



## Fast and Reproducible PCR-Based Approach for DNA Library Screening

Mohamed A. Ghazy\* and Ahmed A. Sayed

*Department of Biochemistry, Faculty of Science, Ain Shams University.*

### ARTICLE INFO

#### Article history:

Received 05 May 2015

Accepted 18 June 2015

#### Keywords:

Metagenomics;  
PCR;  
screening;  
fosmid;  
environmental.

### ABSTRACT

Metagenomics helped the understanding of diverse microbial communities in various habitats by skipping the need for culturing. We established fosmid libraries from the environmental DNAs isolated from the sea. Fosmid library was constructed according to the instruction manual (Epicentre). Moreover, we developed a PCR based method for easy and fast screening of the fosmid libraries for genes of interest. In summary bacterial colonies containing fosmids were allowed to grow in 96 well plate (each well contained only one fosmid). Identical copy from each bacterial plate was prepared and the 96 cultures were pooled followed by fosmids isolation (each plate was represented by one DNA sample contained the 96 different fosmids). In this work we have constructed a fosmid library contains 10656 colonies, primers specific for the subtilisin like serine protease gene were used to perform the screening protocol. In the primary screening, PCR was conducted on the pooled fosmids DNAs isolated from the different plates. The primary screening showed 15 plates that might harbor the gene. As a proof of concept, we performed the secondary screening on plate 34 where we found only one fosmid containing the gene. The PCR based result was confirmed biochemically using the protease assay. The reported protocol for fosmid library screening will be helpful for fast screening and for isolation of genes of interest from uncultured microorganisms.

### Introduction

Microorganisms considered one of the early forms of life [1]. Prokaryotes contributed in different forms of life on earth and their significance is not only for being primary producers of nutrients and recyclers for organic matter [2,3] but also represented by the huge genetic repertoire that remains unexplored [4]. Microorganisms have been under study using simple classical culture dependent approaches on media containing the essential nutrients for their survival. 95-99 % of the total microorganisms are inaccessible to the regular genomic studies that require prior isolation and cultivation of the microorganism; this is due to the absence of factor(s) and conditions needed for their growth on the artificial medium. In addition, some microorganisms live in a symbiotic interaction that is difficult to provide in the laboratory [5-8].

Metagenomics, or environmental genomics, is a culture-independent technique that has been developed to study genetic structure and diversity of a microbial community in a particular environment; in other words, it is the

analysis of the collective genomic make-up of organisms in a particular environment such as marine, soil, insects, and human gut [5,8-10]. The study of such uncultured microorganisms provided us with new insights into the metabolic and physiological processes required for their survival. In addition, it is so promising in providing us with unique and original bioactive products and molecules with potential biotechnological applications [11].

When the term metagenomics was first coined it meant to borrow the notion of collective analysis of un-identical items that share some extent of similarity [9,12,13]. Using recent advances in genomics and computational analyses, DNA isolated from microbial community in a particular environment, "environmental DNA", can be analyzed using 16S rRNA-based phylogenetic analysis [14,15] and whole genome shotgun (WGS) sequencing approach [16]. These analyses allowed detailed descriptions of the genetic makeup and diversities of microbial communities in different and unexplored environments [8,9,17]. In addition to that, metagenomics contributed to the identification of many biomolecules from different environments that couldn't be easily

\* Corresponding author.

E-mail address: [hady72@gmail.com](mailto:hady72@gmail.com)

identified before, like: Cold adapted Catalase, Subtilisin, and Cellulase from Antarctic Sea water<sup>[18-20]</sup>.

In addition to mass sequencing of environmental DNA and developing of a database, it is important to construct a metagenomic library to maintain the genetic material and to isolate genes of interest. Many vectors are available for metagenomic library construction; they vary based on their capacity to accommodate different sizes of inserts and their transformation efficiency; for example plasmids, fosmids and bacterial artificial chromosomes. Fosmid vector is usually used because it can harbor up to 40 kb DNA insert, and it has high efficiency of transformation. The recombinant vectors will then be transformed into the appropriate host. Once a metagenomic library is created, successive screening methods (sequence based, or functional based approach) are used to isolate specific genes and to explore functional and taxonomic potential of uncultured microbes<sup>[21-23]</sup>.

