

**A Metagenomics & Phylogenetic Analysis of Woods Hole  
Passage Microorganisms**

Gail P. Ferguson  
University of Edinburgh (UK)

**ABSTRACT**

Within the Oceans, microorganisms are subject to an array of different stresses such as nutrient deprivation, cold and high-pressure stresses. Yet despite this, a milliliter of seawater contains millions of microorganisms. However, less than 1% of these microorganisms have been cultured. Thus, my project aimed to investigate whether novel genes could be isolated from seawater microorganisms, which could confer phenotypes on an *E. coli* host strain. To do this, I constructed a fosmid library in *E. coli* using a genomic preparation of seawater microorganisms from Woods Hole Passage. The fosmid library contained around 44,000 clones, thus covering approx.  $1.8 \times 10^6$  Kb of DNA from seawater microorganisms. A 16S rRNA gene library was also constructed using the seawater microorganism DNA and 87 clones were sequenced and analyzed by ARB. These data showed that the fosmid library is likely to contain a diverse array of DNA from uncultured and cultured bacteria. Thus, in the future, screening of the fosmid library will be conducted to identify novel genes, such as those involved in exopolysaccharide synthesis or antibiotic production. Additionally the fosmid library will also be used to identify clones containing defined genes such as proteorhodopsin.

**PROJECT BACKGROUND**

Despite over 70% of the Earth's surface being covered by ocean, our understanding of the microbial diversity within the oceans is poorly understood. Intriguingly, microorganisms can grow and/or survive in extreme ocean conditions such as elevated hydrostatic pressures and low temperatures. For example, microorganisms in the deepest part of the ocean, which is ~10,000m, will need to cope with an average hydrostatic pressure of ~1000 atmospheres (equivalent to ~1000 bar or 100 MPa) and an average water temperature of ~3°C. Additionally, hydrothermal vents are also present within the oceans. These vents are at extremely high temperatures (350°C) and provide a unique “warm” niche within the cold ocean environment.

The shrimp, *Rimicaris exoculata*, predominates around the warm vent (also known as sulphide chimneys) waters of the Mid-Atlantic Ridge where it ingests sulfide particles (1, 2). Electron microscopy studies have revealed dense populations of bacteria on the surface and inside the shell of *R. exoculata* and 16S rRNA analysis suggests that the dominant population of bacteria on the shrimp form a distinct branch of ε-proteobacteria (1). This finding was unusual since previous hydrothermal vent studies found that γ-proteobacteria usually predominate. However, to date, there has been little success cultivating pure bacterial cultures from *R. exoculata* (1).

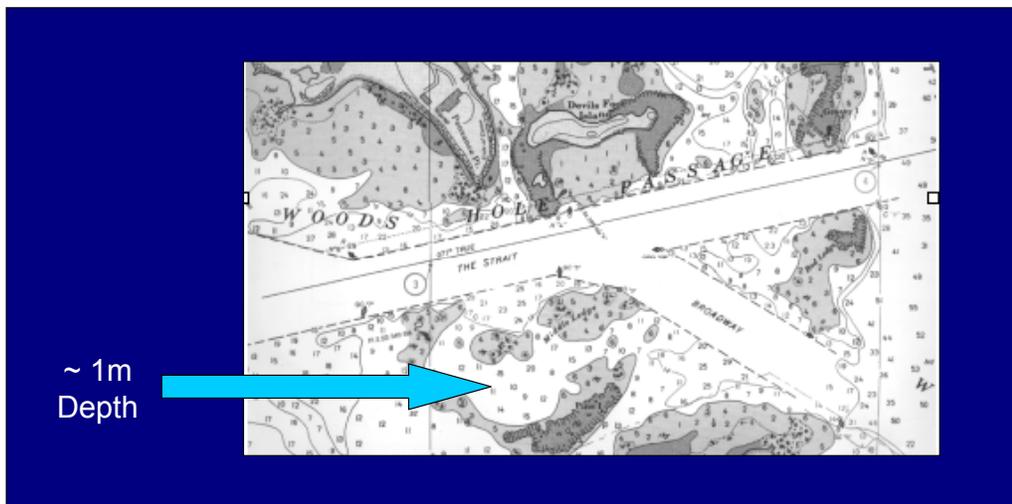
One of the major problems in working with deep-sea microorganisms is in obtaining enough

