GREAT AUSTRALIAN BIGHT
RESEARCH PROGRAM

RESEARCH REPORT SERIES

# Development of a functional genomics assay to measure the abundance of hydrocarbon degrading bacteria in the Great Australian Bight

Sharon Hook, Jodie van de Kamp, Alan Williams, Jason Tanner and Levente Bodrossy

GABRP Research Report Series Number 11

October 2016

bp CSIRO Government of South Australia Primary Industries and Regions SA SARDI South Australian Research and Development Institute THE UNIVERSITY of ADELAIDE Flinders UNIVERSITY

## DISCLAIMER



## COPYRIGHT



## THIS PUBLICATION MAY BE CITED AS:

Hook, S., van de Kamp, J., Williams, A., Tanner, J. and Bodrossy, L. (2016). Development of a functional genomics assay to measure the abundance of hydrocarbon degrading bacteria in the Great Australian Bight. Great Australian Bight Research Program, Great Australian Bight Research Report Series Number 11, 21pp.

## CORRESPONDING AUTHOR

Dr Sharon Hook
CSIRO
e: Sharon.Hook@csiro.au

## FOR FURTHER INFORMATION

www.misa.net.au/GAB

## GREAT AUSTRALIAN BIGHT RESEARCH PROGRAM

The Great Australian Bight Research Program is a collaboration between BP, CSIRO, the South Australian Research and Development Institute (SARDI), the University of Adelaide, and Flinders University. The Program aims to provide a whole-of-system understanding of the environmental, economic and social values of the region; providing an information source for all to use.

# Contents

# Figures



# Tables

# Acknowledgments

Funding for this study was provided by the Great Australian Bight Research Program - a collaboration between BP, CSIRO, the South Australian Research and Development Institute (SARDI), the University of Adelaide, and Flinders University. The Program aims to provide a whole-of-system understanding of the environmental, economic and social values of the region; providing an information source for all to use.

Thanks are due to CSIRO colleagues for help with various aspects of the project, but particularly Franzis Althaus for figures, and the participants on two seagoing surveys for obtaining samples: from 2013 (*RV Southern Surveyor*) the ship's crew and officers, especially the Master Mike Watson, and scientists, especially Mark Green, Mark Lewis, Matt Sherlock and Jeff Cordell, and from 2015 (*RV Investigator*), the ship's crew and officers, especially the Master John Highton and scientists, especially Mark Green, Phil de Boer and Matt Sherlock.

# Executive summary

Being able to quantify hydrocarbon loss to microbial degradation is important for both the risk assessment for marine oil wells, and to determine the amount of petroleum released into the environment either from a spill or from routine discharges. With that in mind, we used previously reported next generation sequencing results to design quantitative PCR primers specific to the microbial taxa indigenous to the Great Australian Bight (GAB). These primers were intended for use in functional gene assays that would quantify the changes in abundance of bacteria involved in alkane, methane or aromatic hydrocarbon degradation. Using these primers, we were able to measure the abundance of hydrocarbon degrading bacteria from environmental samples collected from two surveys in the GAB, the first on the RV Southern Surveyor in April 2013, and the second on RV Investigator in December 2015. Although these bacteria were rare (between $10^{-5}$ to 0.1% of the total bacterial population), they were present at every station and every sample we analysed. Although the previously conducted sequencing efforts were instrumental in identifying the sequences of the selected genes in bacteria indigenous to the GAB, the tool is not adequately high throughput (e.g. low cost, quick turnaround) for routine monitoring. By contrast, these qPCR assays can be used as both a high throughput screening tool to measure the spatial and temporal duration of impact from routine discharges of petroleum and to monitor environmental fate and persistence following an oil spill, especially if coupled with a Fluidigm system or similar technology.

# 1 Introduction

When oil is released into the environment, its composition and concentration changes through a process known as weathering (RSC, 2015). Different hydrocarbons within oil weather differently, and consequently experience a different environmental fate and persistence. Some of these compounds will evaporate, some will be subject to photolysis, and others will adhere to sediment particles and sink (reviewed in RSC, 2015). Ultimately, much of the oil that is released will be degraded by bacteria. Oil degrading bacteria have been detected in all environments studied to date (Atlas, 1995a, b), including the Great Australian Bight (GAB) (Hook et al., 2016 a,b). Microbial degradation is thought to be the fate of most oil released into the environment via natural processes such as seeps (Atlas and Hazen, 2011; Head et al., 2006), and the fate of approximately half of the oil released from the *Deepwater Horizon* wellhead blowout (Kleindienst et al., 2015).

When oil is accidently released into the environment, being able to measure the rates of oil degradation is useful when planning the response (Dubinsky et al., 2013; Hazen et al., 2010; Lamendella et al., 2014; Lu et al., 2012). Knowing rates of oil degradation enables accurate predictions of fate and persistence, and assists assessment of environmental risk. It also assists calculation of how much oil was released into the environment. Finally, being able to measure rates of oil degradation enables determination of the efficacy of response technologies deployed in the event of a spill, such as chemical oil dispersion or bioremediation.

Exploitation of the petroleum resources in the GAB carries with it the risk of an oil spill, either via a well head blowout or from transport of the resources. Companies involved in petroleum resource development will also be required to monitor the impacts of any routine discharges, such as those due to any minor leaks around the well head. A high throughput functional genomics based assay to monitor changes in the abundance of hydrocarbon degrading bacteria will be useful in either scenario. The abundance of functional genes associated with hydrocarbon degradation increases in samples with a source of oil, and returns to “background” levels once that oil source is degraded (Chakraborty et al., 2012; Dubinsky et al., 2013; Lu et al., 2012). Consequently, we have designed a series of functional gene-based assays to quantify the abundance of bacteria in the environment with the capacity to degrade specific fractions of oil. Since the diversity of hydrocarbon degrading gene sequences from hydrocarbon degrading bacteria in environmental samples collected from the GAB could not be sufficiently described using sequences in NCBI’s GenBank, we first used massively parallel next generation sequencing to determine the gene variants indigenous to the region, as described in a previous report (Hook et al., 2016a). In this study, we describe quantitative PCR assays that we use to measure the abundance of different variants of these genes in environmental samples. These assays can be used as a tool to monitor both the environmental fate of routine discharges of oil as well as petroleum compounds released during a spill.

# 2 Materials and methods

## 2.1 Study sites and sampling design

Sediment samples were collected from two different research surveys: a survey from 3-22 April 2013 conducted on the *RV Southern Surveyor*, and a second survey from 30 November to 21 December 2015 conducted on the *RV Investigator.* Sampling locations are shown below in Figure 1. During each survey, samples were collected using a 6-core multicorer from KC (Denmark) that was controlled from the vessel and allowed reliable collection of sediment samples at depths between 200 and 3,000 m (Sherlock et al., 2014). A 30 mm diameter minicore was extracted from each of 3 cores at each location for microbial analysis. Sediment from the top 20 mm of each minicore was collected into a DNA free tube and frozen. Samples were shipped frozen to CSIRO laboratories for DNA extraction.

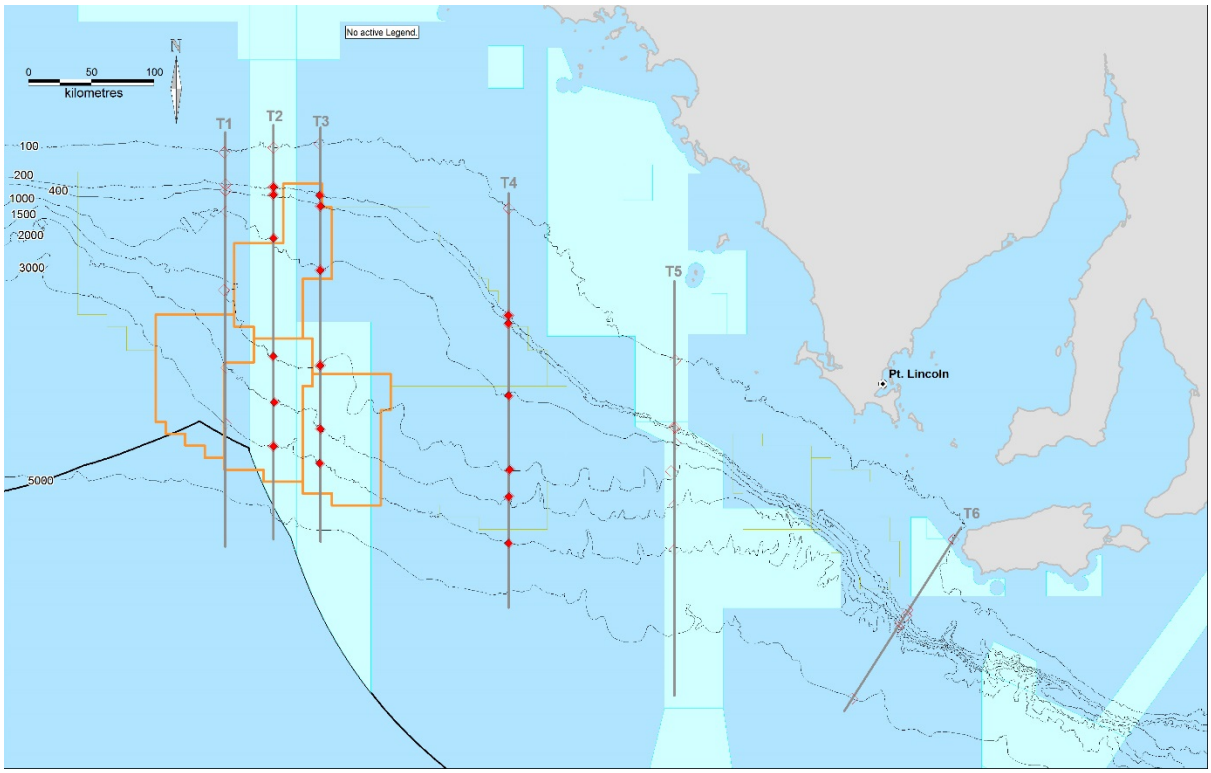



**Figure 1. Transects and sampling locations used in the 2013 and 2015 surveys. Sampling locations used for development of the functional assays are shown in red. BP's lease blocks are shown in orange, and the light blue shading indicates Commonwealth marine reserves.**