Library screening is the rate-limiting step in genomic analysis. Traditional methods for library screening included; Hybridization or PCR based methods<sup>[24, 25]</sup>. Library screening by hybridization involves immobilization of recombinant DNA molecules on solid supports, followed by hybridization with specific probes to detect specific sequences of interest. This approach is laborious and requires several screening steps to avoid the false positives. Over the past decade, a number of new strategies have been implemented. Such strategies include PCR-based methods, which greatly decreased time and cost requirements of library screening. A single PCR is used to determine insert presence, without the need for purification or hybridization. PCR screening methods involve initial large-scale screening using the pooled library, followed by several rounds of screening of smaller pools from which a positive signal can be isolated<sup>[26]</sup>. PCR based screening, is quick and more specific, but requires sequence information and primer generation for each gene<sup>[27]</sup>.

The goal of the work described here is to develop a fast PCR based method for DNA library screening. In this study, we developed a fast and reproducible PCR based method for fosmid library screening in 96 well plate format, which requires only two PCR reactions. This approach was successful in isolating a subtilisin like serine protease, which was confirmed biochemically using the protease assay. The proposed protocol will facilitate the DNA library screening comparing the traditional screening methods.

## Materials and Methods

### A) Genomic DNA Isolation:

Purification of DNA from the environmental sample was conducted as described in Rusch et al.,<sup>[22]</sup> with a modification of CTAB treatment as described in the DOE-Joint Genome Institute, standard operating procedure for "Bacterial genomic DNA isolation using CTAB"<sup>[28]</sup>. DNA quality and quantity were assessed by gel visualization and measuring concentration using

NanoDrop 2000 (Thermo Scientific, USA).

### B) DNA library construction:

DNA fosmid library construction was conducted using Control Fosmid Library Production Kit according to the instruction manual (Epicentre). Briefly, 3 µg of the isolated DNA was sheared, followed by end-repaired step using the End-Repair enzyme mix that generates 5'-phosphorylated DNA with blunt ends ready for ligation. The end-repaired DNA was then size fractionated using 1% (w/v) low melting agarose gel electrophoresis; 30–50 kb DNA fragments was excised from the gel, purified using Gelase enzyme followed by ethanol precipitation. The DNA pellet was dissolved in 30 µl of TE buffer and quantified using a NanoDrop 2000. Size fractionated DNA was ligated into CopyControl pCC1FOS Vector using Fast-Link DNA Ligase. The cloning product was transformed into EPI300 using MaxPlax Lambda Packaging Extracts; in which 10 µl of the phage prep incubated for 20 minutes at 37°C with 100 µl of EPI300 host cells. The infected EPI300 cells were spread on LB plates supplemented by 12.5 µg/ml chloramphenicol, cells allowed to grow at 37°C overnight. The colonies were then picked individually, each colony transferred into one well of 96 well ELISA plates, that contain LB supplemented by 12.5 µg/ml chloramphenicol. Cells were allowed to grow in the plates at 37°C overnight. The library produced 10,656 clones distributed in 111 ELISA plates. A duplicate of the library plates and a glycerol stock was generated and stored in the -80°C.

### C) DNA library preparation for fast screening:

To perform the fast library screening, the bacterial content of each ELISA plate (contains 96 different Fosmids) was pooled into a single tube and allowed to grow overnight at 37°C in LB media supplemented by 12.5 µg/ml chloramphenicol. The total Fosmid DNA (96/plate) was isolated using the promega Wizard® Plus SV Minipreps DNA Purification System, resulting in preparation of 111 separate DNA samples.

### D) DNA library fast screening:

DNA library fast PCR screening was performed using specific primers for the gene of interest. In this work we used Primer3 plus<sup>[29]</sup> to design new primers specific for subtilisin like serine protease gene based on assembly of several sequencing reads specific for the environmental sample; SR-F (TCCCCTTCGAATCCTCTAGC), SR-R (TACGCACCTACGGACTCTCC). The PCR using SR-F and SR-R generated a 500 bp amplicon. In the first PCR round, the pooled DNA samples were used as a template DNA (111 samples, each representing one plate) using Dream Taq (Fermentas, USA). In case of a positive results [a PCR signal representing a plate number], a second colony PCR was performed using the same primer set and the 96 bacterial colonies of the positive signal plate. A positive PCR signal pointed to the exact positive fosmid, which it could be easily isolated.

