S rRNA gene are typically lower than the genomic G+C contents by 2.5% to 8.1% depending on fragment size. Buoyant density of unlabeled DNA with a 67% G+C content in CsCl would have a calculated buoyant density of ca. 1.73 g/ml and DNA with 55 – 60% G+C content (DNA fragment encompassing the 16S rRNA gene of *R. erythropolis* PR4 of 2,500 to 21,500 bps (Table 15)) ca. 1.715 g/ml to 1.720 g/ml (Osterman 1984). These values are in good agreement with those observed in the ^{14}N-endpoint gradients.

Following RDX degradation, bacteria were isolated from the active soil samples on minimal media plates with nitrite as a sole nitrogen source. Based on partial 16S rRNA gene sequences (1417 bps), two of the isolates were identified as strains of *Rhodococcus erythropolis*. The two variants, *Rhodococcus* sp. EG2A and EG2B, had different morphologies, but their partial 16S sequence was identical. The 16S sequences of these isolates were between 97 and 99% identical to the *Rhodococcus* sp. clones. However, most of the mismatches between the isolates and the clone libraries occurred within a 26 bp region of the gene where 9 bps in the isolate sequences indicated sequence variation between 16S rRNA gene copies. The secondary base peaks are unlikely to be due to contamination as the same secondary base spectrum was observed in the sequences of both strains and in sequences generated from DNA extracted from cultures at different times originating from different colonies. Variations between the isolate and clone library sequences outside of this region were between 1 and 3 bps (See Appendix C for further explanation). TRFLP analysis of the isolated *Rhodococcus* sp. generated a RF at 154 bp.

Another strain, *Williamsia* sp. EG1, was also isolated from the Eglin soils with RDX as sole nitrogen source. All of these strains were able to use RDX as a sole nitrogen source and possessed the genes flanking *xplA* identified in *Microbacterium* sp. MA1 and *Rhodococcus rhodochrous* 11Y. Partial sequencing (3166 bps in EG2 and 4000 bps in EG1) of the *xplB*, *xplA* and flanking genes revealed >99.9% identity between these fragments and those in MA1.

Several other strains were isolated, putatively identified as: two *Arthrobacter* spp., a *Siphonobacter* sp., a *Burkholderia* sp., a *Rhizobium* sp., a *Pseudomonas* sp. and a *Stenotrophomonas* sp. based on partial 16S rRNA gene sequences (682 – 940 bps). None of the additional isolated populations were able to degrade RDX or NDAB (4-nitro-2,4-diazabutanal; data not shown), but with the exception of the *Stenotrophomonas* sp., they were all capable of utilizing nitrite. Three of the strains, the *Rhizobium* sp., *Burkholderia* sp., and *Siphonobacter* sp. were able to fix nitrogen (data not shown). The 16S rRNA sequence of one of the *Arthrobacter* strains was 99% identical to a clone from the inactive soil library and the *Rhizobium* sp. was 99% identical to a *Rhizobium* sp. cloned from the gradients.

Isolated RDX-degraders have been shown to assimilate nitrogen through nitrite liberated during RDX-degradation (Coleman, Nelson et al. 1998, Seth-Smith, Rosser et al. 2002, Halasz, Manno et al. 2010). Therefore, we examined cultures of *Rhodococcus* sp. EG2B for nitrite accumulation during RDX-degradation to determine if nitrite uptake by populations could account for any of the ^{15}N-assimilation.

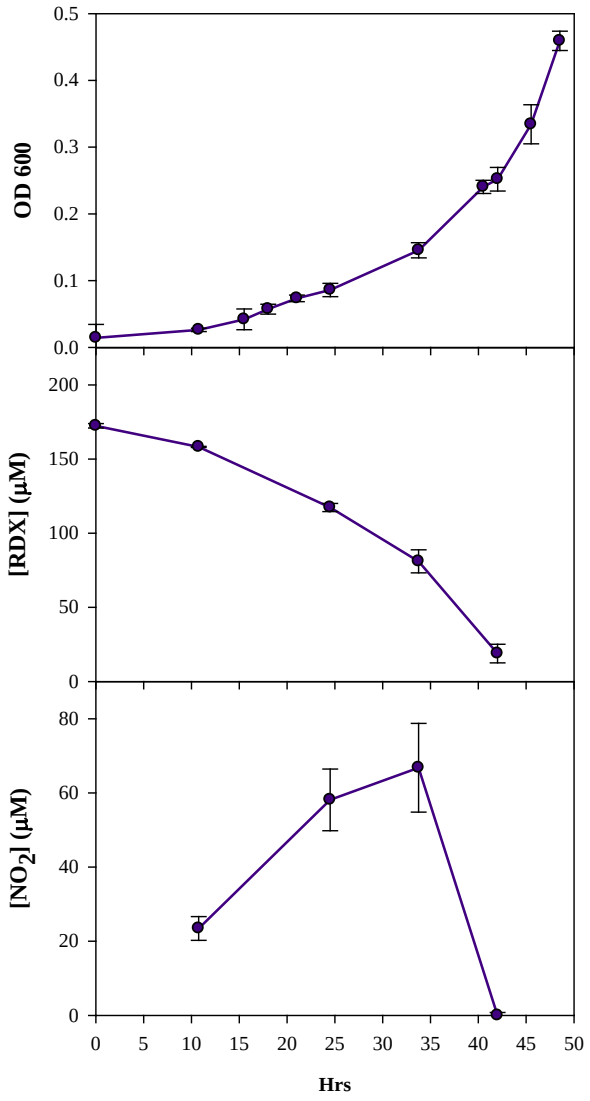


Figure 51. Nitrite production in cultures of *Rhodococcus* sp. EG2B degrading RDX. Biological replicate cultures of *R. erythropolis* EG2B were grown on RDX enrichment media supplied with ca. 40 ppm RDX. Following three transfers, optical density (600 nm; top panel); RDX (middle panel) and nitrite (bottom panel) concentrations were measured over time.

Time course analysis of RDX degradation by *R. erythropolis* EG2B showed nitrite accumulation during RDX degradation until ca. 50% degradation at which point the nitrite was rapidly consumed (Figure 51). This observation indicates that during this time some nitrite could be available to other populations in the soil.

4. Discussion

Microbial degradation of RDX may be exploited to prevent contamination of groundwater with RDX from military range soils. Therefore, the identification of microbial populations and enzymatic pathways relevant to RDX degradation in range soils would be valuable tools to assess *in situ* RDX-degradation potential at military training ranges. However, neither RDX

degradation nor the isolation of RDX-degraders has been reported in military training range soils.

Aerobic RDX degradation was observed in a few (< 30%) of the surface soils extracted from a heavily used Eglin AFB training range. Similarly, the *xplA* gene was detected in a few of the soils. This indicates that RDX-degradation potential was present at this location, but was heterogeneously distributed. Spatial heterogeneity of munitions concentrations in contaminated soils has previously been reported (Clausen, Robb et al. 2004, U.S. Environmental Protection Agency 2005), but correlations between RDX concentrations and degradation potential in the soils were not explored. Consistent with previous surface soils, activity was contingent upon the addition of labile carbon sources to the samples (Shull, Speitel et al. 1999). Thus, biostimulation (Hopkins, Semprini et al. 1993) may be an effective strategy for containing RDX in military range soils.

Supplied with simple carbon sources, RDX-degradation initiated after approximately five days and was completed within eight days. The delay in rapid degradation suggests that either: the RDX-degraders were initially at low abundance, these populations were initially inactive and needed to recover, or that competing nitrogen sources needed to be consumed prior to the initiation of RDX-degradation. The fact that samples negative for RDX degradation yielded much less DNA suggests that one or both of the first two scenarios occurred.

Three *xplA*-bearing RDX-degraders, *Williamsia* sp. EG1 and *Rhodococcus erthyropolis* EG2A and EG2B were isolated from the Eglin AFB soil microcosms. ^{15}N-stable isotope probing analysis confirmed that a *Rhodococcus* sp. closely related to the isolated strains was involved, if not solely responsible, for the RDX-degradation. Previous studies have shown that the *xplA* gene is localized to mobile genetic elements (Indest, Jung et al. 2010) and has been disseminated to RDX contaminated sites worldwide (Seth-Smith, Edwards et al. 2008). This study confirms that *xplA* persists in military range soils as well.

Intra-genomic variation in G+C content (Table 15) is important to consider in SIP analyses. A previous SIP experiment looked for evidence of a known *xplA*-bearing organism in bands extracted from CsCl gradients containing *xplA* (Roh, Yu et al. 2009). The analysis presented here shows that this strategy may not be effective, especially when the separated DNA fragments are less than 20 kbps (Table 15).

In SIP analysis, it is important to consider that cellular material may be labeled through cross-feeding instead of the metabolism of the target compound (Manefield, Griffiths et al. 2007). Analysis of these gradients, revealed a number of organisms that assimilated ^{15}N label in addition to the *xplA*-bearing *Rhodococcus* strains. Without isolation, it is difficult to ascertain the role these organisms played in the RDX degradation. In pure RDX-degrading cultures, nitrite was shown to be released from the *Rhodococcus* cells prior to its uptake (Figure 51). Thus, nitrite assimilation may be one mechanism of cross-feeding in these microcosms. 15N-assimilation into some populations could also have occurred through the metabolism of RDX-degradation products other than nitrite. Past studies have identified NDAB, MEDINA (methylenedinitramine) and formamide to be nitrogenous degradation products of RDX-degradation (Ye, Singh et al. 2004, Crocker, Indest et al. 2006), including *xplA*-mediated degradation (Jackson, Rylott et al. 2007, Halasz, Manno et al. 2010). The fact that a number of the strains that assimilated ^{15}N-nitrogen (Table 14) were possibly present in the inactive soils seems to suggest that cross-feeding may have occurred in this study. However, 16S rRNA

sequences are not always indicative of genetic makeup (Thompson, Crocker et al. 2005) and the *xplA*-gene, as with other genes involved in xenobiotic degradation (van der Meer, de Vos et al. 1992, Trefault, de la Iglesia et al. 2004), is associated with mobile elements (Indest, Jung et al. 2010). Therefore, it is possible that these populations have acquired the ability to degrade RDX.

This research is the first to demonstrate RDX degradation in military range soils; an important observation that validates microbial degradation as a potential RDX containment strategy for ranges. The involvement of *xplA*-bearing strains in the degradation makes the gene a potential biomarker to monitor RDX degradation activity and it expands the known environments where this gene persists. In devising a bioremediation strategy for ranges, it is important to note the requirement of supplemental carbon to promote RDX degradation *in situ*.

G. Cocultures of *Rhodococcus* sp. EG2B with military range soil isolates. (Stuart Strand, David Stahl, Peter Andeer, University of Washington, Seattle)

1. Introduction

As shown in the previous section, stable isotope probing (SIP) analysis of RDX-degrading range soils from Eglin Air Force Base (AFB) was conducted and partial 16S rRNA gene sequences of several bacteria assimilating ^{15}N-nitrogen were identified. In follow up analyses, several populations were isolated from these soils including the RDX-degrading *Rhdococcus* sp. variants, EG2A and EG2B. The remaining isolated populations were unable to grow on RDX, but all but one could grow on nitrite. Cocultures were developed with the RDX-degrading EG2B and each of four other isolated strains to examine their effects on EG2B growth rates and RDX-degradation.

2. Materials and methods

a) Isolation of microbes

End point soil slurry samples (Table 12), ca. 100 µl, were mixed in 10 ml of 0.1% sodium pyrophosphate used to create serial dilutions in enrichment media with 100 ppm and 200 ppm nitrite instead of RDX. The two lowest dilutions that grew were plated on nitrite minimal media plates and colonies formed were used to inoculate nitrite minimal media. Cultures were streaked for isolation on nitrite minimal media plates, nutrient agar, 10x diluted nutrient agar, R2A and LB agar. All agar was supplied by Difco (Becton Dickinson; Franklin, NJ). 16S rRNA genes were sequenced to identify the isolated populations as described above.

b) Growth experiments

Growth experiments were conducted in RDX enrichment media (4 mls) with vitamins in 15 ml polystyrene culture tubes. RDX was replaced with (filter sterilized): 1 mM sodium nitrite (Sigma Aldrich; St. Louis, MO), 100 µM NDAB (4-nitro-2,4-diazabutanal; SRI International; Menlo Park, CA) or 1 mM Hi-Di™ formamide (Applied Biosystems/Ambion; Austin, TX). Cultures were shaken at 30°C and 200 rpm.

c) Coculture experiments

Initially, RDX enrichment media (4 ml) with vitamins (0.4x) in polystyrene culture tubes (15 ml) were inoculated with colonies of *Rhodococcus erythropolis* EG2B alone or with a colony of one of the following: a *Pseudomonas* sp., a *Burkholderia* sp., a *Rhizobium* sp., a *Stenotrophomonas* sp., a *Siphonobacter* sp., and two *Arthrobacter* spp. grown on R2A or LB agar, all cultures were started in duplicate. Optical density and RDX concentrations were monitored and cultures were transferred in late exponential or early stationary phase at a 1/20 inoculum. Cultures chosen for direct comparisons were restarted from glycerol stocks as above and transferred three times prior to analyses to stabilize the communities.

For colony counts, samples from the replicate cultures were serially diluted in enrichment media and plated (20 μl of 10^4 and 10^5 dilutions) onto LB agar (*Stenotrophomonas* sp. cocultures) or R2A agar (*Rhodococcus* sp. cultures with *Pseudomonas* sp., *Burkholderia* sp. and *Rhizobium* sp.) at each dilution.

d) Nitrite assays

Nitrite concentrations were analyzed using the Griess assay (Green, Wagner et al. 1982). Samples (10 μl) or sodium nitrite (10 μl, 0.5 – 250 μM) were mixed with ddH_2O (70 μl) in clear 96 well polystyrene plates. Griess reagent (20 μl; (Green, Wagner et al. 1982)) was added and incubated at room temperature for 10 min and absorbances were read in a TECAN Infite500 (TECAN; Mannedorf, Switzerland) at 540 nm. Sample concentrations were determined based on sodium nitrite standard curves.

3. Results

The identified aerobic RDX-degrader, *Rhodococcus* sp. EG2B with the *xplA*-gene, isolated from Eglin AFB soils was cultured alone and in cocultures with four other bacteria isolated from the soils. Stable cultures were established in RDX enrichment media through multiple late log phase transfers. The stable cultures were used to compare growth rates, nitrite and RDX concentrations in EG2B monocultures with the four cocultures. While nitrite accumulation decreased in three of the cocultures (Figure 52), only the cocultures with *Pseudomonas* sp. showed faster RDX degradation than monocultures. Colony counts indicated that EG2B grew faster in the cocultures with the *Pseudomonas* sp., potentially explaining the increased degradation rates.

In addition to the *Rhodococcus* spp., seven isolates were tested for growth on nitrite, RDX and NDAB (4-nitro-2,4-diazabutanal) as sole nitrogen sources. The *Stenotrophomonas* sp. was the only isolate not capable of using nitrite as a sole nitrogen source. Three of the populations, a *Burkholderia* sp., a *Rhizobium* sp., and a *Siphonobacter* sp. were identified as diazotrophs based on growth in nitrogen free media. However, their growth was faster in the presence of nitrite, and nitrite was not detected in stationary phase cultures. Only the *Rhodococcus* spp. were able to grow on RDX as a sole nitrogen source and none of the isolates could grow on NDAB. NDAB concentration measurements were attempted using reverse phase HPLC, but quantification was not accurate except in uninoculated enrichment media.

Four of the seven strains: a *Burkholderia* sp., a *Rhizobium* sp., a *Stenotrophomonas* sp., and a *Pseudomonas* sp. were chosen for coculture experiments with *Rhodococcus* sp. EG2B to compare RDX and nitrite degradation rates with EG2B monocultures. The *Rhizobium* sp. was

chosen because it was the only population identified as a likely ^{15}N-assimilating population from the SIP analysis with a partial 16S rRNA gene sequence 99% identical to a *Rhizobium* sp. clone library sequence (272 bp) that shifted 0.01 g/ml in the ^{15}N gradients (Table 14). The *Pseudomonas* sp. was chosen because RDX degradation rates appeared faster in preliminary coculture experiments. The *Burkholderia* sp. was chosen to compare with the *Rhizobium* sp. because they are both diazatrophs but the *Burkholderia* sp. grew much faster on nitrite in monocultures (data not shown). The *Stenotrophomonas* sp. was selected as a control organism because it could not grow on nitrite. Table 16 lists the nitrogen sources tested for each of the five organisms used in the coculture experiments.

Table 16. Growth of strains used in cocultures on different nitrogen sources.

Strain	Nitrogen Source			
	N_2 - fixation	NO_2^-	NDAB	Formamide[1]
Rhodoccocus sp. EG2B	-	+	-	Yes
Burkholderia sp.	+	+	-	N.D.
Rhizobium sp.	+	+	-	No
Pseudomonas sp.	-	+	-	No
Stenotrophomonas sp.	-	-	-	N.T.

1 – N.D. – 'Not determined'. Specific growth rate was not measured so it is unknown if it used formamide as a sole nitrogen source or grew through fixation. N.T. – "Not tested"

Growth was fastest in EG2B cocultures with the *Pseudomonas* sp. and the *Burkholderia* sp. Growth rates in EG2B- *Stenotrophomonas* sp. cocultures were almost identical to monocultures and EG2B-*Rhizobium* sp. cocultures grew the slowest. (Figure 52A).

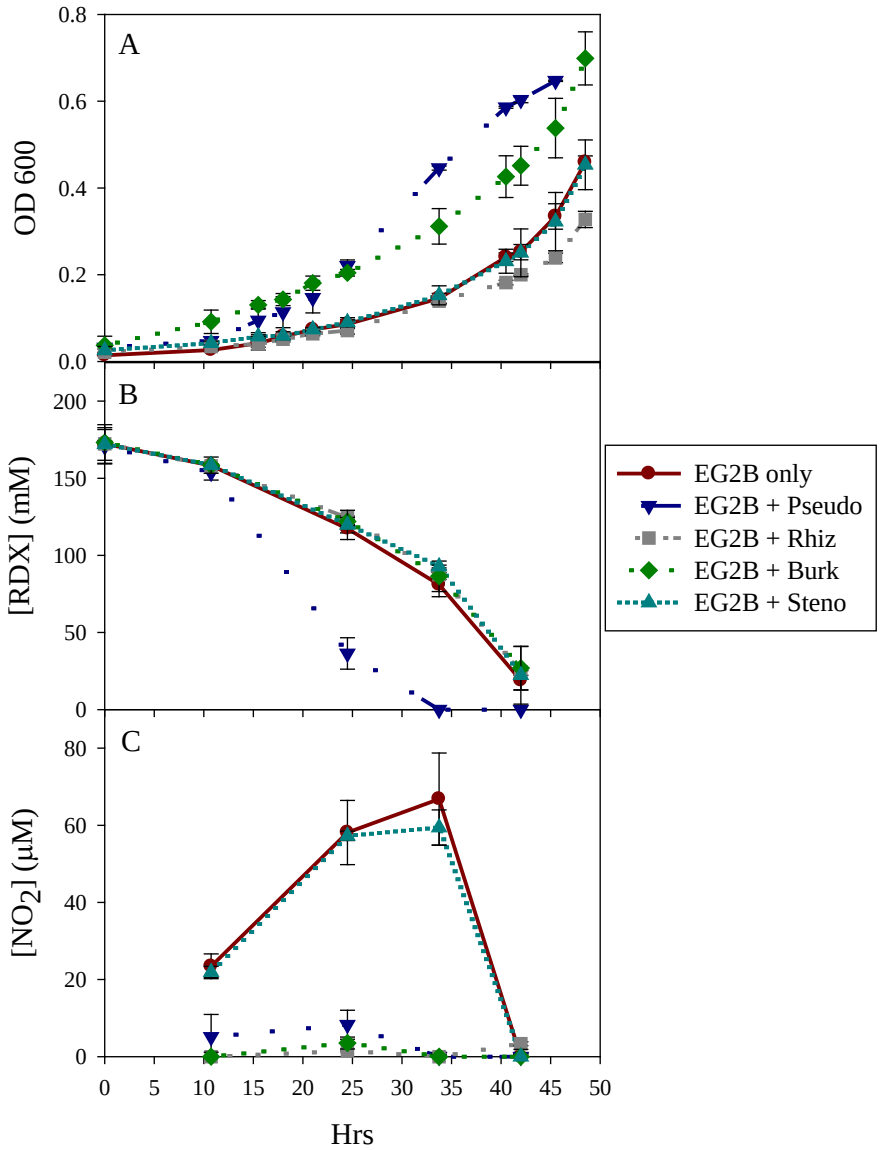


Figure 52. RDX nitrite concentrations in *Rhodococcus erythropolis* EG2B in monoculture and cocultures.

Growth curves of EG2B (solid line) and co-cultures of EG2B with each of the four other organisms (dashed lines) (A). RDX degradation rates of the cultures (B). Degradation with the *Pseudomonas* sp. showed significantly faster degradation than the other cultures. Nitrite measurements of the cultures (C). Nitrite concentrations were much lower in the cocultures than in the monocultures except when paired with *Stenotrophomonas* sp. Error bars represent standard deviation of biological replicate cultures.

RDX degradation rates were approximately the same between EG2B monocultures and in coculture with three of the four species (Figure 52B). Despite faster growth in the *Burkholderia* sp.-EG2B cocultures, RDX degradation rates did not change. However, RDX degradation rates increased markedly in EG2B-*Pseudomonas* sp. cocultures with 130 – 147 µM of RDX degradation in 24.5 hours compared with 49 – 63 µM degradation in the other cultures.

Nitrite accumulated to the same extent in EG2B-*Stenotrophomonas* sp. cocultures as in EG2B monocultures. But nitrite levels were much lower in the other cocultures, indicating that nitrite

produced by EG2B was quickly consumed by the *Pseudomonas* sp., the *Rhizobium* sp., and the *Burkholderia* sp. (i.e. by cross-feeding).

Culture dilutions were plated at 24 hours to measure EG2B growth in the cultures. EG2B colony counts were highest in cultures with *Pseudomonas* sp. indicating that EG2B grew faster with *Pseudomonas* sp. than it did in the other cultures (Table 17).

Table 17. Colony counts and the ratio of the colony counts of added strains to EG2B in the cultures at 24.5 hours[a]

Culture	OD 600 at 24.5 hours	EG2B colony counts	Added strain colony counts	Ratio of colony counts of added strain to EG2B
EG2B monoculture	0.086 (9.9 E-3)	5.6 E5 (1.3 E5)		
EG2B, Pseudomonas sp.	0.220 (0.01)	2.0 E6 (3.4 E5)	1.0 E6 (6.8 E5)	0.50 (0.32)
EG2B, Burkholderia sp.	0.205 (7.8 E-3)	4.5 E5 (1.8 E5)	9.4 E5 (5.3 E5)	2.02 (1.00)
EG2B, *Rhizobium* sp.	0.071 (8.1 E-3)	3.4 E5 (1.7 E5)	1.8 E6 (2.8 E5)	6.14 (2.38)
EG2B, Stenotrophomonas sp.	0.091 (0.01)	9.8 E5 (3.1 E5)	6.5 E4 (9.1 E4)	0.09 (0.09)

a. Average values of plate counts (10^4 and 10^5 dilution of each culture, 4 total plates each). Standard deviations are in parentheses. Times and optical density values correspond with those in Figure 52.

4. Discussion

The experiment presented here indicated that EG2B growth and RDX degradation were impacted by coculture with other bacteria from the community. While the differences of EG2B growth and RDX degradation between three of the cocultures and EG2B monocultures were negligible, EG2B grew faster and degraded RDX faster when grown with the *Pseudomonas* sp. Nitrogenous compounds (i.e. nitrite, nitrate, and ammonia) have been reported to inhibit RDX degradation rates by the *xplA*-bearing *Gordonia* sp. KTR9 (Indest, Jung et al. 2010). However, nitrite cross-feeding appeared to occur between EG2B and three of the four populations. Therefore, nitrite depletion in the cultures may not be responsible for the increased degradation rates. However, the cross-feeding of nitrite may account for the ^{15}N-assimilation by some of the labeled populations identified in Table 14.

The known *xplA*-degradation product, NDAB, was not measurable in the cultures using our current HPLC configuration; however none of the strains were capable of growing on NDAB as a sole nitrogen source. Recent aerobic resting cell assays with the *xplA*-bearing *Rhodococcus*

sp. DN22 demonstrated that while NDAB was the primary degradation product of RDX, methylenediamine (MEDINA) was also detected at low levels (10% of NDAB concentrations) (Halasz, Manno et al. 2010). This suggests that *xplA*-mediated degradation of RDX occurs through a combination of the single and double denitration reactions (Jackson, Rylott et al. 2007). It is unknown if this effect occurs in all of the xplA-bearing isolates or not, and whether growth conditions could impact this ratio. Formamide, another intermediate detected in DN22, could be used as a nitrogen source by EG2B but the *Pseudomonas* sp. was not capable of using formamide as a nitrogen source either (Table 16).

The coculture experiments conducted here demonstrated that cross-feeding of nitrite occurred when EG2B was cultivated with populations capable of metabolizing nitrite, possibly explaining the ^{15}N – assimilation into some of the populations identified in the SIP gradients. A number of experiments could be conducted to elucidate why the growth and RDX degradation rates of EG2B were faster in cocultures with the *Pseudomonas* sp. Coculture experiments with the *Pseudomonas sp.* and other known *xplA*-bearing isolates would be useful to determine if the increased RDX degradation rate is specific to EG2B. More complex communities of three or more populations could be used to determine if the increased RDX degradation and growth rates of EG2B in the presence of the *Pseudomonas* sp. persisted when another strain (i.e the *Rhizobium* sp.), was present. Finally, more rigorous examination of the metabolites needs to be conducted to ascertain metabolic profile variations under different culture conditions.

V. Conclusions and Implications for Future Research

Our findings support the prevalence of Rhodococcus in RDX degradation in training range soils, while suggesting that RDX degradation may also occur as the result of Gram negative bacterial activity, resulting in assimilation of nitrogen derived from RDX. Another interpretation of the SIP results is that Gram negative bacteria may be efficient scavengers of nitrite produced by the RDX degraders.

An important observation was that RDX degradation potential and the occurrence of xplA was highly heterogeneous in samples taken from the target area at Eglin training range (Table 10). More than half of the samples were unable to degrade RDX even with added carbon, and xplA was not detected in four of the soils. These results suggest that a large fraction of training range soils lack the genetic potential to degrade RDX. Munitions particulates deposited on soils that lack bacteria able express XplA are more likely to leach RDX into the subsurface compared to soils with high levels of xplA bearing bacteria. The heterogeneity of RDX degrading potential is similar to the heterogeneity of munitions particulates around targets (Clausen, Robb et al. 2004, U.S. Environmental Protection Agency 2005). A possible explanation of xplA heterogeneity is that the gene accumulates near munitions particulates due to the growth of bacteria carrying xplA, utilizing nitrogen from RDX.

The second observation of importance for understanding limitations on bacterial degradation in training range soils is that soil samples incubated without added carbon were unable to degrade RDX, while the addition of carbon substrates stimulated RDX degradation in some soil samples. Thus the presence of both xplA or other RDX biological degradation mechanism and carbon substrates were necessary for RDX degradation in Eglin training range soil samples. This observation is consistent with the role of xplA as a nitrogen releasing mechanism in bacteria isolates growing on RDX.

The finding that carbon addition is necessary for RDX degradation is consistent with our original hypothesis that carbon from plant root exudates could drive RDX degradation in soils. Unfortunately, our efforts to label soil communities that obtained carbon in rhizospheres from plant exudates failed, so we were not able to identify microbial communities that obtained nitrogen from RDX and carbon from root exudates. Also our efforts to associate model plant root exudate profiles with root colonization were confounded by high variability of root exudate compositions between replicates cultured under apparently identical conditions.

We were not able to induce RDX degradation by the known RDX degrader, R. rhodochrous 11Y when it was used to inoculate soil growing Arabidopsis plants, suggesting that the bacterium was not able to utilize exudates of this plant.

The idea that plants could provide sufficient carbon to drive RDX degradation as a nitrogen source for bacterial growth has problems for its use as a remediation strategy. Many training ranges are in arid regions with sparse plant growth and low plant productivity. Disruption of soils and plant populations near active targets limit plant productivity immediately after target use when leaching potential is high. Plant growth will be inhibited by the phytotoxicity of TNT. Training in winter is likely to result in contamination and leaching under environmental conditions in which plant activity is nonexistent or minimal. Apparently, since RDX has leached from verdant training ranges despite the presence of plants, the provision of plant exudate carbon to plant rhizospheres is an insufficient condition for significant degradation of RDX.

Our analysis of Eglin training range soils show that training range soil samples taken near a target site have potential for in-situ remediation of RDX if two conditions are met: presence of RDX degrading bacteria and sufficient labile carbon to drive bacterial growth using RDX as a nitrogen source. These necessary conditions could be provided by technological means in several ways.

Bioaugmentation of training range soils with bacteria carrying RDX degradation genes (e.g., xplA/B). Bioaugmentation, or the application of suspensions of live bacteria carrying the genetic potential for RDX degradation, could increase the likelihood that munitions particles would be located in soil containing bacteria able to degrade RDX. We have in hand several bacterial species able to rapidly degrade RDX. These species are native to soils and we have demonstrated their activity in soil suspensions when added to soils and provided sufficient carbon.

Bioaugmentation could be accomplished by broadcasting bacterial suspensions over the training range by ground machinery, manned or remote controlled, or by aerial application. This method would expensive and wasteful, since most training range area is not exposed to significant munitions particulates. Application of bioaugmentation suspensions could be restricted to the vicinity of active target areas. This refined method would be less wasteful and potentially less expensive than broadcasting, but would require frequent sorries of application equipment onto the training range and logistical coordination with training exercises to identify recently used targets. A significant drawback to this method is that the frequent presence of remediation equipment on or over training ranges may degrade the training mission.

Bioaugmentation may be able to satisfy the necessary condition for RDX degradation that the genetic potential for RDX degradation be present in the training range soil, but without

sufficient labile carbon, RDX degradation will be nil, as demonstrated by our experiments. Perhaps, with RDX degrading bacteria supplied to plant rhizospheres by bioaugmentation, vegetated soils in the vicinity of targets would be able to better degrade RDX, resulting in sufficient reduction of RDX in leachates to prevent levels of RDX in groundwaters exceeding regulation. Alternatively, labile carbon could be added to the soil technologically to stimulate the growth of bacteria using RDX as a nitrogen source. The addition of carbon or nutrients to stimulate biological degradation of pollutants is known as biostimulation and is frequently used to promote reductive dechlorination in groundwaters polluted with chlorinated hydrocarbons. Biostimulation has a record of successful applications in aquifers demonstrated to have the genetic potential for full reductive dechlorination of trichloroethylene (i.e., the presence of Dehalococcoides spp.). Carbon sources, such as glucose, glycerol, and succinate, or mixed carbon compounds such as molasses, could be applied with or in parallel with bioaugmentation suspensions using ground or aerial equipment, manned or unmanned.

As with bioaugmentation for reductive dechlorination, we would expect that most of the biostimulant carbon would be wasted by uptake for metabolisms unrelated to pollutant degradation. Degradation of RDX leaching from munitions particulates during intermittent rainfall events may be maximized by employing slow-release carbon sources (e.g., encapsulated or polymerized).

An unconventional method for delivery of bioaugmentation suspensions and biostimulation carbon sources would be to package the components into shells capable of firing from the various weapons used on the training range. These shells could be fired at the target during training excerises, dispersing the bioaugmentation/biostimulant mixtures over the vicinity of the target. This ballistic approach may be less expensive than other methods for application of bioaugmentation/stimulation suspensions.

The practicality and efficacy of bioaugmentation and/or biostimulation of RDX degradation in situ on training ranges depends on several parameters which should be explored by experimentation. These include:

- Viability of RDX-degrading bacterial suspensions in the soil environment
    - Which species would be optimal? Rhodococcus, Microbacterium, and Mycobacteria species are easily cultured in the lab and degrade RDX rapidly. All were isolated from soils.
    - Which of those species survive best in soil environments?
    - Can encapsulation or other biopreservative methods enhance the longevity of RDX degrading populations grown in the lab and applied to soils?
- Can RDX-degrading bacteria be grown to large biomasses while retaining the RDX-degradation phenotype?
    - Can RDX or a surrogate be used as a nitrogen source to select for RDX degradation in large biomass cultures?
- What is the best delivery method for the bacteria: spraying of a wet solution or a dry powder?
- How long will labile organic carbon applied to soil surfaces last?
    - Can encapsulation increase the sustainable availability of organic carbon

- Develop an empirical model of RDX degradation potential based on the findings of experiments described abovel. Use the model to estimate the practical timeframe for persistence of bioaugmentation/stimulation. Would in situ remediation potential be sufficient to degrade RDX leached from munitions particles by rainfall or snow melt at a given site?

These considerations suggest that alternative in situ RDX bioremediation methods to prevent training range munitions contamination of groundwaters may be possible.

VI. Appendices

A. Eglin Air Force Base sample information

This appendix contains the descriptions and images of the soil samples taken at Eglin Air Force Base used in the SIP analyses.

Table A.1. Soil samples taken at Eglin Air Force Base at C-52N Cat's Eye.

Sample #	Sample Name and Description	Plant - Plant Name
1	Inner pit1_ bag	
2	Inner pit1_tube 12	*Paspalum notatum* (bahiagrass)
3	Inner pit1_tube 13	
4	Inner pit1_tube 11	Unidentified plants
5	Pit1_edge_tube 10	Near: *Eupatorium compositifolium* (yankeeweed)
6	Pit1_edge_tube 6	*Quercus geminata* (sand live oak)
7	Pit1_edge_tube 7	Near: *Yucca filamentosa* (yucca)
8	Site 2_#9	Unidentified weed
9	Site 2_#8	*Eupatorium compositifolium* (yankeeweed)
10	Site 2 mound	
11	Site 2 middle	
12	Site 3 tube 4	*Cynodon dactylon* (bermudagrass)
13	Site 3 tube 5	*Cynodon dactylon* (bermudagrass)
14	2nd pit	*Cynodon dactylon* (bermudagrass)
15	Site 4 tube 1	*Paspalum notatum* (bahiagrass)
16	Site 4 tube 2	
17	Site 4 tube 3	Unidentified plant
18	Site 4 patch	*Cynodon dactylon* (bermudagrass) or bahiagrass
19	Pit 3 grass	Unidentified plant
20	Plant near red patch	*Quercus geminata* (sand live oak)
21	Red patch	
22	Near Patch	*Schizachyrium tenerum* (creeping bluestem)
23	Pit 3 edge	

Images of the above soil samples (except for Sample #23 which could not be paired with an image) are shown below. In a few instances the exact sampling location is circled if it is visible but not obvious.

Figure A.1. Sample 1 taken from the interior of pit #1 (Site 1).

Figure A.2. Sample 2 taken from bahiagrass rhizosphere and surrcounding soil inside pit #1 (Site 1).

Figure A.3. Sample 3 taken from the interior of pit #1 (Site 1).

Figure A.4. Sample 4 taken from within a patch of unidentified grasses inside pit #1 (Site 1).

Figure A.5. Sample 5 taken outside pit #1 near a yankeeweed bush.

Figure A.6. Sample 6 taken at the base of a sand live oak outside of pit #1.

Figure A.7. Sample 7 taken outside of pit #1 near a yucca.

Figure A.8. Sample 8 taken beneath an unknown weed at Site #2.

Figure A.9. Sample 9 taken from beneath the inside of a yankeeweed bush at Site #2.

Figure A.10. Sample 10 was taken from a barren mound at Site #2.

Figure A.11. Sample 11 taken from the middle of an area with a lot of debris at Site #2.

Figure A.12. Sample 12 was taken from the middle of bermudagrass in a grassland area (Site #3).

Figure A.13. Sample 13 taken from bermudagrass grassland area (Site #3).

Figure A.14. Sample 14 taken from the middle of bermudagrass patch in pit #2.

Figure A.15. Sample 15 taken from a patch of bahiagrass at Site 4.

Figure A.16. Sample 16 was taken from a moist soil patch at Site #4.

Figure A.17. Sample 17 was taken within an unidentified plant at Site #4.

Figure A.18. Sample 18 was taken from a patch of Bermuda grass or bahiagrass at Site #4.

Figure A.19. Sample 19 was taken from within unidentified Pit #3.

Figure A.20. Sample 20 was taken near a sand live oak close to a red soil patch.

Figure A.21. Sample 21 was taken from a patch of dry red soil.

Figure A.22. Sample 22 was taken from a patch of bluestem grass near the red soil patch.