## 2.2 DNA extractions

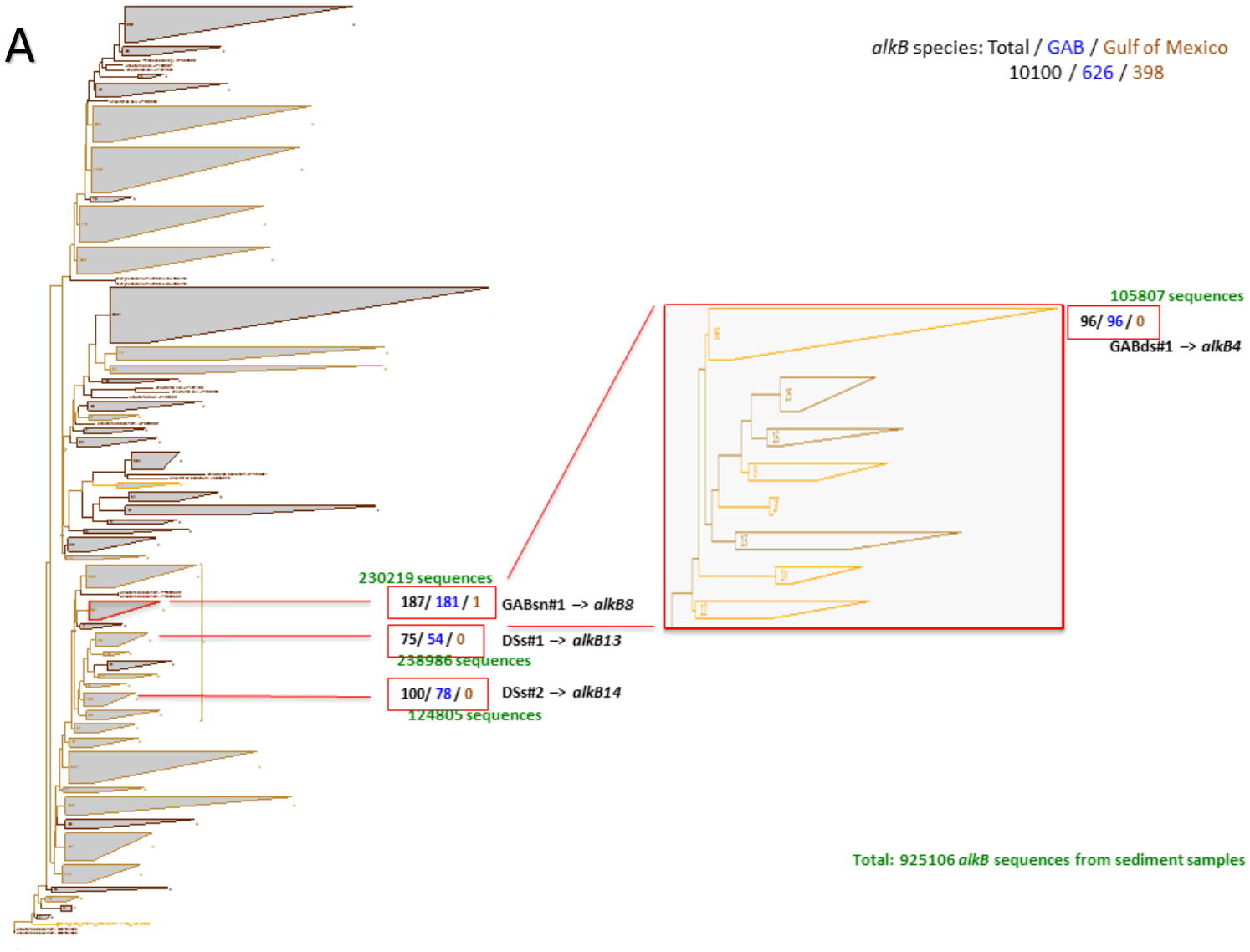
DNA was extracted as described previously (Hook et al., 2016a). DNA was extracted from 10 g of sediment using the MoBio Powermax kit (Mo Bio Laboratories Inc, USA), with the following modifications: after adding the lysis solution (C1), samples were incubated for 10 minute at 70 °C, and

all other incubation times were extended to 30 minutes. Once DNA was eluted, the samples were concentrated to dry pellets in a “speed vac” vacuum concentrator, and then washed in 100% ethanol to remove excess salts. A NanoDrop™ 8000 Spectrophotometer (Thermo Scientific™) was used to verify the quality and quantity of all DNA. DNA was aliquoted into multiple tubes, vacuum dried and stored at -20 °C.

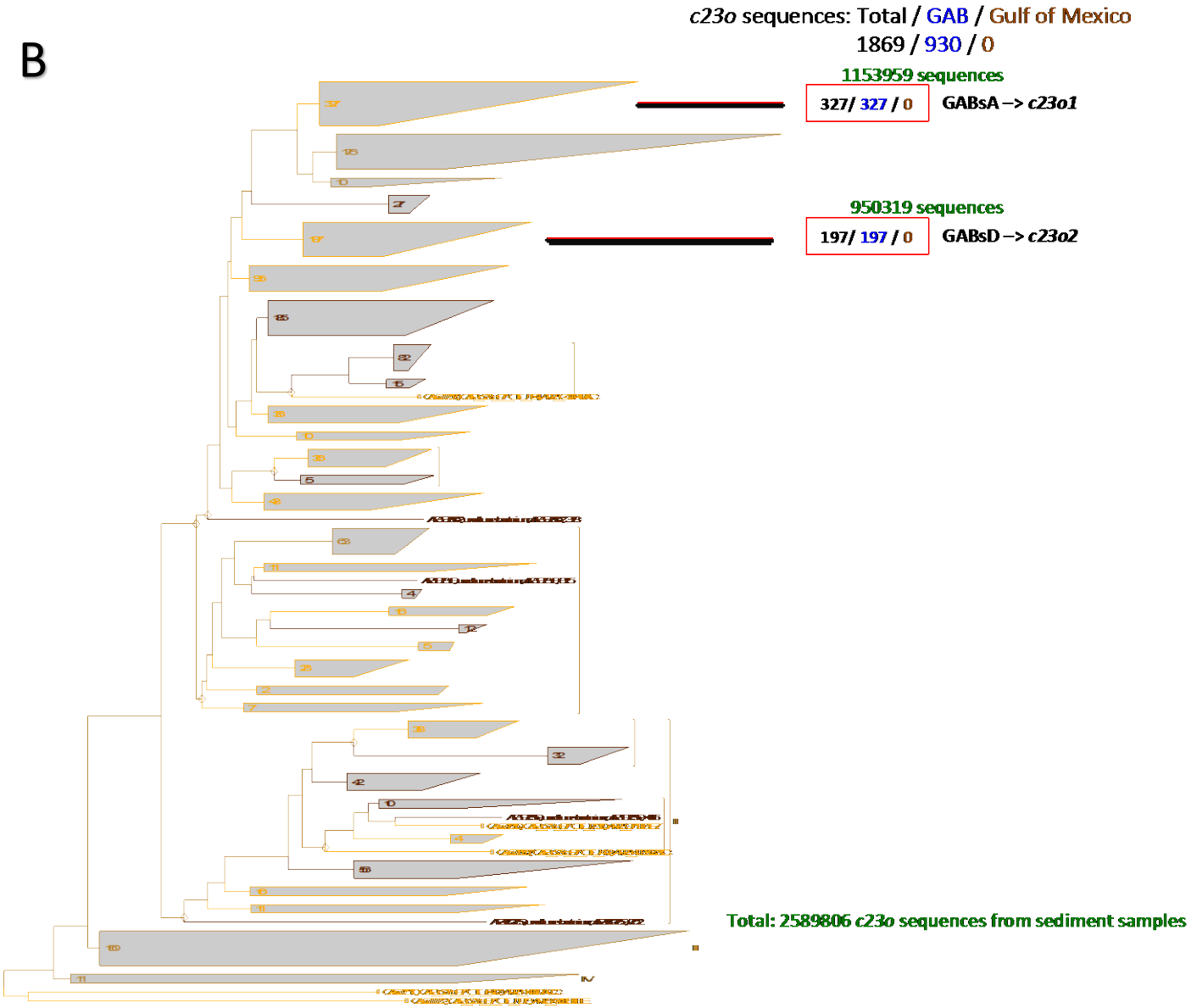
## 2.3 PCR primer design

Databases for *alkB*, *pmoA* and *c23o* that contain all publicly available sequences obtained from the NCBI GenBank databases, including those deposited after the *DeepWater Horizon* wellhead blowout, as well as the representative sequences of OTU's generated from Illumina next generation sequencing of GAB sediment samples (Hook et al., 2016a) were constructed. Sequences in these databases were aligned and phylogenetic trees were created using the ARB phylogenetic software package (http://www.arb-home.de/). Regions of the phylogenetic trees selected for primer design are illustrated in Figure 2 (a-c). Primer pairs were designed to target specific clades of abundant OTUs detected in the GAB sediments (Hook et al., 2016a). The numbers of OTU's in each clade (the diversities within the clades) are given in the numbers in blue next to each clade name. These numbers can be compared to how the same clades have been represented in previous studies as a whole (in black) and specifically in the Gulf of Mexico (in brown). The numbers of amplicons from bacteria (i.e. the relative abundance of each taxon) in each clade are in green. For *alkB*, 3 major clades containing between them 50% of the novel diversity discovered within the GAB sediment samples were selected, as shown in Figure 2A. These clades also contained >60% of all the *alkB* containing bacteria detected in the GAB sediments. For *pmoA* (Figure 2B), clade GABds#1 was chosen as it contained most of the *pmoA* sequences found in the GAB. The Deep Sea #2 clade was also chosen as it has been known for a longer time and includes sequences from a range of deep sea environments. This group was relatively abundant in samples from the GAB. For *c23o*, clades GABsA and GABsD (shown in figure 2C) contained >50% of the novel diversity of *c23o* discovered in the GAB. They also contained >90% of all the *c23o* containing bacteria detected in the GAB sediments.