### E) Protease Assay:

The 96 bacterial colonies of the plate that gave a positive PCR signal in the PCR screening were allowed to grow on LB agar plates supplemented by 1% w/v skim milk, 0.1% arabinose and chloramphenicol 40 µg/ml. Plates were incubated at 37°C for 24-48 hours. Positive clones that showed proteolytic activity formed a clear zone/halo around the colony<sup>[30]</sup>.

### Results

#### Library Quality and Diversity

To examine the quality and the diversity of the library, 24 different fosmid clones were randomly selected and subjected to *Bam*HI digestion; which is flanking the 40Kb insert on the copy control vector. The data presented in **figure 1** show that each recombinant fosmid contains insert with a different digestion pattern, suggesting the diversity of the library.

#### Fast Library Screening for Serine protease

To test the fast PCR screening protocol that we proposed, 111 PCR reactions were performed using PCR primers generated specific to the serine protease. Ten µl PCR was performed using 1 µl of the pooled DNA as a template in presence of 0.1 µM of each primer, 0.1 mM

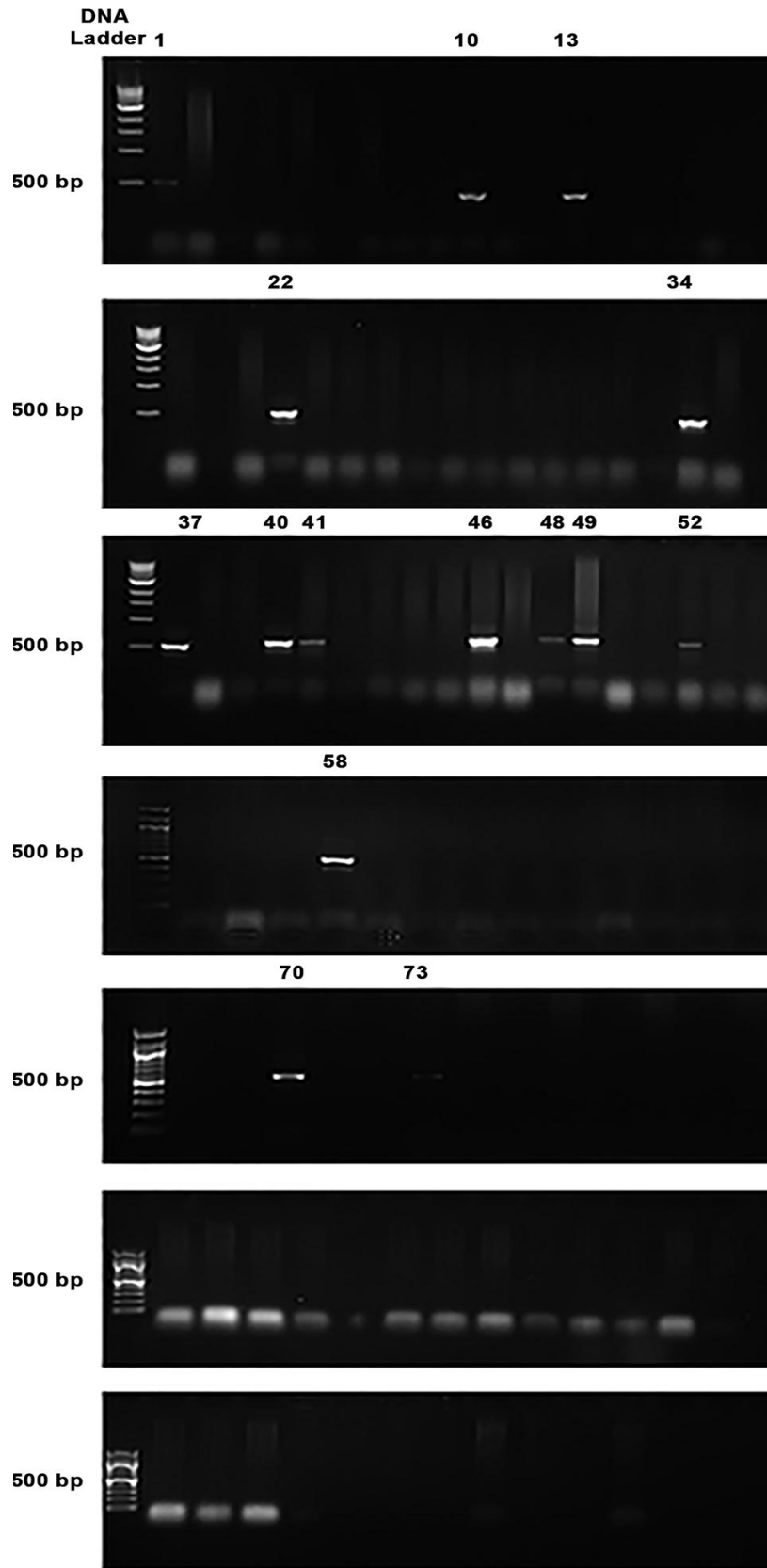
dNTPs, 1X polymerase buffer and 1 unit of the Taq polymerase. The reactions were performed at 94 degrees for 5 minutes, then 30 cycles of 3 temperature settings (94 degrees for 1 minute, 55 degrees for 1 minute, 72 degrees for 1 minute), followed by 72 degrees for 7 minutes. The PCR products were analyzed by gel electrophoresis. The data presented in **figure 2** showed that 15 plates (1, 10, 13, 22, 34, 37, 40, 41, 46, 48, 49, 52, 58, 70 and 73) might harbor potential serine protease genes. In order to identify the exact fosmid that harbor the potential gene, we decided to proceed with plate #34. A second 10 µl colony PCR was performed on the 96 colonies of plate #34 using same primer set, dNTPs, polymerase buffer and enzyme concentrations following the same thermal cycle protocol used in the earlier PCR. The data presented in **figure 3**, showed that only one fosmid (B8) gave a positive PCR signal.