**Table 1: Details on the isolation and storage of the deep-sea shrimp and water sample.**

Sample	Ocean (Location)	Depth (m)	Water Temperature	Date	Special Details	Sample Storage
<i>R. exoculata</i> (~3 Shrimp)	Mid-Atlantic Ridge (Snake Pit)	3486	Warm/Hot	6/19/93	Hydrothermal Vent, Frozen Immediately	-20°C
Water (1.8 L)	North Atlantic Ocean	4280	Cold	1996	Transferred Immediately into Sterile Chilled Bottle	3°C

material for research. However, I was exceptionally lucky in obtaining two deep-sea samples (Table. 1) from Dr Carl Wirsén and Stephen Molyneux (Biology Dept., WHOI). Although, I had obtained these

Fig. 1. Collection of Sea Water for Project



samples, I decided to focus my three-week project investigating the metagenomics and phylogenetics of a seawater sample obtained from Woods Hole Passage (Fig. 1), thus enabling myself to gain the necessary experience in environmental techniques, on a widely accessible source of material. Additionally, the

results of the sea water analysis may also be interesting in its own right, since less than 1% of seawater microorganisms have been cultured and thus a metagenomic investigation could reveal novel genes in some of these microorganisms.

**RESEARCH QUESTION**

Thus, the aim of my 3-week research project was to determine whether novel genes could be isolated from seawater microorganisms, which could alter the phenotype on an *E. coli* host strain.

## SPECIFIC OBJECTIVES

- 1. Metagenomics.** To construct a fosmid library from seawater microorganism, in an *E. coli* host strain, & then screen the library for fosmids, which alter one or more phenotype of the host strain.
- 2. Phylogenetics.** To investigate the bacterial diversity of the seawater microorganisms used in construction of the fosmid library.

## MATERIALS AND METHODS

### Isolation of seawater microorganisms, genomic DNA extraction and purification.

Sea water (38 L) was collected from a depth of ~ 1 meter from the ocean (Fig. 1) and then filtered using a gas powered filtering device (Millipore-YY30 142 36). The seawater was initially filtered through 3 layers of Whatman Filter No 1 (pore size 11  $\mu\text{m}$ ), to remove large particles. The seawater microorganisms were then collected on two, 0.1  $\mu\text{m}$  pore size filters (Millipore Durapore-VVLP14250). The filters were used directly for genomic DNA extraction following a previously published CTAB method (3) with slight modifications. After the additional of isopropanol, the DNA was precipitated overnight at room temperature. The DNA pellet was obtained by centrifugation in a bench top centrifuge at 4000 rpm for 1 h. After solubilization of the DNA pellet in TE buffer (pH 7.5), the DNA was further purified by the Wizard DNA clean up system (Promega-A7280).

### Construction and analysis of the fosmid library.

A fosmid library was constructed with the sea water genomic DNA using the Epicentre kit (CCFOS110). After shearing and for size separation, the DNA was analyzed on a 1% (w/v) large agarose gel, run overnight at 36 V. Since exposure of the DNA to UV is thought to dramatically reduce the efficiency of the packaging reaction, the precise position of the size-fractionated and end-repaired DNA on the gel was determined using the method involving staining the standard 36 Kb DNA marker and then using this to determine the appropriate location of the sample DNA on the gel (refer to kit protocol). To validate the size of inserts in the fosmid library, the fosmids from several induced clones (5ml, induced for 6 h with arabinose inducing solution) were extracted using a standard alkaline lysis procedure and digested with Not I. The digests were then analyzed by pulse-field gel electrophoresis (PFGE). Undigested fosmid DNA was also sized by PFGE using the BAC-Tracker Supercoiled DNA Ladder (Epicentre-BT010950). The clones from the libraries were combined and stored as glycerol stocks at  $-80^{\circ}\text{C}$ .