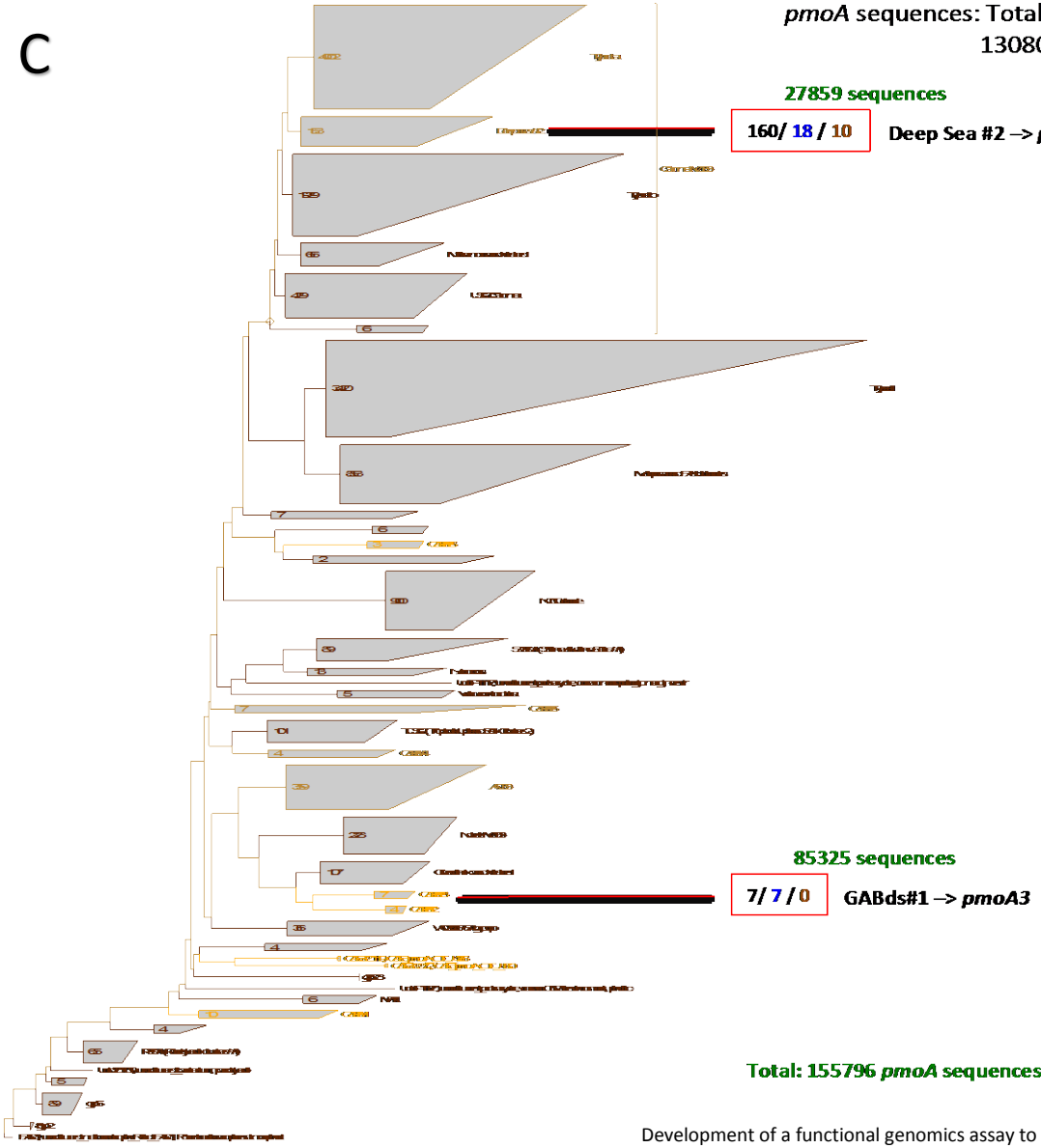
Primer pairs were designed to match common criteria for quantitative PCR (e.g. Mackay, 2004). These were: 1) Similar annealing temperatures, 2) amplicon length between 70 and 300 nucleotides, 3) maximized specificity and coverage for the targeted clade. Primer details are given in Table 1. Generic (broad specificity, targeting known diversity of the functional marker genes) qPCR primers from the literature were adapted and tested as well.


A
alkB species: Total / GAB / Gulf of Mexico
10100 / 626 / 398
105807 sequences
96/ 96 / 0
GABds#1 –> alkB4
230219 sequences
187/ 181 / 1
GABsn#1 –> alkB8
75/ 54 / 0
DSs#1 –> alkB13
238986 sequences
100/ 78 / 0
DSs#2 –> alkB14
124805 sequences
Total: 925106 alkB sequences from sediment samples

B


c23o sequences: Total / GAB / Gulf of Mexico
1869 / 930 / 0
1153959 sequences
327/ 327 / 0
GABsA –> c23o1
950319 sequences
197/ 197 / 0
GABsD –> c23o2
Total: 2589806 c23o sequences from sediment samples

C



*pmoA* sequences: Total / GAB / Gulf of Mexico
13080 / 91 / 39

27859 sequences

160/ 18 / 10 Deep Sea #2 –> *pmoA1*

85325 sequences

7/ 7 / 0 GABds#1 –> *pmoA3*

Total: 155796 *pmoA* sequences from sediment samples

**Figure 2. Clades selected for designing *alkB* (panel A), *pmoA* (panel B), and *c23o* (panel C) qPCR primers. The relative abundance of each taxon is given in the numbers in blue next to each clade name. These abundances can be compared to their representation in the literature as a whole (in black) and to their abundance in the Gulf of Mexico (in brown). The numbers of amplicons from bacteria in each clade are in green. Clades outlined in light brown contain sequences that have only been found in the Great Australian Bight. Clades outlined in dark brown contain sequences that have been found in the GAB as well as sequences that have been described in other studies. Clades outlined in black only contain sequences from studies in the literature and do not have members that were identified in this study.**

**Table 1. Primers used in qPCR reactions**

| Gene | Assay | Forward Primer (5'-3') | Reverse Primer (5'-3') | Annealing Temp (°C) | Acquisition Temp (°C) | Amplicon Size (bp) |
|---|---|---|---|---|---|---|
| alkB | Broad[1] | AAYACNGCNCAYGARCTNGGNCAYAA | GCRTGRTGRTCNGARTGNCGYTG | 55 | 81 | 548 |
| | 4 – GABds#1 | CTTTTCCTGGTGGGTGCTNAC | GGTTCGAGTATTTGTGNTTNGC | 55 | 81 | 136 |
| | 8 – GABsn#1 | ACCAGACTGGAAACCACNYT | GTTCAATCATCCAGGCCCG | 55 | 84 | 203 |
| | 13 – DSs#1 | GUCGGRAUCUGCCCRUAUG | AACCRAACTGCAATAYYGCAGA | 55 | 75 | 80 |
| | 14 – DSs#2 | GCTGGATGATGATTCCRTTYYT | AACAGCACCAGRTTGCTGA | 55 | 83 | 183 |
| | | | | | | |
| pmoA | Broad[2] | GGNGACTGGGACTTCTGG | GAASGCNGAGAAGAASGC | 59 | 84 | 531 |
| | 1 – Deep Sea #2 | GTTAYTTCAAYTTYTGGGGWTGGAC | ACCTTTCTCRATCATMCKRATGTAYTC | 59 | 84 | 288 |
| | 3 – GABds#1 | GGTGGGTCGTGGTRACACCC | GAACGTMRCCGTGATCCAYTTT | 59 | 84 | 125 |
| | | | | | | |
| c23o | Broad[3] | AAGAGGCATGGGGGCGCACCGGTTCGATCA | CCAGCAAACACCTCGTTGCGGTTGCC | 57 | 85 | 363 |
| | 1 - GABsA | TTYCTGACCGTGTCNATGAARGC | TCGTGCAYGGNNATMAGGTC | 57 | 83 | 128 |
| | 2 - GABsD | GAGCTGACCGARAARGTNACNT | GCATGNCCRACRTCNTCCCA | 57 | 83 | 132 |

1. Primers from Kloos et al., 2006
2. Primers from Holmes et al., 1995
3. Primers from Sei et al., 1999

## 2.4 Detection and quantification of *alkB*, *c23o* and *pmoA* genes via qPCR

SybrGreen-based quantitative real-time PCR (qPCR) was performed with the ABI 7500 Real Time PCR System (Applied Biosystems, Foster City, CA, USA). PCR reactions were a total of 15 µL consisting of 7.5 µL of 2x SensiFAST™ SYBR® Lo-Rox Mix (Bioline, London, UK), 0.1 µM each of forward and reverse primers and 1 µL template DNA (environmental template DNA was at a concentration of 10 ng/µL). The temperature at which acquisition of the fluorescence signal occurred was optimized for each assay (as listed in Table 1) to decrease the influence of primer dimers. Annealing temperatures were also gene specific (Table 1). After an initial activation of 95 °C x 10 min, cycling parameters consisted of 40 cycles of 95 °C for 40 sec, followed by annealing for 30 sec, extension at 72 °C for 60 sec and acquisition for 30 sec. Dissociation curves were run at 95 °C x 15 sec, 60 °C x 60 sec, 95 °C x 30 sec and 60 °C x 15 sec.

DNA standards used in qPCR assays consisted of serial dilutions of purified PCR products derived from GAB sediment samples for each of the gene fragments. Trial PCRs were performed for each primer pair on 10 ng of sediment DNA from Transect 2 and 4. Positive amplifications for each gene were pooled and run on a 1% agarose gel in 1 x TAE buffer. Gene fragments of the correct size were excised with sterile scalpels and gel purified using the Wizard® SV Gel and PCR Clean-Up System (Promega Corporation, Madison, Wisconsin, USA) according to the manufacturers' instructions. The concentration of purified PCR products was measured and serial dilutions made from $10^{-1}$ to $10^{-8}$ ng/µL. Each qPCR run consisted of a five-point standard curve, environmental DNA samples and no template controls run in triplicate. Standard curves were linear ($r^2 > 0.98$) and amplification efficiency was >83. Results were analysed using the ABI 7500 Software package V2.3 (Applied Biosystems, Foster City, CA, USA). Final copy numbers of each gene in environmental samples were calculated assuming 100% DNA extraction efficiency, and the results are expressed as copy numbers per gram of sediment.

## 2.5 Statistical Analysis

All statistical analyses were carried out in SigmaPlot (version 12.5) and R (version 3.3.0). Data were log transformed to meet the assumption of normality, then 2 factor ANOVA tests were run with depth and transect as fixed factors, and $\alpha$=0.05. Pairwise comparisons were carried out via Tukey's HSD test, where required, to identify stations that had significantly different abundances of copies for *alkB, pmoA,* or *c23o*. Differences between samples collected at the same station in different years were determined via a t-test ($\alpha$=0.05). For each sample, the results of the triplicate qPCR runs were averaged for all analyses.