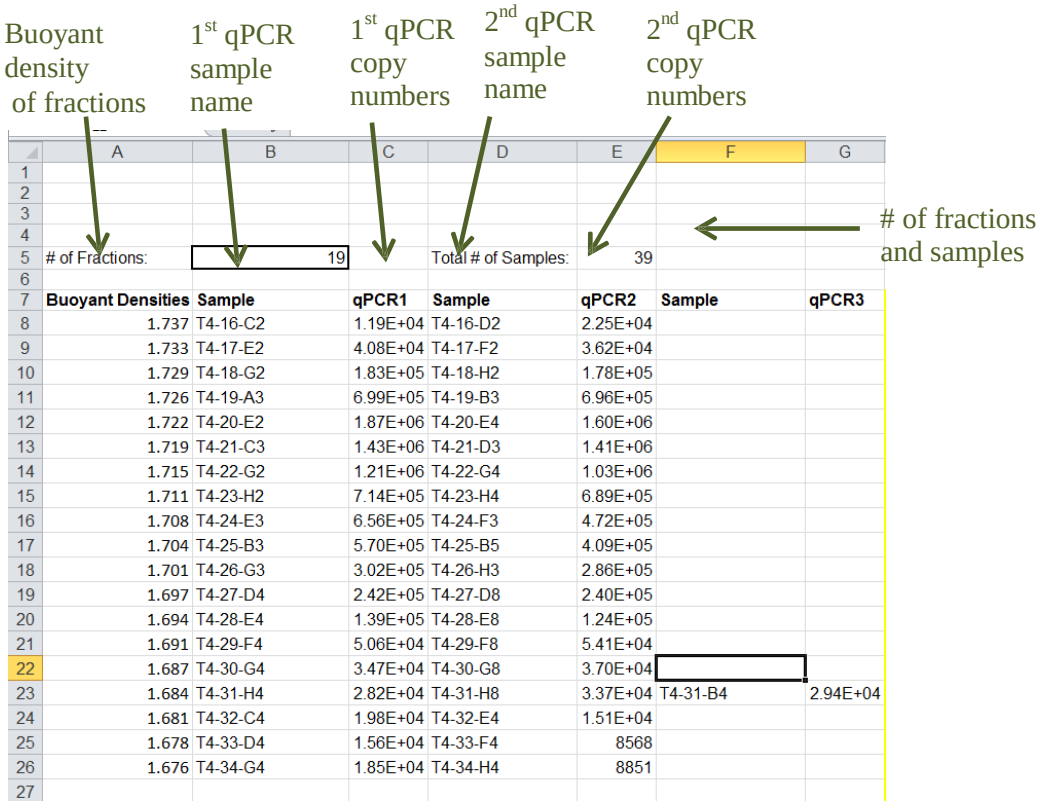
B. SIP analysis workflow

These pages describe the worksheets used to analyze the qPCR-TRFLP data generated for the density gradients. Three worksheets were used in total. In the descriptions, green labels indicate data that the user inputs into the sheet. Blue labels indicate output from the Visual Basic scripts.

Worksheet #1 – RF quantification

Filename = " Updated_format_and_bin". To run: use the command='ctrl+y'

There are several sheets in this worksheet, but only a few are currently in use (others either haven't been integrated or were made for troubleshooting purposes. The relevant pages are: 'qPCR_data_in', 'rflp_in' and 'output_and_review'. 'Sorting' and 'calc and organize' are used in the scripts but don't need to be used directly by the user.



	A	B	C	D	E	F	G
1							
2							
3							
4							
5	# of Fractions:	19		Total # of Samples:	39		
6							
7	**Buoyant Densities**	**Sample**	**qPCR1**	**Sample**	**qPCR2**	**Sample**	**qPCR3**
8	1.737	T4-16-C2	1.19E+04	T4-16-D2	2.25E+04		
9	1.733	T4-17-E2	4.08E+04	T4-17-F2	3.62E+04		
10	1.729	T4-18-G2	1.83E+05	T4-18-H2	1.78E+05		
11	1.726	T4-19-A3	6.99E+05	T4-19-B3	6.96E+05		
12	1.722	T4-20-E2	1.87E+06	T4-20-E4	1.60E+06		
13	1.719	T4-21-C3	1.43E+06	T4-21-D3	1.41E+06		
14	1.715	T4-22-G2	1.21E+06	T4-22-G4	1.03E+06		
15	1.711	T4-23-H2	7.14E+05	T4-23-H4	6.89E+05		
16	1.708	T4-24-E3	6.56E+05	T4-24-F3	4.72E+05		
17	1.704	T4-25-B3	5.70E+05	T4-25-B5	4.09E+05		
18	1.701	T4-26-G3	3.02E+05	T4-26-H3	2.86E+05		
19	1.697	T4-27-D4	2.42E+05	T4-27-D8	2.40E+05		
20	1.694	T4-28-E4	1.39E+05	T4-28-E8	1.24E+05		
21	1.691	T4-29-F4	5.06E+04	T4-29-F8	5.41E+04		
22	1.687	T4-30-G4	3.47E+04	T4-30-G8	3.70E+04		
23	1.684	T4-31-H4	2.82E+04	T4-31-H8	3.37E+04	T4-31-B4	2.94E+04
24	1.681	T4-32-C4	1.98E+04	T4-32-E4	1.51E+04		
25	1.678	T4-33-D4	1.56E+04	T4-33-F4	8568		
26	1.676	T4-34-G4	1.85E+04	T4-34-H4	8851		
27							

Figure B.1. Screenshot of 'qPCR_data_in' where qPCR data is entered.

On the 'qPCR_data_in' sheet, shown above, the user inputs: the fraction buoyant densities, the qPCR sample names, the qPCR copy number for the corresponding sample, the total number of fractions and the total number of samples. Copy numbers and sample descriptions are repeated for each

replicate. There are spots for up to five replicates; however the script is only constructed at the moment for 1 – 3 replicates.

Then, the RFLP data from DAx is pasted onto the 'rflp_in' sheet. The sheet inputs the sample names from the 'qPCR_data_in' sheet and displays them in the boxed cells. This indicates where the RFLP data corresponding to each qPCR data is to be pasted.

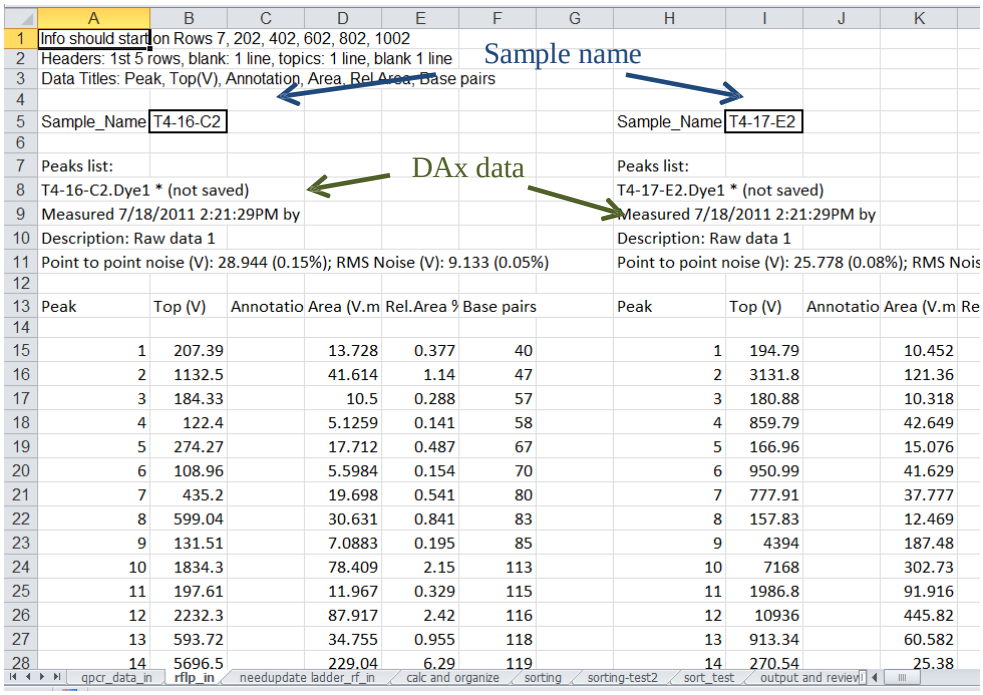
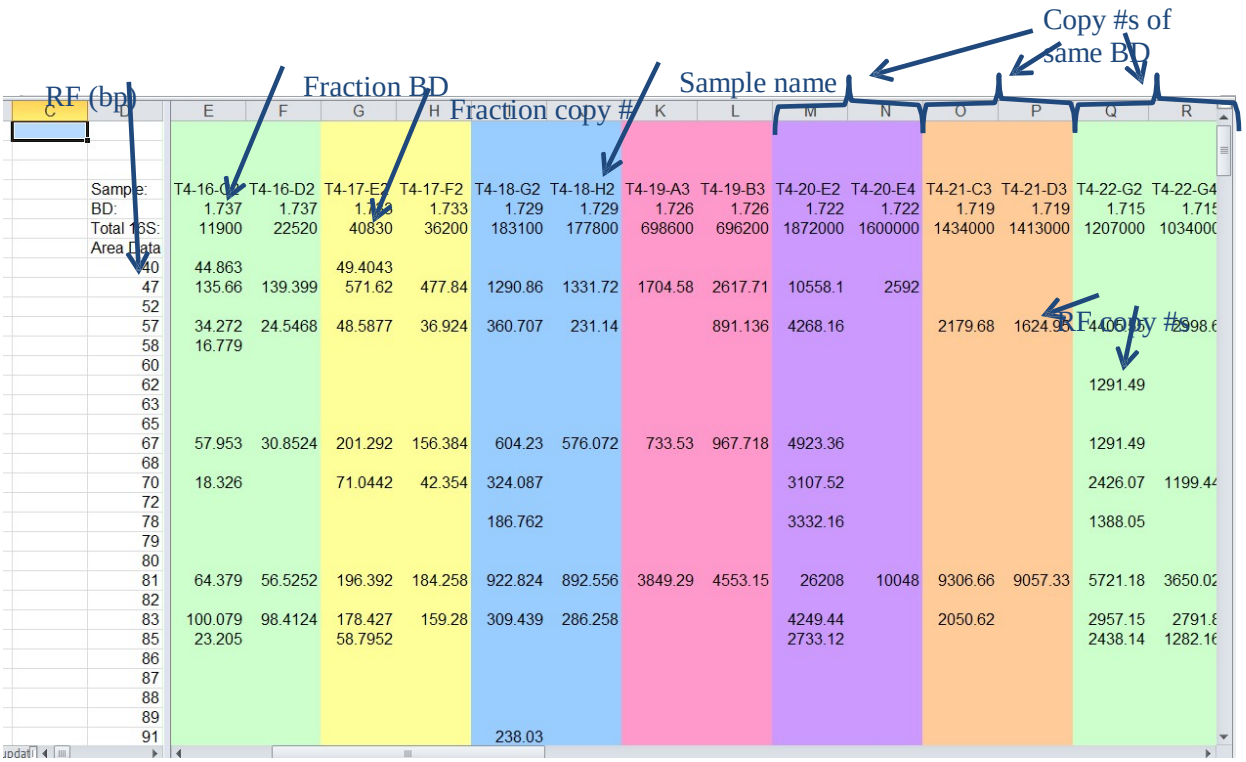


Figure B.2. Screenshot of 'rflp_in' where DAx RFLP data is entered under the appropriate sample name.

To run the script, use the shortcut 'ctrl+xxx' or goto script 'xxx' in the VBscript and run. The script calculates relative heights (relative area is already in the DAx output) organizes the data, indicates if there are duplicate entries for a bp. If so, the user is to input a replacement for the second data point. For example, if a RFLP data set contains two values for 40 bp, the user should look at the next available bp and enter this number (i.e. 43). After the script is finished, it may be necessary to determine if bps should be summed, realigned, etc. and possibly rerun the script after changing the input data. The output is in the 'output_and_review' sheet.

Below, is an example of the output worksheet. The copy numbers calculated from relative area are displayed below, data from relative height are displayed below the relative area data in the same

format. The samples are color-coded by fraction so replicate samples are colored the same to help review the data. The user can then go through and bin samples accordingly by comparing replicate samples and RF values across the gradient.



	E	F	G	H	K	L	M	N	O	P	Q	R
Sample:	T4-16-D2	T4-16-D2	T4-17-E2	T4-17-F2	T4-18-G2	T4-18-H2	T4-19-A3	T4-19-B3	T4-20-E2	T4-20-E4	T4-21-C3	T4-21-D3	T4-22-G2	T4-22-G4
BD:	1.737	1.737	1.733	1.733	1.729	1.729	1.726	1.726	1.722	1.722	1.719	1.719	1.715	1.715
Total 16S:	11900	22520	40830	36200	183100	177800	698600	696200	1872000	1600000	1434000	1413000	1207000	1034000
Area Data														
40	44.863		49.4043											
47	135.66	139.399	571.62	477.84	1290.86	1331.72	1704.58	2617.71	10558.1	2592				
52														
57	34.272	24.5468	48.5877	36.924	360.707	231.14		891.136	4268.16		2179.68	1624.9	4406.6	2998.6
58	16.779													
60														
62													1291.49	
63														
65														
67	57.953	30.8524	201.292	156.384	604.23	576.072	733.53	967.718	4923.36				1291.49	
68														
70	18.326		71.0442	42.354	324.087				3107.52				2426.07	1199.44
72														
78					186.762				3332.16				1388.05	
79														
80														
81	64.379	56.5252	196.392	184.258	922.824	892.556	3849.29	4553.15	26208	10048	9306.66	9057.33	5721.18	3650.02
82														
83	100.079	98.4124	178.427	159.28	309.439	286.258			4249.44		2050.62		2957.15	2791.8
85	23.205		58.7952						2733.12				2438.14	1282.16
86														
87														
88														
89														
91					238.03									

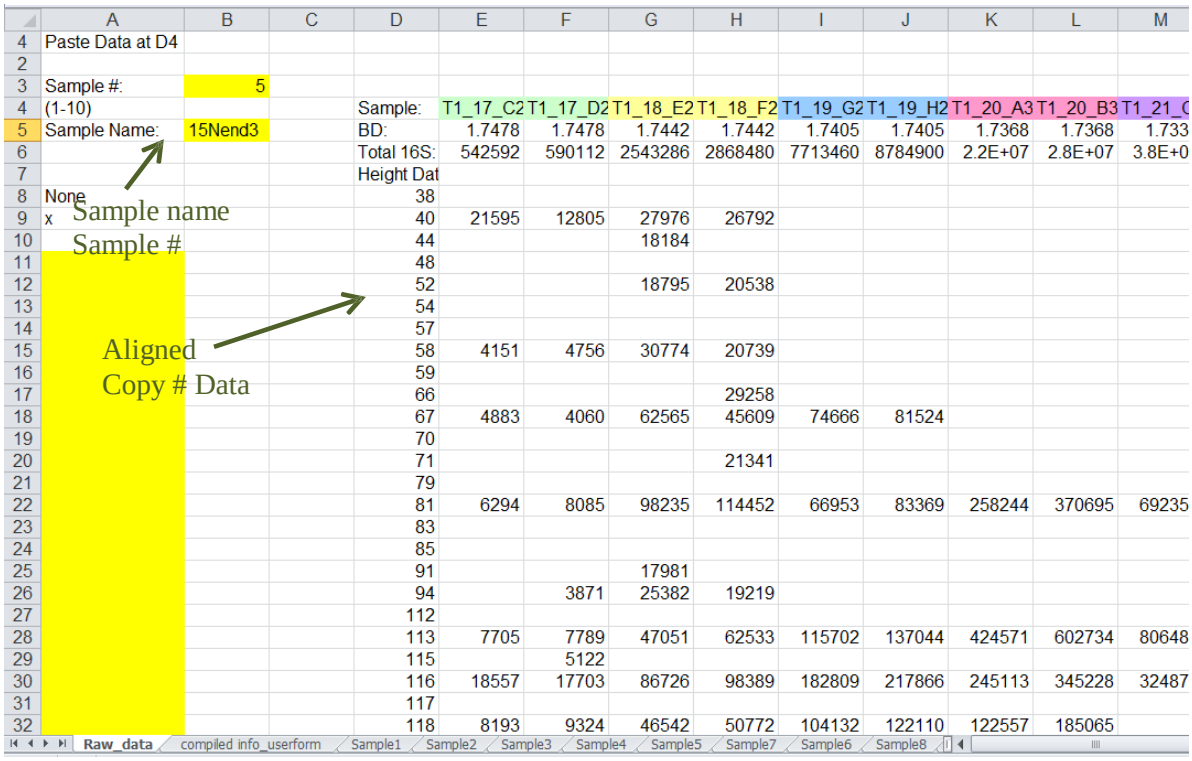
Figure B.3. Screenshot of the output from Worksheet #1. Replicate samples are the same color.

After reviewing and binning the data, save the file and copy the data to paste into the next sheet (either the data calculated based on ‘relative area’, ‘relative height’ or all of it). Open the sheet ‘xxxxx’ for the next portion.

Worksheet 2: RF average values and standard deviation calculations.

Filename=”Updated_avg_sheet”. To run: Run module 1 (Sub Macro18 ( ))

The data from the ‘output and review’ sheet are pasted onto the ‘Raw_data’ sheet starting at cell D4. Also, the user indicates the sample # (1 – 10) and the sample name. The sample # corresponds to the sheet where the output data will be recorded. If there is already data on the sheet, an error message will appear directing you that the script will not run. This is to prevent erasing data on accident.



	A	B	C	D	E	F	G	H	I	J	K	L	M
4	Paste Data at D4												
2													
3	Sample #:	5											
4	(1-10)			Sample:	T1_17_C2	T1_17_D2	T1_18_E2	T1_18_F2	T1_19_G2	T1_19_H2	T1_20_A3	T1_20_B3	T1_21_C
5	Sample Name:	15Nend3		BD:	1.7478	1.7478	1.7442	1.7442	1.7405	1.7405	1.7368	1.7368	1.733
6				Total 16S:	542592	590112	2543286	2868480	7713460	8784900	2.2E+07	2.8E+07	3.8E+0
7				Height Dat									
8	None			38									
9	x			40	21595	12805	27976	26792					
10				44			18184						
11				48									
12				52			18795	20538					
13				54									
14				57									
15				58	4151	4756	30774	20739					
16				59									
17				66				29258					
18				67	4883	4060	62565	45609	74666	81524			
19				70									
20				71				21341					
21				79									
22				81	6294	8085	98235	114452	66953	83369	258244	370695	69235
23				83									
24				85									
25				91			17981						
26				94		3871	25382	19219					
27				112									
28				113	7705	7789	47051	62533	115702	137044	424571	602734	80648
29				115		5122							
30				116	18557	17703	86726	98389	182809	217866	245113	345228	32487
31				117									
32				118	8193	9324	46542	50772	104132	122110	122557	185065	

Figure B.4. Screenshot of the input sheet of Worksheet #2. Output from Worksheet 1 is pasted into the 'Raw_data' sheet along with the sample name and a number specifying the sheet to paste the output.

On the sample output sheet, the following information is recorded: for each RF in each fraction, the average copy number values, the standard deviation values and the number of samples. (Next three images).

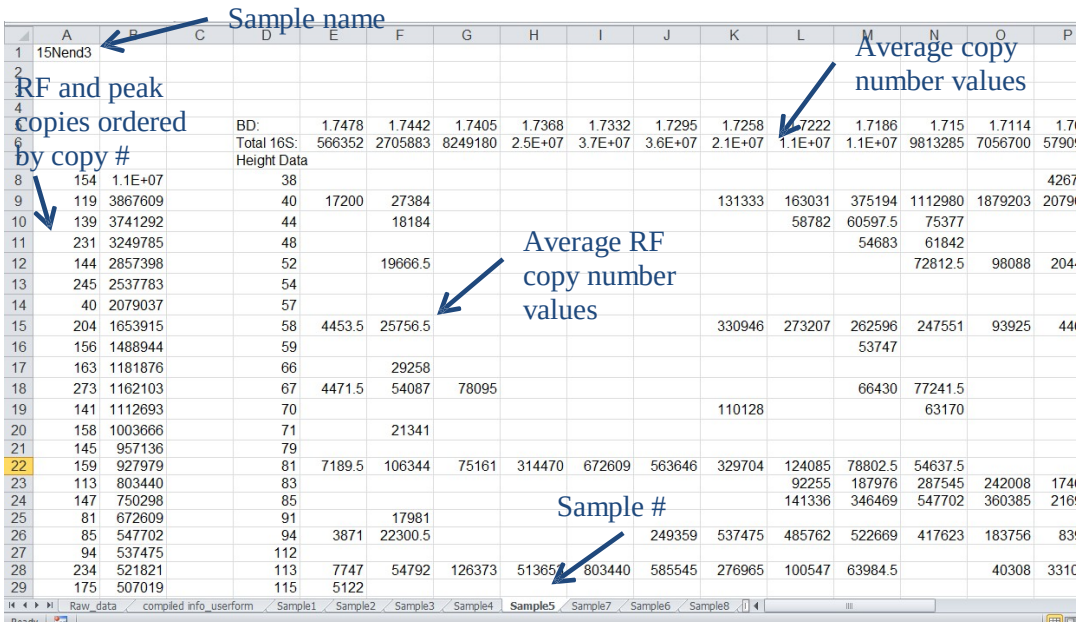


Figure B.5. Screenshot of Worksheet #2 output. The left portion of the sheet is shown which reports the average values.

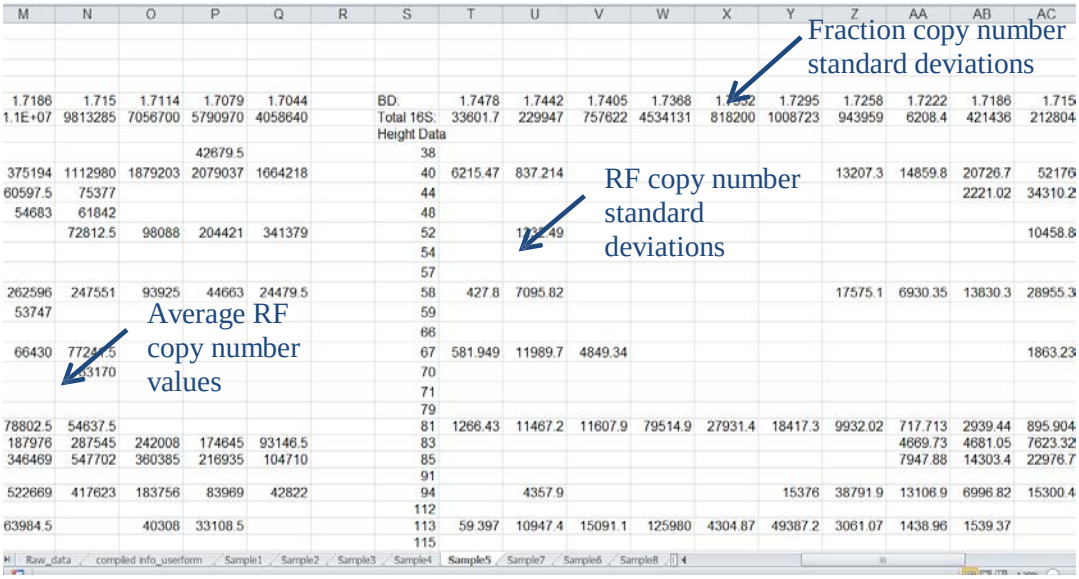
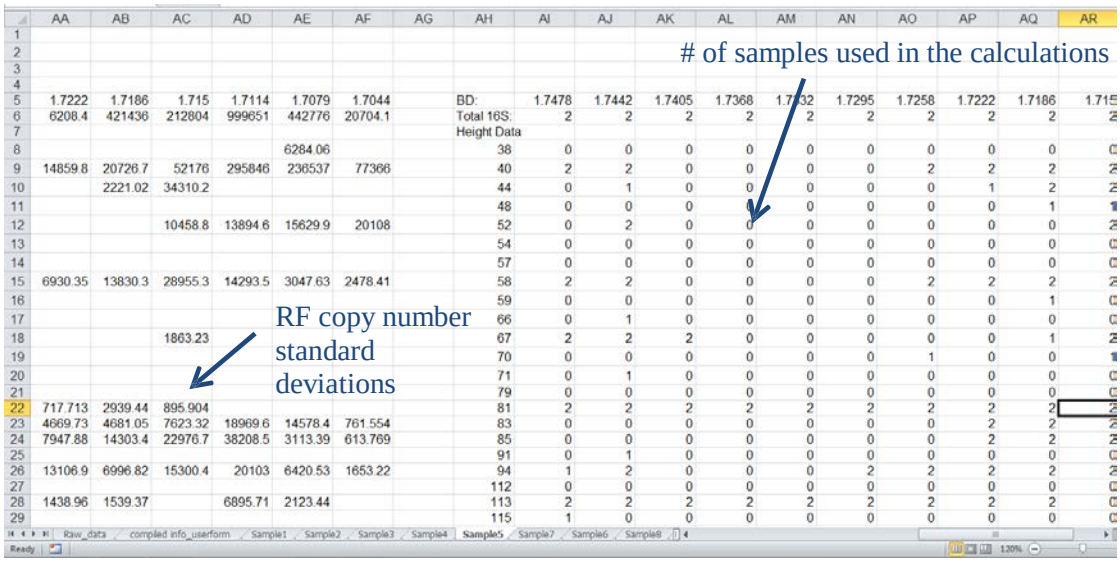


Figure B.6. Screenshot of Worksheet #2 output. The middle portion of the sheet is shown which reports the standard deviations.

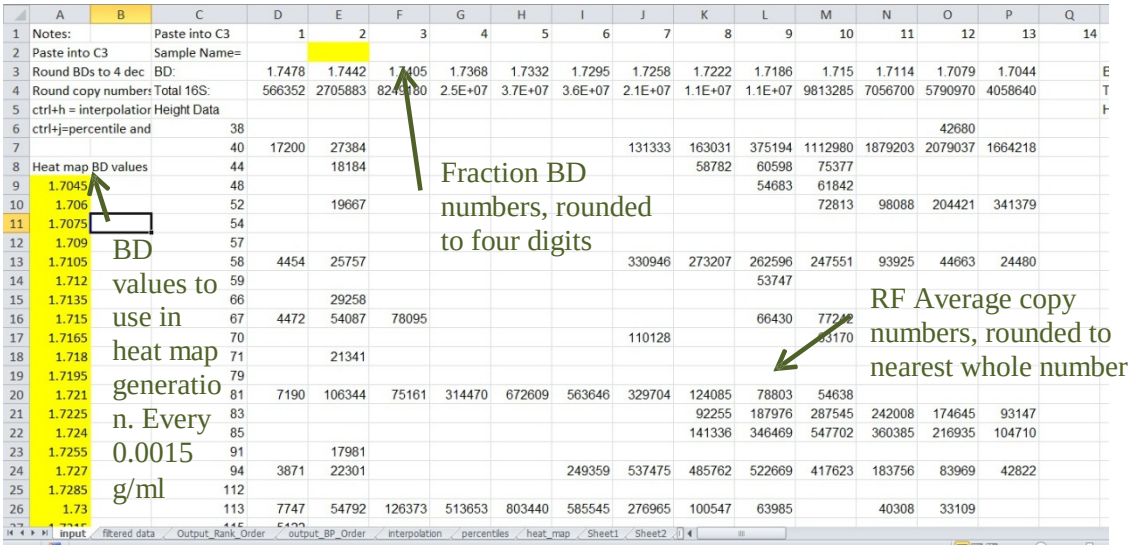


	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR
5	1.7222	1.7186	1.715	1.7114	1.7079	1.7044		BD:	1.7478	1.7442	1.7405	1.7368	1.7332	1.7295	1.7258	1.7222	1.7186	1.715
6	6208.4	421436	212804	999651	442776	20704.1		Total 16S:	2	2	2	2	2	2	2	2	2	2
7								Height Data										
8					6284.06			38	0	0	0	0	0	0	0	0	0	0
9	14859.8	20726.7	52176	295846	236537	77366		40	2	2	0	0	0	0	2	2	2	2
10		2221.02	34310.2					44	0	1	0	0	0	0	0	1	2	2
11								48	0	0	0		0	0	0	0	1	1
12			10458.8	13894.6	15629.9	20108		52	0	2	0	0	0	0	0	0	0	2
13								54	0	0	0	0	0	0	0	0	0	0
14								57	0	0	0	0	0	0	0	0	0	0
15	6930.35	13830.3	28955.3	14293.5	3047.63	2478.41		58	2	2	0	0	0	0	2	2	2	2
16								59	0	0	0	0	0	0	0	0	1	0
17								66	0	1	0	0	0	0	0	0	0	0
18			1863.23					67	2	2	2	0	0	0	0	0	1	2
19								70	0	0	0	0	0	0	1	0	0	1
20								71	0	1	0	0	0	0	0	0	0	0
21								79	0	0	0	0	0	0	0	0	0	0
22	717.713	2939.44	895.904					81	2	2	2	2	2	2	2	2	2	2
23	4669.73	4681.05	7623.32	18969.6	14578.4	761.554		83	0	0	0	0	0	0	0	2	2	2
24	7947.88	14303.4	22976.7	38208.5	3113.39	613.769		85	0	0	0	0	0	0	0	2	2	2
25								91	0	1	0	0	0	0	0	0	0	0
26	13106.9	6996.82	15300.4	20103	6420.53	1653.22		94	1	2	0	0	0	2	2	2	2	2
27								112	0	0	0	0	0	0	0	0	0	0
28	1438.96	1539.37		6895.71	2123.44			113	2	2	2	2	2	2	2	2	2	0
29								115	1	0	0	0	0	0	0	0	0	0

Figure B.7. Screenshot of Worksheet #2 output. The right portion of the sheet is shown which reports the number of samples used in each calculation.

Worksheet #3. Calculation of p values for RF copy numbers, and generation of heat maps. Filename=' Updated_statistical_heatmap.xlsm'. To run: use the command = 'ctrl+h'

Paste the data from the second worksheet (Average copy numbers, standard deviations and number of samples used in the calculations into the sheet labeled 'Input'. For the script to run properly, round the BD values to four decimal places and the copy numbers to the nearest whole number. If errors occur when the script is run, make sure the columns are wide enough so the data is visible (the script uses the VB 'find' function which requires that the data doesn't have too many decimal places and that the number is visible. In the yellow boxes to the left, list BD values for the heat map between the BD values that have data (make sure the values are not outside the range of the data values) in ascending order.



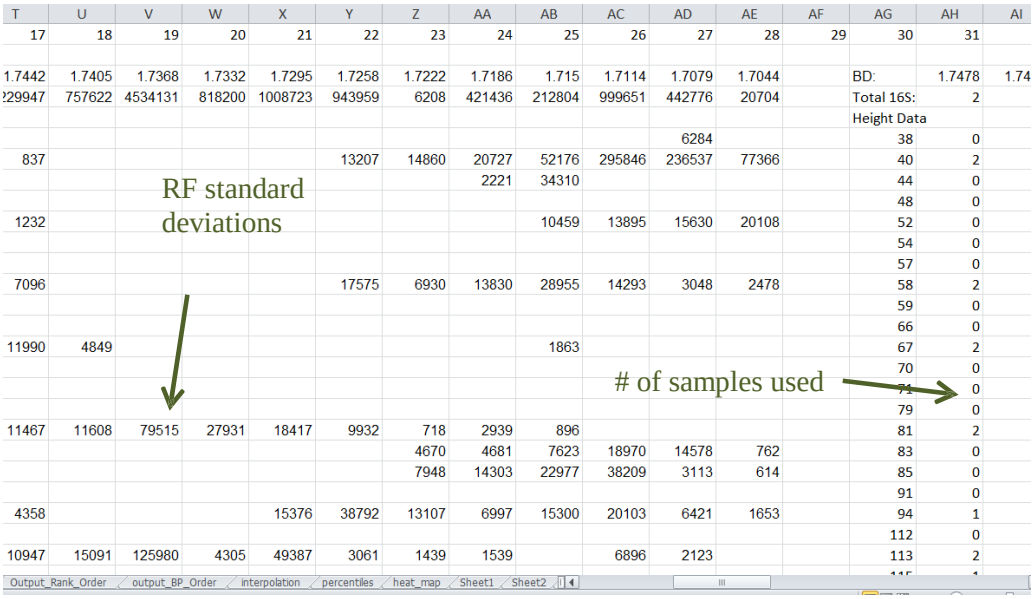
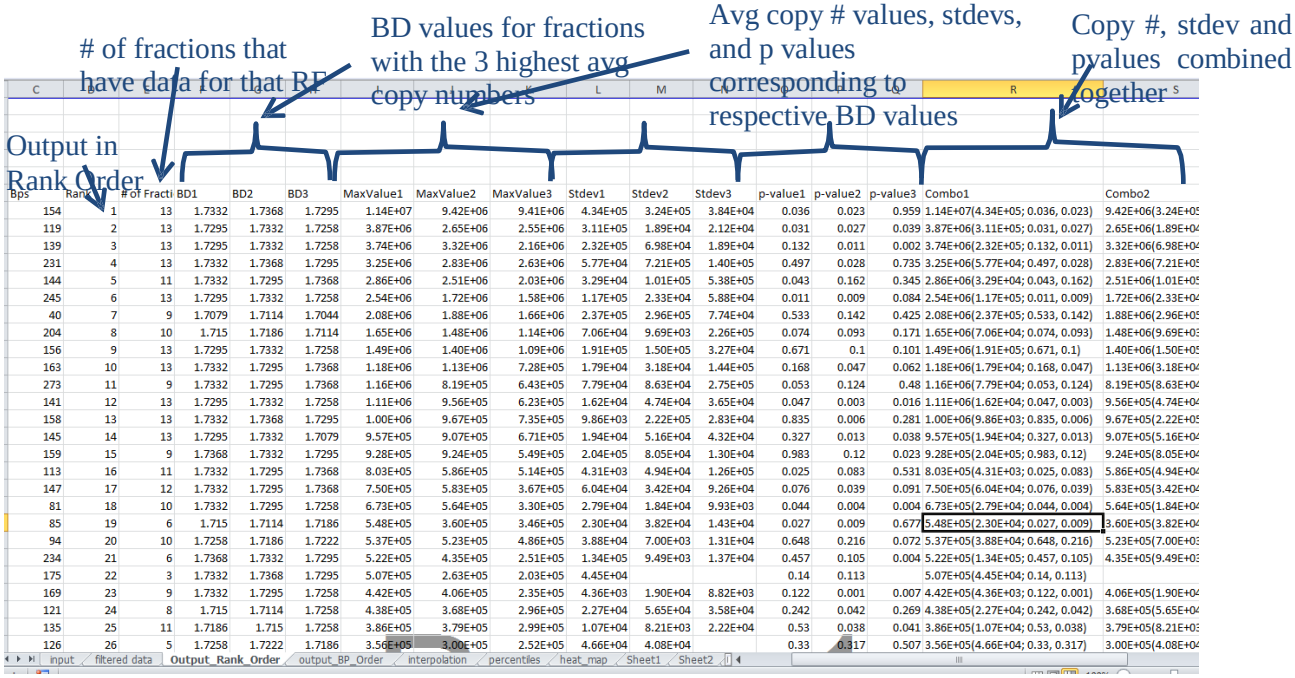


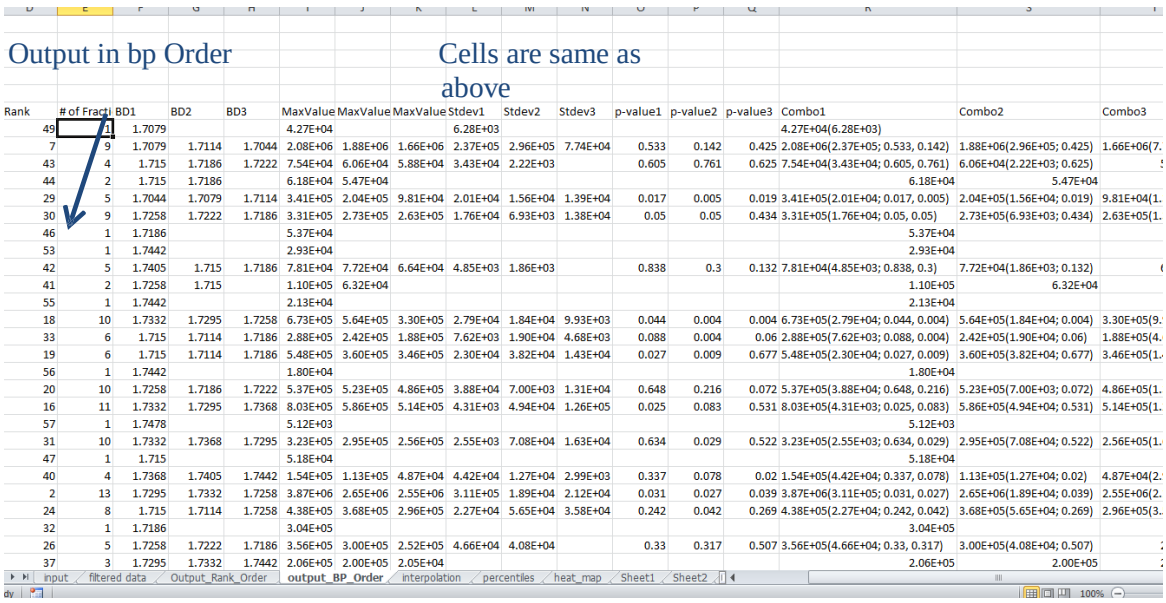
Figure B.8. Screenshots of sheet 'Input' in Worksheet 3 with data copied from Worksheet 2 and rounded. The top image shows the average data and the bottom image the standard deviation and number of samples used. The BD values for the heat maps have been entered into the yellow boxed cells (top image) in ascending order based on the data set.