### Bacterial 16S rRNA gene analysis

The purified genomic DNA was used for PCR following a standard procedure (buffer contained 1.5 mM  $\text{Mg}^{2+}$ ;  $50^{\circ}\text{C}$  annealing) and the bacterial specific 16S rRNA gene primers (8F and 1492R). The PCR reaction was then used directly to clone into the TOPO vector (Invitrogen) and the plasmids extracted from 87 clones and sequenced. The sequences were then analyzed by ARB.

## RESULTS

**Metagenomics-Construction of a seawater microorganism fosmid library in *E. coli*.** To investigate whether there are novel genes in microorganisms from the ocean, which could potentially confer novel phenotypes upon an *E. coli* host strain, a fosmid library was constructed. To do this, the genomic DNA was isolated and purified from a filter, which contained microorganisms from 38 L of seawater. The purified DNA was then sheared to generate large fragments (~40 Kb), the sheared DNA was then end-repaired and the higher molecular weight fragments purified and quantified. Ultimately, 38 L of seawater

**Table 2. Packaged Phage Titering**

Insert DNA Sample Used in Ligation Reaction With pCC1FOS	Titer of Packaged Phage ( $\times 10^4 \text{ cfu ml}^{-1}$ )*
Marine Microorganism Size Fractioned End Repaired Genomic DNA Sample	28
36 Kb Size Standard DNA	560
No Insert DNA (Control)	0

\*Titering was based on the ability of the *E. coli* host to grow on chloramphenicol-containing agar.

gave rise to approx. 3.8  $\mu\text{g}$  of higher molecular weight DNA fragments. The repaired DNA was then ligated into the fosmid vector (used 0.4  $\mu\text{g}$  of DNA), packaged by phage (the phage packages fosmids with inserts ~40 Kb) and then transduced into an *E. coli* host strain (for further details refer to the Epicentre protocol). The phage titer was determined prior to plating the entire library (Table 2). In addition to using my prepared DNA sample, a control insert of 36 Kb (supplied in the kit) and a no insert control were included in the

ligation, packaging and transduction experiments (Table 2). Since, in the absence of insert, the plasmid is too small to be packaged by the phage, whereas the 36 Kb insert control ligation, produces fragments of the optimal size for phage packaging, the results in Table 2 were as expected. Additionally, since our sample DNA contained a mixture of DNA fragment sizes, I was not surprised that the titer of phage was reduced relative to the 36 Kb DNA control ligation, since not all fragments will be the optimum size to be packaged. Based on the initial titering of the packaged phage, a fosmid library of ~44,000 clones was constructed. If each clone contains ~40 Kb of DNA, then these data suggest that I have approx.  $1.76 \times 10^6$  Kb of seawater microorganisms DNA in my clone library.