# 3 Results

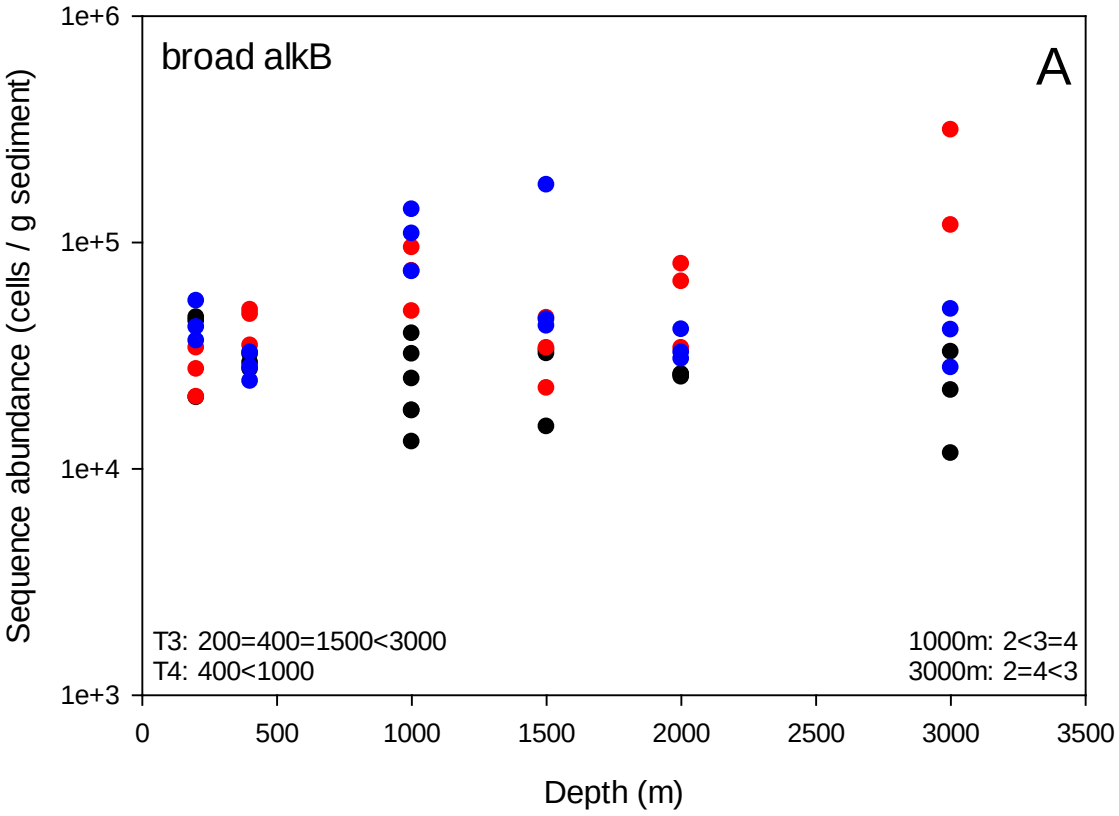
In presenting the results, we assumed that each cell had a single copy of the selected genes, as is normal for bacteria, allowing us to directly convert copies per mL sediment to cells per gram sediment. In general, bacteria with genes encoding enzymes involved in alkane degradation (*alkB*), methane degradation (*pmoA*) and aromatic hydrocarbon degradation (*c23o*) were detected at every site, even though these areas did not have detectable levels of hydrocarbons (Ahmed et al., 2014). As a consequence of the lack of detectable hydrocarbons, the differences between sites (analysed both as depth and transect) were subtle, but occasionally statistically significant. The significance of these differences in abundance with regards to the environmental processes ongoing in the GAB is difficult to interpret.

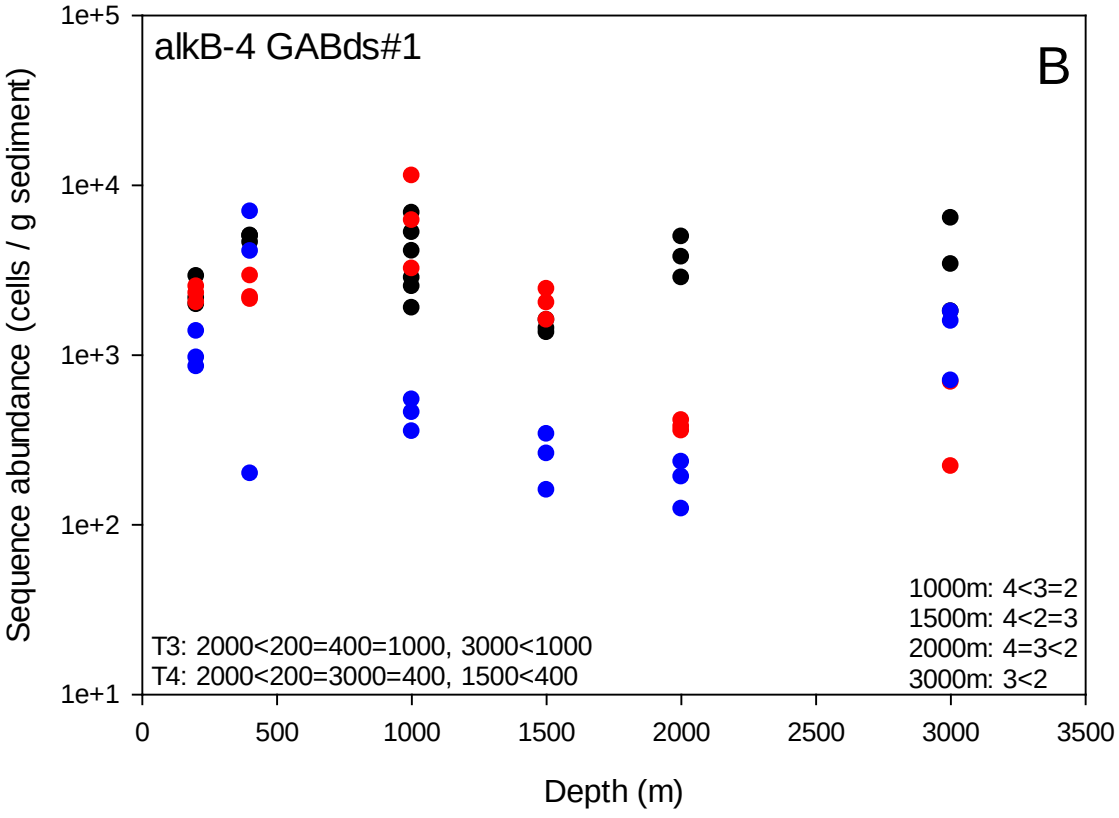
## *3.1 alkB*

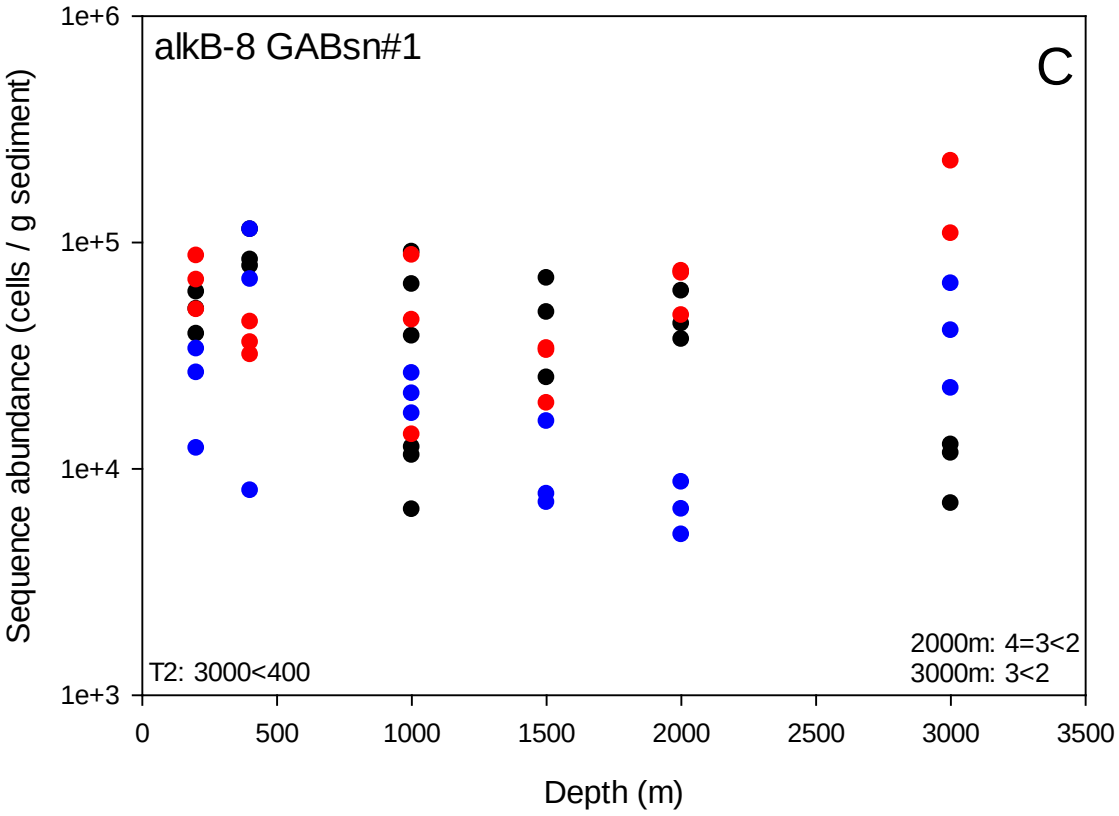
Trends in abundance of *alkB* are shown in Figure 3. Microbial cells with the capacity to degrade alkanes are comparatively abundant, as measured with the broad primers (panel A) and GAB specific primers from clade 8 (panel C). Both primer sets detected between $10^4$ and $10^5$ cells containing the gene for *alkB*. This is slightly different from the original sequencing results, which detected 230,219 sequences in clade 8 (Figure 2A) out of about one million total *alkB* sequences. For reference, a gram of sediment contains $10^8$ - $10^9$ bacterial cells as a generalisation (Bargiela et al., 2015). The *alkB* 4 primer pair, targeting a subgroup of those targeted by the *alkB* 8 pair, also showed abundant copies of this gene, as shown in Figure 3B. The original sequencing efforts detected 105,807 copies of *alkB* that aligned to this clade. By contrast, the primers designed based on sequences that were abundant in other regions as well as in the GAB, *alkB* 13 and *alkB* 14, detected far fewer copies of this gene in the same samples (Figure 3 D and E). However, the previously conducted sequencing results indicated that *alkB* 13 and 14 had abundances comparable to *alkB* 8 and *alkB* 4, respectively (Figure 2A). For all primers, the number of gene copies was significantly influenced by an interaction between transect and depth (ANOVA: $F_{10,38}>2.16$, $p<0.042$).