The data is outputted into two sheets: 'Output_Rank_Order' and 'Output_bp_order'. The data is the same on both pages but sorted differently. For each RF, the average copy numbers are compared and the three highest average values are recovered, numbered in descending order along with the corresponding BD values and standard deviations. p values are (p value1 = t-test between copy #1

and copy #2, p value 2 = t-test between copy #1 and copy #3 and p value 3 = t-test between copy 2 and copy 3. The last three columns are the above data combined together for copy into Table.



Bps	Rank	# of Fracti	BD1	BD2	BD3	MaxValue1	MaxValue2	MaxValue3	Stdev1	Stdev2	Stdev3	p-value1	p-value2	p-value3	Combo1	Combo2
154	1	13	1.7332	1.7368	1.7295	1.14E+07	9.42E+06	9.41E+06	4.34E+05	3.24E+05	3.84E+04	0.036	0.023	0.959	1.14E+07(4.34E+05; 0.036, 0.023)	9.42E+06(3.24E+05
119	2	13	1.7295	1.7332	1.7258	3.87E+06	2.65E+06	2.55E+06	3.11E+05	1.89E+04	2.12E+04	0.031	0.027	0.039	3.87E+06(3.11E+05; 0.031, 0.027)	2.65E+06(1.89E+04
139	3	13	1.7295	1.7332	1.7258	3.74E+06	3.32E+06	2.16E+06	2.32E+05	6.98E+04	1.89E+04	0.132	0.011	0.002	3.74E+06(2.32E+05; 0.132, 0.011)	3.32E+06(6.98E+04
231	4	13	1.7332	1.7368	1.7295	3.25E+06	2.83E+06	2.63E+06	5.77E+04	7.21E+05	1.40E+05	0.497	0.028	0.735	3.25E+06(5.77E+04; 0.497, 0.028)	2.83E+06(7.21E+05
144	5	11	1.7332	1.7295	1.7368	2.86E+06	2.51E+06	2.03E+06	3.29E+04	1.01E+05	5.38E+05	0.043	0.162	0.345	2.86E+06(3.29E+04; 0.043, 0.162)	2.51E+06(1.01E+05
245	6	13	1.7295	1.7332	1.7258	2.54E+06	1.72E+06	1.58E+06	1.17E+05	2.33E+04	5.88E+04	0.011	0.009	0.084	2.54E+06(1.17E+05; 0.011, 0.009)	1.72E+06(2.33E+04
40	7	9	1.7079	1.7114	1.7044	2.08E+06	1.88E+06	1.66E+06	2.37E+05	2.96E+05	7.74E+04	0.533	0.142	0.425	2.08E+06(2.37E+05; 0.533, 0.142)	1.88E+06(2.96E+05
204	8	10	1.715	1.7186	1.7114	1.65E+06	1.48E+06	1.14E+06	7.06E+04	9.69E+03	2.26E+05	0.074	0.093	0.171	1.65E+06(7.06E+04; 0.074, 0.093)	1.48E+06(9.69E+03
156	9	13	1.7295	1.7332	1.7258	1.49E+06	1.40E+06	1.09E+06	1.91E+05	1.50E+05	3.27E+04	0.671	0.1	0.101	1.49E+06(1.91E+05; 0.671, 0.1)	1.40E+06(1.50E+05
163	10	13	1.7332	1.7295	1.7368	1.18E+06	1.13E+06	7.28E+05	1.79E+04	3.18E+04	1.44E+05	0.168	0.047	0.062	1.18E+06(1.79E+04; 0.168, 0.047)	1.13E+06(3.18E+04
273	11	9	1.7332	1.7295	1.7368	1.16E+06	8.19E+05	6.43E+05	7.79E+04	8.63E+04	2.75E+05	0.053	0.124	0.48	1.16E+06(7.79E+04; 0.053, 0.124)	8.19E+05(8.63E+04
141	12	13	1.7295	1.7332	1.7258	1.11E+06	9.56E+05	6.23E+05	1.62E+04	4.74E+04	3.65E+04	0.047	0.003	0.016	1.11E+06(1.62E+04; 0.047, 0.003)	9.56E+05(4.74E+04
158	13	13	1.7332	1.7368	1.7295	1.00E+06	9.67E+05	7.35E+05	9.86E+03	2.22E+05	2.83E+04	0.835	0.006	0.281	1.00E+06(9.86E+03; 0.835, 0.006)	9.67E+05(2.22E+05
145	14	13	1.7295	1.7332	1.7079	9.57E+05	9.07E+05	6.71E+05	1.94E+04	5.16E+04	4.32E+04	0.327	0.013	0.038	9.57E+05(1.94E+04; 0.327, 0.013)	9.07E+05(5.16E+04
159	15	9	1.7368	1.7332	1.7295	9.28E+05	9.24E+05	5.49E+05	2.04E+05	8.05E+04	1.30E+04	0.983	0.12	0.023	9.28E+05(2.04E+05; 0.983, 0.12)	9.24E+05(8.05E+04
113	16	11	1.7332	1.7295	1.7368	8.03E+05	5.86E+05	5.14E+05	4.31E+03	4.94E+04	1.26E+05	0.025	0.083	0.531	8.03E+05(4.31E+03; 0.025, 0.083)	5.86E+05(4.94E+04
147	17	12	1.7332	1.7295	1.7368	7.50E+05	5.83E+05	3.67E+05	6.04E+04	3.42E+04	9.26E+04	0.076	0.039	0.091	7.50E+05(6.04E+04; 0.076, 0.039)	5.83E+05(3.42E+04
81	18	10	1.7332	1.7295	1.7258	6.73E+05	5.64E+05	3.30E+05	2.79E+04	1.84E+04	9.93E+03	0.044	0.004	0.004	6.73E+05(2.79E+04; 0.044, 0.004)	5.64E+05(1.84E+04
85	19	6	1.715	1.7114	1.7186	5.48E+05	3.60E+05	3.46E+05	2.30E+04	3.82E+04	1.43E+04	0.027	0.009	0.677	5.48E+05(2.30E+04; 0.027, 0.009)	3.60E+05(3.82E+04
94	20	10	1.7258	1.7186	1.7222	5.37E+05	5.23E+05	4.86E+05	3.88E+04	7.00E+03	1.31E+04	0.648	0.216	0.072	5.37E+05(3.88E+04; 0.648, 0.216)	5.23E+05(7.00E+03
234	21	6	1.7368	1.7332	1.7295	5.22E+05	4.35E+05	2.51E+05	1.34E+05	9.49E+03	1.37E+04	0.457	0.105	0.004	5.22E+05(1.34E+05; 0.457, 0.105)	4.35E+05(9.49E+03
175	22	3	1.7332	1.7368	1.7295	5.07E+05	2.63E+05	2.03E+05	4.45E+04			0.14	0.113		5.07E+05(4.45E+04; 0.14, 0.113)	
169	23	9	1.7332	1.7295	1.7258	4.42E+05	4.06E+05	2.35E+05	4.36E+03	1.90E+04	8.82E+03	0.122	0.001	0.007	4.42E+05(4.36E+03; 0.122, 0.001)	4.06E+05(1.90E+04
121	24	8	1.715	1.7114	1.7258	4.38E+05	3.68E+05	2.96E+05	2.27E+04	5.65E+04	3.58E+04	0.242	0.042	0.269	4.38E+05(2.27E+04; 0.242, 0.042)	3.68E+05(5.65E+04
135	25	11	1.7186	1.715	1.7258	3.86E+05	3.79E+05	2.99E+05	1.07E+04	8.21E+03	2.22E+04	0.53	0.038	0.041	3.86E+05(1.07E+04; 0.53, 0.038)	3.79E+05(8.21E+03
126	26	5	1.7258	1.7222	1.7186	3.56E+05	3.00E+05	2.52E+05	4.66E+04	4.08E+04		0.33	0.317	0.507	3.56E+05(4.66E+04; 0.33, 0.317)	3.00E+05(4.08E+04



Rank	# of Fracti	BD1	BD2	BD3	MaxValue	MaxValue	MaxValue	Stdev1	Stdev2	Stdev3	p-value1	p-value2	p-value3	Combo1	Combo2	Combo3
49	1	1.7079			4.27E+04			6.28E+03						4.27E+04(6.28E+03)		
7	9	1.7079	1.7114	1.7044	2.08E+06	1.88E+06	1.66E+06	2.37E+05	2.96E+05	7.74E+04	0.533	0.142	0.425	2.08E+06(2.37E+05; 0.533, 0.142)	1.88E+06(2.96E+05; 0.425)	1.66E+06(7.
43	4	1.715	1.7186	1.7222	7.54E+04	6.06E+04	5.88E+04	3.43E+04	2.22E+03		0.605	0.761	0.625	7.54E+04(3.43E+04; 0.605, 0.761)	6.06E+04(2.22E+03; 0.625)	
44	2	1.715	1.7186		6.18E+04	5.47E+04								6.18E+04	5.47E+04	
29	5	1.7044	1.7079	1.7114	3.41E+05	2.04E+05	9.81E+04	2.01E+04	1.56E+04	1.39E+04	0.017	0.005	0.019	3.41E+05(2.01E+04; 0.017, 0.005)	2.04E+05(1.56E+04; 0.019)	9.81E+04(1.
30	9	1.7258	1.7222	1.7186	3.31E+05	2.73E+05	2.63E+05	1.76E+04	6.93E+03	1.38E+04	0.05	0.05	0.434	3.31E+05(1.76E+04; 0.05, 0.05)	2.73E+05(6.93E+03; 0.434)	2.63E+05(1.
46	1	1.7186			5.37E+04									5.37E+04		
53	1	1.7442			2.93E+04									2.93E+04		
42	5	1.7405	1.715	1.7186	7.81E+04	7.72E+04	6.64E+04	4.85E+03	1.86E+03		0.838	0.3	0.132	7.81E+04(4.85E+03; 0.838, 0.3)	7.72E+04(1.86E+03; 0.132)	
41	2	1.7258	1.715		1.10E+05	6.32E+04								1.10E+05	6.32E+04	
55	1	1.7442			2.13E+04									2.13E+04		
18	10	1.7332	1.7295	1.7258	6.73E+05	5.64E+05	3.30E+05	2.79E+04	1.84E+04	9.93E+03	0.044	0.004	0.004	6.73E+05(2.79E+04; 0.044, 0.004)	5.64E+05(1.84E+04; 0.004)	3.30E+05(9.
33	6	1.715	1.7114	1.7186	2.88E+05	2.42E+05	1.88E+05	7.62E+03	1.90E+04	4.68E+03	0.088	0.004	0.06	2.88E+05(7.62E+03; 0.088, 0.004)	2.42E+05(1.90E+04; 0.06)	1.88E+05(4.
19	6	1.715	1.7114	1.7186	5.48E+05	3.60E+05	3.46E+05	2.30E+04	3.82E+04	1.43E+04	0.027	0.009	0.677	5.48E+05(2.30E+04; 0.027, 0.009)	3.60E+05(3.82E+04; 0.677)	3.46E+05(1.
56	1	1.7442			1.80E+04									1.80E+04		
20	10	1.7258	1.7186	1.7222	5.37E+05	5.23E+05	4.86E+05	3.88E+04	7.00E+03	1.31E+04	0.648	0.216	0.072	5.37E+05(3.88E+04; 0.648, 0.216)	5.23E+05(7.00E+03; 0.072)	4.86E+05(1.
16	11	1.7332	1.7295	1.7368	8.03E+05	5.86E+05	5.14E+05	4.31E+03	4.94E+04	1.26E+05	0.025	0.083	0.531	8.03E+05(4.31E+03; 0.025, 0.083)	5.86E+05(4.94E+04; 0.531)	5.14E+05(1.
57	1	1.7478			5.12E+03									5.12E+03		
31	10	1.7332	1.7368	1.7295	3.23E+05	2.95E+05	2.56E+05	2.55E+03	7.08E+04	1.63E+04	0.634	0.029	0.522	3.23E+05(2.55E+03; 0.634, 0.029)	2.95E+05(7.08E+04; 0.522)	2.56E+05(1.
47	1	1.715			5.18E+04									5.18E+04		
40	4	1.7368	1.7405	1.7442	1.54E+05	1.13E+05	4.87E+04	4.42E+04	1.27E+04	2.99E+03	0.337	0.078	0.02	1.54E+05(4.42E+04; 0.337, 0.078)	1.13E+05(1.27E+04; 0.02)	4.87E+04(2.
2	13	1.7295	1.7332	1.7258	3.87E+06	2.65E+06	2.55E+06	3.11E+05	1.89E+04	2.12E+04	0.031	0.027	0.039	3.87E+06(3.11E+05; 0.031, 0.027)	2.65E+06(1.89E+04; 0.039)	2.55E+06(2.
24	8	1.715	1.7114	1.7258	4.38E+05	3.68E+05	2.96E+05	2.27E+04	5.65E+04	3.58E+04	0.242	0.042	0.269	4.38E+05(2.27E+04; 0.242, 0.042)	3.68E+05(5.65E+04; 0.269)	2.96E+05(3.
32	1	1.7186			3.04E+05									3.04E+05		
26	5	1.7258	1.7222	1.7186	3.56E+05	3.00E+05	2.52E+05	4.66E+04	4.08E+04		0.33	0.317	0.507	3.56E+05(4.66E+04; 0.33, 0.317)	3.00E+05(4.08E+04; 0.507)	
37	3	1.7295	1.7332	1.7442	2.06E+05	2.00E+05	2.05E+04							2.06E+05	2.00E+05	

Figure B.9. Screenshots of the output from the first script. Data is displayed on two sheets that are sorted by rank (top sheet), which refers to the RF copy numbers and RF bps (bottom image).

On the same worksheet, heat maps are generated by using the command:'ctrl+j '. First, the script uses linear interpolation to calculate copy numbers for each BD value listed on 'Input' from the average copy numbers. Then, percentiles are calculated from the copy number values using two methods. The first method uses the entire data set for each gradient to calculate the percentiles and the second method only uses the values within each RF for the percentile calculations.

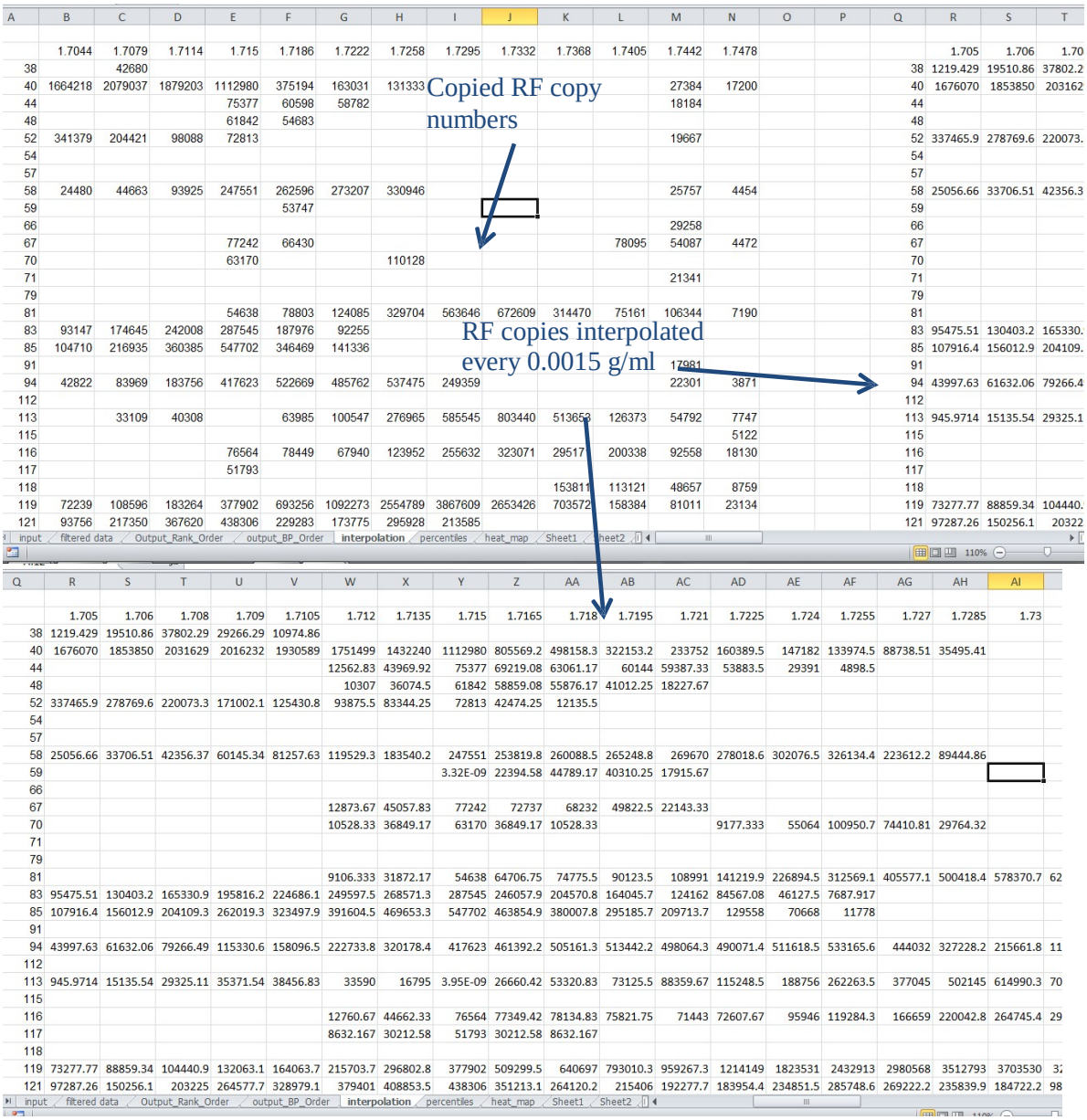


Figure B.10. Interpolated data. The top sheet is the original average value data copied displayed on the left side of the sheet and the right side of the sheet is the interpolated data.

Percentile calculations of RF copy numbers. All copy numbers within the gradient are included in the percentile calculations

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1																			
2																			
3		1.705	1.706	1.708	1.709	1.7105	1.712	1.7135	1.715	1.7165	1.718	1.7195	1.721	1.7225	1.724	1.7255	1.727	1.7285	1.7
4	38	0.023	0.129	0.204	0.169	0.072	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
5	40	0.94	0.948	0.956	0.951	0.951	0.942	0.928	0.91	0.867	0.784	0.695	0.594	0.493	0.472	0.452	0.363	0.196	#N/A
6	44	#N/A	#N/A	#N/A	#N/A	#N/A	0.085	0.225	0.332	0.31	0.293	0.284	0.281	0.263	0.171	0.032	#N/A	#N/A	#N/A
7	48	#N/A	#N/A	#N/A	#N/A	#N/A	0.062	0.197	0.29	0.28	0.273	0.215	0.123	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
8	52	0.706	0.65	0.581	0.504	0.437	0.37	0.35	0.32	0.218	0.081	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
9	54	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
10	57	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
11	58	0.152	0.19	0.217	0.285	0.346	0.429	0.52	0.613	0.623	0.625	0.635	0.64	0.649	0.673	0.7	0.584	0.365	#N/A
12	59	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	0.001	0.14	0.234	0.213	0.119	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
13	66	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
14	67	#N/A	#N/A	#N/A	#N/A	#N/A	0.087	0.236	0.337	0.319	0.307	0.248	0.135	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
15	70	#N/A	#N/A	#N/A	#N/A	#N/A	0.068	0.201	0.294	0.2	0.067	#N/A	#N/A	0.061	0.27	0.388	0.327	0.174	#N/A
16	71	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
17	79	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
18	81	#N/A	#N/A	#N/A	#N/A	#N/A	0.06	0.185	0.269	0.298	0.328	0.367	0.406	0.464	0.589	0.681	0.742	0.787	0.81
19	83	0.374	0.447	0.499	0.541	0.586	0.615	0.636	0.658	0.61	0.56	0.496	0.435	0.352	0.241	0.042	#N/A	#N/A	#N/A
20	85	0.405	0.485	0.559	0.627	0.697	0.739	0.776	0.805	0.772	0.734	0.668	0.567	0.445	0.314	0.077	#N/A	#N/A	#N/A
21	91	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
22	94	0.226	0.289	0.343	0.42	0.49	0.583	0.693	0.746	0.771	0.791	0.796	0.783	0.78	0.793	0.8	0.762	0.701	0.57
23	112	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
24	113	0.018	0.101	0.17	0.195	0.206	0.189	0.114	0.002	0.158	0.258	0.324	0.361	0.419	0.53	0.628	0.727	0.788	0.82
25	115	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
26	116	#N/A	#N/A	#N/A	#N/A	#N/A	0.086	0.233	0.336	0.338	0.341	0.334	0.315	0.318	0.376	0.428	0.502	0.58	0.63
27	117	#N/A	#N/A	#N/A	#N/A	#N/A	0.05	0.179	0.252	0.178	0.049	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
28	118	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
29	119	0.325	0.364	0.394	0.45	0.497	0.576	0.669	0.728	0.792	0.83	0.861	0.886	0.919	0.947	0.965	0.98	0.987	0.9
30	121	0.38	0.477	0.557	0.633	0.703	0.731	0.743	0.757	0.716	0.632	0.574	0.534	0.522	0.596	0.656	0.638	0.6	0.52
31	125	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	0.015	0.44	0.621	0.59	0.389	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
32	126	#N/A	#N/A	#N/A	0.079	0.165	0.275	0.4	0.482	0.539	0.597	0.631	0.654	0.675	0.702	0.717	0.605	0.378	#N/A

input / filtered data / Output_Rank_Order / output_BP_Order / interpolation / **percentiles** / heat_map / Sheet1 / Sheet2

Percentile calculations of RF copy numbers. All copy numbers for that RF are used in the percentile calculations.

AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV
1.745	1.7465			1.705	1.706	1.708	1.709	1.7105	1.712	1.7135	1.715	1.7165	1.718	1.7195	1.721	1.7225	1.724	1.7255	1.
N/A	#N/A		38	0	0.5	1	0.75	0.25	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/
0.153	0.131		40	0.761	0.857	1	0.952	0.904	0.809	0.714	0.666	0.619	0.571	0.523	0.476	0.428	0.38	0.333	0.
0.095	0.037		44	#N/A	#N/A	#N/A	#N/A	#N/A	0.285	0.571	1	0.928	0.857	0.785	0.714	0.642	0.5	0.071	#N/
N/A	#N/A		48	#N/A	#N/A	#N/A	#N/A	#N/A	0	0.333	1	0.833	0.666	0.5	0.166	#N/A	#N/A	#N/A	#N/
0.103	0.038		52	1	0.928	0.857	0.785	0.714	0.642	0.571	0.5	0.428	0.214	#N/A	#N/A	#N/A	#N/A	#N/A	#N/
N/A	#N/A		54	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/
N/A	#N/A		57	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/
0.134	0.082		58	0.238	0.285	0.333	0.38	0.428	0.523	0.571	0.666	0.714	0.761	0.809	0.857	0.904	0.952	1	0.
N/A	#N/A		59	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	0	0.5	1	0.75	0.25	#N/A	#N/A	#N/A	#N/
0.141	0.069		66	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/
0.221	0.138		67	#N/A	#N/A	#N/A	#N/A	#N/A	0	0.384	0.923	0.846	0.692	0.538	0.153	#N/A	#N/A	#N/A	#N/
N/A	#N/A		70	#N/A	#N/A	#N/A	#N/A	#N/A	0.222	0.555	0.777	0.444	0.111	#N/A	#N/A	0	0.666	1	0.
0.109	0.043		71	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/
N/A	#N/A		79	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/
0.351	0.22		81	#N/A	#N/A	#N/A	#N/A	#N/A	0	0.043	0.13	0.173	0.217	0.391	0.478	0.521	0.608	0.695	0.
N/A	#N/A		83	0.214	0.357	0.5	0.571	0.714	0.857	0.928	1	0.785	0.642	0.428	0.285	0.142	0.071	0	#N/
N/A	#N/A		85	0.142	0.285	0.357	0.5	0.642	0.785	0.928	1	0.857	0.714	0.571	0.428	0.214	0.071	0	#N/
0.094	0.036		91	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/
0.122	0.065		94	0.25	0.291	0.333	0.416	0.458	0.541	0.583	0.666	0.75	0.875	0.958	0.833	0.791	0.916	1	0.
N/A	#N/A		112	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/
0.229	0.149		113	0.035	0.071	0.214	0.285	0.321	0.25	0.107	0	0.178	0.392	0.464	0.5	0.571	0.642	0.678	(
0.02	0.029		115	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/
0.335	0.235		116	#N/A	#N/A	#N/A	#N/A	#N/A	0	0.043	0.304	0.347	0.391	0.217	0.13	0.173	0.434	0.521	0.
N/A	#N/A		117	#N/A	#N/A	#N/A	#N/A	#N/A	0.25	0.75	1	0.5	0	#N/A	#N/A	#N/A	#N/A	#N/A	#N/
0.21	0.145		118	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/
0.306	0.227		119	0.071	0.107	0.178	0.25	0.321	0.357	0.392	0.428	0.5	0.571	0.607	0.642	0.714	0.75	0.821	0.
N/A	#N/A		121	0.052	0.157	0.368	0.631	0.789	0.894	0.947	1	0.842	0.578	0.421	0.315	0.21	0.473	0.736	0.
N/A	#N/A		125	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	0	0.5	1	0.75	0.25	#N/A	#N/A	#N/A	#N/
N/A	#N/A		126	#N/A	#N/A	#N/A	0	0.076	0.153	0.307	0.384	0.461	0.538	0.692	0.769	0.846	0.923	1	0.

filtered data / Output_Rank_Order / output_BP_Order / interpolation / **percentiles** / heat_map / Sheet1 / Sheet2

Figure B.11. Percentiles calculated for the interpolated data. The left side of the sheet (top image) displays percentiles calculated from all of the values in the gradient while the percentiles on the right side of the sheet (bottom image) were calculated based on the other values in the RF.

Heat maps generated based on the two percentile calculations. The one on the left uses percentile values calculated from all copy number values. The right map is based on percentile calculations for each RF.

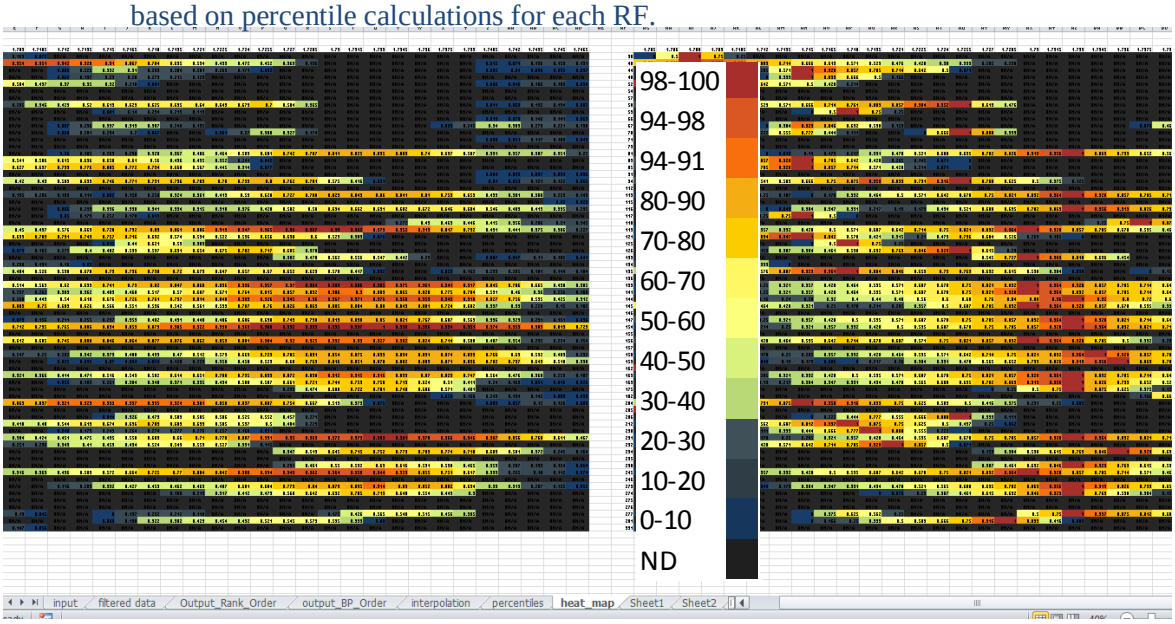


Figure B.12. Heat maps generated from the percentile data (Figure B.11). The left heat map corresponds to the top image in Figure B.11 and the right map from bottom image in Figure B.11. The overlaid image (near center) shows the legend for the heat map colors.