#### Fast Library Screening Conformation

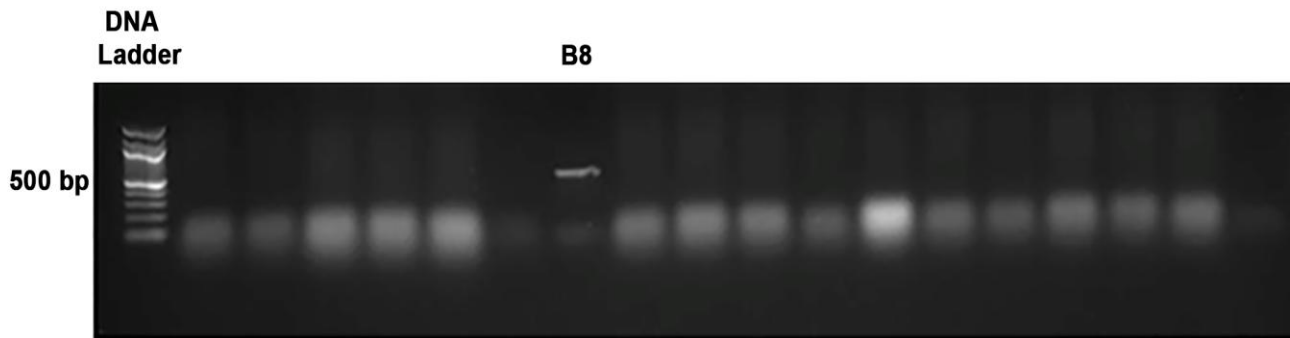
In order to confirm the previous result, we chose to perform a functional assay on plate #34 by testing for serine protease presence using the skim-milk agar procedure. The data presented in **figure 4** showed that only one clone gave a positive lysis indicating the presence of a functional serine protease within this plate.