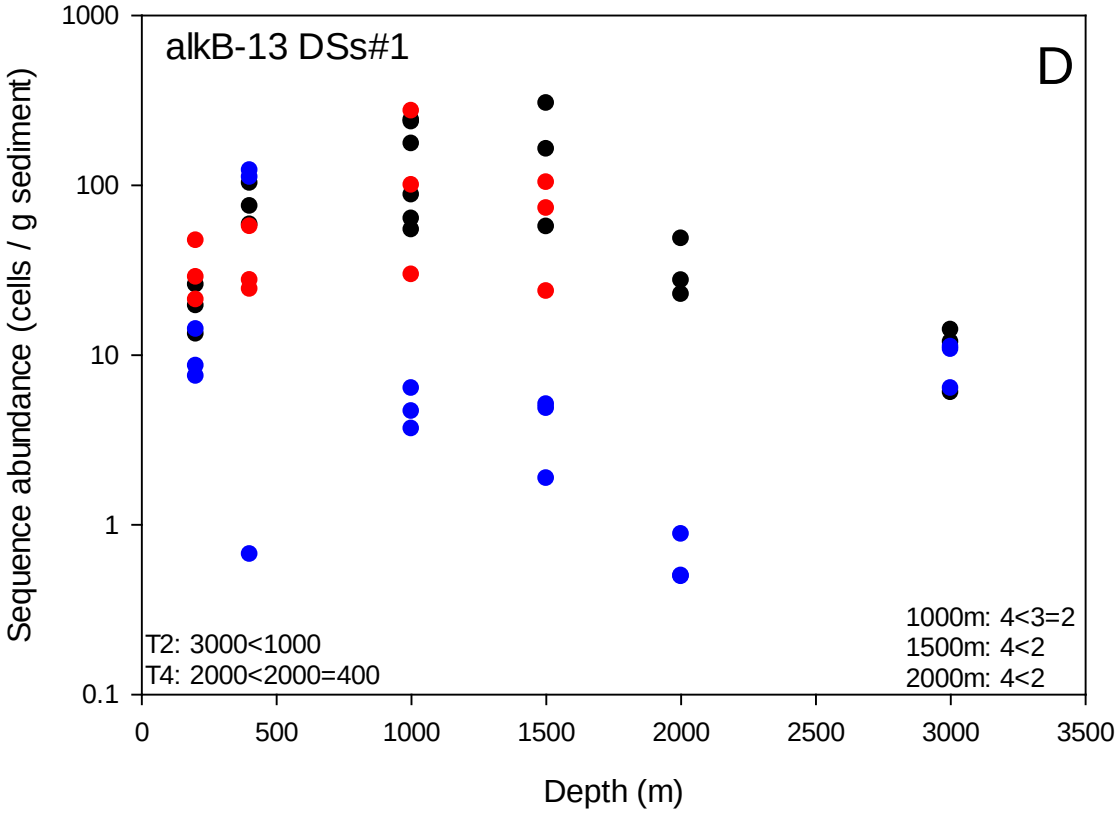
Comparisons between results from high throughput sequencing and the qPCR findings need to be interpreted with caution. The high throughput sequencing results are based on a generic PCR, which is prone to bias. PCR biases have been reported to cause an order or more of magnitude over- or underrepresentation of relative abundances of subgroups (see Acinas et al., 2005; Engelbrektson et al., 2010; O'Donnell et al., 2016). Quantitative PCR based numbers are generally more reliable, especially those focusing on smaller, more closely related groups. It should also be noted that the qPCR methods developed here are aimed at detecting substantial changes in the abundance of hydrocarbon degrading bacterial clades in case of increasing hydrocarbon loads. While a qPCR method can also be inaccurate, reporting up to 2-3x higher or lower abundance than the actual abundance, this inaccuracy is consistent and would thus not impact on the quantification of changes.

Although small differences were occasionally detected between different stations, no statistically significant differences were detected between samples taken at transect 2, 1,000 m depth in different years (broad *alkB:* $t_4 = 2.088$, $p = 0.105$; *alkB4:* $t_4 = 1.608$, $p = 0.183$; *alkB8:* $t_4 = 2.562$, $p = 0.0625$; *alkB13*: $t_4 = 0.616$, $p = 0.571$; *alkB14*: $t_4 = 0.616$, $p = 0.293$).


broad alkB
A
Sequence abundance (cells / g sediment)
1e+6
1e+5
1e+4
1e+3
T3: 200=400=1500<3000
T4: 400<1000
1000m: 2<3=4
3000m: 2=4<3
0
500
1000
1500
2000
2500
3000
3500
Depth (m)


alkB-4 GABds#1
B
Sequence abundance (cells / g sediment)
1e+5
1e+4
1e+3
1e+2
1e+1
T3: 2000<200=400=1000, 3000<1000
T4: 2000<200=3000=400, 1500<400
1000m: 4<3=2
1500m: 4<2=3
2000m: 4=3<2
3000m: 3<2
0
500
1000
1500
2000
2500
3000
3500
Depth (m)


alkB-8 GABsn#1
C
Sequence abundance (cells / g sediment)
1e+6
1e+5
1e+4
1e+3
T2: 3000<400
2000m: 4=3<2
3000m: 3<2
0
500
1000
1500
2000
2500
3000
3500
Depth (m)


alkB-13 DSs#1
D
Sequence abundance (cells / g sediment)
1000
100
10
1
0.1
T2: 3000<1000
T4: 2000<2000=400
1000m: 4<3=2
1500m: 4<2
2000m: 4<2
0
500
1000
1500
2000
2500
3000
3500
Depth (m)

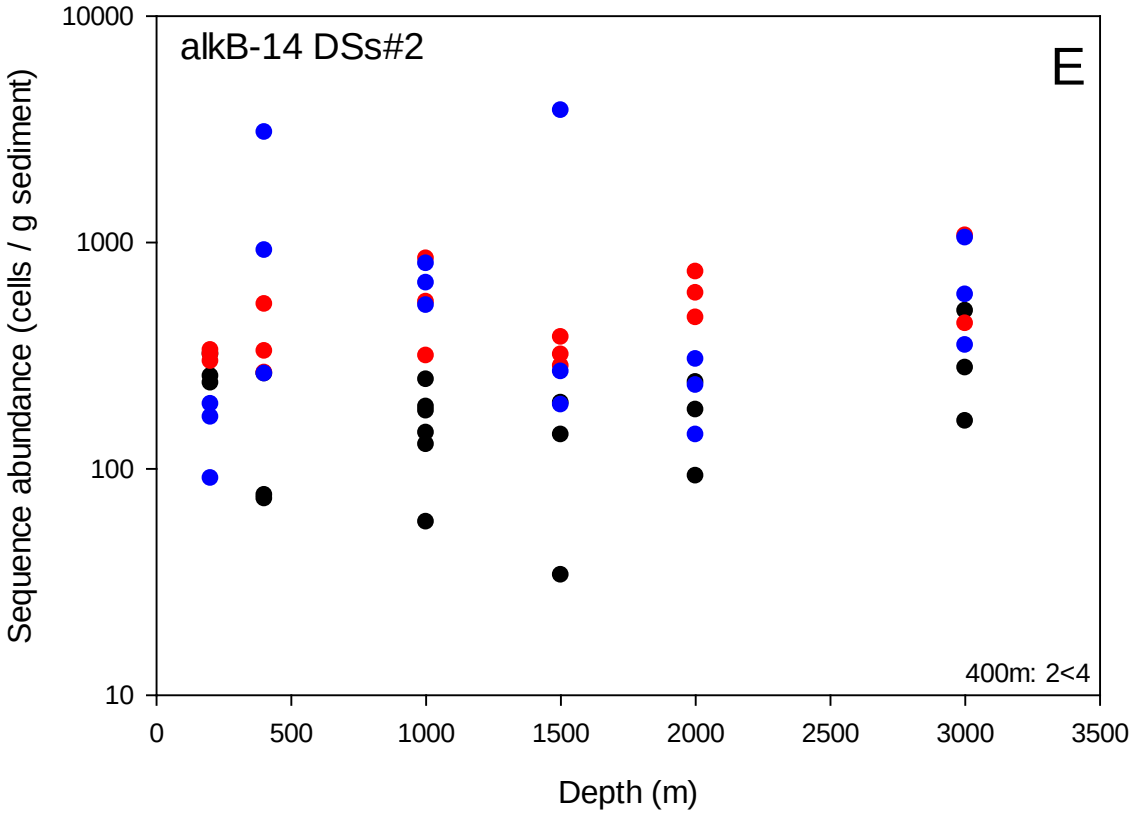
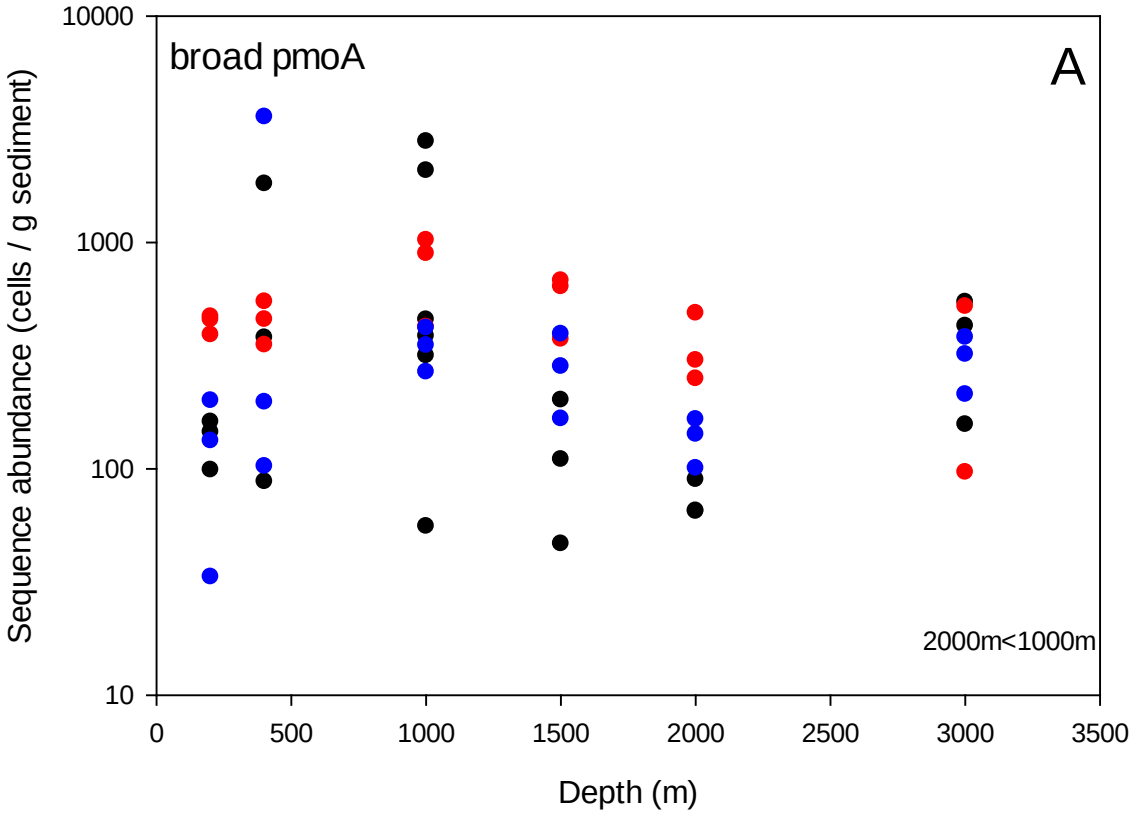
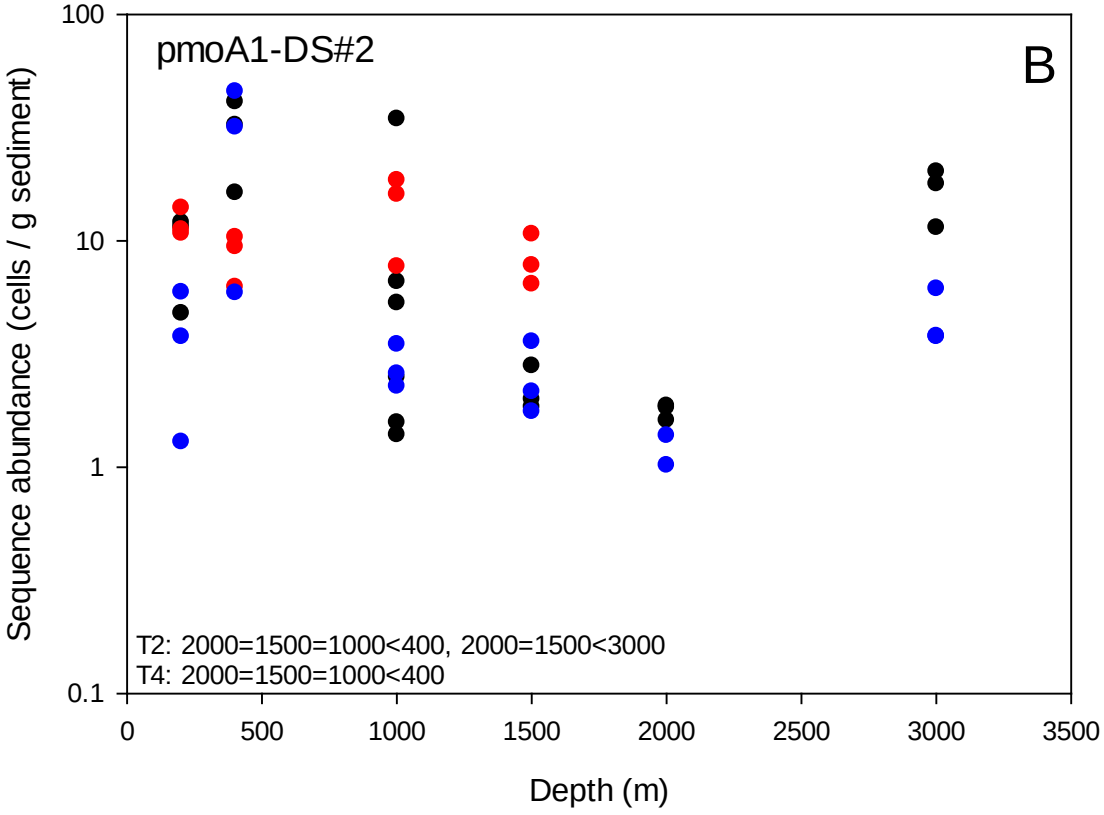