B.1. Script 1: RF quantification

```
Sub Macro1()
'
' Macro1 Macro
' Macro recorded 8/29/2010 by  Peter Andeer
'
' Keyboard Shortcut: Ctrl+y
' updated 9/15/10
'count number of fractions
Dim countfrac1 As Integer
Sheets("qpcr_data_in").Activate
Range("A8").Select
Range(Selection, Selection.End.(xlDown)).Select
countfrac1 = Selection.Count 'counts the number of fractions
'copy bd, name and qpcr data for samples With qpcr values to calc sheet
Dim sampledata, samplepasterow, samplepastecol, samplenum, sortnum As Integer
sampledata = 8
samplepasterow = 3
samplepastecol = 2
samplenum = 1

For sampledata = 8 To countfrac1 + 8
If samplenum = 51 Or samplenum = 101 Then 'move down 200 rows after 50 and 100 samples
GoTo 25
Else

5
'qPCR data1
Sheets("qpcr_data_in").Activate
ActiveSheet.Cells(sampledata, 3).Select
End If
If Selection = "" Then
GoTo 50
Else
Sheets("qpcr_data_in").Activate
Cells(sampledata, 1).Copy
Sheets("calc and organize").Activate
Cells(samplepasterow, samplepastecol).Select
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
    :=False, Transpose:=True
Sheets("sorting").Activate
sortnum = 4 + (2 * samplenum)
Cells(4, sortnum).Select
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
    :=False, Transpose:=True

Sheets("qpcr_data_in").Activate
Range(Cells(sampledata, 2), Cells(sampledata, 3)).Copy
Sheets("calc and organize").Activate
Cells(samplepasterow + 1, samplepastecol).Select
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
    :=False, Transpose:=True
Sheets("sorting").Activate
sortnum = 4 + (2 * samplenum)
Cells(2, sortnum).Select
```

```
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
        :=False, Transpose:=True

End If
'locate rflp data for each sample

Dim locate
Sheets("qpcr_data_in").Activate
ActiveSheet.Cells(sampledata, 2).Select
locate = Selection.Value
Worksheets("rflp_in").Select
With ActiveSheet.Range("A5:IV1002")
Dim Name1, nameadd1, i, j
Set Name1 = .Find(locate, LookIn:=xlValues)
Name1.Select
End With
'select data
ActiveCell.Offset(10, 0).Select
If Selection = "" Then
GoTo 10
Else
End If
 i = ActiveCell.Row
 j = ActiveCell.Column
 Range(Cells(i, j), Cells(i, j + 4)).Select
Range(Selection, Selection.End.(xlDown)).Select

Selection.Copy
Sheets("calc and organize").Activate
Cells(samplepasterow, samplepastecol).Select
ActiveCell.Offset(4, 0).Select
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
        :=False, Transpose:=False
'rearrange data
Cells(samplepasterow + 4, samplepastecol + 4).Select
Range(Selection, Selection.End.(xlDown)).Select
Selection.Cut
Cells(samplepasterow + 4, samplepastecol - 1).Select
ActiveSheet.Paste
Cells(samplepasterow + 4, samplepastecol + 2).Select
Range(Selection, Selection.End.(xlDown)).Select
Selection.Delete
'calculate relative heights
Dim heightnum As Integer
Dim heightrange1 As Range
Dim summed
Dim div1, amt
Dim calc_cells
Cells(samplepasterow, samplepastecol).Select
ActiveCell.Offset(4, 0).Select
Range(Selection, Selection.End.(xlDown)).Select
Set heightrange = Selection
heightnum = Selection.Count
Cells(samplepasterow + 4 + heightnum, samplepastecol).Select
summed = Application.WorksheetFunction.Sum(heightrange)
calc_cells = samplepasterow + 4
For calc_cells = samplepasterow + 4 To samplepasterow + 3 + heightnum
amt = Cells(calc_cells, samplepastecol).Value
div1 = amt / summed * 100
Cells(calc_cells, samplepastecol + 1).Value = div1
Next calc_cells
```

```
samplenum = samplenum + 1
samplepastecol = samplepastecol + 5
10
'qpcr2 data
Sheets("qpcr_data_in").Activate
ActiveSheet.Cells(sampledata, 5).Select
If Selection = "" Then
GoTo 50
Else End If
Sheets("qpcr_data_in").Activate
Cells(sampledata, 1).Copy Sheets("calc and
organize").Activate Cells(samplepasterow,
samplepastecol).Select
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
    :=False, Transpose:=True
Sheets("sorting").Activate
sortnum = 4 + (2 * samplenum)
Cells(4, sortnum).Select
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
    :=False, Transpose:=True

Sheets("qpcr_data_in").Activate
Range(Cells(sampledata, 4), Cells(sampledata, 5)).Copy
Sheets("calc and organize").Activate
Cells(samplepasterow + 1, samplepastecol).Select
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
    :=False, Transpose:=True
Sheets("sorting").Activate
sortnum = 4 + (2 * samplenum)
Cells(2, sortnum).Select
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
    :=False, Transpose:=True

'locate rflp data for each sample

Sheets("qpcr_data_in").Activate
ActiveSheet.Cells(sampledata, 4).Select
locate = Selection.Value
Worksheets("rflp_in").Select
With ActiveSheet.Range("A5:IV1002")
Dim Name2, nameadd2, o, k
Set Name2 = .Find(locate, LookIn:=xlValues)
Name2.Select
End With
'select data
ActiveCell.Offset(10, 0).Select
If Selection = "" Then
GoTo 20
Else: End If

 o = ActiveCell.Row
 k = ActiveCell.Column
 Range(Cells(o, k), Cells(o, k + 4)).Select
Range(Selection, Selection.End.(xlDown)).Select

Selection.Copy
Sheets("calc and organize").Activate
Cells(samplepasterow, samplepastecol).Select
ActiveCell.Offset(4, 0).Select
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
    :=False, Transpose:=False
```

```
'rearrange data
Cells(samplepasterow + 4, samplepastecol + 4).Select
Range(Selection, Selection.End.(xlDown)).Select
Selection.Cut
Cells(samplepasterow + 4, samplepastecol - 1).Select
ActiveSheet.Paste
Cells(samplepasterow + 4, samplepastecol + 2).Select
Range(Selection, Selection.End.(xlDown)).Select
Selection.Delete
Dim heightnum2 As Integer
Dim heightrange2 As Range
Dim summed2
Dim div2, amt2

Cells(samplepasterow, samplepastecol).Select
ActiveCell.Offset(4, 0).Select
Range(Selection, Selection.End.(xlDown)).Select
Set heightrange2 = Selection
heightnum2 = Selection.Count
Cells(samplepasterow + 4 + heightnum2, samplepastecol).Select
summed2 =  Application.WorksheetFunction.Sum(heightrange2)
calc_cells = samplepasterow + 4
For calc_cells = samplepasterow + 4 To samplepasterow + 3 + heightnum2
amt2 = Cells(calc_cells, samplepastecol).Value
div2 = amt2 / summed2 * 100
Cells(calc_cells, samplepastecol + 1).Value = div2
Next calc_cells
samplenum = samplenum + 1
samplepastecol = samplepastecol + 5

20
'qpcr3 data
Sheets("qpcr_data_in").Activate
ActiveSheet.Cells(sampledata, 7).Select
If Selection = "" Then
GoTo 50
Else End If
Sheets("qpcr_data_in").Activate
Cells(sampledata, 1).Copy
Sheets("calc and organize").Activate
Cells(samplepasterow, samplepastecol).Select
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
    :=False, Transpose:=True
Sheets("sorting").Activate
sortnum = 4 + (2 * samplenum)
Cells(4, sortnum).Select
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
    :=False, Transpose:=True

Sheets("qpcr_data_in").Activate
Range(Cells(sampledata, 6), Cells(sampledata, 7)).Copy
Sheets("calc and organize").Activate
Cells(samplepasterow + 1, samplepastecol).Select
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
    :=False, Transpose:=True
Sheets("sorting").Activate
sortnum = 4 + (2 * samplenum)
Cells(2, sortnum).Select
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
    :=False, Transpose:=True
```

```
Sheets("qpcr_data_in").Activate
ActiveSheet.Cells(sampledata, 6).Select
locate = Selection.Value
Worksheets("rflp_in").Select
With ActiveSheet.Range("A5:IV1002")
Dim Name3, nameadd3, p, l
Set Name3 = .Find(locate, LookIn:=xlValues)
Name3.Select
End With
'select data
ActiveCell.Offset(10, 0).Select
If Selection = "" Then
GoTo 50
Else: End If
 p = ActiveCell.Row
 l = ActiveCell.Column
 Range(Cells(p, l), Cells(p, l + 4)).Select
Range(Selection, Selection.End.(xlDown)).Select

Selection.Copy
Sheets("calc and organize").Activate
Cells(samplepasterow, samplepastecol).Select
ActiveCell.Offset(4, 0).Select
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
      :=False, Transpose:=False
'rearrange data
Cells(samplepasterow + 4, samplepastecol + 4).Select
Range(Selection, Selection.End.(xlDown)).Select
Selection.Cut
Cells(samplepasterow + 4, samplepastecol - 1).Select
ActiveSheet.Paste
Cells(samplepasterow + 4, samplepastecol + 2).Select
Range(Selection, Selection.End.(xlDown)).Select
Selection.Delete
Dim heightnum3 As Integer
Dim heightrange3 As Range
Dim summed3
Dim div3, amt3

Cells(samplepasterow, samplepastecol).Select
ActiveCell.Offset(4, 0).Select
Range(Selection, Selection.End.(xlDown)).Select
Set heightrange3 = Selection
heightnum3 = Selection.Count
Cells(samplepasterow + 4 + heightnum3, samplepastecol).Select
summed3  =  Application.WorksheetFunction.Sum(heightrange3)
calc_cells = samplepasterow + 4
For calc_cells = samplepasterow + 4 To samplepasterow + 3 + heightnum3
amt3 = Cells(calc_cells, samplepastecol).Value
div3 = amt3 / summed3 * 100
Cells(calc_cells, samplepastecol + 1).Value = div3
Next calc_cells
samplenum = samplenum + 1
samplepastecol = samplepastecol + 5
GoTo 50

25
samplepasterow = samplepasterow + 200
samplepastecol = 2
GoTo 5
50
Next sampledata
```

```
Dim count1, count2, count3, countqpcr1, countpcr2, countpcr3, countpcr As Integer
'count # of samples With qpcr data
Sheets("qpcr_data_in").Activate
countpcr1 = Application.WorksheetFunction.Count(Range(Cells(8, 3), Cells(7 + countfrac1, 3)))
countpcr2 = Application.WorksheetFunction.Count(Range(Cells(8, 5), Cells(7 + countfrac1, 5)))
countpcr3 = Application.WorksheetFunction.Count(Range(Cells(8, 7), Cells(7 + countfrac1, 7)))
countpcr = countpcr1 + countpcr2 + countpcr3
Cells(5, 5).Select
ActiveCell.Value = countpcr
'copy bp_data to a new sheet "sorting"
Dim bpsamples, bprows1, bpcolumns1 As Integer
Dim bpcalc_rw, bpsamples1, bpcalc_rw2
bpcalc_rw = 7
bpcalc_rw2 = 207
bpsamples = 1
bprows1 = 5
bpsamples1 = 1
For bpsamples = 1 To countpcr
Sheets("calc and organize").Activate
bpcolumns1 = (bpsamples - 1) * 5 + 1
Cells(bpcalc_rw, bpcolumns1).Select

If Selection = "" Then
GoTo 75
Else
Range(Selection, Selection.End.(xlDown)).Select
Selection.Copy
Sheets("sorting").Activate
End If

Cells(bprows1, 1).Select
ActiveSheet.Paste
bprows1 = bprows1 + Selection.Count
GoTo 80
75
If countpcr < 100 Then
Cells(207, 1).Select
bpcalc_rw2 = 207
Else
Cells(407, 1).Select
bpcalc_rw2 = 407
End If
If Selection = "" Then
GoTo 80
End If
bpcolumns1 = (bpsamples1 - 1) * 5 + 1
Cells(bpcalc_rw2, bpcolumns1).Select
If Selection = "" Then
GoTo 80
End If
Range(Selection, Selection.End.(xlDown)).Select
Selection.Copy
Sheets("sorting").Activate
bpsamples1 = bpsamples1 + 1

Cells(bprows1, 1).Select
ActiveSheet.Paste
bprows1 = bprows1 + Selection.Count
80
Next bpsamples
'sort all basepair values in ascEnding order
Sheets("sorting").Activate
```

```
Cells(5, 1).Select
Range(Selection, Selection.End.(xlDown)).Select
Selection.Sort Key1:=Range("a5"), Order1:=xlAscEnd.ing, Header:=xlGuess, _
    OrderCustom:=1, MatchCase:=False, Orientation:=xlTopToBottom, _
    DataOption1:=xlSortNormal
'eleminate redundant samples
Dim bprows2, bpvalue, bpentry, bptotal, bpnewtotal As Integer
bpentry = 5

Sheets("sorting").Activate
Cells(5, 1).Select
Range(Selection, Selection.End.(xlDown)).Select
bptotal = Selection.Count + 5
Do While bpentry < bptotal + 1
    Do
    Cells(bpentry, 1).Select
    bpvalue = Selection.Value
    If bpvalue = "" Then
    Exit Do
    End If
    Cells(bpentry + 1, 1).Select
    If Selection.Value = bpvalue Then
    bprows = bpentry + 1
    Rows(bprows).Select
    Selection.Delete Shift:=xlUp
    Else: Exit Do
    End If
    Loop
bpentry = bpentry + 1
Loop
'move bp lists to d4 and copy it 5 rows below the first list
Cells(5, 1).Select
Range(Selection, Selection.End.(xlDown)).Select
bpnewtotal = Selection.Count
Selection.Cut
Cells(5, 4).Select
ActiveSheet.Paste
Cells(5, 4).Select
Range(Selection, Selection.End.(xlDown)).Select
Selection.Copy
Cells(9 + bpnewtotal, 4).Select
ActiveSheet.Paste Cells(2,
4).Select ActiveCell.Value =
"sample"
Cells(3, 4).Select
ActiveCell.Value = "qPCR"
Cells(4, 4).Select
ActiveCell.Value = "BD"
ActiveCell.Offset(bpnewtotal + 2, 0).Select
ActiveCell.Value = "sample"
ActiveCell.Offset(1, 0).Select
ActiveCell.Value = "qPCR"
ActiveCell.Offset(1, 0).Select
ActiveCell.Value = "BD"

'transfer rflp data to sorting sheet
Dim totalsamp, currsamp, currcol, currdata As Integer
Dim current, calc_row
currsamp = 1
calc_row = 4
Sheets("qpcr_data_in").Activate
Cells(5, 5).Select
```

```
totalsamp = Selection.Value
For currsamp = 1 To totalsamp - 1
currcol = 4 + (2 * currsamp)
Sheets("sorting").Activate
Cells(2, currcol).Select
current = Selection.Value
Range(Selection, Selection.End.(xlDown)).Select
Selection.Copy
ActiveCell.Offset(bpnewtotal + 4, 0).Select
ActiveSheet.Paste
Sheets("calc and organize").Activate
With ActiveSheet.Range("A1:IV1002")
Dim rflp_curr
Set rflp_curr = .Find(current, LookIn:=xlValues)
rflp_curr.Select
End With
If Selection = current Then
ActiveCell.Offset(3, -1).Select
If Selection = "" Then
GoTo 100
End If
'copy and paste bps
Range(Selection, Selection.End.(xlDown)).Select
Selection.Copy
Sheets("sorting").Activate
Cells(5, currcol - 1).Select
ActiveSheet.Paste
ActiveCell.Offset(bpnewtotal + 4, 0).Select
ActiveSheet.Paste
'copy and paste rel area to top set
Sheets("calc and organize").Activate
'Cells(4, currdata).Select
rflp_curr.Select
ActiveCell.Offset(3, 2).Select
Range(Selection, Selection.End.(xlDown)).Select
Selection.Copy
Sheets("sorting").Activate
Cells(5, currcol).Select
ActiveSheet.Paste
'copy and paste rel height to bottom set
Sheets("calc and organize").Activate
'Cells(4, currdata).Select
rflp_curr.Select

ActiveCell.Offset(3, 1).Select
Range(Selection, Selection.End.(xlDown)).Select
Selection.Copy
Sheets("sorting").Activate
Cells(bpnewtotal + 9, currcol).Select
ActiveSheet.Paste
Else: GoTo 100

End If
100
Next currsamp

'organize samples
Dim relarea1, relarea1b, relheight1, refrow1, refrow2, sampbp, sampbppos, samprow, samptotal, samp2row, sampnum1 As
Integer
Dim ref1
sampbp = 1
relarea1b = 0
```

```
For sampbp = 1 To countpcr 'cycle through sampledata

sampbppos = sampbp * 2 + 3 'column where bp data is located
samprow = 5 'fix starting row at 5
'count no of bps in data
Cells(samprow, sampbppos).Select
relarea1 = Selection
refrow1 = 5
relheight1 = 1
sampnum1 = 1
For samprow = 5 To 4 + bpnewtotal 'cycle through data for each bp in set
Cells(samprow, sampbppos).Select
110
relarea1 = Selection
   If relarea1 = "" Then
      GoTo 130
      End If
If relarea1 = relarea1b Then
bpdata = Application.InputBox(prompt:="Duplicate bp, enter alt number", Type:=1)
Cells(samprow, sampbppos).Value = bpdata
Cells(samprow, sampbppos).Font.Bold = True
Cells(samprow, sampbppos + 1).Font.Bold = True
relarea1 = bpdata
End If
   Do While refrow1 < 4 + bpnewtotal 'cycle through ref bps
      ref1 = Cells(refrow1, 4).Value
            If relarea1 = ref1 Then
      Exit Do
      End If
     refrow1 = refrow1 + 1
   Loop
   samp2row = samprow + 4 + bpnewtotal
   refrow2 = refrow1 + 4 + bpnewtotal

   If samprow = refrow1 Then
   GoTo 120
   End If
   If ActiveCell.Offset(1, 0).Value = "" Then
   Range(Cells(samprow, sampbppos), Cells(samprow, sampbppos + 1)).Select
   Selection.Cut
   Cells(refrow1, sampbppos).Select 'realign data
   ActiveSheet.Paste
   Range(Cells(samp2row, sampbppos), Cells(samp2row, sampbppos + 1)).Select
   Selection.Cut
   Cells(refrow2, sampbppos).Select 'realign data
   ActiveSheet.Paste
   GoTo 130
End If
   Range(Cells(samprow, sampbppos), Cells(samprow, sampbppos + 1)).Select 'select data
   Range(Selection, Selection.End.(xlDown)).Select
   Selection.Cut
   Cells(refrow1, sampbppos).Select 'realign data
   ActiveSheet.Paste

   Range(Cells(samp2row, sampbppos), Cells(samp2row, sampbppos + 1)).Select 'select data
   Range(Selection, Selection.End.(xlDown)).Select
   Selection.Cut
   Cells(refrow2, sampbppos).Select 'realign data
   ActiveSheet.Paste
      GoTo 130
120
sampnum1 = sampnum1 + 1
```

```
130
relarea1b = relarea1
samprow = refrow1
refrow1 = refrow1 + 1
Next samprow

Next sampbp
'delete bp info for the cells
Dim sampbppos2 As Integer
sampbp = 1
sampbppos2 = 5
For sampbp = 1 To countpcr 'cycle through sampledata
Columns(sampbppos2).Select
Selection.ClearContents
Selection.Delete Shift:=xlToLeft

sampbppos2 = sampbppos2 + 1
Next sampbp
'delete bp info for the cells
'setup qpcr data
Dim q_area_loc, q_he_loc, total_bps, bplist, q_samplecount As Integer
Dim q_samp_row1, q_samp_row2, q_BD_row1, q_BD_row2, q_tot_row1, q_tot_row2, q_area_row, q_he_row,
q_tot_values As Integer
'count bps and headers
Cells(2, 4).Select
Range(Selection, Selection.End.(xlDown)).Select
total_bps = Selection.Count

'copy and paste bp list and write headers
q_area_loc = total_bps * 2 + 4
Cells(q_area_loc, 4).Select

ActiveCell.Value = "Sample:" q_samp_row1
= Selection.Row ActiveCell.Offset(1,
0).Value = "BD:" ActiveCell.Offset(2,
0).Value = "Total 16S:" ActiveCell.Offset(3,
0).Value = "Area Data" Cells(5, 4).Select
Range(Selection, Selection.End.(xlDown)).Copy
Cells(q_area_loc, 4).Select

ActiveCell.Offset(4, 0).Select
ActiveSheet.Paste
Cells(q_area_loc, 4).Select
Range(Selection, Selection.End.(xlDown)).Select
Selection.Copy
ActiveCell.Offset(total_bps + 2, 0).Select

ActiveSheet.Paste q_samp_row2 =
Selection.Row ActiveCell.Offset(3, 0).Value =
"Height Data" q_BD_row1 = q_samp_row1 +
1
q_BD_row2 = q_samp_row2 + 1
q_tot_row1 = q_samp_row1 + 2
q_tot_row2 = q_samp_row2 + 2
q_area_row = q_samp_row1 + 4
q_he_row = q_samp_row2 + 4
'copy sample info from the top sets for the bottom
Cells(2, 5).Select
Range(Selection, Selection.End.(xlToRight)).Select
q_samplecount = Selection.Count
```

```
Selection.Copy
Cells(q_samp_row1, 5).Select
ActiveSheet.Paste
Cells(q_samp_row2, 5).Select
ActiveSheet.Paste

Cells(4, 5).Select
Range(Selection, Selection.End.(xlToRight)).Copy
Cells(q_BD_row1, 5).Select
ActiveSheet.Paste
Cells(q_BD_row2, 5).Select
ActiveSheet.Paste

Cells(3, 5).Select
Range(Selection, Selection.End.(xlToRight)).Copy
Cells(q_tot_row1, 5).Select
ActiveSheet.Paste
Cells(q_tot_row2, 5).Select
ActiveSheet.Paste
'multiply percentages by total 16S numbers
'area data first
Dim total_bps2, curr_areas, area_val1, area_val2, area_mult, he_val1, he_val2, he_mult
Dim curr_samp, sample_col, samp_area_1st, samp_he_1st, area1, he1 As Integer
Cells(5, 4).Select
Range(Selection, Selection.End.(xlDown)).Select
total_bps2 = Selection.Count
samp_area_1st = q_area_row
samp_he_1st = q_he_row
curr_samp = 1
area1 = 5
he1 = area1 + 4 + total_bps2

For curr_samp = 1 To q_samplecount
   sample_col = curr_samp + 4
      For curr_values = 1 To total_bps2
      area_val1 = Cells(area1, sample_col).Value
      area_val2 = Cells(q_tot_row1, sample_col).Value
      area_mult = (area_val1 / 100) * area_val2
      If area_mult = 0 Then
      area_mult = ""
      End If
      Cells(q_area_row, sample_col).Value = area_mult
            q_area_row = q_area_row + 1
      he_val1 = Cells(he1, sample_col).Value
      he_val2 = Cells(q_tot_row2, sample_col).Value
      he_mult = (he_val1 / 100) * he_val2
      If he_mult = 0 Then
      he_mult = ""
      End If
      Cells(q_he_row, sample_col).Value = he_mult
            q_he_row = q_he_row + 1
     area1 = area1 + 1
      he1 = he1 + 1
      Next curr_values
   q_area_row = samp_area_1st
   q_he_row = samp_he_1st
   area1 = 5
   he1 = area1 + 4 + total_bps2
   Next curr_samp
 Cells(5, 4).Select
Range(Selection, Selection.End.(xlDown)).Select
Range(Selection, Selection.End.(xlToRight)).Copy
```

```
Cells(q_area_row, 4).Select
Selection.PasteSpecial Paste:=xlPasteFormats, Operation:=xlNone, _
      SkipBlanks:=False, Transpose:=False
   Application.CutCopyMode = False

Cells(5, 4).Select
Range(Selection, Selection.End.(xlDown)).Select
Range(Selection, Selection.End.(xlToRight)).Copy
Cells(q_he_row, 4).Select
Selection.PasteSpecial Paste:=xlPasteFormats, Operation:=xlNone, _
      SkipBlanks:=False, Transpose:=False
   Application.CutCopyMode = False
'paste qpcr data into a new sheet for review
Cells(q_samp_row1, 4).End.(xlDown).Select
Selection.End.(xlDown).Select
Selection.End.(xlDown).Select
last_Row = Selection.Row
Range(Cells(q_samp_row1, 4), Cells(last_Row, 4)).Select
Range(Selection, Selection.End.(xlToRight)).Copy
Sheets("output and review").Activate
Cells(4, 4).Select
ActiveSheet.Paste
Dim bd_check1, bd, check2
Dim bdcol_1, bdcol_2, bdcol_store, col, col_color, curr_color As Integer

bdcol_1 = 5
bdcol_2 = 6
curr_color = 35
Sheets("output and review").Select
Do

   Cells(5, bdcol_1).Select
   bd_check1 = Selection.Value
   If bd_check1 = "" Then
   Exit Do
   End If

   Do
   Cells(5, bdcol_2).Select
   bd_check2 = Selection.Value
   If bd_check2 = bd_check1 Then
   bdcol_store = bdcol_2
   bdcol_2 = bdcol_2 + 1
   End If
   Loop Until bd_check2 < bd_check1
      col_color = curr_color
      col = bdcol_1
         For col = bdcol_1 To bdcol_store
         Columns(col).Select
         With Selection.Interior
         .ColorIndex = col_color
         .Pattern = xlSolid
         End With
         Next col bdcol_1
   = bdcol_2 bdcol_2 =
   bdcol_2 + 1
   curr_color = curr_color + 1
   If curr_color > 40 Then
   curr_color = 35
   End If
   Loop
   With ActiveWindow
```

```
        .SplitColumn = 4
        .SplitRow = 0
    End With
    ActiveWindow.Panes(1).Activate
    Cells(1, 1).Select
    ActiveWindow.Panes(2).Activate
    Cells(8, 5).Select
End Sub
```

B.2. Script 2: RF average values and standard deviation calculations.

```
Sub Macro18()
'
' Macro18 Macro
' Macro recorded 9/8/2010 by  Peter Andeer
'
Dim nosample, nolabel, pgfull, datapage, dataname, realpg, Msg
'
Sheets("Raw_data").Activate
Cells(3, 2).Select
If Selection = "" Then
nosample = Application.InputBox(prompt:="No sample # entered. Enter below", Type:=1)
Selection = nosample
End If
datapage = Selection
Cells(5, 2).Select
If Selection = "" Then
nolabel = Application.InputBox(prompt:="No label name. Enter below", Type:=2)
Selection = nolabel
End If
dataname = Selection

If datapage = 1 Then
Sheets("Sample1").Select
ElseIf datapage = 2 Then
Sheets("Sample2").Select
ElseIf datapage = 3 Then
Sheets("Sample3").Select
ElseIf datapage = 4 Then
Sheets("Sample4").Select
ElseIf datapage = 5 Then
Sheets("Sample5").Select
ElseIf datapage = 6 Then
Sheets("Sample6").Select
ElseIf datapage = 7 Then
Sheets("Sample7").Select
ElseIf datapage = 8 Then
Sheets("Sample8").Select
ElseIf datapage = 9 Then
Sheets("Sample9").Select
ElseIf datapage = 10 Then
Sheets("Sample10").Select

End If
realpg = ActiveSheet.Name
Cells(1, 1).Select
If Selection = "" Then
Cells(1, 1).Value = dataname
Else
Msg = "Page isn't cleared, check data"
MsgBox (Msg)
GoTo 500
End If
```

```
Dim total_num, bp_num, area_Endrw
Dim he_Endrw, he_strw, stdev_col, fst_bd, avg_col
Dim lst_bd, fst_bd_col, lst_bd_col, bprw1, bprw2
Dim avg_rw1, avg_rw2, std_rw1, std_rw2
Dim current_bd, nxt_bd, total_samp, total_frac, current_samp, current_frac As Integer
Dim bdchk
Dim sample_no_col
Sheets("Raw_data").Activate
Cells(5, 4).Select
Range(Selection, Selection.End(xlDown)).Select
total_num = Selection.Count
bp_num = total_num - 3
Selection.Copy
Sheets(realpg).Select
Cells(5, 4).Select
ActiveSheet.Paste
Selection.End(xlDown).Select
area_enrw = Selection.Row
avg_col = Selection.Column
Selection.Offset(3, 0).Select
he_strw = Selection.Row
ActiveSheet.Paste
Selection.End(xlDown).Select
he_enrw = Selection.Row
Cells(he_strw + 2, 4).Value = "Height Data"
Sheets("Raw_data").Activate
current_frac = 5
total_frac = 0
Do Until Cells(5, current_frac).Value = ""
Cells(5, current_frac).Select
current_bd = Selection.Value
nxt_bd = Selection.Offset(0, 1).Value
If nxt_bd < current_bd Then
total_frac = total_frac + 1
End If
current_frac = current_frac + 1
Loop

Sheets(realpg).Select
Range(Cells(5, 4), Cells(he_enrw, 4)).Copy
Cells(5, 6 + total_frac).Select
ActiveSheet.Paste
stdev_col = Selection.Column
Cells(5, 8 + (2 * total_frac)).Select
ActiveSheet.Paste
sample_no_col = Selection.Column

Sheets("Raw_data").Select

current_frac = 1

fst_bd_col = Cells(5, 5).Column
lst_bd_col = Cells(5, 6).Column
fst_bd = Cells(5, fst_bd_col).Value
bdchk = Cells(5, 6).Value
For current_frac = 1 To total_frac
Sheets("Raw_data").Select
fst_bd = Cells(5, fst_bd_col).Value
   bdchk = Cells(5, lst_bd_col).Value
Do Until fst_bd > bdchk
   fst_bd = Cells(5, fst_bd_col).Value
   bdchk = Cells(5, lst_bd_col).Value
```

```
    If fst_bd = bdchk Then
    lst_bd_col = lst_bd_col + 1
    End If
    Loop
    lst_bd_col = lst_bd_col - 1
avg_col = avg_col + 1
stdev_col = stdev_col + 1
sample_no_col = sample_no_col + 1
bprw1 = 6
bprw2 = he_strw + 1
avg_rw1 = 6
std_rw1 = 6
avg_rw2 = he_strw + 1
std_rw2 = he_strw + 1

Dim fracrange1, fracrange2, fracrange3, fracrange4, avg1, avg2, std1, std2, samp_no1, samp_no2
Sheets(realpg).Cells(avg_rw1 - 1, avg_col).Value = fst_bd
Sheets(realpg).Cells(avg_rw2 - 1, avg_col).Value = fst_bd
Sheets(realpg).Cells(std_rw1 - 1, stdev_col).Value = fst_bd
Sheets(realpg).Cells(std_rw2 - 1, stdev_col).Value = fst_bd
Sheets(realpg).Cells(avg_rw1 - 1, sample_no_col).Value = fst_bd
Sheets(realpg).Cells(avg_rw2 - 1, sample_no_col).Value = fst_bd

For bprw1 = 6 To area_enrw
    fracrange1 = Sheets("Raw_data").Range(Cells(bprw1, fst_bd_col), Cells(bprw1, lst_bd_col))
        On Error GoTo errorfix
        avg1 = Application.WorksheetFunction.Average(fracrange1)
    Sheets(realpg).Cells(avg_rw1, avg_col).Value = avg1
    On Error GoTo errorfix2
    std1 = Application.WorksheetFunction.StDev(fracrange1)
    Sheets(realpg).Cells(std_rw1, stdev_col).Value = std1
    fracrange2 = Sheets("Raw_data").Range(Cells(bprw2, fst_bd_col), Cells(bprw2, lst_bd_col))
        On Error GoTo errorfix3
        avg2 = Application.WorksheetFunction.Average(fracrange2)
    Sheets(realpg).Cells(avg_rw2, avg_col).Value = avg2
    On Error GoTo errorfix4
    std2 = Application.WorksheetFunction.StDev(fracrange2)
    Sheets(realpg).Cells(std_rw2, stdev_col).Value = std2
    samp_no1 = Application.WorksheetFunction.Count(fracrange1)
    samp_no2 = Application.WorksheetFunction.Count(fracrange2)
    Sheets(realpg).Cells(avg_rw1, sample_no_col).Value = samp_no1
    Sheets(realpg).Cells(avg_rw2, sample_no_col).Value = samp_no2
200
    If bprw1 = 6 Then
    avg_rw1 = avg_rw1 + 1
    std_rw1 = std_rw1 + 1
    avg_rw2 = avg_rw2 + 1
    std_rw2 = std_rw2 + 1
    bprw2 = bprw2 + 1
    bprw1 = bprw1 + 1
    End If
    avg_rw1 = avg_rw1 + 1
    std_rw1 = std_rw1 + 1
    avg_rw2 = avg_rw2 + 1
    std_rw2 = std_rw2 + 1
    bprw2 = bprw2 + 1

 Next bprw1
On Error GoTo 0

fst_bd_col = lst_bd_col + 1
lst_bd_col = fst_bd_col + 1
```

```
Next current_frac
GoTo 500
errorfix: avg1 = ""
Resume Next
errorfix2: std1 = ""
Resume Next
errorfix3: avg2 = ""
Resume Next
errorfix4: std2 = ""
Resume Next

500
Dim bprow, max1, value_range
Sheets(realpg).Select
Cells(8, 4).Select
Range(Selection, Selection.End(xlDown)).Copy

Cells(8, 1).Select
ActiveSheet.Paste
bprow = 8
For bprow = 8 To area_enrw
value_range = Range(Cells(bprow, 4), Cells(bprow, 4 + total_frac))
max1 = Application.WorksheetFunction.Max(value_range)
Cells(bprow, 2) = max1
Next bprow
Range("A8").Select
   Range(Selection, Selection.End(xlToRight)).Select
   Range(Selection, Selection.End(xlDown)).Select
   Application.CutCopyMode = False
   Selection.Sort Key1:=Range("B8"), Order1:=xlDescEnding, Header:=xlGuess, _
      OrderCustom:=1, MatchCase:=False, Orientation:=xlTopToBottom, _
      DataOption1:=xlSortNormal
Range("A8").Select
Sheets("compiled info_userform").Select
With ActiveSheet.Range("c4:l4")
Dim dat_stats
Set dat_stats = .Find(realpg, LookIn:=xlValues)
dat_stats.Select
End With
Selection.Offset(1, 0).Select
Selection.Value = total_frac
Selection.Offset(1, 0).Select
Selection.Value = bp_num
Selection.Offset(1, 0).Select
Selection.Value = 4 + total_frac
Selection.Offset(1, 0).Select
Selection.Value = area_enrw
Selection.Offset(1, 0).Select
Selection.Value = 6 + total_frac
Selection.Offset(1, 0).Select
Selection.Value = 6 + 2 * total_frac
Selection.Offset(1, 0).Select
Selection.Value = he_strw
Selection.Offset(1, 0).Select
Selection.Value = he_enrw

End Sub
```