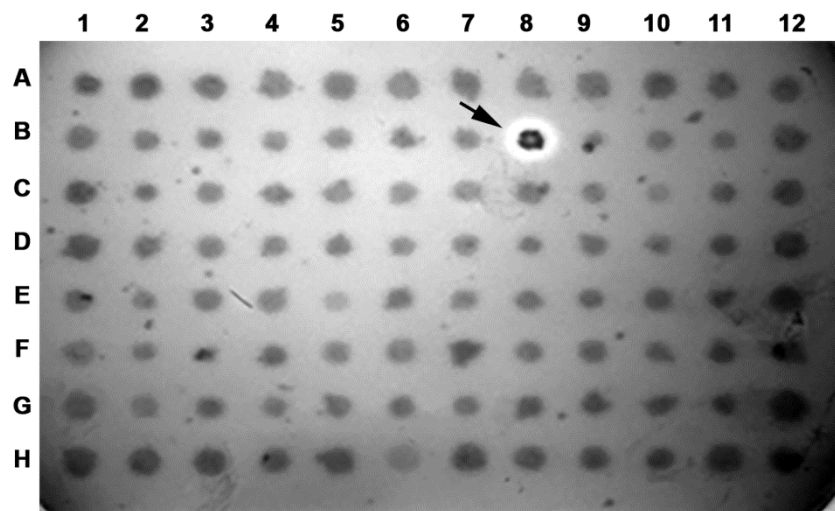
**Fig. 1:** *Bam*HI digestion of the 24 recombinant fosmid clones from randomly-picked colonies. 1 µg of each fosmid was subjected to *Bam*HI digestion in the recommended digestion buffer for 2 hours, followed by addition of the 6X loading buffer and analyzed on 0.8% agarose gel electrophoresis.



**Fig. 2: Screening of the 111 plates of the fosmid library.** Total DNA of each plate was subjected to PCR. The products of the amplification were analyzed on 0.8% agarose gel electrophoresis. The numbers on the top of each lane represent the plate number.



**Fig. 3: Screening of the plate number 34.** Each colony (showing results of only 18) of the plate # 34 was subjected to colony PCR. The products of the amplification were analyzed on 0.8% agarose gel electrophoresis. The number on the top of any lane represents the colony number.



**Fig. 4: Functional screening of the plate number 34.** The bacterial colonies of the plate were grown on skim milk agar for 24-48 hours at 37°C. Arrowheads indicate the positive clones.

## Discussion

Metagenomics is a powerful tool that allows us to study the uncultured microbial communities. It is based on the analysis of the extracted DNA from environmental samples that allows the identification of novel metabolic and physiological processes of the different microorganisms inhabiting this environment. By incorporating all the data obtained from this powerful genomics tool, the mode of living and interactions of these microbial communities can be explored and determined. Furthermore, new biological molecules can be discovered and therefore be of great potential to the therapeutic and biotechnological applications<sup>[31,32]</sup>. Metagenomics projects with different purposes construct libraries of cloning vectors having the environmental DNA as inserts. These can be screened afterwards for genes of interest. Direct sequencing of the collected DNA is another tool in order to obtain metagenomics datasets. Such techniques would provide us with an insight about the biochemical pathways of these microbes.

In this work, we constructed a fosmid library from the total genomic DNA isolated from the environmental sample following the manufacture protocol; the library consisted of 10,656 clones; which is corresponding to  $4.2 \times 10^8$  bp. In order to test the quality and the diversity of the library, we used a simple restriction digestion approach, the data presented in **figure 1** showed different restriction pattern for the different fosmids, suggesting each fosmid carrying a different 40 kb insert. Further end terminal sequencing of the fosmids should be performed to confirm this idea.

The library screening for any gene of interest can be performed using the hybridization method, which is time consuming and require several screenings to confirm the positive signal. In this paper we used a modified PCR based approach, in which, by using only 2 PCR's, we can identify the fosmid that carry the gene of interest out of 10,656 fosmid.

PCR primers specific for the subtilisin like serine proteases was generated, and used in this work. In the first PCR reaction, we used total library DNA. Fifteen

plates produced positive signals, indicating the presence of serine protease genes in the 15 different plates. To identify which fosmid out of the 96 ones in each plate contained the gene of interest, we performed the second PCR using the same primers and DNA templates of the 96 samples of the plate. The results showed that only one fosmid produced a positive signal. The same plate was used to confirm the presence of the serine protease gene using a biochemical approach, where the skim-milk agar plate method was used and showed an identical result to the PCR approach.

The work described here provides an easy and fast strategy for screening libraries. This strategy should consistently yield positive results with the use of good reagents and careful technique. Further analysis of the isolated fosmid using sequencing of the area covering the gene of interest should be performed to confirm the presence of the gene and to identify its start and stop codons, as well as the ribosomal binding site and the -10 and -35 boxes.