**Figure 3. Number of copies of *alkB* detected with the different primer sets in sediment samples collected from the GAB. Transect 2 is shown in black, transect 3 in red, and transect 4 in blue. Inequalities in the bottom left of each panel indicate statistically significant differences between depths within a transect, while those in the bottom right indicate differences between transects within a depth (ANOVA followed by a Tukey's HSD pairwise comparison, p<0.05).**

## *3.2 pmoA*

The abundance of cells containing the gene sequence for *pmoA* is shown in figure 4. In comparison to *alkB*, cells containing *pmoA* were much less abundant, with roughly $10^2$ to $10^3$ cells per gram sediment measured if the broad *pmoA* (Figure 4A) or the primers designed for clade *pmoA* GABDS#1 (*pmoA*3, Figure 4C) are used. By contrast, the primers for the *pmoA* DS#1 clade (*pmoA*1) detected between 1 to 100 cells per gram sediment. In the original sequencing efforts, members of the GABDS#1 clade were more abundant (83525 copies) than members of the DS#1 clade (27859 copies) (Figure 2B). For the broad *pmoA* primers (figure 4 A), the number of gene copies detected varied between transects (ANOVA: $F_{2,10}$=3.32, p=0.047) and depths ($F_{5,10}$=3.24, p=0.016), although there was no interaction ($F_{10,34}$=0.88, p=0.56). For the primers specific to a single clade (Figure 4B and 4C), there were interactions between transect and depth (ANOVA: $F_{10,38}$>3.03, p<0.012).


broad pmoA
A
Sequence abundance (cells / g sediment)
10000
1000
100
10
0
500
1000
1500
2000
2500
3000
3500
Depth (m)
2000m<1000m


pmoA1-DS#2
B
Sequence abundance (cells / g sediment)
100
10
1
0.1
0
500
1000
1500
2000
2500
3000
3500
Depth (m)
T2: 2000=1500=1000<400, 2000=1500<3000
T4: 2000=1500=1000<400

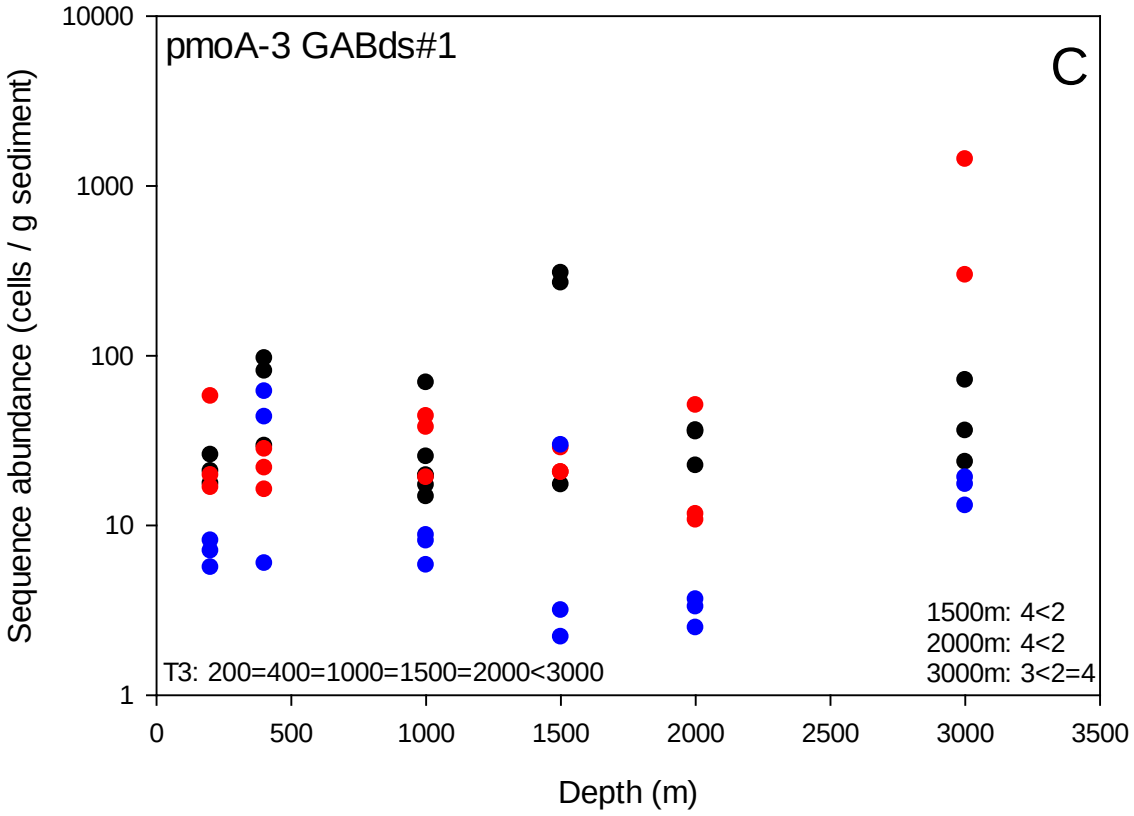
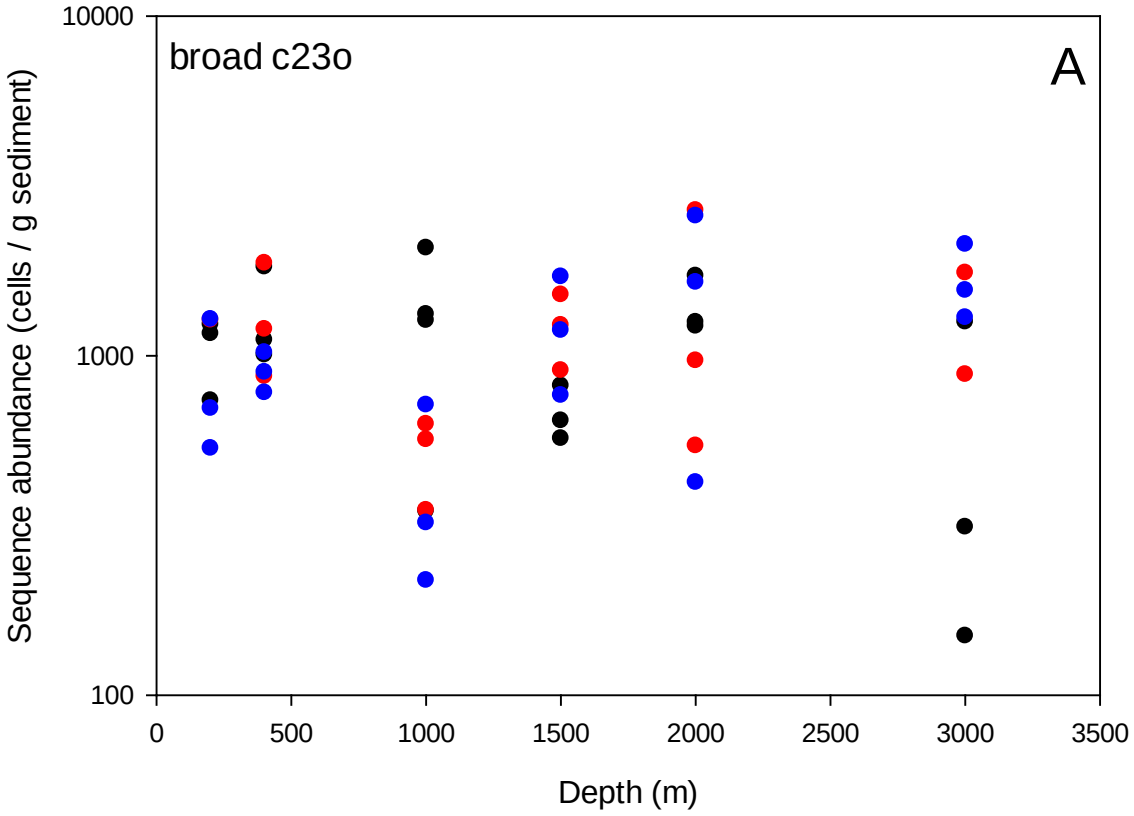