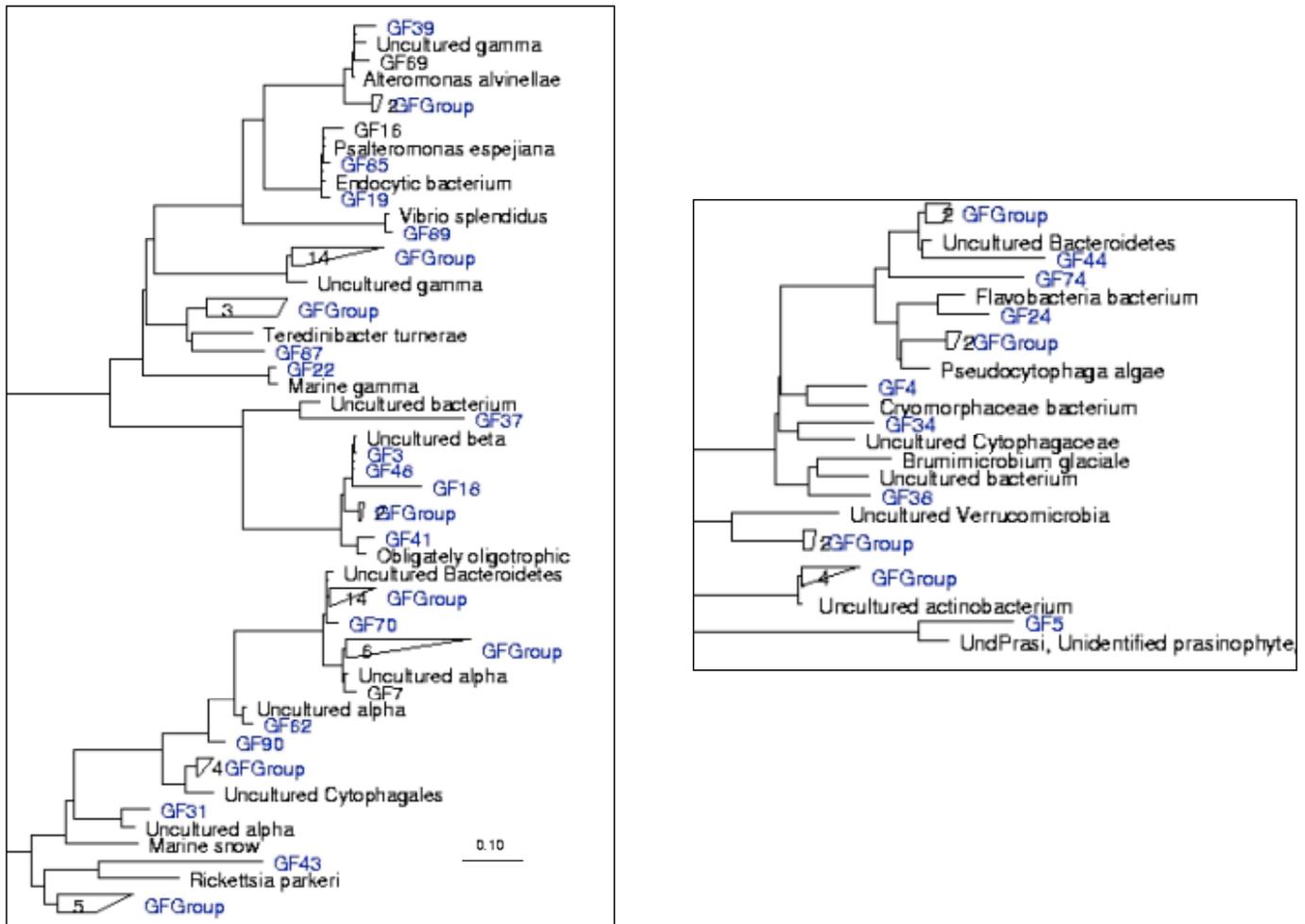
To confirm that the fosmid library contained large fragments of DNA (~40 Kb), five fosmids were extracted following a standard alkaline lysis method and analyzed by PFGE before and after Not I digestion (Fig. 2). As a control, a fosmid clone, prepared using the control 36Kb insert DNA, was also analyzed. These results showed that all undigested clones were of a similar size and contained a large DNA insert similar to that of the control fosmid clone (Fig. 2). In contrast, Not I digestion of the fosmid clones gave rise to different patterns by PFGE. The fosmid DNA vector has two Not I sites, either side of the multiple cloning site, and hence if the cloned fragment contained no Not I site then the digest would be expected to reveal two fragments (fosmid backbone and insert). For two of the fosmid clones (3 & 4), this profile was obtained and revealed that they did indeed have large DNA inserts. In contrast, digestion of the three other fosmid clones revealed either 3 or 4 fragments, suggesting that the cloned fragments contained internal Not I sites. Unfortunately, due to the absence of a proper DNA ladder, the correct sizing of these inserts was not possible. However, based upon the digested control fosmid clone, and the undigested results, these results do suggest that the inserts are around the expected size. Thus, these data provide strong evidence that I have constructed a successful fosmid library from my seawater microorganism genomic preparation. The fosmid clones were pooled into LB, centrifuged and then re-suspended in 14 ml LB. 1ml aliquots were prepared as glycerol stocks and stored at  $-80^{\circ}\text{C}$  until required.