B..3. Script 3: p values for RF values

```
Sub Macro1()
```

```
'
' Macro1 Macro
'
' Keyboard Shortcut: Ctrl+h
'
    Dim BD_start_val
    Dim current_frac
    Dim BD_start
    Dim Total_frac
    Dim BD_start2
    Dim BD_start2_val
    Dim BD_start3
    Dim BD_start3_val
    Dim BD_end
    Dim BD_end2
    Dim bp_now
    Dim bp_frac_count As Integer
    Dim bp_max
    Dim bp_row As Integer
    Dim fracs_vals
    Dim val_range, val_range2, bd_range, bd_range2, val_range1b, val_range2b, val_range3c
    Dim col, col2, val_col, std_col
    Dim filt_row
    Dim rep
    Dim bp_numb As Integer
    Dim bp_range As Range
    Dim bd_vals As Range
    Dim bd_filt As Range

    'determine the number of fractions and location of start and
    'end columns For averages and std deviations and sample nos
   Application.ScreenUpdating = False
    Sheets("input").Select
    Range("D3").Select
    BD_start_val = Selection.Value
    BD_start = Selection.Column
    current_frac = BD_start
    Total_frac = 0
    Do Until Cells(3, current_frac).Value = ""
    Cells(3, current_frac).Select
    current_bd = Selection.Value
    nxt_bd = Selection.OffSet(0, 1).Value
    If nxt_bd < current_bd Then
    Total_frac = Total_frac + 1
    BD_end = Selection.Value
    End If
    current_frac = current_frac + 1
    Loop
    Selection.End(xlToRight).Select
    Selection.OffSet(0, 1).Select
    BD_start2_val = Selection.Value
    If BD_start2_val = BD_start_val Then
    BD_start2 = Selection.Column
    bd_range2 = Range(Cells(3, BD_start2), Cells(3, BD_start2 + Total_frac - 1))
    Else: GoTo 5000
    End If
    Selection.OffSet(0, Total_frac - 1).Select
    BD_end2 = Selection.Value
    If BD_end2 = BD_end Then

    Else: GoTo 5000
    End If
```

```
Selection.End(xlToRight).Select
Selection.OffSet(0, 1).Select
BD_start3_val = Selection.Value
If BD_start3_val = BD_start_val Then

BD_start3 = Selection.Column
bd_range3 = Range(Cells(3, BD_start3), Cells(3, BD_start3 + Total_frac - 1))
'If BD values don't match (i.e. the values are wrong) end the program
Else: GoTo 5000
End If
   'copy the BD range To the filtered data (calculation) sheet
Set bd_vals = Range(Cells(3, BD_start), Cells(3, BD_start + Total_frac - 1))
bd_vals.Copy
Worksheets("filtered data").Select
Set bd_filt = Range(Cells(6, 25), Cells(6, 25 + Total_frac - 1))
bd_filt.PAsteSpecial PAste:=xlPAsteValues, Operation:=xlNone, SkipBlanks:=False, Transpose:=False
'input column headers
Cells(6, 3).Select
Selection.Value = "Bps"
Cells(6, 4).Select
Selection.Value = "Rank"
Cells(6, 5).Select
Selection.Value = "# of Fractions"
Cells(6, 6).Select
Selection.Value = "BD1"
Cells(6, 7).Select
Selection.Value = "BD2"
Cells(6, 8).Select
Selection.Value = "BD3"
Cells(6, 9).Select
Selection.Value = "MaxValue1"
Cells(6, 10).Select
Selection.Value = "MaxValue2"
Cells(6, 11).Select
Selection.Value = "MaxValue3"
Cells(6, 12).Select
Selection.Value = "Stdev1"
Cells(6, 13).Select
Selection.Value = "Stdev2"
Cells(6, 14).Select
Selection.Value = "Stdev3"
Cells(6, 15).Select
Selection.Value = "No1"
Cells(6, 16).Select
Selection.Value = "No2"
Cells(6, 17).Select
Selection.Value = "No3"
Cells(6, 18).Select
Selection.Value = "ttest1"
Cells(6, 19).Select
Selection.Value = "ttest2"
Cells(6, 20).Select
Selection.Value = "ttest3"
Cells(6, 21).Select
Selection.Value = "p-value1"
Cells(6, 22).Select
Selection.Value = "p-value2"
Cells(6, 23).Select
Selection.Value = "p-value3"
```

```
'determine the Total number of bps in the sheet and the value of the initial bp
Worksheets("input").Select
Cells(6, 3).Select
bp_now = Selection.Value
Cells(6, 3).Select
Selection.End(xlDown).Select
bp_numb = Selection.Row
'bp_range = Range(Selection, Selection.End(xlDown))

'bp_numb = Application.WorksheetFunction.Count(bp_range)

filt_row = 7
Dim fracs_array()
Dim fracs2_array()
Dim fracs3_array()
Dim dest_array()
Dim dest2_array()
Dim dest3_array()
Dim BD1()
Dim stdev1()
Dim BD2()
Dim BD3()
Dim stdev2()
Dim stdev3()
Dim max1col, max2col, max3col
Dim max1add, max2add, max3add
Dim  samp_no1()
Dim  samp_no2()
Dim samp_no3()

ReDim fracs_array(6 To bp_numb)
ReDim fracs2_array(6 To bp_numb)
ReDim fracs3_array(6 To bp_numb)
ReDim dest_array(6 To bp_numb)
ReDim dest2_array(6 To bp_numb)
ReDim dest3_array(6 To bp_numb)
ReDim BD1(6 To bp_numb)
ReDim stdev1(6 To bp_numb)
ReDim BD2(6 To bp_numb)
ReDim BD3(6 To bp_numb)
ReDim stdev2(6 To bp_numb)
ReDim stdev3(6 To bp_numb)

ReDim samp_no1(6 To bp_numb)
ReDim samp_no2(6 To bp_numb)
ReDim samp_no3(6 To bp_numb)

'calculations, cycling through To find the 3 highest copy number values and std deviations
n = 6
bp_row = 6
For bp_row = 6 To bp_numb
bp_row = n
Cells(bp_row, 3).Select
bp_now = Selection.Value
Worksheets("input").Select

Set fracs_array(n) = Range(Cells(bp_row, BD_start), Cells(bp_row, BD_start + Total_frac - 1))
fracs_vals = Application.WorksheetFunction.Count(fracs_array(n))
If fracs_vals = 0 Then GoTo 2000
```

```
    fracs_array(n).Copy

    Worksheets("filtered data").Select
    Set dest_array(n) = Range(Cells(filt_row, 25), Cells(filt_row, 25 + Total_frac - 1))
    dest_array(n).PAsteSpecial PAste:=xlPAsteValues, Operation:=xlNone, SkipBlanks:=False, Transpose:=False

    Worksheets("input").Select
    Set fracs2_array(n) = Range(Cells(bp_row, BD_start2), Cells(bp_row, BD_start2 + Total_frac - 1))
    fracs2_array(n).Copy
    Worksheets("filtered data").Select
    Set dest2_array(n) = Range(Cells(filt_row + 1, 25), Cells(filt_row + 1, 25 + Total_frac - 1))
    dest2_array(n).PAsteSpecial PAste:=xlPAsteValues, Operation:=xlNone, SkipBlanks:=False, Transpose:=False
    Worksheets("input").Select
    Set fracs3_array(n) = Range(Cells(bp_row, BD_start3), Cells(bp_row, BD_start3 + Total_frac - 1))
    fracs3_array(n).Copy

    Worksheets("filtered data").Select
    Set dest3_array(n) = Range(Cells(filt_row + 2, 25), Cells(filt_row + 2, 25 + Total_frac - 1))
    dest3_array(n).PAsteSpecial PAste:=xlPAsteValues, Operation:=xlNone, SkipBlanks:=False, Transpose:=False

    If fracs_vals = 1 Then GoTo 500
    If fracs_vals = 2 Then GoTo 1000
    GoTo 1500
500
    Cells(filt_row, 3).Select
    Selection.Value = bp_now
    Cells(filt_row, 5).Select
    Selection.Value = fracs_vals
    Cells(filt_row, 9).Select
    Selection.Value = WorksheetFunction.Large(dest_array(n), 1)
    Cells(filt_row, 9).Select
    max1 = Selection.Value
    Set max1add = Range(Cells(filt_row, 25), Cells(filt_row, 25 + Total_frac - 1)).Find(max1)
    max1add.Select
    max1col = Selection.Column
    Cells(6, max1col).Select BD1(n) =
    Selection.Value Cells(filt_row + 1,
    max1col).Select stdev1(n) =
    Selection.Value Cells(filt_row + 2,
    max1col).Select samp_no1(n) =
    Selection.Value Cells(filt_row,
    6).Select Selection.Value = BD1(n)
    Cells(filt_row, 12).Select
    Selection.Value = stdev1(n)
    Cells(filt_row, 15).Select
    Selection.Value = samp_no1(n)

    Range(Cells(filt_row + 1, 25), Cells(filt_row + 2, 25 + Total_frac - 1)).Clear
    filt_row = filt_row + 1

    GoTo 2000

1000
    Cells(filt_row, 3).Select
    Selection.Value = bp_now
    Cells(filt_row, 5).Select
    Selection.Value = fracs_vals
    Cells(filt_row, 9).Select
    Selection.Value = WorksheetFunction.Large(dest_array(n), 1)
    Cells(filt_row, 9).Select
    max1 = Selection.Value
```

```
  Set max1add = Range(Cells(filt_row, 25), Cells(filt_row, 25 + Total_frac - 1)).Find(max1)
  max1add.Select
  max1col = Selection.Column
  Cells(6, max1col).Select BD1(n) =
  Selection.Value Cells(filt_row + 1,
  max1col).Select stdev1(n) =
  Selection.Value Cells(filt_row + 2,
  max1col).Select samp_no1(n) =
  Selection.Value Cells(filt_row,
  15).Select Selection.Value =
  samp_no1(n)

  Cells(filt_row, 6).Select
  Selection.Value = BD1(n)
  Cells(filt_row, 12).Select
  Selection.Value = stdev1(n)

  Cells(filt_row, 10).Select
  Selection.Value = WorksheetFunction.Large(dest_array(n), 2)
  Cells(filt_row, 10).Select
  max2 = Selection.Value
  Set max2add = Range(Cells(filt_row, 25), Cells(filt_row, 25 + Total_frac - 1)).Find(max2)
  max2add.Select
  max2col = Selection.Column
  Cells(6, max2col).Select
  BD2(n) = Selection.Value
  Cells(filt_row + 1, max2col).Select
  stdev2(n) = Selection.Value
  Cells(filt_row + 2, max2col).Select
  samp_no2(n) = Selection.Value
  Cells(filt_row, 7).Select
  Selection.Value = BD2(n)
  Cells(filt_row, 13).Select
  Selection.Value = stdev2(n)
  Cells(filt_row, 16).Select
  Selection.Value = samp_no2(n)

  Range(Cells(filt_row + 1, 25), Cells(filt_row + 2, 25 + Total_frac - 1)).Clear
  filt_row = filt_row + 1

  GoTo 2000

1500
  Dim ttest1, ttest2, ttest3, nvals1, nvals2, nvals3, dfree1, dfree2, dfree3, stdfact1, stdfact2, stdfact3, pval1, pval2
  Cells(filt_row, 3).Select
  Selection.Value = bp_now
  Cells(filt_row, 5).Select
  Selection.Value = fracs_vals
  Cells(filt_row, 9).Select
  Selection.Value = WorksheetFunction.Large(dest_array(n), 1)
  Cells(filt_row, 9).Select
  max1 = Selection.Value
  Set max1add = Range(Cells(filt_row, 25), Cells(filt_row, 25 + Total_frac - 1)).Find(max1)
  max1add.Select
  max1col = Selection.Column
  Cells(6, max1col).Select
  BD1(n) = Selection.Value
  Cells(filt_row + 1, max1col).Select
  stdev1(n) = Selection.Value
  Cells(filt_row + 2, max1col).Select
  samp_no1(n) = Selection.Value
```

```
Cells(filt_row, 6).Select
Selection.Value = BD1(n)
Cells(filt_row, 12).Select
Selection.Value = stdev1(n)
Cells(filt_row, 15).Select
Selection.Value = samp_no1(n)

Cells(filt_row, 10).Select
Selection.Value = WorksheetFunction.Large(dest_array(n), 2)
Cells(filt_row, 10).Select
max2 = Selection.Value
Set max2add = Range(Cells(filt_row, 25), Cells(filt_row, 25 + Total_frac - 1)).Find(max2)
max2add.Select
max2col = Selection.Column
Cells(6, max2col).Select
BD2(n) = Selection.Value
Cells(filt_row + 1, max2col).Select
stdev2(n) = Selection.Value
Cells(filt_row + 2, max2col).Select
samp_no2(n) = Selection.Value
Cells(filt_row, 7).Select
Selection.Value = BD2(n)
Cells(filt_row, 13).Select
Selection.Value = stdev2(n)
Cells(filt_row, 16).Select
Selection.Value = samp_no2(n)

Cells(filt_row, 11).Select
Selection.Value = WorksheetFunction.Large(dest_array(n), 3)
Cells(filt_row, 11).Select
max3 = Selection.Value
Set max3add = Range(Cells(filt_row, 25), Cells(filt_row, 25 + Total_frac - 1)).Find(max3)
max3add.Select
max3col = Selection.Column
Cells(6, max3col).Select
BD3(n) = Selection.Value
Cells(filt_row + 1, max3col).Select
stdev3(n) = Selection.Value
Cells(filt_row + 2, max3col).Select
samp_no3(n) = Selection.Value

Cells(filt_row, 8).Select
Selection.Value = BD3(n)
Cells(filt_row, 14).Select
Selection.Value = stdev3(n)
Cells(filt_row, 17).Select
Selection.Value = samp_no3(n)

 nvals1 = (samp_no1(n) + samp_no2(n)) / (samp_no1(n) * samp_no2(n))
If stdev1(n) = "" Then stdev1(n) = 0
If stdev2(n) = "" Then stdev2(n) = 0
If stdev3(n) = "" Then stdev3(n) = 0

stdfact1 = ((samp_no1(n) - 1) * ((stdev1(n)) ^ 2) + (samp_no2(n) - 1) * ((stdev2(n)) ^ 2))
dfree1 = (samp_no1(n) + samp_no2(n) - 2)
If stdfact1 = 0 Then GoTo 1800
ttest1 = (max1 - max2) / (((stdfact1 / dfree1) * nvals1) ^ 0.5)
pval1 = Application.WorksheetFunction.T_Dist_2T(ttest1, dfree1)
Cells(filt_row, 18).Select
Selection.Value = ttest1
Cells(filt_row, 21).Select
Selection.Value = pval1
```

```
    nvals2 = (samp_no1(n) + samp_no3(n)) / (samp_no1(n) * samp_no3(n))
    stdfact2 = ((samp_no1(n) - 1) * ((stdev1(n)) ^ 2) + (samp_no3(n) - 1) * ((stdev3(n)) ^ 2))
    dfree2 = (samp_no1(n) + samp_no3(n) - 2)
    If stdfact2 = 0 Then GoTo 1800
    ttest2 = (max1 - max3) / (((stdfact2 / dfree2) * nvals2) ^ 0.5)
    pval2 = Application.WorksheetFunction.T_Dist_2T(ttest2, dfree2)
    Cells(filt_row, 19).Select
    Selection.Value = ttest2
    Cells(filt_row, 22).Select
    Selection.Value = pval2
    nvals3 = (samp_no2(n) + samp_no3(n)) / (samp_no2(n) * samp_no3(n))
    stdfact3 = ((samp_no2(n) - 1) * ((stdev2(n)) ^ 2) + (samp_no3(n) - 1) * ((stdev3(n)) ^ 2))
    dfree3 = (samp_no2(n) + samp_no3(n) - 2)
    If stdfact3 = 0 Then GoTo 1800
    ttest3 = (max2 - max3) / (((stdfact3 / dfree3) * nvals3) ^ 0.5)
    pval3 = Application.WorksheetFunction.T_Dist_2T(ttest3, dfree3)
    Cells(filt_row, 20).Select
    Selection.Value = ttest3
    Cells(filt_row, 23).Select
    Selection.Value = pval3

1800
    Range(Cells(filt_row + 1, 25), Cells(filt_row + 2, 25 + Total_frac - 1)).Clear
    filt_row = filt_row + 1

    GoTo 2000

2000

Sheets("input").Select
Cells(bp_row, 3).Select
Selection.OffSet(1, 0).Select

n = n + 1
Next bp_row
'Rank bp by order of largest peaks
 Dim rank_range As Range
 Dim sel_range As Range
 Dim bp_new
 Dim Tot_area As Range
 Dim  bpcurr_end  As  Integer
 Dim bpcurr_now  As  Integer
 Dim bpcurr_row As Integer
 Dim curr_max1, curr_max2, curr_max3, curr_std1, curr_std2, curr_std3, curr_pval1, curr_pval2, curr_pval3

Worksheets("input").Select
Range(Cells(3, 4), Cells(3, 4 + Total_frac - 1)).Select
Selection.Copy
Worksheets("interpolation").Select
Cells(3, 2).Select
ActiveSheet.PAste
Worksheets("input").Select
Range(Cells(6, 3), Cells(bp_row, 4 + Total_frac - 1)).Select
Selection.Copy
Worksheets("interpolation").Select
Cells(4, 1).Select
ActiveSheet.PAste

Worksheets("filtered data").Select
bp_new = 7
Cells(6, 3).Select
Selection.End(xlDown).Select
```

```
bp_row = Selection.Row
Set sel_range = Range(Cells(7, 9), Cells(bp_row, 9))
For bp_new = 7 To bp_row
Cells(bp_new, 4).Select
Selection.Value = Application.WorksheetFunction.Rank(Cells(bp_new, 9), sel_range)
Next bp_new
Range("C6").Select
   Range(Selection, Selection.End(xlToRight)).Select
   Range(Selection, Selection.End(xlDown)).Select
   Selection.Copy
   Sheets("output_Rank_Order").Select
   Range("C6").Select
   ActiveSheet.PAste
   Columns("O:T").Select
   Application.CutCopyMode = False
   Selection.Delete ShIft:=xlToLeft
   Range("R6").Select
   Selection.Value = "Combo1"
   Range("S6").Select
   Selection.Value = "Combo2"
   Range("T6").Select
   Selection.Value = "Combo3"

   Range("C6").Select
   Selection.End(xlDown).Select
   bpcurr_end = Selection.Row
   bpcurr_row = 7
   For bpcurr_row = 7 To bpcurr_end
   Cells(bpcurr_row, 9).Select
   curr_max1 = Selection.Value
   curr_max1 = Format(curr_max1, "scientIfic")
   Cells(bpcurr_row, 9).Value = curr_max1
   Cells(bpcurr_row, 10).Select
   curr_max2 = Selection.Value
   curr_max2 = Format(curr_max2, "scientIfic")
   Cells(bpcurr_row, 10).Value = curr_max2
   Cells(bpcurr_row, 11).Select
   curr_max3 = Selection.Value
   curr_max3 = Format(curr_max3, "scientIfic")
   Cells(bpcurr_row, 11).Value = curr_max3
   Cells(bpcurr_row, 12).Select
   curr_std1 = Selection.Value
   curr_std1 = Format(curr_std1, "scientIfic")
   Cells(bpcurr_row, 12).Value = curr_std1
   Cells(bpcurr_row, 13).Select
   curr_std2 = Selection.Value
   curr_std2 = Format(curr_std2, "scientIfic")
   Cells(bpcurr_row, 13).Value = curr_std2
      Cells(bpcurr_row, 14).Select
   curr_std3 = Selection.Value
   curr_std3 = Format(curr_std3, "scientIfic")
   Cells(bpcurr_row, 14).Value = curr_std3
   Cells(bpcurr_row, 15).Select
   curr_pval1 = Selection.Value
      If curr_pval1 = "" Then

   curr_pval1 = ""
   Else
   curr_pval1 = Round(curr_pval1, 3)
   Cells(bpcurr_row, 15).Value = curr_pval1
   End If
```

```
Cells(bpcurr_row, 16).Select
curr_pval2 = Selection.Value
If curr_pval2 = "" Then
curr_pval2 = ""
Else
curr_pval2 = Round(curr_pval2, 3)
Cells(bpcurr_row, 16).Value = curr_pval2
End If
Cells(bpcurr_row,   17).Select
curr_pval3  =  Selection.Value
If curr_pval3 = "" Then
curr_pval3 = ""
Else

curr_pval3 = Selection.Value
curr_pval3 = Round(curr_pval3, 3)
End If
Cells(bpcurr_row, 17).Value = curr_pval3

Dim maxTotal
Dim maxTotal2
Dim maxTotal3
If curr_std1 = "" Then
maxTotal = curr_max1
ElseIf curr_pval1 = "" Then
maxTotal = curr_max1 & "(" & curr_std1 & ")"
ElseIf curr_pval2 = "" Then
maxTotal = curr_max1 & "(" & curr_std1 & "; " & curr_pval1 & ")"
Else
maxTotal = curr_max1 & "(" & curr_std1 & "; " & curr_pval1 & ", " & curr_pval2 & ")"
End If
If curr_max2 = "" Then
maxTotal2 = ""
ElseIf curr_std2 = "" Then
maxTotal2 = curr_max2
ElseIf curr_pval3 = "" Then
maxTotal2 = curr_max2 & "(" & curr_std2 & ")"
Else
maxTotal2 = curr_max2 & "(" & curr_std2 & "; " & curr_pval3 & ")"
End If
If curr_max3 = "" Then
maxTotal3 = ""
ElseIf curr_std3 = "" Then
maxTotal3 = curr_max3
Else
maxTotal3 = curr_max3 & "(" & curr_std3 & ")"
End If
Cells(bpcurr_row, 18) = maxTotal
Cells(bpcurr_row, 19) = maxTotal2
Cells(bpcurr_row, 20) = maxTotal3
Next bpcurr_row
Range("C6").Select
Range(Selection, Cells(6, 20)).Select
Range(Selection, Selection.End(xlDown)).Select
Selection.Copy
Sheets("Output_BP_Order").Select
Range("C6").Select
ActiveSheet.PAste

Sheets("Output_Rank_Order").Select
Range("C7").Select
Range(Selection, Cells(7, 20)).Select
```

```
    Range(Selection, Selection.End(xlDown)).Select
    Set Tot_area = Selection
    Set rank_range = Range(Cells(7, 4), Cells(bp_row, 4))
    Tot_area.Sort key1:=rank_range, order1:=xlAscending
    Range(Selection, Cells(6, 20)).Select
    Range(Selection, Selection.End(xlDown)).Select

5000


End Sub
```

B.4. Script 4: Data interpolation and heat map generation

```
Sub Interpolate()
'
' Interpolate Macro
'
' Keyboard Shortcut: Ctrl+j
'
Dim bd_int()
Dim bd
Dim x As Integer
Dim fnl_bd, Total_vals
Dim rw_now
Dim rw_lst
Dim frac_lst
Dim bd_exp As Range
Dim alldata

'determine number of intermediate BD values

Sheets("input").Select Cells(9,
1).Select
Selection.End(xlDown).Select
fnl_bd = Selection.Row
Total_vals = fnl_bd - 8

'sort data by ascending BD values
Sheets("interpolation").Select
Cells(4, 1).Select
Selection.End(xlDown).Select
rw_lst = Selection.Row
Cells(3, 2).Select
Selection.End(xlToRight).Select
frac_lst = Selection.Column
Set bd_exp = Range(Cells(3, 2), Cells(3, frac_lst))
Range(Cells(3, 2), Cells(rw_lst, frac_lst)).Select
Set alldata = Selection
alldata.Sort key1:=bd_exp, order1:=xlAscending, Header:=xlNo, Orientation:=xlLeftToRight
'copy intermediate bd values
Dim newbdst, newbdfin
Dim y1()
Dim xdiff()
Dim yrange()
Dim xrange()
Dim inter_bp()
Dim bdexp_val()
Dim bdexp_col()
Dim bdint_col()
Dim curr_col
Dim int_col
```

```
Dim xfst()
Dim xlst()
Dim xfst_col()
Dim xlst_col()
Dim yfst()
Dim ylst()
Dim yvals()
'paste new bds and determine first and last columns
Sheets("input").Select
Range(Cells(9, 1), Cells(fnl_bd, 1)).Select
Selection.Copy
Sheets("interpolation").Select
Cells(3, frac_lst + 4).Select
ActiveCell.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks:=False, Transpose:=True
Cells(3, frac_lst + 4).Select
newbdst = Selection.Column
Selection.End(xlToRight).Select
newbdfin = Selection.Column
Range(Cells(4, 1), Cells(rw_lst, 1)).Select
Selection.Copy
Cells(4, newbdst - 1).Select
ActiveCell.PasteSpecial
'determine each bd value and column
ReDim bdexp_col(1 To frac_lst - 1)
ReDim bdexp_val(1 To frac_lst - 1)
curr_col = 2
a = 1
For curr_col = 2 To frac_lst
Cells(3, curr_col).Select
bdexp_val(a) = Selection.Value
bdexp_col(a) = Selection.Column
a = a + 1

Next curr_col
x = 1
ReDim bd_int(1 To Total_vals)
ReDim bdint_col(1 To Total_vals)
ReDim y1(1 To Total_vals)
ReDim xdiff(1 To Total_vals)
ReDim yrange(1 To Total_vals)
ReDim xrange(1 To Total_vals)
ReDim xfst(1 To Total_vals)
ReDim xlst(1 To Total_vals)
ReDim xfst_col(1 To Total_vals)
ReDim xlst_col(1 To Total_vals)
ReDim yfst(1 To Total_vals)
ReDim ylst(1 To Total_vals)
ReDim yval(1 To Total_vals)
int_col = newbdst
For int_col = newbdst To newbdfin

Cells(3, int_col).Select bd_int(x)
= Selection.Value bdint_col(x) =
Selection.Column b = 1
For b = 1 To frac_lst - 1
If bdexp_val(b) = bd_int(x) Then
xdiff(x) = 0
Exit For
EndIf
If bdexp_val(b) < bd_int(x) Then
xfst(x) = bdexp_val(b)
```

```
xfst_col(x) = bdexp_col(b)
EndIf
If bdexp_val(b) > bd_int(x) Then
xlst(x) = bdexp_val(b)
xlst_col(x) = bdexp_col(b)
Exit For
EndIf
Next b
If bdexp_val(b) = bd_int(x) Then
xdiff(x) = 0
xrange(x) = 1
GoTo 100
EndIf
xdiff(x) = bd_int(x) - xfst(x)
xrange(x) = xlst(x) - xfst(x)

100
x = x + 1
Next int_col

'interpolate values
Dim currcol2
Dim vals_range
r = 4
For r = 4 To rw_lst
Application.ScreenUpdating = False
currcol2 = newbdst
c = 1
For currcol2 = newbdst To newbdfin
Cells(r, currcol2).Select
yfst(c) = Cells(r, xfst_col(c)).Value
If yfst(c) = "" Then yfst(c) = 0
ylst(c) = Cells(r, xlst_col(c)).Value
If ylst(c) = "" Then ylst(c) = 0
yrange(c) = ylst(c) - yfst(c)
yval(c) = yfst(c) + (xdiff(c) * (yrange(c) / xrange(c)))
If yval(c) = 0 Then yval(c) = ""
Selection.Value = yval(c)
c = c + 1
Next currcol2
Next r
'calculate percentiles and paste on the next sheet
Range(Cells(4, newbdst), Cells(rw_lst, newbdfin)).Select
vals_range = Selection.Value
Range(Cells(3, newbdst - 1), Cells(rw_lst, newbdfin)).Select
Selection.Copy
Sheets("percentiles").Select
Cells(3, 1).Select
ActiveCell.PasteSpecial

Dim per2nd_col
Dim per1lstcol
Dim per2lstcol
Dim per2val()
Dim per1val()
Dim rwper
Dim val2()
Dim currper2()
Dim per2col
Dim bdint_tot
Dim per1col
Dim val1()
```

```
ReDim currper2(4 To rw_lst)
bdint_tot = newbdfin - newbdst + 1
Cells(3, (newbdfin - newbdst + 4)).Select
per2nd_col = ActiveCell.Column + 1
ActiveCell.PasteSpecial
Cells(3, per2nd_col).Select
Selection.End(xlToRight).Select
per2lstcol = ActiveCell.Column
Cells(3, 2).Select
Selection.End(xlToRight).Select
per1lstcol = ActiveCell.Column

ReDim val2(1 To bdint_tot)
ReDim per2val(1 To bdint_tot)
ReDim per1val(1 To bdint_tot)

e = 4
For rwper = 4 To rw_lst
currper2(e) = Range(Cells(rwper, 2), Cells(rwper, per1lstcol)).Value
d = 1
per2col = per2nd_col
For per2col = per2nd_col To per2lstcol

Cells(rwper, per2col).Select
val2(d) = Selection.Value
If val2(d) = "" Then
Selection.Value = "#N/A"

GoTo 500
EndIf
per2val(d) = Application.WorksheetFunction.PercentRank_Inc(currper2(e), val2(d))
Selection.Value = per2val(d)
d = d + 1
500
Next per2col
e = e + 1
Next rwper

ReDim val1(1 To bdint_Tot)

rwper = 4

For rwper = 4 To rw_lst
d = 1
per1col = 2
For per1col = 2 To per1lstcol

Cells(rwper, per1col).Select
val1(d) = Selection.Value
If val1(d) = "" Then
Selection.Value = "#N/A"

GoTo 1000
EndIf
per1val(d) = Application.WorksheetFunction.PercentRank_Inc(vals_range, val1(d))
Selection.Value = per1val(d)
d = d + 1
1000
Next per1col

Next rwper
```

```
Cells.Select
Selection.Copy
Sheets("heat_map").Select
Range("A1").Select
   ActiveSheet.Paste
Dim bps, fracs
  'Dim bps_array()
  'Dim fracs_array()
  Dim n As Integer
  Dim f As Integer
  Dim perc

  'ReDim bps_array(6 To 6 + bps)
  'ReDim fracs_array(40 To 40 + fracs)
  n = 4
  f = 2

  For n = 4 To rw_lst
  For f = 2 To per1lstcol
   Cells(n, f).Select
  perc = Selection.Value

    If Application.WorksheetFunction.IsNA(perc) = True Then
      With Selection.Interior
      .Pattern = xlSolid
      .PatternColorIndex = xlAuTomatic
      .Color = 2500134
      .TintAndShade = 0
      .PatternTintAndShade = 0
   EndWith
   GoTo 3000
  EndIf

  If Selection.Value >= 0 And Selection.Value < 0.1 Then
  With Selection.Interior
      .Pattern = xlSolid
      .PatternColorIndex = xlAuTomatic
      .Color = 6896921
      .TintAndShade = 0
      .PatternTintAndShade = 0
   EndWith
  EndIf

  If Selection.Value >= 0.1 And Selection.Value < 0.2 Then
  With Selection.Interior
      .Pattern = xlSolid
      .PatternColorIndex = xlAuTomatic
      .Color = 5392439
      .TintAndShade = 0
      .PatternTintAndShade = 0
   EndWith
  EndIf
   If Selection.Value >= 0.2 And Selection.Value < 0.3 Then
  With Selection.Interior
      .Pattern = xlSolid
      .PatternColorIndex = xlAuTomatic
      .Color = 5853248
      .TintAndShade = 0
      .PatternTintAndShade = 0
   EndWith
  EndIf
```

```
    If Selection.Value >= 0.3 And Selection.Value < 0.4 Then
  With Selection.Interior
      .Pattern = xlSolid
      .PatternColorIndex = xlAuTomatic
      .Color = 7723448
      .TintAndShade = 0
      .PatternTintAndShade = 0
   EndWith
  EndIf
   If Selection.Value >= 0.4 And Selection.Value < 0.5 Then
  With Selection.Interior
      .Pattern = xlSolid
      .PatternColorIndex = xlAuTomatic
      .Color = 8319705
      .TintAndShade = 0
      .PatternTintAndShade = 0
   EndWith
  EndIf
  If Selection.Value >= 0.5 And Selection.Value < 0.6 Then
  With Selection.Interior
      .Pattern = xlSolid
      .PatternColorIndex = xlAuTomatic
      .Color = 8257532
      .TintAndShade = 0
      .PatternTintAndShade = 0
   EndWith
  EndIf
    If Selection.Value >= 0.6 And Selection.Value < 0.7 Then
  With Selection.Interior
      .Pattern = xlSolid
      .PatternColorIndex = xlAuTomatic
      .Color = 5238527

      .TintAndShade = 0
      .PatternTintAndShade = 0
   EndWith
  EndIf
  If Selection.Value >= 0.7 And Selection.Value < 0.8 Then
  With Selection.Interior
      .Pattern = xlSolid
      .PatternColorIndex = xlAuTomatic
      .Color = 52476

      .TintAndShade = 0
      .PatternTintAndShade = 0
   EndWith
  EndIf
  If Selection.Value >= 0.8 And Selection.Value < 0.9 Then
  With Selection.Interior
      .Pattern = xlSolid
      .PatternColorIndex = xlAuTomatic
      .Color = 1355764

      .TintAndShade = 0
      .PatternTintAndShade = 0
   EndWith
  EndIf

 If Selection.Value >= 0.9 And Selection.Value < 0.94 Then
  With Selection.Interior
      .Pattern = xlSolid
      .PatternColorIndex = xlAuTomatic
```

```
        .Color = 881143

        .TintAndShade = 0
        .PatternTintAndShade = 0
    EndWith
   EndIf
If Selection.Value >= 0.94 And Selection.Value < 0.98 Then
   With Selection.Interior
        .Pattern = xlSolid
        .PatternColorIndex = xlAuTomatic
        .Color = 2382042

        .TintAndShade = 0
        .PatternTintAndShade = 0
    EndWith
   EndIf
  If Selection.Value >= 0.98 Then
   With Selection.Interior
        .Pattern = xlSolid
        .PatternColorIndex = xlAuTomatic
        .Color = 2895784
        .TintAndShade = 0
        .PatternTintAndShade = 0
    EndWith
   EndIf

 3000

  Next f
  Dim g As Integer
  g = per2nd_col

For g = per2nd_col To per2lstcol
 Cells(n, g).Select
   perc = Selection.Value

     If Application.WorksheetFunction.IsNA(perc) = True Then
        With Selection.Interior
        .Pattern = xlSolid
        .PatternColorIndex = xlAuTomatic
        .Color = 2500134
        .TintAndShade = 0
        .PatternTintAndShade = 0
    EndWith
    GoTo 4000
   EndIf

   If Selection.Value >= 0 And Selection.Value < 0.1 Then
   With Selection.Interior
        .Pattern = xlSolid
        .PatternColorIndex = xlAuTomatic
        .Color = 6896921
        .TintAndShade = 0
        .PatternTintAndShade = 0
    EndWith
   EndIf

   If Selection.Value >= 0.1 And Selection.Value < 0.2 Then
   With Selection.Interior
        .Pattern = xlSolid
        .PatternColorIndex = xlAuTomatic
        .Color = 5392439
```

```
    .TintAndShade = 0
    .PatternTintAndShade = 0
 EndWith
EndIf
 If Selection.Value >= 0.2 And Selection.Value < 0.3 Then
With Selection.Interior
    .Pattern = xlSolid
    .PatternColorIndex = xlAuTomatic
    .Color = 5853248
    .TintAndShade = 0
    .PatternTintAndShade = 0
 EndWith
EndIf
   If Selection.Value >= 0.3 And Selection.Value < 0.4 Then
With Selection.Interior
    .Pattern = xlSolid
    .PatternColorIndex = xlAuTomatic
    .Color = 7723448
    .TintAndShade = 0
    .PatternTintAndShade = 0
 EndWith
EndIf
 If Selection.Value >= 0.4 And Selection.Value < 0.5 Then
With Selection.Interior
    .Pattern = xlSolid
    .PatternColorIndex = xlAuTomatic
    .Color = 8319705
    .TintAndShade = 0
    .PatternTintAndShade = 0
 EndWith
EndIf
If Selection.Value >= 0.5 And Selection.Value < 0.6 Then
With Selection.Interior
    .Pattern = xlSolid
    .PatternColorIndex = xlAuTomatic
    .Color = 8257532
    .TintAndShade = 0
    .PatternTintAndShade = 0
 EndWith
EndIf
  If Selection.Value >= 0.6 And Selection.Value < 0.7 Then
With Selection.Interior
    .Pattern = xlSolid
    .PatternColorIndex = xlAuTomatic
    .Color = 5238527

    .TintAndShade = 0
    .PatternTintAndShade = 0
 EndWith
EndIf
If Selection.Value >= 0.7 And Selection.Value < 0.8 Then
With Selection.Interior
    .Pattern = xlSolid
    .PatternColorIndex = xlAuTomatic
    .Color = 52476

    .TintAndShade = 0
    .PatternTintAndShade = 0
 EndWith
EndIf
If Selection.Value >= 0.8 And Selection.Value < 0.9 Then
With Selection.Interior
```

```
        .Pattern = xlSolid
        .PatternColorIndex = xlAuTomatic
        .Color = 1355764

        .TintAndShade = 0
        .PatternTintAndShade = 0
    EndWith
   EndIf

If Selection.Value >= 0.9 And Selection.Value < 0.94 Then
   With Selection.Interior
        .Pattern = xlSolid
        .PatternColorIndex = xlAuTomatic
        .Color = 881143

        .TintAndShade = 0
        .PatternTintAndShade = 0
    EndWith
   EndIf
If Selection.Value >= 0.94 And Selection.Value < 0.98 Then
   With Selection.Interior
        .Pattern = xlSolid
        .PatternColorIndex = xlAuTomatic
        .Color = 2382042

        .TintAndShade = 0
        .PatternTintAndShade = 0
    EndWith
   EndIf
  If Selection.Value >= 0.98 Then
   With Selection.Interior
        .Pattern = xlSolid
        .PatternColorIndex = xlAuTomatic
        .Color = 2895784
        .TintAndShade = 0
        .PatternTintAndShade = 0
    EndWith
   EndIf

4000

  Next g

g = per2nd_col + 1

  Next n
  EndSub
```

C. Eglin SIP data

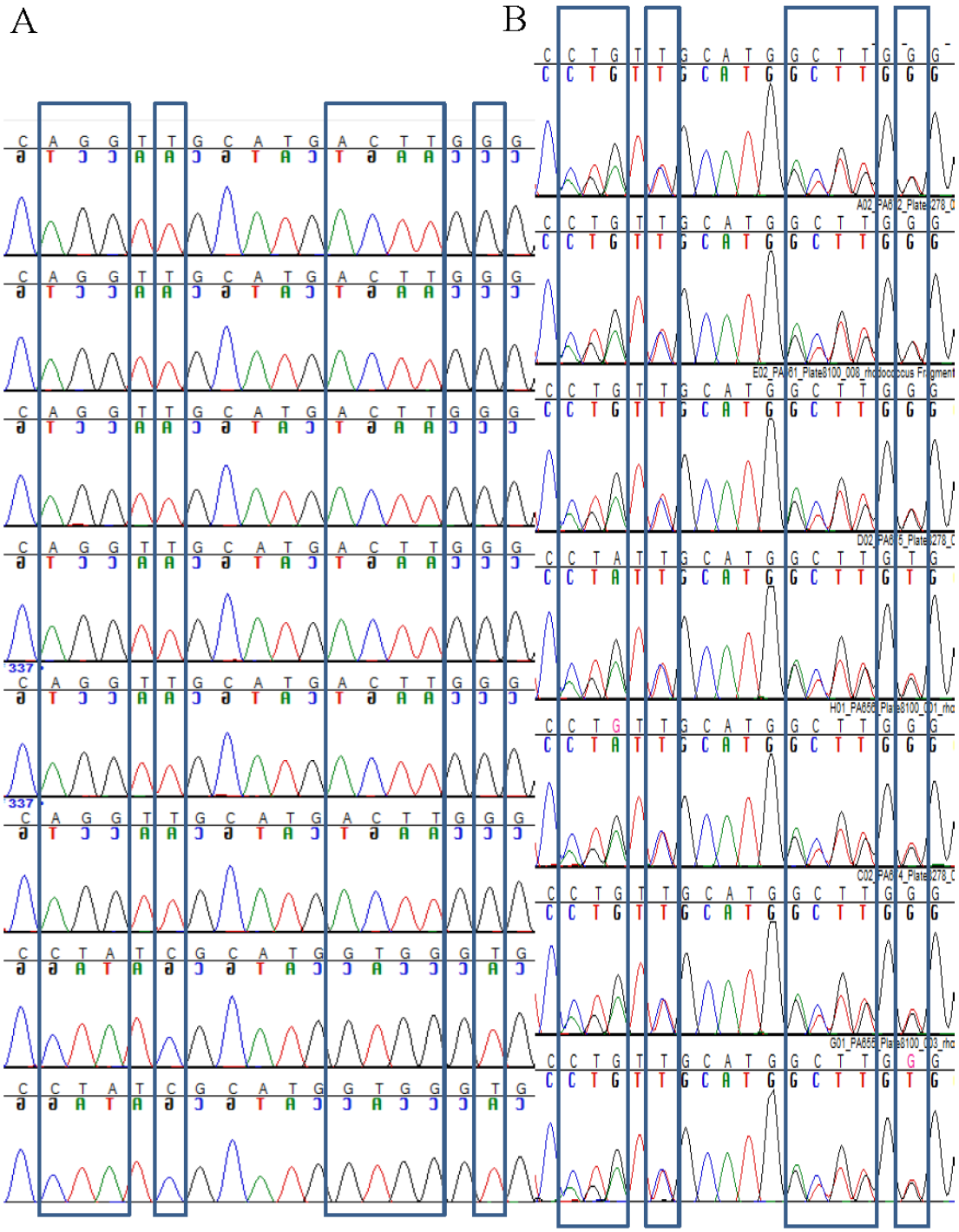


Figure C.l. An 18 bp segment of DNA in the 16S rRNA genes of *Rhodococcus* **sp. sequences.** (A) Alignment of sequences from the clone library sequences listed in Table B.2. (B) Alignment of sequences from Rhodococcus sp. EG2A and EG2B. Boxes indicate regions of variation (A) or multiple bp peaks (B). Isolate sequences (B) were from multiple colonies taken at different times.

C.1 Populations identified in clone library sequences and their RFs.

Table C.1. Summary of clone library sequences with RFs of up to 150 bp.

RF (bp)	Population[1]	Var.[2]	Relation To Others (<99%)[3]	^{14}N1			^{15}N1	^{15}N2	^{15}Ninactive		
				Clone Group[4]							
				1	2	3	1	1	1	2	3
40	*Chitinophaga* sp.				1						
67	Actinobacteria						1				
94	*Stenotrophomonas* sp.	A								1	2
94	*Stenotrophomonas* sp.	B								1	
116	Alphaprotebacteria		84% with Alphaproteobacteria (Bosea)								1
119	*Mesorhizobium* sp.	A	98% F, G	1		1			1		
119	*Mesorhizobium* sp.	B	98% E - G	1							
119	*Mesorhizobium* sp.	C	98% D - G		1						
119	*Mesorhizobium* sp.	D	98% C, E, F, 97% G						1		
119	*Mesorhizobium* sp.	E	98% B, C, D, 97% F, G						1		
119	*Mesorhizobium* sp.	F	98% A - D, 97% E				2				
119	*Mesorhizobium* sp.	G	98% A - C, 97% D, E					1			
135	*Burkholderia* sp.								1		
139	*Terrabacter* sp.	A	96% B,C,D, 94% E				2				
139	*Terrabacter* sp.	B	97%C, 96% A,D, 94% E		1						
139	*Terrabacter* sp.	C	97% B, D, 96% A, 94% E	1							
139	*Terrabacter* sp.	D	97% C, 96% A,B, 94% E	1							
141	*Terrabacter* sp.	E	94% A - D			1					
144	Alphaproteobacteria (*Bosea* sp.?)	A		1	1	1					
144	Alphaproteobacteria (*Bosea* sp.?)	B	98% E		1						
144	Alphaproteobacteria (*Bosea* sp.?)	C	98% E			1					
144	Alphaproteobacteria (*Bosea* sp.?)	D	98% E			1					
144	Alphaproteobacteria (*Bosea* sp.?)	E	98% B - D								1
144	*Caulobacter* sp.		92% with Alphaproteobacteria (Bosea)	1							
145	*Bradyrhizobium* sp.	B	98% A					1			
146	*Bradyrhizobium* sp.	A	98% B						1		
150	*Rhodanobacter* sp.	A	98%D	1	1						
150	*Rhodanobacter* sp.	B	98% A,C						1		
150	*Rhodanobacter* sp.	C	98% B							1	
150	*Rhodanobacter* sp.	D									1

1. Putative identification of clones based on partial 16S rRNA sequences
2. Sequence variations. Sequences have at least 1 bp mismatch with the other clones of the same genera.
3. Sequence identity between the clones of the same genera. If a percentage is not listed, the identity is at least 99%.

4. Clone group. Each group represents pooled clones that were sequenced after TRFLP analysis was performed on them.

Table C.2. Summary of clone library sequences with RFs of greater than 150 bp.

RF (bp)	Population[1]	Var.[2]	Relation To Others (<99%)[3]	^{14}N1			^{15}N1	^{15}N2	^{15}Ninactive		
				Clone Group[4]							
				1	2	3	1	1	1	2	3
154	*Rhodococcus* sp.	A	98% B, 97% E, 96% C	1	1	1					
154	*Rhodococcus* sp.	B	98% A, 97% E, 96% C	1		1					
154	*Rhodococcus* sp.	C	97% D, 96% A,B					2			
154	*Rhodococcus* sp.	D	97% C, E			1					
154	*Rhodococcus* sp.	E	97% A, B, D					1			
155	*Arthrobacter* sp.	B	96% A, C, 75% D							1	
156	*Arthrobacter* sp.	A	96% B, 95% C, 76% D							1	
156	*Arthrobacter* sp.	D	76% A, 75% B, C							1	
156	*Streptomyces* sp.							1			
158	*Arthrobacter* sp.	C	96% B, 95% A, 75% D				1				
163	Gammaproteobacteria	A	97% F, I	1							1
163	Gammaproteobacteria	B	97% F, 96% I				2				
163	Gammaproteobacteria	C	97% F, I		1						
163	Gammaproteobacteria	D	97% F, I		1						
163	Gammaproteobacteria	E	97% F, I			1					
163	Gammaproteobacteria	F	98% I, 97% A - E, G, H						1		
163	Gammaproteobacteria	G	97% F, 96% I							1	
163	Gammaproteobacteria	H	97% F, I								1
163	Gammaproteobacteria	I	98% F, 97%A, C - F, H, 96% B,G								1
163	Gammaproteobacteria								1		
163	*Pseudomonas* sp.	A				1					
163	*Pseudomonas* sp.	B									1
163	*Pseudomonas* sp.	C									1
163	*Pseudomonas* sp.	D								1	
204	*Janthinobacterium* sp.	A		1						1	
204	*Janthinobacterium* sp.	B								1	
204	*Janthinobacterium* sp.	C									1
204	Betaproteobacteria	C	98% E, D, 91% B, 90% A		1						
204	Betaproteobacteria	D	98% C, 91% B, 90% A			1					
204	Betaproteobacteria	E	98% C, 91% B, 90% A					1			
204	Betaproteobacteria	F	91% A, B							1	
231	*Variovorax* sp.	A			1						
245	*Rhizobium* sp.	A	All within 99%					3	2		1
163?	Betaproteobacteria	A	91% F, 90% C - E				1				
163?	Betaproteobacteria	B	91% C - F					1			
169?	*Ralstonia* sp.				1						

1. Putative identification of clones based on partial 16S rRNA sequences
2. Sequence variations. Sequences have at least 1 bp mismatch with the other clones of the same genera.
3. Sequence identity between the clones of the same genera. If a percentage is not listed, the identity is at least 99%.