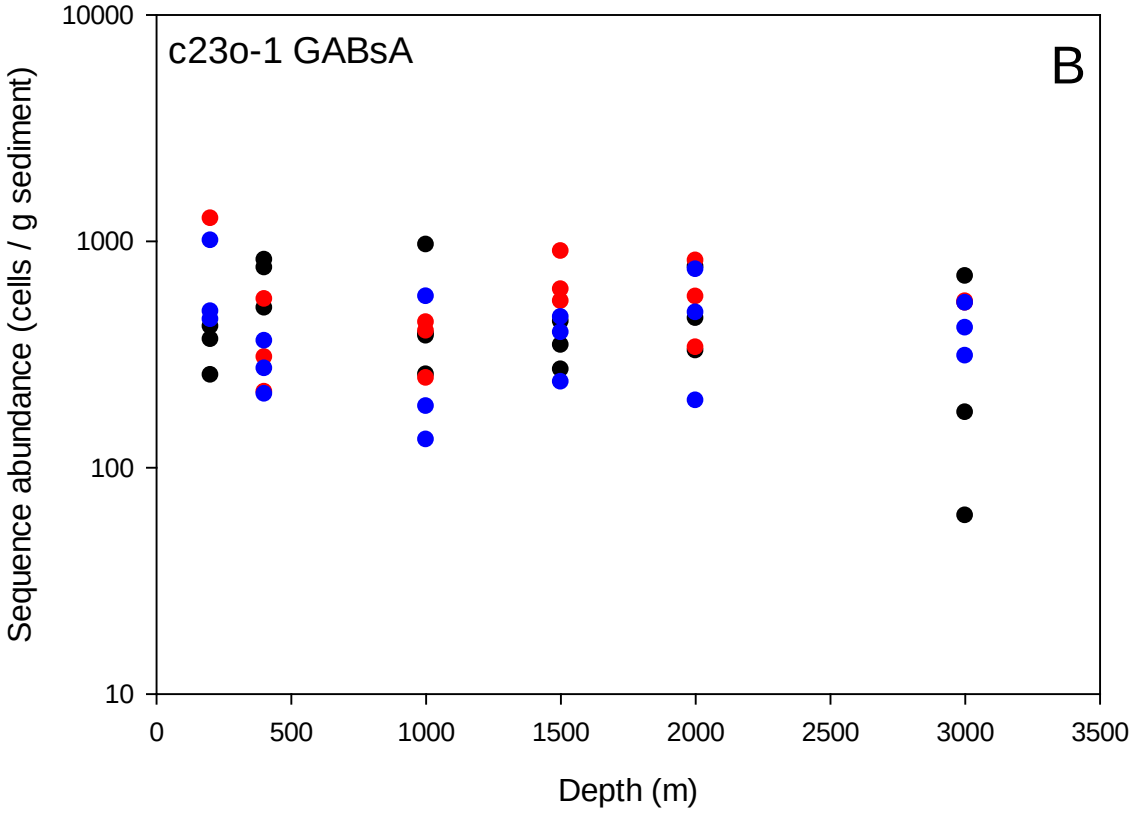

**Figure 4. Number of copies of *pmoA* detected with the different primer sets in sediment samples collected from the GAB. Transect 2 is shown in black, transect 3 in red, and transect 4 in blue. Inequalities in the bottom left of each panel indicate statistically significant differences between depths within a transect, while those in the bottom right indicate differences between transects within a depth except for panel A, where differences are between the main effect of depth (ANOVA followed by a Tukey's HSD pairwise comparison, p<0.05).**

Although small differences in *pmoA* abundance were detected at different stations, no statistically significant differences were detected between samples taken at transect 2, 1,000 m depth from the different voyages (t-test, broad *pmoA:* $t_4 = 1.531$, p = 0.201; *pmoA*1*:* $t_4 = 1.531$, p= 0.484; *pmoA*3*:* $t_4 = 1.193$, p = 0.299).

## *3.3 c23o*

The abundance of cells with a copy of the gene for *c23o* is shown in Figure 5. Like *pmoA*, cells containing *c23o* were less abundant than those containing *alkB,* with roughly 1,000 cells detected per gram sediment as measured by all primer pairs (Figures 5A-5C). Both clades had equivalent abundance as measured through both the sequencing efforts (1.15 M and 950,000, respectively, Figure 2C) and the qPCR assays (Figure 5B and 5C). There were detectable differences in abundance due to the interaction between transect and depth for c23o and c23o2 (ANOVA: $F_{10,34}>2.1$, p<0.046), but no significant effects for c23o1.


broad c23o
A
Sequence abundance (cells / g sediment)
10000
1000
100
0
500
1000
1500
2000
2500
3000
3500
Depth (m)


c23o-1 GABsA
B
Sequence abundance (cells / g sediment)
10000
1000
100
10
0
500
1000
1500
2000
2500
3000
3500
Depth (m)

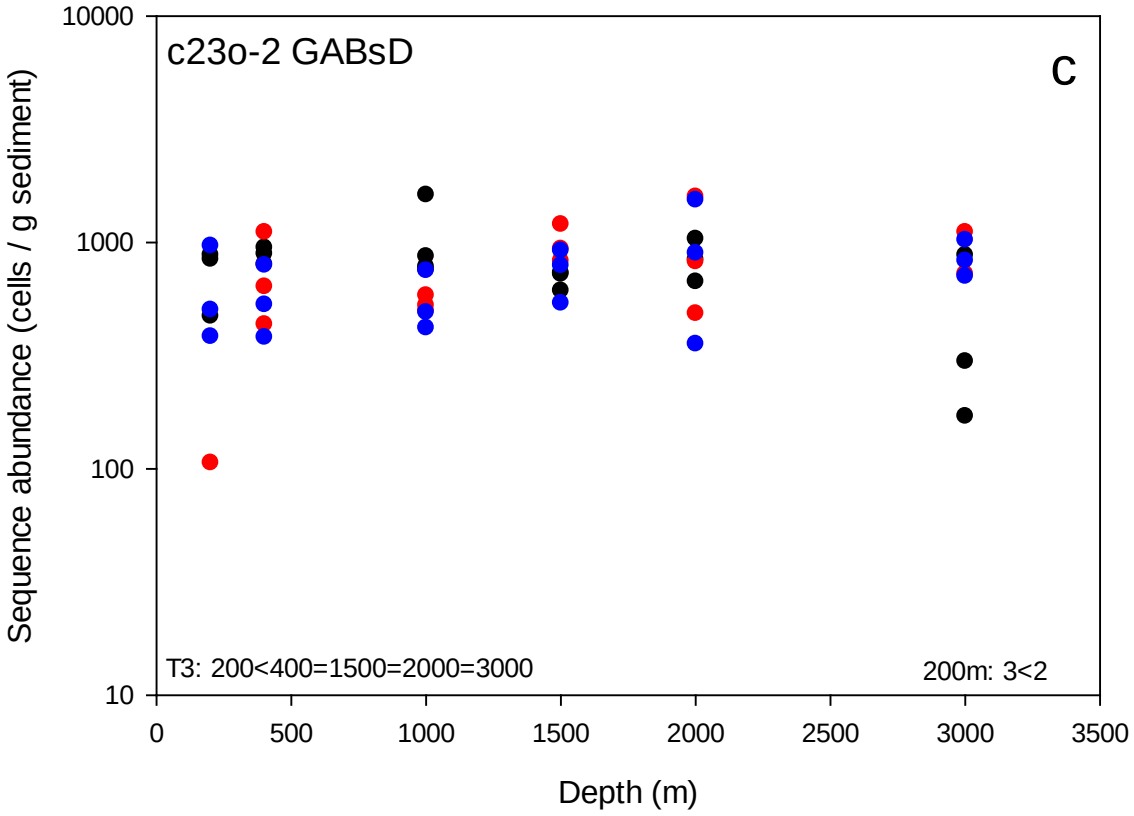



**Figure 5. Number of copies of *c23o* detected with the different primer sets in sediment samples collected from the GAB. Transect 2 is shown in black, transect 3 in red, and transect 4 in blue. Where shown, inequalities in the bottom left of each panel indicate statistically significant differences between depths within a transect, while those in the bottom right indicate differences between transects within a depth (ANOVA followed by a Tukey's HSD pairwise comparison, p<0.05).**

Despite measuring small changes in *c23o* abundance at different stations, no statistically significant differences (t-test, broad *c23o*: $t_4 = -0.704$, $p = 0.520$; *c23o-1*: $t_4 = 0.854$, $p = 0.441$; $t_4 = -0.659$, $p = 0.546$) were detected between samples taken at transect 2, 1,000 m depth in 2013 and 2015.

# 4 Discussion

Our goal was to develop a high throughput functional genomics based screening tool to measure changes in the abundance of hydrocarbon degrading bacteria. The tool has application to monitoring routine discharges during oil and gas production and for evaluating the persistence of oil in the event of an unplanned release.

We focussed our efforts on three genes in three different functional pathways: 1) *alkB*, which is the first step in the breakdown of alkanes (Abbasian et al., 2015), 2) *pmoA*, which is involved in aerobic methane oxidation (Inagaki et al., 2004), and 3) *c23o* which is the conserved ring breaking step in the degradation of different polycyclic aromatic hydrocarbons (PAH) (Meyer et al., 1999). Because of the high levels of diversity in gene sequences from the GAB, we designed primers from the literature as well as from sequencing results we obtained in an earlier portion of this study (Hook et al., 2016a). Using these primers, we were able to identify genes in the chosen oil degrading pathways at every site, indicating the presence of bacteria capable of hydrocarbon degradation. In our analyses, we assumed one gene copy per cell, allowing us to covert between our copy number per ng/DNA added to the reaction to cells/g sediment. The genes, and in turn the bacteria carrying them, were present at $10^1$ to $10^5$ cells per gram sediment, making up $1/10^8$ to $1/10^3$ of the total bacterial community. Previous studies have also identified a low but consistent presence of hydrocarbon degrading bacteria even in environments without a local hydrocarbon source (Atlas, 1995b). While $1/10^8$ to $1/10^3$ is a low proportion of the bacterial community, it should be noted that following the *DeepWater Horizon* wellhead blowout, the bacteria that increased in abundance had also been rare previously (Kleindienst et al., 2015). By using primer pairs targeted to different clades, in the case of increased hydrocarbon load, we will be able to better follow the progress of, and differences in, the types of alkane, methane, and aromatic compound degraders that are abundant as oil weathers.