### Phylogenetics

Having constructed a fosmid library, I wanted to gain some insights into the microbial diversity that had been present in the water sample. This could give valuable clues and/or ideas for future screens from the fosmid library. The 16S rRNA genes were amplified using bacterial-specific primers, cloned into TOPO and then 87 TOPO clones were sequenced and then the sequences were analyzed by ARB (Fig. 3). The ARB results revealed that I had a wide diversity of bacterial 16S rRNA genes in my genomic preparation. Additionally, many of the 16S rRNA genes appeared to be from uncultured bacteria. Thus, these data supported the idea that my fosmid library is likely to contain genes from both cultured and uncultured diverse bacteria.

### DISCUSSION AND FUTURE RESEARCH

In summary, I have constructed a genomic DNA fosmid library from microorganisms from Woods Hole Passage. Preliminary, phylogenetic studies have revealed that the library is likely to contain a diverse array of cultured and uncultured microorganisms. Thus, future studies will be conducted to screen the library for novel genes, which could affect phenotypes of the *E. coli* host strain. For example, can I identify novel exopolysaccharide genes, which could potentially be useful as under water adhesives? Additionally, is it possible to screen for novel antibiotic production? Although, it is expected that there may be problems with gene expression, these experiments can be conducted in the presence of arabinose, which increases the copy number of the fosmid clone from 1 copy to ~10-50 copies. Thus, this may improve the amount of poorly expressed genes. Thus, by these approaches, I may identify novel genes, which affect the phenotypes of *E. coli*. In addition, genes could also potentially be identified on fosmid



**Fig. 3. ARB Analysis of 16S rRNA Gene Clone Library**

clones using defined probes. Such an approach will be used by a fellow student our course, who will use my library to screen for proteorhodopsin genes in seawater microorganisms.

To gain further insights into the diversity of microorganisms in my fosmid library, further TOPO clones from the 16S rRNA will be sequenced. FISH will also be used to investigate diversity within the sample and this will be followed up in collaboration with Mark on my return to my lab.

As mentioned earlier, in the longer-term I would like to conduct a similar study on microorganisms from the deep-sea samples described in Table 1. Since the microorganisms are adapted to life at the real extremes, this future research could reveal even more unusual genes, which could be of great use in biotechnology.

## REFERENCES

1. **Polz, M. F., and C. M. Cavanaugh.** 1995. Dominance of one bacterial phylotype at a Mid-Atlantic Ridge hydrothermal vent site. *Proc Natl Acad Sci U S A* **92**:7232-6.
2. **Van Dover, C. L., E. Z. Szuts, S. C. Chamberlain, and J. R. Cann.** 1989. A novel eye in 'eyeless' shrimp from hydrothermal vents of the Mid-Atlantic Ridge. *Nature* **337**:458-60.
3. **Zhou, J., M. A. Bruns, and J. M. Tiedje.** 1996. DNA recovery from soils of diverse composition. *Appl Environ Microbiol* **62**:316-22.

## ACKNOWLEDGEMENTS

I am indebted to the great team of directors, faculty and teaching assistants of the summer of '05. Thanks to Bill and Tom for your insights and encouragement and especially to Tom for prompting me to move "out with" my comfort zone with and learn ARB. Thanks to Mark, Amaya, Adam, Dion, Laurie, Ben, Aurora and my fellow students for all your support, hard work and encouragement-you kept me going! I would also like to thank the Howard Hughes Medical Institute, S.O. Mast Memorial Fund and Gordon & Betty Moore Foundation for their sponsorship. Thanks also to the University of Edinburgh and CSEC for additional funding.