4. Clone group. Each group represents pooled clones that were sequenced after TRFLP analysis was performed on them.

Table C.3. Populations isolated from Eglin soils and their corresponding RFs.

RF	Isolated Organisms
154	*Rhodococcus* sp. EG2A, EG2B
129/135	*Burkholderia* sp.
112	*Siphonobacter* sp.
50	*Pseudomonas* sp.
156	*Arthrobacter* sp.
157	*Arthrobacter* sp.
245	*Rhizobium* sp.
123	*Williamsia* sp. EG1

Table C.4. Example and descriptions of RF copy number data shown in Tables C.5 – C.20.

Bps	BD1 (g/ml)[a]	BD2(g/ml)	BD3(g/ml)
154	1.731	1.735	1.739
	5.23E+07(1.49E+07; 0.432, 0.143)[b]	4.25E+07(1.00E+07; 0.136)[c]	2.69E+07(3.60E+06)[d]

a. Buoyant density of fraction with highest average copy numbers for the specific restriction fragment.
b. Number of copy numbers/ fraction. In parentheses: (standard deviation of copy numbers; p-value calculated from t-tests performed between highest and second highest average copy numbers, p-value calculated from t-tests performed between the highest and third highest average copy numbers).
c. Number of copy numbers/fraction. In parentheses: (standard deviation of copy numbers; p-value calculated from t-tests performed between the second and third highest copy numbers)
d. Third highest copy numbers with standard deviation in parentheses.

Table C.5. Twenty RFs with the highest copy numbers in the ^{15}N1 endpoint gradient.

Bps	BD1 (g/ml)	BD2(g/ml)	BD3(g/ml)
154	1.731	1.735	1.739
	5.23E+07(1.49E+07; 0.432, 0.143)	4.25E+07(1.00E+07; 0.136)	2.69E+07(3.60E+06)
245	1.731	1.728	1.724
	1.56E+07(6.31E+06; 0.443, 0.166)	1.12E+07(1.56E+06; 0.045)	6.01E+06(3.70E+05)
119	1.731	1.728	1.724
	1.48E+07(3.58E+06; 0.588, 0.141)	1.30E+07(1.56E+06; 0.062)	8.72E+06(2.88E+05)
163	1.731	1.727	1.735
	1.23E+07(3.41E+06; 0.218, 0.112)	7.88E+06(8.19E+05; 0.636)	7.06E+06(2.03E+06)
139	1.731	1.727	1.735
	1.06E+07(2.65E+06; 0.22, 0.077)	7.20E+06(5.70E+05; 0.305)	5.60E+06(1.70E+06)
144	1.731	1.739	1.735
	1.03E+07(2.48E+06; 0.896, 0.653)	9.99E+06(8.79E+05; 0.682)	9.23E+06(2.17E+06)
112	1.716	1.720	1.724
	9.79E+06(1.62E+06; 0.072, 0.046)	5.70E+06(2.84E+05; 0.053)	4.55E+06(2.71E+05)
94	1.724	1.727	1.720
	6.23E+06(3.97E+05; 0.033, 0.009)	3.55E+06(5.78E+05; 0.557)	3.26E+06(9.98E+03)
40	1.712	1.716	1.708
	6.17E+06(1.01E+06; 0.393, 0.048)	5.30E+06(8.67E+05; 0.13)	3.67E+06(3.10E+05)
143	1.727	1.724	1.720
	5.84E+06	5.09E+06(1.90E+05; 0.003)	1.33E+06(2.27E+05)
156	1.724	1.716	1.727
	5.61E+06(2.42E+05; 0.107, 0.075)	4.86E+06(2.90E+05; 0.477)	4.58E+06(3.44E+05)
204	1.716	1.724	1.720
	5.56E+06(2.81E+05; 0.037, 0.068)	4.54E+06(4.65E+04; 0.287)	3.99E+06(5.42E+05)
231	1.731	1.739	1.735
	5.04E+06(1.91E+06; 0.585, 0.321)	4.16E+06(3.71E+05; 0.494)	3.55E+06(1.01E+06)
85	1.716	1.712	1.720
	4.58E+06(5.24E+05; 0.005, 0.019)	2.19E+06(2.01E+05; 0.192)	1.93E+06(4.83E+04)
145	1.727	1.731	1.724
	4.54E+06(3.73E+06; 0.657, 0.36)	3.16E+06(6.04E+05; 0.056)	1.43E+06(2.63E+04)
113	1.731	1.735	1.739
	4.36E+06(1.53E+06; 0.139, 0.176)	2.44E+06(6.86E+05; 0.567)	2.11E+06(1.61E+05)
141	1.731	1.727	1.724
	3.55E+06(6.13E+05; 0.214, 0.088)	2.70E+06(2.65E+05; 0.149)	1.90E+06(4.17E+05)
80	1.731	1.735	1.739
	3.24E+06	1.51E+06(4.61E+05)	9.92E+05
273	1.731	1.735	1.739
	2.87E+06(8.13E+05; 0.854, 0.218)	2.67E+06(1.16E+06; 0.405)	1.84E+06(1.20E+05)
147	1.731	1.727	1.735
	2.85E+06(5.79E+05; 0.231, 0.193)	2.03E+06(3.52E+05; 0.88)	1.95E+06(5.88E+05)

Table C.6. RFs with 21st to 40th copy numbers in the ^{15}N1 endpoint gradient.

Bps	BD1 (g/ml)	BD2(g/ml)	BD3(g/ml)
232	1.724	1.716	1.720
	2.68E+06(3.50E+05; 0.018, 0.033)	8.41E+05(1.84E+04; 0.36)	4.48E+05(4.71E+05)
81	1.731	1.727	1.739
	2.63E+06	1.44E+06(2.45E+05)	1.01E+06
135	1.724	1.727	1.731
	2.49E+06(4.75E+05; 0.263, 0.122)	1.93E+06(2.09E+05; 0.204)	1.54E+06(2.11E+05)
274	1.739	1.731	1.724
	2.33E+06(2.10E+05; 0.28, 0.01)	1.91E+06(3.45E+05; 0.047)	8.11E+05(4.89E+04)
159	1.739	1.743	1.723
	2.17E+06(8.34E+05; 0.733, 0.352)	1.94E+06(4.22E+04; 0.008)	1.46E+06(4.39E+04)
146	1.727	1.708	1.705
	2.04E+06	1.05E+06	9.83E+05(8.84E+04)
121	1.716	1.712	1.720
	1.90E+06(9.77E+04; 0.006, 0.028)	1.22E+06(1.09E+05; 0.025)	3.64E+05(3.56E+05)
158	1.724	1.727	1.731
	1.86E+06(2.40E+04; 0.435, 0.015)	1.75E+06(1.51E+05; 0.1)	5.88E+05
281	1.724	1.727	1.720
	1.47E+06(1.43E+05; 0.252, 0.015)	1.07E+06(3.20E+05; 0.14)	5.13E+05(8.28E+04)
57	1.724	1.727	1.716
	1.37E+06(1.42E+05; 0.069, 0.059)	9.65E+05(7.24E+04; 0.351)	8.40E+05(1.27E+05)
235	1.739	1.731	1.743
	1.26E+06(1.78E+05; 0.463, 0.118)	1.01E+06	4.33E+05(4.03E+05)
157	1.743	1.731	1.739
	1.20E+06(1.23E+04; 0.025, 0.004)	8.15E+05	7.96E+05(3.37E+04)
83	1.712	1.716	1.708
	1.11E+06(8.21E+04; 0.061, 0.004)	8.77E+05(9.93E+04; 0.06)	6.00E+05(1.42E+04)
206	1.731	1.727	1.739
	1.10E+06(3.77E+05; 0.755, 0.388)	9.94E+05(2.06E+05; 0.272)	4.41E+05
132	1.724	1.727	1.739
	1.07E+06	2.38E+05	7.65E+04
212	1.724	1.716	1.720
	1.03E+06(6.82E+04; 0.86, 0.116)	1.02E+06(1.61E+03; 0.065)	8.51E+05(6.28E+04)
169	1.731	1.727	1.724
	1.02E+06(9.17E+03; 0.025, 0.181)	7.57E+05(5.92E+04; 0.664)	6.66E+05(2.48E+05)
171	1.731	1.727	1.739
	1.02E+06(2.36E+04; 0.038, 0.003)	5.95E+05(1.19E+05; 0.456)	5.15E+05(3.29E+04)
115	1.724	1.727	1.731
	1.02E+06(1.60E+05; 0.769, 0.703)	9.75E+05(7.22E+04; 0.869)	9.61E+05(8.61E+04)
234	1.720	1.731	1.735
	9.98E+05	6.45E+05	5.43E+05(1.81E+05)

Table C.7. Twenty RFs with the highest copy numbers in the ^{15}N2 endpoint gradient.

Bps	BD1 (g/ml)	BD2(g/ml)	BD3(g/ml)
154	1.734	1.730	1.738
	5.67E+07(6.58E+05; 0.062, 0.001)	4.10E+07(7.21E+06; 0.02)	1.61E+07(1.89E+06)
119	1.730	1.726	1.734
	3.09E+07(2.15E+06; 0.005, 0.003)	1.96E+07(2.78E+06; 0.187)	1.61E+07(2.94E+05)
207	1.730	1.726	1.723
	1.87E+07(1.71E+06; 0.372, 0.06)	1.71E+07(2.13E+06; 0.258)	1.49E+07(4.43E+05)
246	1.730	1.726	1.723
	1.86E+07(2.14E+06; 0.056, 0.013)	1.38E+07(2.21E+06; 0.106)	9.99E+06(5.40E+05)
144	1.734	1.730	1.738
	1.71E+07(3.17E+05; 0, 0.001)	1.05E+07(2.75E+05; 0.005)	8.80E+06(2.10E+05)
139	1.730	1.734	1.726
	1.43E+07(1.74E+06; 0.087, 0.008)	1.08E+07(8.84E+05; 0.077)	8.64E+06(9.10E+05)
163	1.730	1.734	1.726
	8.59E+06(5.16E+05; 0.033, 0.004)	7.05E+06(2.64E+05; 0.063)	5.21E+06(8.38E+05)
83	1.708	1.711	1.715
	8.35E+06(1.09E+05; 0.007, 0)	5.26E+06(3.49E+05; 0.003)	3.12E+06(2.22E+05)
231	1.734	1.730	1.738
	7.10E+06(6.56E+05; 0.066, 0.019)	5.48E+06(6.15E+05; 0.031)	3.70E+06(1.20E+05)
204	1.719	1.723	1.715
	6.85E+06(3.04E+05; 0.007, 0)	5.21E+06(2.08E+05; 0.001)	1.91E+06(3.38E+05)
113	1.734	1.7303	1.738
	6.36E+06(5.97E+05; 0.077, 0.01)	4.81E+06(6.59E+05; 0.01)	1.82E+06(2.13E+05)
156	1.719	1.723	1.726
	5.55E+06(3.33E+05; 0.085, 0.008)	4.91E+06(7.58E+04; 0.057)	3.34E+06(6.96E+05)
145	1.730	1.708	1.726
	5.45E+06(4.74E+05; 0.02, 0.001)	2.71E+06(9.26E+05; 0.73)	2.50E+06(3.70E+05)
141	1.730	1.726	1.734
	4.29E+06(5.35E+05; 0.079, 0.031)	3.09E+06(7.14E+05; 0.517)	2.69E+06(2.20E+05)
40	1.708	1.711	1.715
	3.57E+06(2.96E+05; 0.025, 0.002)	2.13E+06(1.34E+05; 0.048)	1.68E+06(1.60E+05)
249	1.730	1.726	1.723
	3.51E+06(3.42E+05; 0.045, 0.007)	2.53E+06(4.80E+05; 0.049)	1.08E+06(5.20E+05)
147	1.730	1.734	1.726
	3.34E+06(7.07E+05; 0.385, 0.018)	2.73E+06(5.26E+05; 0.057)	1.60E+06(3.36E+05)
273	1.734	1.730	1.726
	3.23E+06(8.11E+04; 0.011, 0.008)	2.13E+06(2.59E+05; 0.047)	1.37E+06(3.92E+05)
135	1.719	1.723	1.715
	2.10E+06(2.38E+05; 0.062, 0)	1.45E+06(2.59E+05; 0.009)	6.14E+05(2.74E+04)
159	1.738	1.730	1.734
	1.93E+06(1.24E+05; 0.167, 0.014)	1.32E+06(4.46E+05; 0.565)	1.10E+06(6.00E+04)

Table C.8. RFs with 21st to 40th copy numbers in the ^{15}N2 endpoint gradient.

Bps	BD1 (g/ml)	BD2(g/ml)	BD3(g/ml)
274	1.730	1.734	1.726
	1.60E+06(1.64E+05; 0.818, 0.014)	1.49E+06(7.92E+05; 0.335)	9.87E+05(1.94E+05)
80	1.734	1.730	1.726
	1.60E+06(5.32E+04; 0.897, 0.002)	1.59E+06(1.17E+05; 0.001)	7.16E+05(1.02E+05)
116	1.734	1.726	1.730
	1.46E+06(1.25E+03; 0.198, 0.143)	8.11E+05(5.25E+05; 0.937)	7.78E+05(4.62E+05)
85	1.711	1.719	1.715
	1.45E+06(1.08E+05; 0.059, 0.073)	1.14E+06(1.15E+05; 0.765)	1.11E+06(1.51E+05)
118	1.730	1.711	
	1.42E+06	5.32E+04	
126	1.723	1.719	1.726
	1.30E+06(3.64E+05; 0.511, 0.184)	1.01E+06(4.61E+05; 0.509)	7.76E+05(3.21E+05)
115	1.730	1.726	1.734
	1.28E+06(2.82E+05; 0.238, 0.188)	7.39E+05	6.56E+05(5.74E+05)
158	1.730	1.723	1.738
	1.15E+06(1.07E+06; 0.881, 0.888)	1.02E+06(2.97E+05; 0.996)	1.02E+06(7.47E+05)
224	1.734		
	1.04E+06		
58	1.723	1.719	1.730
	1.03E+06(1.64E+04; 0.709, 0.014)	1.02E+06(3.78E+04; 0.003)	5.89E+05(1.14E+05)
47	1.734	1.730	1.738
	9.40E+05(1.09E+05; 0.514, 0.037)	8.96E+05(1.91E+04; 0.001)	5.03E+05(5.49E+04)
142	1.730	1.726	1.723
	9.17E+05(6.27E+05; 0.812, 0.404)	7.20E+05	4.14E+05(4.29E+05)
157	1.726	1.719	
	8.93E+05	2.66E+04	
272	1.730	1.734	1.726
	8.87E+05(1.94E+05; 0.6, 0.138)	7.49E+05	3.49E+05
283	1.734	1.730	1.738
	8.86E+05(8.46E+05; 0.316, 0.37)	3.48E+05(5.07E+04; 0.135)	1.93E+05(1.26E+05)
238	1.734	1.730	1.726
	8.32E+05(4.83E+04; 0.777, 0.165)	7.66E+05(2.83E+05; 0.62)	6.72E+05(1.13E+05)
183	1.734		
	8.11E+05		
237	1.730	1.734	1.726
	7.79E+05(3.48E+05; 0.567, 0.32)	5.05E+05	4.68E+05(5.59E+04)
212	1.723	1.719	1.726
	7.42E+05(1.24E+05; 0.527, 0.028)	5.53E+05(3.45E+05; 0.295)	2.27E+05(5.70E+03)
133	1.730	1.726	1.734
	7.32E+05(3.24E+05; 0.146, 0.254)	3.94E+05(2.58E+04; 0.909)	3.91E+05(3.86E+03)

Table C.9. Twenty RFs with the highest copy numbers in the ^{15}N3 endpoint gradient.

Bps	BD1 (g/ml)	BD2(g/ml)	BD3(g/ml)
154	1.733	1.737	1.730
	1.14E+07(4.34E+05; 0.036, 0.023)	9.42E+06(3.24E+05; 0.959)	9.41E+06(3.84E+04)
119	1.730	1.733	1.726
	3.87E+06(3.11E+05; 0.031, 0.027)	2.65E+06(1.89E+04; 0.039)	2.55E+06(2.12E+04)
139	1.730	1.733	1.726
	3.74E+06(2.32E+05; 0.132, 0.011)	3.32E+06(6.98E+04; 0.002)	2.16E+06(1.89E+04)
231	1.733	1.737	1.730
	3.25E+06(5.77E+04; 0.497, 0.028)	2.83E+06(7.21E+05; 0.735)	2.63E+06(1.40E+05)
144	1.733	1.730	1.737
	2.86E+06(3.29E+04; 0.043, 0.162)	2.51E+06(1.01E+05; 0.345)	2.03E+06(5.38E+05)
245	1.730	1.733	1.726
	2.54E+06(1.17E+05; 0.011, 0.009)	1.72E+06(2.33E+04; 0.084)	1.58E+06(5.88E+04)
40	1.708	1.711	1.704
	2.08E+06(2.37E+05; 0.533, 0.142)	1.88E+06(2.96E+05; 0.425)	1.66E+06(7.74E+04)
204	1.715	1.719	1.711
	1.65E+06(7.06E+04; 0.074, 0.093)	1.48E+06(9.69E+03; 0.171)	1.14E+06(2.26E+05)
156	1.730	1.733	1.726
	1.49E+06(1.91E+05; 0.671, 0.1)	1.40E+06(1.50E+05; 0.101)	1.09E+06(3.27E+04)
163	1.733	1.730	1.737
	1.18E+06(1.79E+04; 0.168, 0.047)	1.13E+06(3.18E+04; 0.062)	7.28E+05(1.44E+05)
273	1.733	1.730	1.737
	1.16E+06(7.79E+04; 0.053, 0.124)	8.19E+05(8.63E+04; 0.48)	6.43E+05(2.75E+05)
141	1.730	1.733	1.726
	1.11E+06(1.62E+04; 0.047, 0.003)	9.56E+05(4.74E+04; 0.016)	6.23E+05(3.65E+04)
158	1.733	1.737	1.730
	1.00E+06(9.86E+03; 0.835, 0.006)	9.67E+05(2.22E+05; 0.281)	7.35E+05(2.83E+04)
145	1.730	1.733	1.708
	9.57E+05(1.94E+04; 0.327, 0.013)	9.07E+05(5.16E+04; 0.038)	6.71E+05(4.32E+04)
159	1.737	1.733	1.730
	9.28E+05(2.04E+05; 0.983, 0.12)	9.24E+05(8.05E+04; 0.023)	5.49E+05(1.30E+04)
113	1.733	1.730	1.737
	8.03E+05(4.31E+03; 0.025, 0.083)	5.86E+05(4.94E+04; 0.531)	5.14E+05(1.26E+05)
147	1.733	1.730	1.737
	7.50E+05(6.04E+04; 0.076, 0.039)	5.83E+05(3.42E+04; 0.091)	3.67E+05(9.26E+04)
81	1.733	1.730	1.726
	6.73E+05(2.79E+04; 0.044, 0.004)	5.64E+05(1.84E+04; 0.004)	3.30E+05(9.93E+03)
85	1.715	1.711	1.719
	5.48E+05(2.30E+04; 0.027, 0.009)	3.60E+05(3.82E+04; 0.677)	3.46E+05(1.43E+04)
94	1.726	1.719	1.722
	5.37E+05(3.88E+04; 0.648, 0.216)	5.23E+05(7.00E+03; 0.072)	4.86E+05(1.31E+04)

Table C.10. RFs with 21st to 40th copy numbers in the ^{15}N3 endpoint gradient.

Bps	BD1 (g/ml)	BD2(g/ml)	BD3(g/ml)
234	1.737	1.733	1.730
	5.22E+05(1.34E+05; 0.457, 0.105)	4.35E+05(9.49E+03; 0.004)	2.51E+05(1.37E+04)
175	1.733	1.737	1.730
	5.07E+05(4.45E+04; 0.14, 0.113)	2.63E+05	2.03E+05
169	1.733	1.730	1.726
	4.42E+05(4.36E+03; 0.122, 0.001)	4.06E+05(1.90E+04; 0.007)	2.35E+05(8.82E+03)
121	1.715	1.711	1.726
	4.38E+05(2.27E+04; 0.242, 0.042)	3.68E+05(5.65E+04; 0.269)	2.96E+05(3.58E+04)
135	1.719	1.715	1.726
	3.86E+05(1.07E+04; 0.53, 0.038)	3.79E+05(8.21E+03; 0.041)	2.99E+05(2.22E+04)
126	1.726	1.722	1.719
	3.56E+05(4.66E+04; 0.33, 0.317)	3.00E+05(4.08E+04; 0.507)	2.52E+05
212	1.719	1.715	1.722
	3.47E+05(2.65E+04; 0.164, 0.025)	3.05E+05(8.69E+03; 0.008)	2.30E+05(3.02E+03)
274	1.733	1.730	1.737
	3.43E+05(2.73E+03; 0.067, 0.052)	3.16E+05(1.02E+04; 0.072)	1.63E+05(6.05E+04)
52	1.704	1.708	1.711
	3.41E+05(2.01E+04; 0.017, 0.005)	2.04E+05(1.56E+04; 0.019)	9.81E+04(1.39E+04)
58	1.726	1.722	1.719
	3.31E+05(1.76E+04; 0.05, 0.05)	2.73E+05(6.93E+03; 0.434)	2.63E+05(1.38E+04)
116	1.733	1.737	1.730
	3.23E+05(2.55E+03; 0.634, 0.029)	2.95E+05(7.08E+04; 0.522)	2.56E+05(1.63E+04)
125	1.719		
	3.04E+05		
83	1.715	1.711	1.719
	2.88E+05(7.62E+03; 0.088, 0.004)	2.42E+05(1.90E+04; 0.06)	1.88E+05(4.68E+03)
238	1.733	1.737	1.730
	2.67E+05(2.11E+03; 0.394, 0.021)	2.20E+05(6.23E+04; 0.768)	1.91E+05
281	1.730	1.726	1.722
	2.24E+05(2.81E+04; 0.231, 0.042)	1.89E+05(9.36E+03; 0.013)	1.30E+05(2.29E+03)
277	1.733	1.737	1.719
	2.19E+05	1.68E+05(6.23E+04)	5.35E+04
133	1.730	1.733	1.744
	2.06E+05	2.00E+05	2.05E+04
232	1.722	1.719	1.715
	2.05E+05(1.82E+04; 0.518, 0.038)	1.95E+05(3.91E+03; 0.006)	1.39E+05(5.09E+03)
206	1.726	1.719	1.722
	2.04E+05(3.98E+04; 0.673, 0.347)	1.77E+05	1.69E+05(9.54E+03)
118	1.737	1.741	1.744
	1.54E+05(4.42E+04; 0.337, 0.078)	1.13E+05(1.27E+04; 0.02)	4.87E+04(2.99E+03)

Table C.11. Twenty RFs with the highest copy numbers in the ^{14}N1 endpoint gradient.

Bps	BD1 (g/ml)	BD2(g/ml)	BD3(g/ml)
154	1.719	1.715	1.723
	4.85E+07(1.06E+07; 0.133, 0.054)	2.99E+07(1.66E+06; 0.034)	1.60E+07(3.33E+06)
119	1.719	1.715	1.711
	2.51E+07(7.26E+05; 0.294, 0.014)	2.26E+07(2.41E+06; 0.046)	1.26E+07(2.00E+06)
144	1.719	1.723	1.715
	2.44E+07(2.41E+06; 0.06, 0.027)	1.70E+07(1.17E+06; 0.072)	1.12E+07(2.01E+06)
163	1.719	1.723	1.715
	2.34E+07(2.21E+06; 0.027, 0.034)	1.41E+07(3.66E+05; 0.784)	1.38E+07(1.33E+06)
273	1.719	1.715	1.723
	1.12E+07(3.17E+06; 0.175, 0.114)	5.96E+06(1.68E+06; 0.541)	5.05E+06(5.74E+05)
139	1.730	1.723	1.719
	8.77E+06(0.00E+00; 0.006, 0.007)	5.90E+06(3.02E+05; 0.454)	5.59E+06(3.69E+05)
246	1.719	1.715	1.723
	8.17E+06(6.48E+05; 0.025, 0.015)	5.03E+06(2.97E+05; 0.046)	2.31E+06(8.07E+05)
40	1.708	1.704	1.711
	7.14E+06(7.86E+05; 0.046, 0.015)	4.94E+06(1.08E+06; 0.214)	3.52E+06(7.68E+05)
122	1.708	1.704	1.711
	6.11E+06(5.79E+05; 0.001, 0.003)	2.68E+06(4.68E+05; 0.148)	1.89E+06(3.97E+05)
231	1.719	1.715	1.723
	5.66E+06(8.49E+05; 0.059, 0.062)	3.28E+06(1.07E+05; 0.447)	2.88E+06(5.83E+05)
146	1.704	1.708	1.701
	5.03E+06(6.08E+05; 0.265, 0.009)	3.78E+06(1.55E+06; 0.259)	2.17E+06(1.98E+05)
113	1.719	1.723	1.715
	4.92E+06(2.10E+05; 0.015, 0.045)	3.69E+06(3.60E+04; 0.15)	2.55E+06(7.06E+05)
204	1.719	1.715	1.723
	4.02E+06(4.43E+05; 0.059, 0.059)	2.77E+06(7.25E+04; 0.204)	2.02E+06(5.67E+05)
156	1.723	1.719	1.726
	3.53E+06(2.45E+06; 0.704, 0.47)	2.77E+06(2.00E+05; 0.042)	2.00E+06(1.17E+05)
206	1.719	1.715	1.723
	3.46E+06(3.64E+05; 0.048, 0.052)	2.07E+06(2.66E+05; 0.591)	1.85E+06(4.00E+05)
159	1.723	1.730	1.719
	3.29E+06(1.95E+05; 0.028, 0.576)	2.49E+06(0.00E+00; 0.865)	2.16E+06(2.42E+06)
322	1.719	1.715	1.723
	3.27E+06(9.58E+04; 0.013, 0.044)	2.32E+06(1.25E+05; 0.135)	1.24E+06(6.13E+05)
158	1.723	1.719	1.726
	2.98E+06(7.64E+03; 0.69, 0.008)	2.12E+06(2.64E+06; 0.776)	1.51E+06(1.92E+05)
141	1.723	1.719	1.726
	2.93E+06(7.50E+03; 0.881, 0.513)	2.86E+06(5.85E+05; 0.82)	2.73E+06(3.49E+05)
278	1.719	1.704	1.730
	2.56E+06	8.56E+05(7.41E+05)	5.83E+05(0.00E+00)

Table C.12. RFs with 21st to 40th copy numbers in the ^{14}N1 endpoint gradient.

Bps	BD1 (g/ml)	BD2(g/ml)	BD3(g/ml)
147	1.730	1.726	1.723
	2.55E+06(0.00E+00; 0.368, 0.002)	1.71E+06(1.03E+06; 0.743)	1.44E+06(6.46E+04)
135	1.719	1.715	1.723
	2.52E+06(3.04E+05; 0.105, 0.038)	1.52E+06(3.95E+05; 0.526)	1.28E+06(1.78E+05)
116	1.730	1.719	1.726
	2.40E+06(0.00E+00; 0.115, 0.008)	2.03E+06(1.90E+05; 0.055)	1.36E+06(1.36E+05)
115	1.718	1.715	1.723
	1.87E+06(6.94E+04; 0.274, 0.085)	1.42E+06(4.23E+05; 0.788)	1.24E+06
85	1.708	1.704	1.719
	1.67E+06(6.03E+05; 0.32, 0.398)	1.22E+06(3.23E+05; 0.515)	9.28E+05
274	1.719	1.723	1.715
	1.58E+06(2.26E+05; 0.169, 0.072)	1.07E+06(2.56E+05; 0.477)	8.81E+05(1.65E+05)
58	1.719	1.723	1.715
	1.41E+06(2.66E+05; 0.29, 0.224)	9.50E+05(3.74E+05; 0.842)	8.68E+05(3.53E+05)
94	1.719	1.715	1.723
	1.39E+06(1.79E+05; 0.395, 0.213)	1.11E+06(3.23E+05; 0.479)	7.67E+05(4.52E+05)
83	1.704	1.708	1.701
	1.38E+06(1.66E+05; 0.471, 0.025)	1.28E+06(1.45E+05; 0.032)	6.56E+05(2.32E+05)
133	1.719	1.715	1.723
	1.20E+06(3.66E+05; 0.63, 0.224)	1.02E+06(2.58E+05; 0.277)	7.49E+05(3.76E+04)
131	1.708	1.711	1.715
	9.88E+05(2.05E+05; 0.187, 0.172)	7.07E+05(1.19E+05; 0.837)	6.73E+05(1.68E+05)
281	1.719	1.715	1.723
	9.76E+05(3.69E+05; 0.342, 0.447)	5.49E+05(3.20E+05; 0.83)	4.42E+05
169	1.719	1.723	1.715
	9.51E+05(2.15E+05; 0.59, 0.164)	7.65E+05(3.55E+05; 0.629)	6.22E+05(1.84E+04)
271	1.719	1.726	1.762
	9.32E+05	2.10E+05	9.25E+03
89	1.723	1.704	1.711
	9.21E+05	3.68E+05(7.54E+04)	2.83E+05
81	1.723	1.719	1.715
	8.47E+05(7.96E+05; 0.798, 0.668)	6.82E+05(9.10E+04; 0.513)	5.58E+05(2.03E+05)
70	1.723	1.719	1.708
	8.06E+05(8.19E+05; 0.821, 0.68)	6.29E+05(5.30E+05; 0.94)	6.05E+05(1.35E+05)
118	1.723	1.726	1.704
	7.94E+05	6.66E+05(1.11E+05; 0.046)	2.26E+05(1.61E+05)
237	1.726	1.723	1.708
	7.73E+05	6.07E+05	8.17E+04
238	1.719		
	7.31E+05		

Table C.13. Twenty RFs with the highest copy numbers in the ^{15}N inactive gradient.

Bps	BD1 (g/ml)	BD2(g/ml)	BD3(g/ml)
156	1.717	1.712	1.710
	3.60E+06(2.83E+05; 0.288, 0.006)	2.30E+06(1.25E+06; 0.266)	9.46E+05(6.73E+04)
163	1.712	1.717	1.710
	2.99E+06(6.69E+05; 0.907, 0.06)	2.92E+06(3.33E+05; 0.017)	1.14E+06(5.04E+04)
204	1.712	1.717	1.710
	2.50E+06(6.98E+05; 0.399, 0.109)	1.97E+06(1.48E+05; 0.016)	1.14E+06(1.02E+04)
144	1.717	1.712	1.710
	2.03E+06(4.41E+05; 0.145, 0.031)	1.29E+06(7.58E+04; 0.004)	2.95E+05(3.58E+04)
40	1.703	1.707	1.699
	1.80E+06(2.64E+05; 0.801, 0.057)	1.75E+06(6.50E+04; 0.006)	1.05E+06(3.98E+04)
119	1.712	1.717	1.710
	1.71E+06(5.54E+05; 0.32, 0.152)	1.19E+06(5.62E+04; 0.013)	8.19E+05(2.45E+04)
94	1.712	1.717	1.710
	1.49E+06(3.84E+05; 0.318, 0.096)	1.11E+06(1.35E+05; 0.047)	6.76E+05(3.12E+04)
139	1.717	1.712	1.720
	1.34E+06(6.65E+04; 0.262, 0.064)	7.96E+05(4.92E+05; 0.919)	7.52E+05(2.11E+05)
159	1.712	1.710	1.707
	1.25E+06(3.13E+05; 0.07, 0.06)	4.13E+05(1.09E+05; 0.469)	3.00E+05(1.44E+05)
154	1.717	1.712	1.720
	9.22E+05(4.16E+05; 0.424, 0.236)	4.74E+05(4.78E+05; 0.884)	4.17E+05(8.90E+04)
146	1.712	1.740	1.737
	9.04E+05	4.17E+03	1.21E+03
136	1.712	1.710	1.707
	7.23E+05	2.85E+05(5.55E+02; 0.021)	1.92E+05(1.97E+04)
326	1.699	1.696	1.693
	5.94E+05(1.19E+04; 0.001, 0.001)	3.55E+05(0.00E+00; 0.003)	1.54E+05(1.63E+04)
232	1.712	1.710	1.717
	5.83E+05(2.37E+05; 0.285, 0.204)	3.36E+05(4.50E+04; 0.349)	2.48E+05(9.29E+04)
145	1.720	1.710	1.725
	5.82E+05(1.06E+05; 0.065, 0.052)	2.49E+05(6.71E+04; 0.956)	2.45E+05(3.92E+04)
275	1.717	1.712	1.710
	5.81E+05(3.07E+05; 0.676, 0.271)	3.71E+05	2.54E+05(7.38E+03)
273	1.712	1.710	1.717
	5.73E+05(3.01E+05; 0.325, 0.528)	2.98E+05(1.44E+04; 0.177)	2.35E+05
52	1.699	1.696	1.703
	5.65E+05(6.89E+03; 0.006, 0.089)	5.05E+05(0.00E+00; 0.18)	3.95E+05(7.68E+04)
134	1.712	1.707	1.703
	5.63E+05(6.74E+05; 0.795, 0.6)	4.22E+05(3.15E+04; 0.044)	2.69E+05(3.51E+04)
282	1.712	1.717	1.710
	5.38E+05(3.12E+05; 0.833, 0.354)	4.75E+05(2.07E+05; 0.32)	2.60E+05(1.04E+05)

Table C.14. RFs with 21st to 40th copy numbers in the ^{15}N inactive gradient.

Bps	BD1 (g/ml)	BD2(g/ml)	BD3(g/ml)
147	1.703	1.717	1.699
	5.33E+05(6.46E+04; 0.49, 0.199)	4.51E+05	4.39E+05(2.70E+04)
278	1.699	1.712	1.717
	5.28E+05(5.80E+03; 0.819, 0.048)	4.73E+05(3.03E+05; 0.934)	4.34E+05
83	1.707	1.699	1.703
	5.13E+05(8.59E+04; 0.096, 0.086)	3.30E+05(9.92E+03; 0.332)	3.19E+05(7.45E+03)
279	1.712	1.710	1.720
	4.99E+05	9.74E+04(5.93E+03)	6.02E+04
246	1.712	1.707	1.699
	4.91E+05(8.27E+04; 0.158, 0.136)	9.22E+04	2.49E+04
245	1.710	1.707	1.696
	4.70E+05(5.48E+04; 0.107, 0.007)	7.40E+04	2.49E+04(0.00E+00)
268	1.712	1.717	1.710
	4.64E+05	1.66E+05	8.31E+04
161	1.717	1.712	1.710
	4.60E+05	3.99E+05	2.02E+05
58	1.712	1.717	1.710
	4.47E+05(2.11E+05; 0.768, 0.21)	3.96E+05(4.00E+04; 0.023)	1.73E+05(2.68E+04)
135	1.717	1.725	1.720
	4.27E+05(4.88E+04; 0.044, 0.019)	1.13E+05(8.38E+04; 0.713)	8.36E+04(4.83E+04)
230	1.712	1.717	1.707
	3.58E+05(1.47E+05; 0.416, 0.11)	2.27E+05(1.07E+05; 0.177)	7.16E+04(1.04E+04)
160	1.712	1.710	1.696
	3.58E+05	1.83E+05	2.38E+04(0.00E+00)
277	1.696	1.693	1.720
	3.45E+05(0.00E+00; 0.002)	1.46E+05(1.32E+04)	7.78E+04
122	1.712	1.703	1.737
	3.28E+05(1.20E+05; 0.28, 0.269)	1.63E+04	1.46E+03
262	1.712	1.710	1.720
	3.20E+05	1.75E+04	9.92E+03(4.27E+03)
115	1.717	1.712	1.710
	2.99E+05(4.37E+04; 0.254, 0.027)	2.50E+05(1.89E+03; 0)	1.15E+05(2.62E+03)
241	1.712		
	2.67E+05		
87	1.707	1.699	1.703
	2.44E+05	1.46E+05(4.88E+03; 0.207)	1.26E+05(1.42E+04)
236	1.712	1.686	
	2.41E+05	2.06E+03	
220	1.712	1.693	
	2.38E+05	5.30E+03	

Table C.15. Twenty RFs with the highest copy numbers in the ^{15}N1 midpoint gradient.