Using these primers, we did measure some statistically significant differences in the numbers of bacterial cells carrying these marker genes at different stations. Although the differences in abundance at different stations are statistically significant, the magnitude of the differences is subtle. The small magnitude of differences is not surprising because no source of natural hydrocarbon leakages has been detected in the study area. We could not determine the significance, if any, of these small differences in the abundance of bacterial cells to ecological processes in the GAB with the geochemical data currently available. As these differences are idiosyncratic and vary between the different genes, and even between different primers for the same gene, there were no consistent spatial differences in the abundances of genes related to hydrocarbon degradation. It is thus unlikely that they are related to the presence of a specific source of hydrocarbons, and they may instead be related to some other aspect of environmental variation.

This study was conducted using three different sets of primers: Universal primers whose design was obtained from the literature (Kloos et al., 2006, Holmes et al., 1995, Sei et al., 1999), regionally-specific primers we designed based on our previous work with Illumina based next generation sequencing, and primers designed for clades that were made up of sequences we obtained from the literature, as well as sequences derived from our previous Illumina based assays. The universal primers typically detected the most cells, as would be expected since they target a larger amplicon and were used to select the region targeted for sequencing. However, for *pmoA*, these broad primers lacked the sensitivity to detect subtle changes in the abundance of cells. The regionally specific primers could detect nearly as many cells as the universal primers, despite being designed for a single clade, but detected more statistically significant changes in abundance (Figures 4A-C). Their use shows promise to develop a monitoring tool with a greater dynamic range (i.e. applicable at both higher and lower abundances) than the universal primers. Also, because of the shorter amplicon size, these assays are likely to perform better if used in a high throughput scenario.

Cells containing *alkB* were more abundant than cells containing either *c23o* or *pmoA*. Water samples from the Gulf of Mexico collected outside the hydrocarbon plume had a greater relative abundance of *alkB* than *c23o*, however, *pmoA* had equivalent abundance to *alkB* in those samples (Rivers et al., 2013).

As detailed in our previous report (Hook et al., 2016a), amplicon based next generation sequencing was a highly effective approach to determine the diversity of gene sequences in the GAB. However, we would not recommend using this approach in either routine monitoring or in oil spill response, as it has higher costs per sample, the assays take longer to run, and the data takes longer to analyse. By contrast, the qPCR assays can be run very quickly and cheaply, and can be coupled with a microfluidics device such as the Fluidigm© system (https://www.fluidigm.com/products/biomark-hd-system), which can run 96 primer pairs against 96 samples in a single afternoon, to make them high throughput. The regionally specific primers, designed using outputs of our previous sequencing work, provide both the sensitivity and dynamic range to be utilised in any monitoring efforts that may be needed in the GAB.

# References


Abbasian, F., Lockington, R., Mallavarapu, M., Naidu, R., 2015. A Comprehensive Review of Aliphatic Hydrocarbon Biodegradation by Bacteria. Applied Biochemistry and Biotechnology 176, 670-699.

Acinas, S.G., Sarma-Rupavtarm, R., Klepac-Ceraj, V., Polz, M.F., 2005. PCR-induced sequence artifacts and bias: Insights from comparison of two 16S rRNA clone libraries constructed from the same sample. Applied and Environmental Microbiology 71, 8966-8969.

Ahmed M, Ross A, Stalvies C, Armand S, Fuentes D, Gong S, Sestak S. 2014. Organic geochemical study of seawater, seabed sediment and headspace gas samples to monitor baseline hydrocarbon levels around BP permits in the Great Australian Bight, Australia. Report to the BP-GAB alliance.

Atlas, R.M., 1995a. Bioremediation of petroleum pollutants. International Biodeterioration & Biodegradation. 35, 317-327.

Atlas, R.M., 1995b. Petroleum biodegradation and oil spill bioremediation. Marine pollution bulletin 31, 178-182.

Atlas, R.M., Hazen, T.C., 2011. Oil biodegradation and bioremediation: a tale of the two worst spills in U.S. history. Environmental science & technology 45, 6709-6715.

Bargiela, R., Mapelli, F., Rojo, D., Chouaia, B., Tornés, J., Borin, S., et al. (2015). Bacterial population and biodegradation potential in chronically crude oil-contaminated marine sediments are strongly linked to temperature. Nature Publishing Group, 5, 11651. http://doi.org/10.1038/srep11651.

Chakraborty, R., Borglin, S.E., Dubinsky, E.A., Andersen, G.L., Hazen, T.C., 2012. Microbial Response to the MC-252 Oil and Corexit 9500 in the Gulf of Mexico. Frontiers in microbiology 3, 357.

Dubinsky, E.A., Conrad, M.E., Chakraborty, R., Bill, M., Borglin, S.E., Hollibaugh, J.T., Mason, O.U., Piceno, Y.M., Reid, F.C., Stringfellow, W.T., Tom, L.M., Hazen, T.C., Andersen, G.L., 2013. Succession of Hydrocarbon-Degrading Bacteria in the Aftermath of the Deepwater Horizon Oil Spill in the Gulf of Mexico. Environmental science & technology 47, 10860-10867.

Engelbrektson, A., Kunin, V., Wrighton, K.C., Zvenigorodsky, N., Chen, F., Ochman, H., Hugenholtz, P., 2010. Experimental factors affecting PCR-based estimates of microbial species richness and evenness. Isme Journal 4, 642-647.

Hazen, T.C., Dubinsky, E.A., DeSantis, T.Z., Andersen, G.L., Piceno, Y.M., Singh, N., Jansson, J.K., Probst, A., Borglin, S.E., Fortney, J.L., Stringfellow, W.T., Bill, M., Conrad, M.E., Tom, L.M., Chavarria, K.L., Alusi, T.R., Lamendella, R., Joyner, D.C., Spier, C., Baelum, J., Auer, M., Zemla, M.L., Chakraborty, R., Sonnenthal, E.L., D'Haeseleer, P., Holman, H.Y., Osman, S., Lu, Z., Van Nostrand, J.D., Deng, Y., Zhou, J., Mason, O.U., 2010. Deep-sea oil plume enriches indigenous oil-degrading bacteria. Science 330, 204-208.

Head, I.M., Jones, D.M., Roling, W.F., 2006. Marine microorganisms make a meal of oil. Nature reviews. Microbiology 4, 173-182.

Hook, S., van de Kamp, J., Williams, A., Tanner, J., Bodrossy, L. 2016a. Spatial distribution and diversity in hydrocarbon degrading microbes in the Great Australian Bight I: Functional Gene Distribution. CSIRO report to the Great Australian Bight Research Program.

Hook, S., van de Kamp, J., Williams, A., Tanner, J., Bodrossy, L. 2016b. Spatial distribution and diversity in hydrocarbon degrading microbes in the Great Australian Bight II: Initial analysis of bacterial and archaeal community structure. CSIRO report to the Great Australian Bight Research Program.

Kleindienst, S., Grim, S., Sogin, M., Bracco, A., Crespo-Medina, M., Joye, S.B., 2015. Diverse, rare microbial taxa responded to the Deepwater Horizon deep-sea hydrocarbon plume. The ISME journal.

Kloos, K, Munch, JC, & Schloter, M (2006) A new method for the detection of alkane-monooxygenase homologous genes (alkB) in soils based on PCR-hybridization. Journal of Microbiological Methods, 66: 486-496. doi: 10.1016/j.mimet.2006.01.014.

Lamendella, R., Strutt, S., Borglin, S., Chakraborty, R., Tas, N., Mason, O.U., Hultman, J., Prestat, E., Hazen, T.C., Jansson, J.K., 2014. Assessment of the Deepwater Horizon oil spill impact on Gulf coast microbial communities. Frontiers in microbiology 5, 130.

Lu, Z., Deng, Y., Van Nostrand, J.D., He, Z., Voordeckers, J., Zhou, A., Lee, Y.-J., Mason, O.U., Dubinsky, E.A., Chavarria, K.L., Tom, L.M., Fortney, J.L., Lamendella, R., Jansson, J.K., D'Haeseleer, P., Hazen, T.C., Zhou, J., 2012. Microbial gene functions enriched in the Deepwater Horizon deep-sea oil plume. Isme Journal 6, 451-460.

Mackay, I.M., 2004. Real-time PCR in the microbiology laboratory. Clinical Microbiology and Infection 10, 190-212.

Meyer, S., Moser, R., Neef, A., Stahl, U., Kampfer, P., 1999. Differential detection of key enzymes of polyaromatic-hydrocarbon-degrading bacteria using PCR and gene probes. Microbiology-Uk 145, 1731-1741.

O'Donnell, J.L., Kelly, R.P., Lowell, N.C., Port, J.A., 2016. Indexed PCR Primers Induce Template-Specific Bias in Large-Scale DNA Sequencing Studies. PloS one 11.

Rivers, A.R., Sharma, S., Tringe, S.G., Martin, J., Joye, S.B., Moran, M.A., 2013. Transcriptional response of bathypelagic marine bacterioplankton to the Deepwater Horizon oil spill. The ISME journal 7, 2315-2329.

RSC (2015) *The Behaviour and Environmental Impacts of Crude Oil Released into Aqueous Environments*. Royal Society of Canada, Ottawa, Ontario, Canada. <https://rsc-src.ca/en/ expert-panels/rsc-reports/behaviour-and-environmental-impacts-crude-oil-released-into-aqueous>.

Sei, K, Asano, K, Tateishi, N, Mori, K, Ike, M, & Fujita, M (1999) Design of PCR primers and gene probes for the general detection of bacterial populations capable of degrading aromatic compounds via catechol cleavage pathways. Journal of Bioscience and Bioengineering, 88: 542-550. doi: 10.1016/S1389-1723(00)87673-2.


bp


CSIRO


SOUTH AUSTRALIA
Government
of South Australia
Primary Industries
and Regions SA


SARDI
SOUTH AUSTRALIAN
RESEARCH AND
DEVELOPMENT
INSTITUTE


SUB CRUCE LUMEN
THE UNIVERSITY
of ADELAIDE


Flinders
UNIVERSITY