Bps	BD1 (g/ml)	BD2(g/ml)	BD3(g/ml)
154	1.729	1.732	1.725
	2.72E+07(4.27E+06; 0.059, 0.032)	1.53E+07(1.99E+05; 0.034)	9.73E+06(1.49E+06)
156	1.718	1.722	1.715
	1.56E+07(9.32E+05; 0.045, 0.027)	1.17E+07(7.66E+05; 0.074)	6.72E+06(1.86E+06)
204	1.718	1.722	1.715
	1.38E+07(8.38E+04; 0.001, 0.03)	7.85E+06(2.57E+05; 0.512)	6.87E+06(1.74E+06)
163	1.729	1.722	1.725
	1.01E+07(1.48E+06; 0.389, 0.307)	8.94E+06(2.75E+05; 0.521)	8.25E+06(1.23E+06)
139	1.729	1.732	1.722
	9.73E+06(1.54E+06; 0.214, 0.232)	7.77E+06(6.44E+04; 0.983)	7.76E+06(5.71E+05)
40	1.711	1.708	1.701
	6.84E+06(5.72E+05; 0.996, 0.606)	6.84E+06(1.53E+05; 0.602)	5.83E+06(2.32E+06)
119	1.722	1.718	1.725
	5.74E+06(3.29E+05; 0.072, 0.144)	4.59E+06(3.21E+05; 0.845)	4.47E+06(6.92E+05)
245	1.722	1.718	1.725
	5.31E+06(1.30E+05; 0.033, 0.043)	4.53E+06(1.63E+05; 0.11)	3.36E+06(5.80E+05)
144	1.732	1.729	1.722
	5.14E+06(7.63E+04; 0.128, 0.001)	4.14E+06(5.58E+05; 0.041)	2.18E+06(1.42E+05)
112	1.711	1.715	1.708
	4.89E+06(4.92E+05; 0.402, 0.031)	3.93E+06(1.19E+06; 0.362)	2.95E+06(6.05E+04)
48	1.718		
	4.65E+06		
212	1.718	1.722	1.715
	4.63E+06(1.43E+05; 0.002, 0.024)	2.36E+06(4.84E+04; 0.676)	2.17E+06(5.35E+05)
94	1.718	1.722	1.715
	4.11E+06(3.04E+05; 0.028, 0.11)	2.66E+06(1.76E+05; 0.875)	2.56E+06(7.35E+05)
58	1.718	1.722	1.715
	3.90E+06	2.66E+06(1.49E+05; 0.219)	1.88E+06(6.08E+05)
85	1.711	1.708	1.705
	3.70E+06(3.18E+05; 0.071, 0.015)	2.72E+06(2.26E+05; 0.032)	1.81E+06(7.09E+04)
57	1.718	1.725	1.701
	3.51E+06	1.16E+06(2.27E+05)	2.93E+05
231	1.732	1.729	1.736
	3.49E+06(5.03E+04; 0.067, 0.002)	2.40E+06(4.14E+05; 0.075)	1.37E+06(1.11E+05)
141	1.729	1.722	1.725
	3.11E+06(3.91E+05; 0.373, 0.239)	2.79E+06(8.10E+04; 0.361)	2.53E+06(2.96E+05)
135	1.718	1.722	1.715
	3.06E+06(1.43E+05; 0.011, 0.03)	1.96E+06(7.80E+04; 0.341)	1.67E+06(3.21E+05)
158	1.718	1.722	1.715
	3.00E+06(2.12E+05; 0.05, 0.029)	2.32E+06(6.42E+04; 0.061)	1.40E+06(3.33E+05)

Table C.16. RFs with 21st to 40th copy numbers in the ^{15}N1 midpoint gradient.

Bps	BD1 (g/ml)	BD2(g/ml)	BD3(g/ml)
232	1.718	1.722	1.715
	2.83E+06(1.66E+05; 0.011, 0.026)	1.70E+06(9.26E+03; 0.268)	1.39E+06(2.94E+05)
145	1.701	1.705	1.708
	2.21E+06(2.83E+05; 0.556, 0.069)	2.07E+06(1.14E+05; 0.163)	1.26E+06(5.18E+05)
273	1.732	1.708	1.736
	2.21E+06(8.20E+04; 0.001, 0.001)	5.33E+05(6.48E+03; 0.001)	3.60E+05(3.17E+03)
147	1.732	1.729	1.718
	2.08E+06(1.49E+05; 0.009, 0.081)	9.74E+05(9.39E+03; 0.023)	6.53E+05
274	1.732	1.736	1.740
	1.99E+06(1.01E+05; 0.044, 0.041)	2.06E+05	7.16E+04
83	1.708	1.705	1.701
	1.77E+06(1.26E+05; 0.285, 0.536)	1.60E+06(1.03E+05; 0.958)	1.59E+06(3.37E+05)
157	1.732	1.729	1.725
	1.73E+06(3.04E+04; 0.051, 0.128)	1.36E+06(1.18E+05; 0.472)	1.14E+06(3.27E+05)
121	1.711	1.708	1.705
	1.50E+06(2.81E+04; 0.148, 0.003)	1.35E+06(8.34E+04; 0.021)	9.30E+05(3.15E+04)
80	1.732	1.729	1.722
	1.44E+06(3.42E+03; 0.009, 0.005)	1.15E+06	8.49E+05
81	1.729	1.722	1.711
	1.42E+06	7.52E+05	4.64E+05
134	1.711	1.708	1.705
	1.39E+06(3.23E+04; 0.13, 0.006)	1.23E+06(8.36E+04; 0.043)	9.27E+05(4.01E+04)
113	1.732	1.729	1.725
	1.22E+06(2.01E+04; 0.76, 0.013)	1.19E+06(1.44E+05; 0.041)	5.88E+05(1.01E+05)
143	1.718	1.722	1.725
	1.11E+06(4.76E+03; 0.005, 0.061)	1.04E+06(6.06E+03; 0.085)	6.58E+05(1.67E+05)
159	1.732	1.736	1.740
	1.11E+06(3.94E+03; 0.003, 0)	3.40E+05(6.02E+04; 0.145)	2.35E+05(2.10E+04)
115	1.718	1.715	1.711
	8.43E+05(2.88E+04; 0.291, 0.283)	6.97E+05(1.43E+05; 0.617)	5.74E+05(2.60E+05)
175	1.732	1.736	1.740
	8.38E+05(1.18E+05; 0.039, 0.012)	3.88E+05(5.20E+04; 0.015)	8.44E+04(1.38E+04)
234	1.732	1.729	
	6.03E+05(2.44E+03)	3.91E+05	
79	1.708	1.705	1.711
	5.30E+05(4.71E+04; 0.026, 0.147)	2.94E+05(2.82E+04; 0.836)	2.85E+05
278	1.718	1.708	1.701
	4.71E+05	2.74E+05(1.25E+03; 0.748)	2.50E+05(9.29E+04)
153	1.736	1.743	1.747
	4.55E+05	4.34E+04	8.40E+03

Table C.17. Twenty RFs with the highest copy numbers in the $^{15}N2$ midpoint gradients.

Bps	BD1 (g/ml)	BD2(g/ml)	BD3(g/ml)
154	1.722	1.736	1.732
	2.00E+07(3.63E+06; 0.999, 0.599)	2.00E+07(6.36E+06; 0.661)	1.76E+07(4.82E+06)
204	1.726	1.729	1.719
	1.59E+07(2.31E+06; 0.068, 0.025)	6.71E+06(4.11E+06; 0.766)	5.71E+06(1.60E+05)
207	1.729	1.726	1.719
	1.55E+07(4.81E+06; 0.55, 0.089)	1.31E+07(1.70E+06; 0.034)	6.64E+06(2.08E+05)
40	1.726	1.709	1.706
	1.21E+07(2.15E+06; 0.097, 0.053)	7.52E+06(1.64E+05; 0.021)	5.63E+06(3.58E+05)
163	1.729	1.732	1.722
	1.15E+07(1.27E+06; 0.689, 0.138)	1.11E+07(1.08E+06; 0.16)	9.60E+06(1.31E+05)
156	1.726	1.729	1.719
	8.77E+06(1.62E+06; 0.452, 0.075)	6.67E+06(3.07E+06; 0.437)	4.61E+06(5.50E+05)
119	1.729	1.732	1.722
	8.60E+06(1.55E+06; 0.053, 0.038)	5.49E+06(1.22E+06; 0.342)	4.44E+06(3.74E+05)
139	1.729	1.732	1.722
	6.94E+06(1.15E+06; 0.387, 0.27)	6.24E+06(4.64E+05; 0.278)	5.78E+06(1.23E+05)
245	1.729	1.719	1.726
	6.55E+06(1.47E+06; 0.05, 0.04)	3.07E+06(1.08E+05; 0.295)	2.69E+06(3.66E+05)
83	1.726	1.709	1.706
	6.43E+06(9.56E+05; 0.243, 0.088)	5.30E+06(1.89E+05; 0.054)	4.18E+06(3.37E+05)
144	1.736	1.729	1.732
	5.47E+06(3.96E+05; 0.089, 0.006)	4.08E+06(6.95E+05; 0.854)	4.00E+06(6.29E+04)
85	1.726	1.712	1.709
	3.74E+06(5.61E+05; 0.049, 0.03)	1.83E+06(2.54E+05; 0.192)	1.42E+06(1.62E+05)
231	1.736	1.739	1.722
	3.36E+06(2.09E+05; 0.254, 0.063)	3.07E+06(1.55E+05; 0.117)	2.60E+06(1.90E+05)
145	1.706	1.709	1.726
	3.21E+06(3.22E+05; 0.408, 0.227)	2.97E+06(1.02E+05; 0.311)	2.41E+06(5.75E+05)
135	1.726	1.729	1.719
	3.09E+06(6.47E+05; 0.129, 0.056)	1.67E+06(4.76E+05; 0.333)	1.24E+06(7.00E+04)
274	1.736	1.719	1.716
	2.46E+06(2.07E+04; 0.003, 0.002)	1.06E+06(1.10E+05; 0.177)	8.37E+05(1.04E+05)
141	1.729	1.732	1.722
	2.37E+06(4.52E+05; 0.53, 0.397)	2.16E+06(3.04E+05; 0.636)	2.03E+06(1.78E+05)
212	1.726	1.729	1.716
	2.02E+06(3.16E+05; 0.042, 0.083)	8.17E+05(1.76E+05; 0.866)	7.52E+05(4.49E+05)
273	1.736	1.719	1.716
	1.80E+06(8.28E+04; 0.01, 0.008)	1.06E+06(6.77E+04; 0.254)	9.58E+05(6.41E+04)
58	1.726	1.729	1.719
	1.69E+06(3.50E+05; 0.616, 0.099)	1.42E+06(6.02E+05; 0.36)	9.28E+05(1.00E+05)

Table C.18. RFs with 21st to 40th copy numbers in the ^{15}N2 midpoint gradient.

Bps	BD1 (g/ml)	BD2(g/ml)	BD3(g/ml)
246	1.722	1.742	1.746
	1.69E+06(1.21E+05; 0.059, 0.057)	8.06E+04	2.69E+04
158	1.729	1.726	1.719
	1.54E+06(9.79E+05; 0.665, 0.43)	1.19E+06(1.36E+05; 0.261)	8.60E+05(2.63E+05)
134	1.726	1.709	1.712
	1.49E+06(1.65E+05; 0.024, 0.023)	7.40E+05(2.28E+04; 0.271)	6.65E+05(6.65E+04)
113	1.722	1.732	1.736
	1.44E+06(4.43E+05; 0.297, 0.272)	1.09E+06(2.09E+05; 0.503)	9.72E+05(3.65E+04)
249	1.719	1.716	1.736
	1.32E+06(6.77E+04; 0.071, 0.004)	9.68E+05(1.23E+05; 0.047)	5.83E+05(1.86E+03)
160	1.722		
	1.24E+06		
47	1.722	1.732	1.739
	1.18E+06(1.13E+06; 0.427, 0.383)	3.91E+05(7.39E+04; 0.227)	2.98E+05(2.02E+04)
230	1.716	1.719	1.712
	1.08E+06(6.81E+04; 0.616, 0.01)	8.66E+05(5.04E+05; 0.331)	4.09E+05(6.54E+04)
126	1.719	1.716	1.729
	1.02E+06(6.21E+04; 0.139, 0.234)	8.19E+05(1.00E+05; 0.97)	8.13E+05(1.60E+05)
94	1.726	1.729	1.712
	1.00E+06(1.74E+05; 0.113, 0.053)	5.43E+05(1.65E+05; 0.624)	4.74E+05(4.26E+04)
120	1.722	1.729	1.732
	9.67E+05(3.94E+04; 0.021, 0.023)	5.22E+05(1.30E+05; 0.528)	4.43E+05(1.07E+05)
147	1.736	1.732	1.722
	9.19E+05(2.44E+05; 0.952, 0.891)	9.08E+05(1.50E+05; 0.837)	8.67E+05
213	1.729	1.716	1.712
	8.11E+05(4.23E+05; 0.865, 0.485)	7.17E+05	5.59E+05(7.30E+04)
159	1.736	1.739	1.732
	7.77E+05(2.32E+05; 0.797, 0.206)	7.27E+05(7.02E+04; 0.107)	4.29E+05(1.32E+05)
164	1.726	1.729	
	7.71E+05	5.14E+05	
81	1.732	1.736	1.729
	6.76E+05(1.38E+05; 0.263, 0.253)	5.34E+05(1.44E+03; 0.01)	4.22E+05
157	1.726	1.729	1.719
	6.73E+05(1.21E+05; 0.271, 0.231)	5.31E+05(5.69E+04; 0.174)	2.82E+05
232	1.719	1.712	1.709
	6.36E+05	4.55E+05	1.90E+05
142	1.726	1.729	1.719
	5.80E+05	5.72E+05(2.12E+05; 0.147)	2.64E+05(1.49E+04)
117	1.722	1.732	1.736
	5.72E+05(3.29E+04; 0.646, 0.105)	5.38E+05(8.58E+04; 0.461)	4.83E+05(2.99E+04)

Table C.19. Twenty RFs with the highest copy numbers in the $^{14}N2$ midpoint gradient.

Bps	BD1 (g/ml)	BD2(g/ml)	BD3(g/ml)
154	1.719	1.716	1.712
	2.40E+07(3.25E+06; 0.02, 0.015)	7.73E+06(5.30E+05; 0.053)	5.13E+06(7.04E+05)
119	1.719	1.716	1.712
	1.06E+07(2.56E+06; 0.21, 0.27)	7.27E+06(2.40E+04; 1)	7.27E+06(1.73E+06)
207	1.712	1.716	1.719
	1.03E+07(5.83E+05; 0.036, 0.034)	7.77E+06(3.81E+05; 0.357)	7.17E+06(6.01E+05)
144	1.719	1.716	1.723
	8.90E+06(5.62E+05; 0.008, 0.006)	3.72E+06(3.46E+05; 0.771)	3.62E+06(1.68E+05)
245	1.716	1.712	1.719
	8.30E+06(6.36E+05; 0.054, 0.019)	6.23E+06(3.20E+05; 0.039)	5.10E+06(4.29E+04)
139	1.719	1.716	1.712
	6.35E+06(9.22E+05; 0.17, 0.059)	4.91E+06(2.88E+05; 0.067)	3.47E+06(4.76E+05)
231	1.719	1.723	1.716
	5.77E+06(7.16E+05; 0.1, 0.041)	4.12E+06(3.54E+05; 0.096)	3.16E+06(2.88E+05)
249	1.716	1.719	1.705
	5.59E+06(1.88E+05; 0.002, 0.004)	1.35E+06(1.50E+05; 0.548)	1.16E+06(3.63E+05)
163	1.719	1.716	1.712
	4.40E+06(4.18E+05; 0.051, 0.027)	2.86E+06(3.00E+05; 0.245)	2.44E+06(2.16E+05)
145	1.719	1.716	1.694
	3.48E+06(2.93E+05; 0.019, 0.014)	1.86E+06(1.25E+05; 0.292)	1.72E+06(5.10E+04)
113	1.719	1.716	1.723
	3.14E+06(1.39E+06; 0.158, 0.131)	9.63E+05(9.48E+04; 0.062)	6.99E+05(2.45E+04)
204	1.701	1.705	1.709
	3.00E+06(7.66E+05; 0.503, 0.224)	2.43E+06(6.20E+05; 0.472)	2.03E+06(1.71E+05)
156	1.716	1.705	1.709
	2.48E+06(2.29E+05; 0.1, 0.033)	1.71E+06(2.93E+05; 0.664)	1.61E+06(1.73E+04)
83	1.698	1.694	1.701
	2.21E+06(1.19E+05; 0.216, 0.133)	1.92E+06(1.94E+05; 0.34)	1.57E+06(3.46E+05)
274	1.716	1.723	1.719
	2.09E+06(8.67E+04; 0.01, 0.011)	1.12E+06(1.07E+05; 0.168)	8.26E+05(1.63E+05)
273	1.716	1.719	1.723
	1.84E+06(1.69E+05; 0.123, 0.134)	1.51E+06(6.14E+04; 0.86)	1.49E+06(1.06E+05)
141	1.719	1.716	1.712
	1.72E+06(5.16E+04; 0.286, 0.019)	1.52E+06(1.84E+05; 0.088)	1.02E+06(1.30E+05)
147	1.719	1.716	1.712
	1.39E+06(7.70E+04; 0.193, 0.017)	1.20E+06(1.22E+05; 0.046)	6.73E+05(1.09E+05)
159	1.723	1.716	1.719
	1.31E+06(7.15E+04; 0.049, 0.023)	9.60E+05(8.98E+04; 0.523)	9.03E+05(5.31E+04)
158	1.723	1.709	1.705
	1.07E+06(9.83E+04; 0.027, 0.01)	6.16E+05(4.09E+04; 0.021)	2.91E+05(5.43E+04)

Table C.20. RFs with 21st to 40th copy numbers in the ^{14}N2 midpoint gradients.

Bps	BD1 (g/ml)	BD2(g/ml)	BD3(g/ml)
81	1.719	1.716	1.712
	9.83E+05(6.19E+05; 0.352, 0.223)	4.55E+05(8.74E+03; 0.037)	2.14E+05(6.69E+04)
135	1.705	1.701	1.709
	8.09E+05(2.35E+05; 0.465, 0.343)	6.31E+05(1.55E+05; 0.808)	5.98E+05(5.66E+04)
125	1.709	1.705	1.712
	7.74E+05(4.52E+04; 0.131, 0.137)	6.53E+05(5.16E+04; 0.292)	4.78E+05(1.67E+05)
116	1.719	1.716	1.723
	6.56E+05(2.12E+05; 0.128, 0.112)	2.76E+05(1.49E+04; 0.358)	2.39E+05(4.14E+04)
133	1.719	1.716	1.712
	6.13E+05(2.45E+05; 0.472, 0.131)	4.60E+05(9.79E+03; 0.015)	1.74E+05(4.92E+04)
157	1.712	1.719	1.727
	5.61E+05(7.21E+04; 0.807, 0.057)	5.43E+05(6.21E+04; 0.052)	3.56E+05(9.14E+03)
40	1.698	1.701	1.694
	5.43E+05(7.33E+04; 0.496, 0.409)	5.00E+05(1.33E+04; 0.592)	4.63E+05(8.13E+04)
120	1.709	1.712	1.719
	5.07E+05(4.50E+04; 0.9, 0.06)	4.70E+05(3.72E+05; 0.557)	2.83E+05(6.77E+04)
85	1.701	1.698	1.705
	4.82E+05(1.15E+05; 0.443, 0.213)	4.04E+05(2.41E+04; 0.167)	3.24E+05(4.72E+04)
115	1.719	1.716	1.712
	4.44E+05(1.25E+05; 0.163, 0.085)	2.42E+05(4.19E+04; 0.111)	1.59E+05(9.09E+03)
281	1.705	1.716	1.712
	4.06E+05(1.45E+05; 0.308, 0.088)	2.67E+05(5.78E+03; 0.002)	8.34E+04(8.39E+03)
140	1.716	1.719	1.723
	3.93E+05(2.01E+04; 0.219, 0.004)	3.25E+05	1.50E+05(1.04E+04)
230	1.705	1.698	1.694
	3.76E+05(1.05E+05; 0.075, 0.075)	1.19E+05(8.91E+03; 0.987)	1.19E+05(9.34E+03)
213	1.705	1.701	1.709
	3.73E+05(1.27E+05; 0.746, 0.647)	3.39E+05(2.73E+04; 0.556)	3.25E+05(6.09E+03)
47	1.719	1.723	1.727
	3.56E+05(3.05E+05; 0.358, 0.303)	9.98E+04(2.98E+04; 0.199)	5.99E+04(1.32E+03)
262	1.705	1.716	1.694
	3.26E+05(7.90E+04; 0.238, 0.044)	2.33E+05(5.43E+03; 0.001)	6.92E+04(4.08E+03)
175	1.719	1.723	1.716
	3.03E+05(1.29E+05; 0.65, 0.196)	2.38E+05(1.17E+05; 0.31)	1.22E+05(3.82E+04)
57	1.709	1.705	1.719
	2.99E+05(8.93E+04; 0.693, 0.646)	2.64E+05(6.12E+04; 0.736)	2.31E+05
278	1.694	1.691	1.719
	2.88E+05(1.61E+04; 0.009, 0.068)	1.46E+05(9.95E+03; 0.189)	1.06E+05
239	1.694	1.698	1.705
	2.69E+05(4.77E+03; 0.086, 0.011)	2.30E+05(1.64E+04; 0.023)	9.22E+04(2.55E+04)

D. SigmaPlot stratistical reports

Experiment 1 Effluent

Three Way Analysis of Variance Saturday, July 30, 2011, 5:32:22 PM

Data source: Data 2 in plant anovas.JNB

General Linear Model

Dependent Variable: Data

Normality Test (Shapiro-Wilk) Passed (P = 0.466)

Equal Variance Test: Passed (P = 0.241)

Source of Variation	**DF**	**SS**	**MS**	**F**	**P**
Plant	1	1317708.095	1317708.095	26.721	<0.001
Inoculant	1	107499.112	107499.112	2.180	0.168
Nitrogen	1	198568.617	198568.617	4.027	0.070
Plant x Inoculant	1	2623.463	2623.463	0.0532	0.822
Plant x Nitrogen	1	206711.866	206711.866	4.192	0.065
Inoculant x Nitrogen	1	1374.239	1374.239	0.0279	0.870
Plant x Inoculant x Nitrogen	1	10838.177	10838.177	0.220	0.648
Residual	11	542450.372	49313.670		
Total	18	2728661.552	151592.308		

The difference in the mean values among the different levels of Plant are greater than would be expected by chance after allowing for the effects of differences in Inoculant and Nitrogen. There is a statistically significant difference (P = <0.001). To isolate which group(s) differ from the others use a multiple comparison procedure.

The difference in the mean values among the different levels of Inoculant are not great enough to exclude the possibility that the difference is just due to random sampling variability after allowing for the effects of differences in Plant and Nitrogen. There is not a statistically significant difference (P = 0.168).

The difference in the mean values among the different levels of Nitrogen are not great enough to exclude the possibility that the difference is just due to random sampling variability after allowing for the effects of differences in Plant and Inoculant. There is not a statistically significant difference (P = 0.070).

The effect of different levels of Plant does not depend on what level of Inoculant is present. There is not a statistically significant interaction between Plant and Inoculant. (P = 0.822)

The effect of different levels of Plant does not depend on what level of Nitrogen is present. There is not a statistically significant interaction between Plant and Nitrogen. (P = 0.065)

The effect of different levels of Inoculant does not depend on what level of Nitrogen is present. There is not a statistically significant interaction between Inoculant and Nitrogen. (P = 0.870)

All Pairwise Multiple Comparison Procedures (Holm-Sidak method):
Overall significance level = 0.05

Comparisons for factor: **Plant**

Comparison	Diff of Means	t	Unadjusted P	Critical Level	Significant?
Yes vs. No	591.622	5.169	<0.001	0.050	Yes

Comparisons for factor: **Inoculant**

Comparison	Diff of Means	t	Unadjusted P	Critical Level	Significant?
Yes vs. No	168.981	1.476	0.168	0.050	No

Comparisons for factor: **Nitrogen**

Comparison	Diff of Means	t	Unadjusted P	Critical Level	Significant?
Yes vs. No	229.662	2.007	0.070	0.050	No

Comparisons for factor: **Inoculant within Yes**

Comparison	Diff of Means	t	Unadjusted P	Critical Level	Significant?
Yes vs. No	195.379	1.398	0.190	0.050	No

Comparisons for factor: **Inoculant within No**

Comparison	Diff of Means	t	Unadjusted P	Critical Level	Significant?
Yes vs. No	142.583	0.786	0.448	0.050	No

Comparisons for factor: **Plant within No**

Comparison	Diff of Means	t	Unadjusted P	Critical Level	Significant?
Yes vs. No	565.224	2.939	0.013	0.050	Yes

Comparisons for factor: **Plant within Yes**

Comparison	Diff of Means	t	Unadjusted P	Critical Level	Significant?
Yes vs. No	618.020	4.978	<0.001	0.050	Yes

Comparisons for factor: **Nitrogen within Yes**

Comparison	Diff of Means	t	Unadjusted P	Critical Level	Significant?
Yes vs. No	463.987	3.321	0.007	0.050	Yes

Comparisons for factor: **Nitrogen within No**

Comparison	Diff of Means	t	Unadjusted P	Critical Level	Significant?
No vs. Yes	4.662	0.0257	0.980	0.050	No

Comparisons for factor: **Plant within No**

Comparison	Diff of Means	t	Unadjusted P	Critical Level	Significant?
Yes vs. No	357.297	2.229	0.048	0.050	Yes

Comparisons for factor: **Plant within Yes**

Comparison	Diff of Means	t	Unadjusted P	Critical Level	Significant?
Yes vs. No	825.946	5.054	<0.001	0.050	Yes

Comparisons for factor: **Nitrogen within No**

Comparison	Diff of Means	t	Unadjusted P	Critical Level	Significant?

Yes vs. No	210.557	1.095	0.297	0.050	No

Comparisons for factor: **Nitrogen within Yes**

Comparison	**Diff of Means**	**t**	**Unadjusted P**	**Critical Level**	**Significant?**
Yes vs. No	248.768	2.004	0.070	0.050	No

Comparisons for factor: **Inoculant within No**

Comparison	**Diff of Means**	**t**	**Unadjusted P**	**Critical Level**	**Significant?**
Yes vs. No	149.875	0.935	0.370	0.050	No

Comparisons for factor: **Inoculant within Yes**

Comparison	**Diff of Means**	**t**	**Unadjusted P**	**Critical Level**	**Significant?**
Yes vs. No	188.086	1.151	0.274	0.050	No

Power of performed test with alpha = 0.0500: for Plant : 0.998
Power of performed test with alpha = 0.0500: for Inoculant : 0.160
Power of performed test with alpha = 0.0500: for Nitrogen : 0.344
Power of performed test with alpha = 0.0500: for Plant x Inoculant : 0.0500
Power of performed test with alpha = 0.0500: for Plant x Nitrogen : 0.360
Power of performed test with alpha = 0.0500: for Inoculant x Nitrogen : 0.0500

Least square means for Plant :

Group	**Mean**	**SEM**
Yes	1502.106	69.857
No	910.484	90.658

Least square means for Inoculant :

Group	**Mean**	**SEM**
No	1121.805	96.158
Yes	1290.785	62.070

Least square means for Nitrogen :

Group	**Mean**	**SEM**
No	1091.464	80.131
Yes	1321.126	81.718

Least square means for Plant x Inoculant :

Group	**Mean**	**SEM**
Yes x No	1404.416	111.033
Yes x Yes	1599.795	84.803
No x No	839.193	157.025
No x Yes	981.775	90.658

Least square means for Plant x Nitrogen :

Group	**Mean**	**SEM**
Yes x No	1270.113	96.158
Yes x Yes	1734.099	101.359
No x No	912.815	128.210

No x Yes	908.153	128.210

Least square means for Inoculant x Nitrogen :

Group	**Mean**	**SEM**
No x No	1016.526	135.988
No x Yes	1227.083	135.988
Yes x No	1166.401	84.803
Yes x Yes	1415.169	90.658

Least square means for Plant x Inoculant x Nitrogen :

Group	**Mean**	**SEM**
Yes x No x No	1208.804	157.025
Yes x No x Yes	1600.029	157.025
Yes x Yes x No	1331.421	111.033
Yes x Yes x Yes	1868.169	128.210
No x No x No	824.249	222.067
No x No x Yes	854.137	222.067
No x Yes x No	1001.381	128.210
No x Yes x Yes	962.170	128.210

One Way Analysis of Variance Saturday, July 30, 2011, 7:24:12 PM

Data source: Data 2 in plant anovas.JNB

Dependent Variable: Col 12

Normality Test (Shapiro-Wilk) Passed (P = 0.126)

Equal Variance Test: Passed (P = 0.513)

Group Name	**N**	**Missing**	**Mean**	**Std Dev**	**SEM**
I_C_noN	3	0	130.706	66.958	38.658
I_C_N	3	0	1102.618	182.763	105.518
I_noC_noN	3	0	1001.381	282.539	163.124
I_noC_N	3	0	962.170	35.973	20.769
Noinoc	4	0	874.435	138.185	69.092

Source of Variation	**DF**	**SS**	**MS**	**F**	**P**
Between Groups	4	1836123.529	459030.882	17.099	<0.001
Residual	11	295301.099	26845.554		
Total	15	2131424.628			

The differences in the mean values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P = <0.001).

Power of performed test with alpha = 0.050: 1.000

All Pairwise Multiple Comparison Procedures (Holm-Sidak method):
Overall significance level = 0.05

Comparisons for factor: **Col 11**

Comparison	Diff of Means	t	Unadjusted P	Critical Level	Significant?
I_C_N vs. I_C_noN	971.913	7.265	<0.001	0.005	Yes
I_noC_noN vs. I_C_noN	870.675	6.508	<0.001	0.006	Yes
I_noC_N vs. I_C_noN	831.464	6.215	<0.001	0.006	Yes
Noinoc vs. I_C_noN	743.730	5.943	<0.001	0.007	Yes
I_C_N vs. Noinoc	228.183	1.823	0.096	0.009	No
I_C_N vs. I_noC_N	140.449	1.050	0.316	0.010	No
I_noC_noN vs. Noinoc	126.946	1.014	0.332	0.013	No
I_C_N vs. I_noC_noN	101.237	0.757	0.465	0.017	No
I_noC_N vs. Noinoc	87.734	0.701	0.498	0.025	No
I_noC_noN vs. I_noC_N	39.211	0.293	0.775	0.050	No

Experiment 3 Effluent

Two Way Analysis of Variance Saturday, July 30, 2011, 6:41:37 PM

Data source: Data 3 in plant anovas.JNB

General Linear Model

Dependent Variable: Data

Normality Test (Shapiro-Wilk) Passed (P = 0.245)

Equal Variance Test: Passed (P = 0.407)

Source of Variation	DF	SS	MS	F	P
Plant	1	53313.405	53313.405	18.134	<0.001
Inoculant	1	2229.080	2229.080	0.758	0.399
Plant x Inoculant	1	281.997	281.997	0.0959	0.761
Residual	14	41159.767	2939.983		
Total	17	99287.340	5840.432		

The difference in the mean values among the different levels of Plant is greater than would be expected by chance after allowing for effects of differences in Inoculant. There is a statistically significant difference (P = <0.001). To isolate which group(s) differ from the others use a multiple comparison procedure.

The difference in the mean values among the different levels of Inoculant is not great enough to exclude the possibility that the difference is just due to random sampling variability after allowing for the effects of differences in Plant. There is not a statistically significant difference (P = 0.399).

The effect of different levels of Plant does not depend on what level of Inoculant is present. There is not a statistically significant interaction between Plant and Inoculant. (P = 0.761)

Power of performed test with alpha = 0.0500: for Plant : 0.979
Power of performed test with alpha = 0.0500: for Inoculant : 0.0500
Power of performed test with alpha = 0.0500: for Plant x Inoculant : 0.0500

Least square means for Plant :

Group	Mean	SEM
Yes	283.389	17.500
No	172.856	19.170

Least square means for Inoculant :

Group	Mean	SEM
No	216.822	19.170
Yes	239.424	17.500

Least square means for Plant x Inoculant :

Group	Mean	SEM
Yes x No	276.108	27.111
Yes x Yes	290.671	22.136
No x No	157.536	27.111
No x Yes	188.176	27.111

All Pairwise Multiple Comparison Procedures (Holm-Sidak method):
Overall significance level = 0.05

Comparisons for factor: **Plant**

Comparison	Diff of Means	t	Unadjusted P	Critical Level	Significant?
Yes vs. No	110.534	4.258	<0.001	0.050	Yes

Comparisons for factor: **Inoculant**

Comparison	Diff of Means	t	Unadjusted P	Critical Level	Significant?
Yes vs. No	22.602	0.871	0.399	0.050	No

Comparisons for factor: **Inoculant within Yes**

Comparison	Diff of Means	t	Unadjusted P	Critical Level	Significant?
Yes vs. No	14.563	0.416	0.684	0.050	No

Comparisons for factor: **Inoculant within No**

Comparison	Diff of Means	t	Unadjusted P	Critical Level	Significant?
Yes vs. No	30.640	0.799	0.438	0.050	No

Comparisons for factor: **Plant within No**

Comparison	Diff of Means	t	Unadjusted P	Critical Level	Significant?
Yes vs. No	118.572	3.093	0.008	0.050	Yes

Comparisons for factor: **Plant within Yes**

Comparison	Diff of Means	t	Unadjusted P	Critical Level	Significant?
Yes vs. No	102.495	2.928	0.011	0.050	Yes

One Way Analysis of Variance Saturday, July 30, 2011, 6:45:40 PM

Data source: Data 3 in plant anovas.JNB

Dependent Variable: Data

Normality Test (Shapiro-Wilk) Passed (P = 0.347)

Equal Variance Test: Passed (P = 0.442)

Group Name	N	Missing	Mean	Std Dev	SEM
No	9	0	228.355	74.106	24.702
Yes	9	0	241.790	82.570	27.523

Source of Variation	DF	SS	MS	F	P
Between Groups	1	812.158	812.158	0.132	0.721
Residual	16	98475.182	6154.699		
Total	17	99287.340			

The differences in the mean values among the treatment groups are not great enough to exclude the possibility that the difference is due to random sampling variability; there is not a statistically significant difference (P = 0.721).

Power of performed test with alpha = 0.050: 0.048

The power of the performed test (0.048) is below the desired power of 0.800.
Less than desired power indicates you are less likely to detect a difference when one actually exists. Negative results should be interpreted cautiously.

Experiment 3: RDX Degradation

Kruskal-Wallis One Way Analysis of Variance on Ranks Saturday, July 30, 2011, 7:04:17 PM

Data source: Data 3 in plant anovas.JNB

Dependent Variable: Col 29

Normality Test (Shapiro-Wilk) Failed (P < 0.050)

Group	N	Missing	Median	25%	75%
Plant N	68	0	1382.819	1323.358	1421.944
Plant NoN	72	0	1279.218	1054.826	1415.235
Soil N	26	0	350.173	118.941	426.847
Soil noN	24	0	166.813	-144.992	767.294

H = 110.411 with 3 degrees of freedom. (P = <0.001)

The differences in the median values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P = <0.001)

To isolate the group or groups that differ from the others use a multiple comparison procedure.

All Pairwise Multiple Comparison Procedures (Dunn's Method) :

Comparison	Diff of Ranks	Q	P<0.05
Plant N vs Soil N	104.790	8.264	Yes
Plant N vs Soil noN	103.725	7.944	Yes

Plant N vs Plant NoN	22.739	2.445	No
Plant NoN vs Soil N	82.050	6.521	Yes
Plant NoN vs Soil noN	80.986	6.248	Yes
Soil noN vs Soil N	1.064	0.0684	No

Note: The multiple comparisons on ranks do not include an adjustment for ties.

Kruskal-Wallis One Way Analysis of Variance on Ranks Saturday, July 30, 2011, 7:10:43 PM

Data source: Data 3 in plant anovas.JNB

Dependent Variable: Col 32

Normality Test (Shapiro-Wilk) Failed (P < 0.050)

Group	N	Missing	Median	25%	75%
Plant_noInoc	60	0	1413.527	1268.450	1540.399
Plant_inoc	80	0	1351.525	1075.047	1389.298
Soil_noinoc	26	0	445.854	339.122	750.016
Soil_inoc	24	0	28.586	-190.388	211.087

H = 115.102 with 3 degrees of freedom. (P = <0.001)

The differences in the median values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P = <0.001)

To isolate the group or groups that differ from the others use a multiple comparison procedure.

All Pairwise Multiple Comparison Procedures (Dunn's Method) :

Comparison	Diff of Ranks	Q	P<0.05
Plant_noInoc vs Soil_inoc	120.742	9.091	Yes
Plant_noInoc vs Soil_noinoc	96.085	7.442	Yes
Plant_noInoc vs Plant_inoc	26.837	2.858	Yes
Plant_inoc vs Soil_inoc	93.904	7.337	Yes
Plant_inoc vs Soil_noinoc	69.247	5.578	Yes
Soil_noinoc vs Soil_inoc	24.657	1.584	No

Note: The multiple comparisons on ranks do not include an adjustment for ties.

Plant Health Rank Sum Test

t-test Sunday, July 31, 2011, 9:20:25 PM

Data source: Data 4 in plant anovas.JNB

Normality Test (Shapiro-Wilk) Failed (P < 0.050)

Test execution ended by user request, Rank Sum Test begun

Mann-Whitney Rank Sum Test Sunday, July 31, 2011, 9:20:25 PM

Data source: Data 4 in plant anovas.JNB

Group	**N**	**Missing**	**Median**	**25%**	**75%**
Good Health	64	0	1436.069	1391.454	1540.399
Poor Health	90	0	1268.368	1086.843	1343.960

Mann-Whitney U Statistic= 312.000

T = 7528.000 n(small)= 64 n(big)= 90 (P = <0.001)

The difference in the median values between the two groups is greater than would be expected by chance; there is a statistically significant difference (P = <0.001)

E. List of Technical Publications

1. Andeer, P., S.E. Strand, and D.A. Stahl, *High-Sensitivity Stable-Isotope Probing by a Quantitative Terminal Restriction Fragment Length Polymorphism Protocol.* Applied and Environmental Microbiology, 2012. **78**(1): p. 163-169.
2. Andeer, P.F., et al., *Lateral Transfer of Genes for Hexahydro-1,3,5-Trinitro-1,3,5-Triazine (RDX) Degradation.* Applied and Environmental Microbiology, 2009. **75**(10): p. 3258-3262.
3. Jackson, R.G., et al., *Exploring the biochemical properties and remediation applications of the unusual explosive-degrading P450 system XplA/B.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(43): p. 16822-16827.
4. Rylott, E.L., et al., *The explosive-degrading cytochrome P450 XplA: Biochemistry, structural features and prospects for bioremediation.* Biochimica Et Biophysica Acta-Proteins and Proteomics, 2011. **1814**(1): p. 230-236.
5. Rylott, E.L., A. Lorenz, and N.C. Bruce, *Biodegradation and biotransformation of explosives.* Current Opinion in Biotechnology, 2011. **22**(3): p. 434-440.
6. Sabbadin, F., et al., *The 1.5-angstrom Structure of XplA-heme, an Unusual Cytochrome P450 Heme Domain That Catalyzes Reductive Biotransformation of Royal Demolition Explosive.* Journal of Biological Chemistry, 2009. **284**(41): p. 28467-28475.

F. References

Adamczyk, J., M. Hesselsoe, N. Iversen, M. Horn, A. Lehner, P. H. Nielsen, M. Schloter, P. Roslev and M. Wagner (2003). "The isotope array, a new tool that employs substrate-mediated labeling of rRNA for determination of microbial community structure and function." Applied and Environmental Microbiology **69**(11): 6875-6887.

Addison, S. L., I. R. McDonald and G. Lloyd-Jones (2010). "Stable isotope probing: Technical considerations when resolving N-15-labeled RNA in gradients." Journal of Microbiological Methods **80**(1): 70-75.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." Journal of Molecular Biology **215**: 403-410.

Amann, R. I., B. J. Binder, R. J. Olson, S. W. Chisholm, R. Devereux and D. A. Stahl (1990). "Combination of 16s Ribosomal-Rna-Targeted Oligonucleotide Probes With Flow-Cytometry For Analyzing Mixed Microbial- Populations." Applied and Environmental Microbiology **56**(6): 1919-1925.

Arp, D. J., C. M. Yeager and M. R. Hyman (2001). "Molecular and cellular fundamentals of aerobic cometabolism of trichloroethylene." Biodegradation **12**(2): 81-103.

Bach, H. J., A. Hartmann, J. T. Trevors and J. C. Munch (1999). "Magnetic capture-hybridization method for purification and probing of mRNA for neutral protease of Bacillus cereus." Journal of Microbiological Methods **37**(2): 187-192.

Bergin, C. (2002). Charakterisierung von stabilen Kohlenstoffisotopen in Gewebe und RNA von Muscheln mit endosymbiontischen Bakterien. Diplomarbeit, Ernst-Moritz-Arndt-Universität.

Bernhard, A. E., T. Donn, A. E. Giblin and D. A. Stahl (2005). "Loss of diversity of ammonia-oxidizing bacteria correlates with increasing salinity in an estuary system." Environmental Microbiology **7**(9): 1289-1297.

Bernstein, A., E. Adar, A. Nejidat and Z. Ronen (2011). "Isolation and characterization of RDX-degrading Rhodococcus species from a contaminated aquifer." Biodegradation **22**(5): 997-1005.

Binks, P. R., C. E. French, S. Nicklin and N. C. Bruce (1996). "Degradation of pentaerythritol tetranitrate by Enterobacter cloacae PB2." Appl Environ Microbiol **62**(4): 1214-1219.

Binks, P. R., S. Nicklin and N. C. Bruce (1995). "Degradation of hexahydro-1,3,5-trinitro-1,3,5-triazine (rdx) by Stenotrophomonas-maltophilia PB1." Applied and Environmental Microbiology **61**(4): 1318-1322.

Blattner, F. R., G. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. ColladoVides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau and Y. Shao (1997). "The complete genome sequence of *Escherichia coli* K-12." Science **277**(5331): 1453-&.

Brandis, A. and R. K. Thauer (1981). "Growth of *Desulfovibrio* species on hydrogen and sulfate as sole energy-source." Journal of General Microbiology **126**(SEP): 249-252.

Breslow, N. (1970). "Generalized Kruskal-Wallis test for comparing k samples subject to unequal patterns of censorship." Biometrika **57**(3): 579-&.

Buckley, D. H., V. Huangyutitham, S.-F. Hsu and T. A. Nelson (2007). "Stable isotope probing with ^{15}N achieved by disentangling the dffects of genome G+C content and isotope enrichment on DNA density." Appl. Environ. Microbiol. **73**(10): 3189-3195.

Buckley, D. H., V. Huangyutitham, S. F. Hsu and T. A. Nelson (2007). "Stable isotope probing with N-15(2) reveals novel noncultivated diazotrophs in soil." Applied and Environmental Microbiology **73**(10): 3196-3204.

Cadisch, G., M. Espana, R. Causey, M. Richter, E. Shaw, J. A. W. Morgan, C. Rahn and G. D. Bending (2005). "Technical considerations for the use of N-15-DNA stable-isotope probing for functional microbial activity in soils." Rapid Communications in Mass Spectrometry **19**(11): 1424-1428.
Cao, P. R., H. Bulow, B. Dumas and R. Bernhardt (2000). "Construction and characterization of a catalytic fusion protein system: P-45011(beta)-adrenodoxin reductase-adrenodoxin." Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology **1476**(2): 253-264.
Cirillo, J. D., R. G. Barletta, B. R. Bloom and W. R. Jacobs Jr. (1991). "A novel transposon trap for mycobacteria: isolation and characterization of IS1096." Journal of Bacteriology **173**(24): 7772-7780.
Clausen, J., J. Robb, D. Curry and N. Korte (2004). "A case study of contaminants on military ranges: Camp Edwards, Massachusetts, USA." Environmental Pollution **129**(1): 13-21.
Coleman, N. V., D. R. Nelson and T. Duxbury (1998). "Aerobic biodegradation of hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) as a nitrogen source by a Rhodococcus sp., strain DN22." Soil Biology & Biochemistry **30**(8-9): 1159-1167.
Coleman, N. V., J. C. Spain and T. Duxbury (2002). "Evidence that RDX biodegradation by Rhodococcus strain DN22 is plasmid-borne and involves a cytochrome p-450." Journal of Applied Microbiology **93**(3): 463-472.
Crocker, F. H., K. J. Indest and H. L. Fredrickson (2006). "Biodegradation of the cyclic nitramine explosives RDX, HMX, and CL-20." Appl Microbiol Biotechnol **73**: 274-290.
Culman, S. W., H. G. Gauch, C. B. Blackwood and J. E. Thies (2008). "Analysis of T-RFLP data using analysis of variance and ordination methods: A comparative study." Journal of Microbiological Methods **75**(1): 55-63.
Cupples, A. M., E. A. Shaffer, J. C. Chee-Sanford and G. K. Sims (2007). "DNA buoyant density shifts during N-15-DNA stable isotope probing." Microbiological Research **162**(4): 328-334.
Cuyckens, F. and M. Claeys (2004). "Mass spectrometry in the structural analysis of flavonoids4." J Mass Spectrom. **39**(1): 1-15.
Davison, J. (1999). "Genetic exchange between bacteria in the environment." Plasmid **42**(2): 73-91.
de Boer, H. A., L. J. Comstock and M. Vasser (1983). "The tac promoter: a functional hybrid derived from the trp and lac promoters." Proc Natl Acad Sci U S A **80**(1): 21-25.
de Souza, M. L., J. Seffernick, B. Martinez, M. J. Sadowsky and L. P. Wackett (1998). "The atrazine catabolism genes atzABC are widespread and highly conserved." J. Bacteriol. **180**(7): 1951-1954.
Dehal, P. S., M. P. Joachimiak, M. N. Price, J. T. Bates, J. K. Baumohl, D. Chivian, G. D. Friedland, K. H. Huang, K. Keller, P. S. Novichkov, I. L. Dubchak, E. J. Alm and A. P. Arkin "MicrobesOnline: an integrated portal for comparative and functional genomics." Nucleic Acids Research.
Dong, D. X., A. Yan, H. M. Liu, X. H. Zhang and Y. Q. Xu (2006). "Removal of humic substances from soil DNA using aluminium sulfate." Journal of Microbiological Methods **66**(2): 217-222.
Donnelly, P. K., R. S. Hegde and J. S. Fletcher (1994). "Growth of PCB-degrading bacteria on compounds from photosynthetic plants." Chemosphere **28**(5): 981-988.
Egert, M. and M. W. Friedrich (2003). "Formation of pseudo-terminal restriction fragments, a PCR-related bias affecting terminal restriction fragment length polymorphism analysis of microbial community structure." Applied and Environmental Microbiology **69**(5): 2555-2562.
Ekblad, A. and A. Nordgren (2002). "Is growth of soil microorganisms in boreal forests limited by carbon or nitrogen availability?" Plant and Soil **242**(1): 115-122.
Errington, J., J. Bath and L. J. Wu (2001). "DNA transport in bacteria." Nature Reviews Molecular Cell Biology **2**(7): 538-544.
Felsenstein, J. (1989). "PHYLIP - phylogeny inference package (version 3.2)." Cladistics **5**: 164-166.
Fournier, D., A. Halasz, J. Spain, P. Fiurasek and J. Hawari (2002). "Determination of key metabolites during biodegradation of hexahydro-1,3,5-trinitro-1,3,5-triazine with Rhodococcus sp strain DN22." Applied and Environmental Microbiology **68**(1): 166-172.
Fournier, D., A. Halasz, J. Spain, R. J. Spanggord, J. C. Bottaro and J. Hawari (2004). "Biodegradation of the hexahydro-1,3,5-trinitro-1,3,5-triazine ring cleavage product 4-nitro-2,4-diazabutanal by Phanerochaete chrysospotium." Applied and Environmental Microbiology **70**(2): 1123-1128.

Fournier, D., S. Trott, J. Hawari and J. Spain (2005). "Metabolism of the aliphatic nitramine 4-nitro-2,4-diazabutanal by *Methylobacterium* sp. Strain JS178." Applied and Environmental Microbiology **71**(8): 4199-4202.
Fuchs, B. M., F. O. Glockner, J. Wulf and R. Amann (2000). "Unlabeled helper oligonucleotides increase the in situ accessibility to 16S rRNA of fluorescently labeled oligonucleotide probes." Applied and Environmental Microbiology **66**(8): 3603-3607.
Gerhardt, P., R. G. E. Murray, W. A. Wood and N. R. Krieg, Eds. (1994). Methods for General and Molecular Bacteriology. Washington, D.C., American Society for Microbiology.
Green, L. C., D. A. Wagner, J. Glogowski, P. L. Skipper, J. S. Wishnok and S. R. Tannenbaum (1982). "Analysis of nitrate, nitrite, and N-15 -labeled nitrate in biological-fluids." Analytical Biochemistry **126**(1): 131-138.
Griffiths, R. I., A. S. Whiteley, A. G. O'Donnell and M. J. Bailey (2003). "Influence of depth and sampling time on bacterial community structure in an upland grassland soil." Fems Microbiology Ecology **43**(1): 35-43.
Halasz, A., D. Manno, S. E. Strand, N. C. Bruce and J. Hawari (2010). "Biodegradation of RDX and MNX with *Rhodococcus* sp. Strain DN22: New insights into the degradation pathway." Environmental Science & Technology **44**(24): 9330-9336.
Hawari, J., A. Halasz, T. Sheremata, S. Beaudet, C. Groom, L. Paquet, C. Rhofir, G. Ampleman and S. Thiboutot (2000). "Characterization of metabolites during biodegradation of hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) with municipal anaerobic sludge." Applied and Environmental Microbiology **66**(6): 2652-2657.
Heeb, S., Y. Itoh, T. Nishijyo, U. Schnider, C. Keel, J. Wade, U. Walsh, F. O'Gara and D. Haas (2000). "Small, stable shuttle vectors based on the minimal pVS1 replicon for use in gram-negative, plant-associated bacteria." Molecular Plant-Microbe Interactions **13**(2): 232-237.
Hopkins, G. D., L. Semprini and P. L. Mccarty (1993). "Microcosm and in situ field studies of enhanced biotransformation of trichloroethylene by phenol-utilizing microorganisms." Appl Environ Microbiol **59**(7): 2277-2285.
Horz, H. P., M. T. Yimga and W. Liesack (2001). "Detection of methanotroph diversity on roots of submerged rice plants by molecular retrieval of *pmoA*, *mmoX*, *mxaF*, and 16S rRNA and ribosomal DNA, including *pmoA*-based terminal restriction fragment length polymorphism profiling." Applied and Environmental Microbiology **67**(9): 4177-4185.
Hou, Y. B., H. Zhang, L. Miranda and S. J. Lin (2010). "Serious overestimation in quantitative PCR by circular (supercoiled) plasmid standard: Microalgal pcna as the model gene." Plos One **5**(3).
Hunter, D. J. B., G. A. Roberts, T. W. B. Ost, J. H. White, S. Muller, N. J. Turner, S. L. Flitsch and S. K. Chapman (2005). "Analysis of the domain properties of the novel cytochrome P450RhF." Febs Letters **579**(10): 2215-2220.
Hurst, C. J. and G. R. Knudsen, Eds. (1997). Manual of environmental microbiology. Washington, D.C., ASM Press.
Indest, K. J., F. H. Crocker and R. Athow (2007). "A TaqMan polymerase chain reaction method for monitoring RDX-degrading bacteria based on the xp1A functional gene." Journal of Microbiological Methods **68**(2): 267-274.
Indest, K. J., C. M. Jung, H.-P. Chen, D. Hancock, C. Florizone, L. D. Eltis and F. H. Crocker (2010). "Functional characterization of pGKT2, a 182-kilobase plasmid containing the *xplAB* genes, which are involved in the degradation of hexahydro-1,3,5-trinitro-1,3,5-triazine by *Gordonia* sp. Strain KTR9." Appl. Environ. Microbiol. **76**(19): 6329-6337.
Ito, K., M. Takahashi, T. Yoshimoto and D. Tsuru (1994). "Cloning and High-Level Expression of the Glutathione-Independent Formaldehyde Dehydrogenase Gene from Pseudomonas-Putida." Journal of Bacteriology **176**(9): 2483-2491.
Jackson, R. G., E. L. Rylott, D. Fournier, J. Hawari and N. C. Bruce (2007). "Exploring the biochemical properties and remediation applications of the unusual explosive-degrading P450 system XpIA/B." Proceedings of the National Academy of Sciences of the United States of America **104**: 16822-16827.

Jackson, R. G., E. L. Rylott, D. Fournier, J. Hawari and N. C. Bruce (2007). "Exploring the biochemical properties and remediation applications of the unusual explosive-degrading P450 system XplA/B." Proc.Natl.Acad.Sci.U.S.A **104**(43): 16822-16827.
Jensen, L. E. and O. Nybroe (1999). "Nitrogen availability to Pseudomonas fluorescens DF57 is limited during decomposition of barley straw in bulk soil and in the barley rhizosphere." Applied and Environmental Microbiology **65**(10): 4320-4328.
Joner, E. J., S. C. Corgie, N. Amellal and C. Leyval (2002). "Nutritional constraints to degradation of polycyclic aromatic hydrocarbons in a simulated rhizosphere." Soil Biology & Biochemistry **34**(6): 859-864.
Juhasz, A. L. and R. Naidu (2000). "Bioremediation of high molecular weight polycyclic aromatic hydrocarbons: a review of the microbial degradation of benzo a pyrene." International Biodeterioration & Biodegradation **45**(1-2): 57-88.
Jung, C. M., F. H. Crocker, J. O. Eberly and K. J. Indest (2011). "Horizontal gene transfer (HGT) as a mechanism of disseminating RDX-degrading activity among Actinomycete bacteria." Journal of Applied Microbiology **110**(6): 1449-1459.
Kalyuzhnaya, M. G., A. Lapidus, N. Ivanova, A. C. Copeland, A. C. McHardy, E. Szeto, A. Salamov, I. V. Grigoriev, D. Suciu, S. R. Levine, V. M. Markowitz, I. Rigoutsos, S. G. Tringe, D. C. Bruce, P. M. Richardson, M. E. Lidstrom and L. Chistoserdova (2008). "High-resolution metagenomics targets specific functional types in complex microbial communities." Nature Biotechnology **26**(9): 1029-1034.
Kamilova, F., L. Kravchenko, A. Shaposhnikov, T. Azarova, N. Makarova and B. Lugtenberg (2006). "Organic acids, sugars, and L-tryptophane in exudates of vegetables growing on stonewool and their effects on activities of rhizosphere bacteria." Molecular Plant-Microbe Interactions **19**(3): 250-256.
Karlovsky, P. and A. Decock (1991). "Buoyant density of DNA Hoechst-33258 (bisbenzimide) complexes in CsCl gradients - Hoechst-33258 binds to single AT base-pairs." Analytical Biochemistry **194**(1): 192-197.
Kimura, M. (1980). "A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide-sequences." Journal of Molecular Evolution **16**(2): 111-120.
Koch, B., J. Worm, L. E. Jensen, O. Hojberg and O. Nybroe (2001). "Carbon limitation induces sigma(S)-dependent gene expression in Pseudomonas fluorescens in soil." Applied and Environmental Microbiology **67**(8): 3363-3370.
Lane, D. J. (1991). 16S/23S rRNA sequenciing. Nucleic Acid Techniques in Bacterial Systematics. E. Stakebrandt New York, Wiley**:** 115-175.
Leigh, M. B., J. S. Fletcher, X. O. Fu and F. J. Schmitz (2002). "Root turnover: An important source of microbial substrates in rhizosphere remediation of recalcitrant contaminants." Environmental Science & Technology **36**(7): 1579-1583.
Lucker, S., D. Steger, K. U. Kjeldsen, B. J. MacGregor, M. Wagner and A. Loy (2007). "Improved 16S rRNA-targeted probe set for analysis of sulfate-reducing bacteria by fluorescence in situ hybridization." Journal of Microbiological Methods **69**(3): 523-528.
Ludwig, W., O. Strunk, R. Westram, L. Richter, H. Meier, Yadhukumar, A. Buchner, T. Lai, S. Steppi, G. Jobb, W. Forster, I. Brettske, S. Gerber, A. W. Ginhart, O. Gross, S. Grumann, S. Hermann, R. Jost, A. Konig, T. Liss, R. Lussmann, M. May, B. Nonhoff, B. Reichel, R. Strehlow, A. Stamatakis, N. Stuckmann, A. Vilbig, M. Lenke, T. Ludwig, A. Bode and K. H. Schleifer (2004). "ARB: a software environment for sequence data." Nucleic Acids Research **32**(4): 1363-1371.
Lueders, T., B. Pommerenke and M. W. Friedrich (2004). "Stable-isotope probing of microorganisms thriving at thermodynamic limits: Syntrophic propionate oxidation in flooded soil." Applied and Environmental Microbiology **70**(10): 5778-5786.
Lugtenberg, B. and F. Kamilova (2009). "Plant-Growth-Promoting Rhizobacteria." Annual Review of Microbiology **63**: 541-556.
Lukashin, A. and M. Borodovsky (1998). "GeneMark.hmm: new solutions for gene finding." Nucleic Acids Research **26**(4): 1107-1115.

MacGregor, B. J. and R. Amann (2006). "Single-stranded conformational polymorphism for separation of mixed rRNAS (rRNA-SSCP), a new method for profiling microbial communities." Systematic and Applied Microbiology **29**(8): 661-670.
MacGregor, B. J., H. T. S. Boschker and R. Amann (2006). "Comparison of rRNA and polar-lipid-derived fatty acid biomarkers for assessment of C-13-substrate incorporation by microorganisms in marine sediments." Applied and Environmental Microbiology **72**(8): 5246-5253.
MacGregor, B. J., V. Bruchert, S. Fleischer and R. Amann (2002). "Isolation of small-subunit rRNA for stable isotopic characterization." Environmental Microbiology **4**(8): 451-464.
MacGregor, B. J., D. P. Moser, E. W. Alm, K. H. Nealson and D. A. Stahl (1997). "Crenarchaeota in Lake Michigan sediment." Applied and Environmental Microbiology **63**(3): 1178-1181.
Mahillon, J. and M. Chandler (1998). "Insertion Sequences." Microbiology and Molercular Biology Reviews **62**(3): 725-774.
Manefield, M., R. Griffiths, N. P. McNamara, D. Sleep, N. Ostle and A. Whiteley (2007). "Insights into the fate of a C-13 labelled phenol pulse for stable isotope probing (SIP) experiments." Journal of Microbiological Methods **69**(2): 340-344.
Manefield, M., A. S. Whiteley, R. I. Griffiths and M. J. Bailey (2002). "RNA stable isotope probing, a novel means of linking microbial community function to Phylogeny." Applied and Environmental Microbiology **68**(11): 5367-5373.
Mastrangeli, R., E. Micangeli and S. Donini (1996). "Cloning of murine LAG-3 by magnetic bead bound homologous probes and PCR (GENE-CAPTURE PCR)." Analytical Biochemistry **241**(1): 93-102.
McIlroy, S., K. Porter, R. J. Seviour and D. Tillett (2008). "Simple and Safe Method for Simultaneous Isolation of Microbial RNA and DNA from Problematic Populations." Applied and Environmental Microbiology **74**(21): 6806-6807.
Miller, K. J. and J. M. Wood (1996). "Osmoadaptation by rhizosphere bacteria." Annu Rev Microbiol **50**: 101-136.
Miyatake, T., B. J. MacGregor and H. T. S. Boschker (2009). "Linking Microbial Community Function to Phylogeny of Sulfate-Reducing Deltaproteobacteria in Marine Sediments by Combining Stable Isotope Probing with Magnetic-Bead Capture Hybridization of 16S rRNA." Applied and Environmental Microbiology **75**(15): 4927-4935.
Munro, A. W., J. G. Lindsay, J. R. Coggins, S. M. Kelly and N. C. Price (1996). "Analysis of the structural stability of the multidomain enzyme flavocytochrome P-450 BM3." Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology **1296**(2): 127-137.
Muyzer, G., A. Teske, C. O. Wirsen and H. W. Jannasch (1995). "Phylogenetic-relationships of *Thiomicrospira* species and their identification in deep-sea hydrothermal vent samples by denaturing gradient gel-electrophoresis of 16S rDNA fragments." Archives of Microbiology **164**(3): 165-172.
Narasimhan, K., C. Basheer, V. B. Bajic and S. Swarup (2003). "Enhancement of plant-microbe interactions using a rhizosphere metabolomics-driven approach and its application in the removal of polychlorinated biphenyls." Plant Physiology **132**(1): 146-153.
Nejidat, A., L. Kafka, Y. Tekoah and Z. Ronen (2008). "Effect of organic and inorganic nitrogenous compounds on RDX degradation and cytochrome P-450 expression in *Rhodococcus* strain YH1." Biodegradation **19**(3): 313-320.
Neufeld, J. D., Y. Chen, M. G. Dumont and J. C. Murrell (2008). "Marine methylotrophs revealed by stable-isotope probing, multiple displacement amplification and metagenomics." Environmental Microbiology **10**(6): 1526-1535.
Neufeld, J. D., J. Vohra, M. G. Dumont, T. Lueders, M. Manefield, M. W. Friedrich and J. C. Murrell (2007). "DNA stable-isotope probing." Nature Protocols **2**(4): 860-866.
Nielsen, P. H., K. Andreasen, N. Lee and M. Wagner (1999). "Use of microautoradiography and fluorescent in situ hybridization for characterization of microbial activity in activated sludge." Water Science and Technology **39**(1): 1-9.
Oger, P., A. Petit and Y. Dessaux (1997). "Genetically engineered plants producing opines alter their biological environment." Nature Biotechnology **15**(4): 369-372.

Osterman, L. A. (1984). Methods of protein and nucleic acid research. Berlin; New York, Springer-Verlag.
Padmanabhan, P., S. Padmanabhan, C. DeRito, A. Gray, D. Gannon, J. R. Snape, C. S. Tsai, W. Park, C. Jeon and E. L. Madsen (2003). "Respiration of C-13-labeled substrates added to soil in the field and subsequent 16S rRNA gene analysis of C-13-labeled soil DNA." Applied and Environmental Microbiology **69**(3): 1614-1622.
Pandey, J., K. Ganesan and R. K. Jain (2007). "Variations in T-RFLP profiles with differing chemistries of fluorescent dyes used for labeling the PCR primers." Journal of Microbiological Methods **68**(3): 633-638.
Parsons, J. A., T. L. Bannam, R. J. Devenish and J. I. Rood (2007). "TcpA an FtsK/SpoIIIE homolog, is essential for transfer of the conjugative plasmid pCW3 in *Clostridium perfringens*." Journal of Bacteriology **189**(21): 7782-7790.
Pennington, J. C. and J. M. Brannon (2002). "Environmental fate of explosives." Thermochimica Acta **384**(1-2): 163-172.
Pfennig, N. and K. D. Lippert (1966). "Uber das vitamin B12-beduffnis phototropher *Schwefelbakterien*." Archiv Fur Mikrobiologie **55**(3): 245-&.
Polz, M. F. and C. M. Cavanaugh (1998). "Bias in template-to-product ratios in multitemplate PCR." Applied and Environmental Microbiology **64**(10): 3724-3730.
Pruesse, E., C. Quast, K. Knittel, B. M. Fuchs, W. G. Ludwig, J. Peplies and F. O. Glockner (2007). "SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB." Nucleic Acids Research **35**(21): 7188-7196.
Radajewski, S., P. Ineson, N. R. Parekh and J. C. Murrell (2000). "Stable-isotope probing as a tool in microbial ecology." Nature **403**(6770): 646-649.
Radajewski, S., I. R. McDonald and J. C. Murrell (2003). "Stable-isotope probing of nucleic acids: a window to the function of uncultured microorganisms." Current Opinion in Biotechnology **14**(3): 296-302.
Radajewski, S., G. Webster, D. S. Reay, S. A. Morris, P. Ineson, D. B. Nedwell, J. I. Prosser and J. C. Murrell (2002). "Identification of active methylotroph populations in an acidic forest soil by stable isotope probing." Microbiology-Sgm **148**: 2331-2342.
Rangel-Castro, J. I., K. Killham, N. Ostle, G. W. Nicol, I. C. Anderson, C. M. Scrimgeour, P. Ineson, A. Meharg and J. I. Prosser (2005). "Stable isotope probing analysis of the influence of liming on root exudate utilization by soil microorganisms." Environmental Microbiology **7**(6): 828-838.
Rasmussen, R. (2001). Quantification on the Lightcycler. *Rapid Cycle Real-time PCR, Methods and Applications*. S. Meuer, C. Wittwer and K. Nakagawara. Heidelberg, Springer Press**:** 21-34.
Reasoner, D. J. and E. E. Geldreich (1985). "A new medium for the enumeration and subculture of bacteria from potable water." Applied and Environmental Microbiology **49**(1): 1-7.
Riley, M. (1993). "Functions of the gene-products of *Escherichia coli*." Microbiological Reviews **57**(4): 862-952.
Roh, H., C. P. Yu, M. E. Fuller and K. H. Chu (2009). "Identification of Hexahydro-1,3,5-trinitro-1,3,5-triazine-Degrading Microorganisms via N-15-Stable Isotope Probing." Environmental Science & Technology **43**(7): 2505-2511.
Rosenberger, R. F. and S. R. Elsden (1960). "The yields of *Streptococcus-faecalis* grown in continuous culture." Journal of General Microbiology **22**(3): 726-739.
Rosser, S. J., A. Basran, E. R. Travis, C. E. French and N. C. Bruce (2001). Microbial transformations of explosives. Advances in Applied Microbiology, Vol 49. **49:** 1-35.
Rylott, E. L., R. G. Jackson, J. Edwards, G. L. Womack, H. M. B. Seth-Smith, D. A. Rathbone, S. E. Strand and N. C. Bruce (2006). "An explosive-degrading cytochrome P450 activity and its targeted application for the phytoremediation of RDX." Nature Biotechnology **24**(2): 216-219.
Sabbadin, F., R. Jackson, K. Haider, G. Tampi, J. P. Turkenburg, S. Hart, N. C. Bruce and G. Grogan (2009). "The 1.5-angstrom Structure of XplA-heme, an Unusual Cytochrome P450 Heme Domain That

Catalyzes Reductive Biotransformation of Royal Demolition Explosive." Journal of Biological Chemistry **284**(41): 28467-28475.
Sambrook, J. and D. W. Russel (2001). Molecular cloning: a laboratory manual. Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press.
Savka, M. A., Y. Dessaux, P. Oger and S. Rossbach (2002). "Engineering bacterial competitiveness and persistence in the phytosphere." Molecular Plant-Microbe Interactions **15**(9): 866-874.
Schwartz, E. (2007). "Characterization of growing microorganisms in soil by stable isotope probing with (H2O)-O-18." Applied and Environmental Microbiology **73**(8): 2541-2546.
Schwarz, S., T. Waschkowitz and R. Daniel (2006). "Enhancement of gene detection frequencies by combining DNA-based stable-isotope probing with the construction of metagenomic DNA libraries." World Journal of Microbiology & Biotechnology **22**(4): 363-367.
Seth-Smith, H. M., S. J. Rosser, A. Basran, E. R. Travis, E. R. Dabbs, S. Nicklin and N. C. Bruce (2002). "Cloning, sequencing, and characterization of the hexahydro-1,3,5-trinitro-1,3,5-triazine degradation gene cluster from Rhodococcus rhodochrous." Appl.Environ.Microbiol. **68**(10): 4764-4771.
Seth-Smith, H. M. B., J. Edwards, S. J. Rosser, D. A. Rathbone and N. C. Bruce (2008). "The explosive-degrading cytochrome P450 system is highly conserved among strains of Rhodococcus spp." Applied and Environmental Microbiology **74**(14): 4550-4552.
Seth-Smith, H. M. B., S. J. Rosser, A. Basran, E. R. Travis, E. R. Dabbs, S. Nicklin and N. C. Bruce (2002). "Cloning, sequencing, and characterization of the hexahydro-1,3,5-trinitro-1,3,5-triazine degradation gene cluster from Rhodococcus rhodochrous." Applied and Environmental Microbiology **68**(10): 4764-4771.
Shull, T. L., G. E. Speitel and D. C. McKinney (1999). Bioremediaiton of RDX in the vadose zone beneath the Pantex Plant. The Amarillo National Resource Center for Plutonium Report. Amarillo, U.S. Department of Energy**:** 7-8.
Suzuki, M. T. and S. J. Giovannoni (1996). "Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR." Applied and Environmental Microbiology **62**(2): 625-630.
Thompson, K. T., F. H. Crocker and H. L. Fredrickson (2005). "Mineralization of the cyclic nitramine explosive hexahydro-1,3,5-trinitro-1,3,5-triazine by Gordonia and Williamsia spp." Applied and Environmental Microbiology **71**(12): 8265-8272.
Tokala, R. K., J. L. Strap, C. M. Jung, D. L. Crawford, M. H. Salove, L. A. Deobald, J. F. Bailey and M. J. Morra (2002). "Novel plant-microbe rhizosphere interaction involving Streptomyces lydicus WYEC108 and the pea plant (Pisum sativum)." Appl Environ Microbiol **68**(5): 2161-2171.
Top, E. M., D. Springael and N. Boon (2002). "Catabolic mobile genetic elements and their potential use in bioaugmentation of polluted soils and waters." Fems Microbiology Ecology **42**(2): 199-208.
Trefault, N., R. de la Iglesia, A. M. Molina, M. Manzano, T. Ledger, D. Pérez-Pantoja, M. A. Sánchez, M. Stuardo and B. González (2004). "Genetic organization of the catabolic plasmid pJP4 from *Ralstonia eutropha* JMP134 (pJP4) reveals mechanisms of adaptation to chloroaromatic pollutants and evolution of specialized chloroaromatic degradation pathway." Environmental Microbiology **6**(7): 655-668.
U.S. Environmental Protection Agency (2005). Handbook on the management of munitions response actions. Washington D.C., Office of Solid Waste and Emergency Response.
U.S. Environmental Protection Agency (2006). 2006 edition of the drinking water standards and health advisories. Washington, D.C., Office of Water.
United States Environmental Protection Agency (USEPA) (1994). Method 8330: Nitroaromatics and nitramines by high performance liquid chromatography (HPLC). Washington, D.C.**:** 1-21.
van der Meer, J. R., W. M. de Vos, S. Harayama and J. B. Zehnder (1992). "Molecular mechanisms of genetic adaptation to xenobiotic compounds." Microbiological Reviews **56**(4): 677-694.
Van Eerd, L. L., R. E. Hoagland, R. M. Zablotowicz and J. C. Hall (2003). "Pesticide metabolism in plants and microorganisms." Weed Science **51**(4): 472-495.
Villacieros, M., C. Whelan, M. Mackova, J. Molgaard, M. Sánchez-Contreras, J. Lloret, D. Aguirre de Cárcer, R. I. Oruezábal, L. Bolaños, T. Macek, U. Karlson, D. N. Dowling, M. Martín and R. Rivilla (2005). "Polychlorinated biphenyl rhizoremediation by Pseudomonas fluorescens F113 derivatives, using

a Sinorhizobium meliloti nod system to drive bph gene expression." Appl Environ Microbiol **71**(5): 2687-2694.
von Wintzingerode, F., U. B. Gobel and E. Stackebrandt (1997). "Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis." Fems Microbiology Reviews **21**(3): 213-229.
Walsh, M. E., S. Taylor, A. D. Hewitt, M. R. Walsh, C. A. Ramsey and C. M. Collins (2010). "Field observations of the persistence of Comp B explosives residues in a salt marsh impact area." Chemosphere **78**(4): 467-473.
Wawrik, B., A. V. Callaghan and D. A. Bronk (2009). "Use of inorganic and organic nitrogen by *Synechococcus* spp. and diatoms on the West Florida Shelf as measured using stable isotope probing." Applied and Environmental Microbiology **75**(21): 6662-6670.
Yang, Y., X. Wang, P. Yin, W. Li and P. Zhou (1983). "Studies on three strains of Corynebacterium degrading cyclotrimethylene-trinitroamine (RDX)." Acta. Microbiol. Sinica **23**: 251-256.
Ye, J., A. Singh and O. P. Ward (2004). "Biodegradation of nitroaromatics and other nitrogen-containing xenobiotics." World J Microbial Biotechnol **20**: 117-135.
Yokota, A., M. Takeuchi and N. Weiss (1993). "Proposal of two new species in the genus *Microbacterium*: *Microbacterium dextranolyticum* sp. nov. and *Microbacterium aurum* sp. nov." International Journal of Systematic Bacteriology **43**: 549-554.
Yu, C. P., R. Ahuja, G. Sayler and K. H. Chu (2005). "Quantitative molecular assay for fingerprinting microbial communities of wastewater and estrogen-degrading consortia." Applied and Environmental Microbiology **71**(3): 1433-1444.