#### Acknowledgment

The authors would like to thank Dr. Hamza El-Dorry and the members of the Biology Department at the School of Science and Engineering at the American University in Cairo for their help and support during the course of this work.

#### References

- 1) **Schopf, J. and Packer, B. (1987).** Early Archean (3.3-billion to 3.5-billion-year-old) microfossils from Warrawoona Group, Australia. *Science*, 237: 70-73.
- 2) **Whitman, W. B., Coleman, D. C. and Wiebe, W. J. (1998).** Prokaryotes : The unseen majority. *Proc. Natl. Acad. Sci.*, 95: 6578–6583.
- 3) **Wooley, J. C., Godzik, A. and Friedberg, I. (2010).** A primer on metagenomics. *PLoS Comput. Biol.*, 6, e1000667.
- 4) **Simon, C. and Daniel, R. (2011).** Metagenomic analyses: past and future trends. *Appl Environ Microbiol.*, 77: 1153–61.
- 5) **Kennedy, J., Marchesi, J. R. and Dobson, A. D. (2008).** Marine metagenomics: strategies for the discovery of novel enzymes with biotechnological applications from marine environments. *Microb Cell Fact.* 7:27. doi: 10.1186/1475-2859-7-27.
- 6) **Ainsworth, T. D., Thurber, R. V. and Gates, R. D. (2010).** The future of coral reefs: a microbial perspective. *Trends. Ecol. Evol.*, 25: 233-240.
- 7) **Schloss, P. D. and Handelsman, J. (2003).** Biotechnological prospects from metagenomics. *Curr. Opin. Biotechnol.*, 14: 303–310.
- 8) **Sleator, R. D., Shortall, C. and Hill, C. (2008).** Metagenomics. *Lett. Appl. Microbiol.*, 47: 361-366.
- 9) **Singh, J., Behal, A., Singla, N., Joshi, A., Birbian, N., Singh, S., Bali, V. and Batra, N. (2009).** Metagenomics: Concept, methodology, ecological inference and recent advances. *Biotechnol. J.*, 4: 480-494.
- 10) **Li, X. and Qin, L. (2005).** Metagenomics-based drug discovery and marine microbial diversity. *Trends Biotechnol.*, 23: 539-543.
- 11) **Lorenz, P. and Eck, J. (2005).** Metagenomics and industrial applications. *Nat. Rev. Microbiol.*, 3: 510-516.
- 12) **Schmidt, T. M., DeLong, E. F. and Pace, N. R. (1991).** Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. *J. Bacteriol.*, 173: 4371-4378.
- 13) **Handelsman, J. (2004).** Metagenomics: application of genomics to uncultured microorganisms. *Microbiol. Mol. Biol. Rev.*, 68: 669-685.
- 14) **DeLong, E. F., Wu, K. Y., Prezelin, B. B. and Jovine, R. V. (1994).** High abundance of Archaea in Antarctic marine picoplankton. *Nature*, 371: 695-697.
- 15) **Fuhrman, J. A., McCallum, K. and Davis, A. A. (1993).** Phylogenetic diversity of subsurface marine microbial communities from the Atlantic and Pacific oceans. *Appl. Environ. Microbiol.*, 59: 1294–1302.
- 16) **Venter, J. C., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A., Wu, D., Paulsen, I., Nelson, K. E., Nelson, W., Fouts, D. E., Levy, S., Knap, A. H., Lomas, M. W., Neelson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., Baden-Tillson, H., Pfannkoch, C., Rogers, Y. H. and Smith, H. O. (2004).** Environmental genome shotgun sequencing of the Sargasso Sea. *Science*, 304: 66-74.
- 17) **Spear, J. R., Ley, R. E., Berger, A. B. and Pace, N. R. (2003).** Complexity in natural microbial ecosystems: the Guerrero Negro experience. *Biol. Bull.*, 204: 168-173.
- 18) **Lorentzen, M. S., Moe, E., Jouve, H. M. and Willassen, N. P. (2006).** Cold adapted features of *Vibrio salmonicida* catalase: characterisation and comparison to the mesophilic counterpart from *Proteus mirabilis*. *Extremophiles*, 10: 427-440.
- 19) **Davail, S., Feller, G., Narinx, E. and Gerday, C. (1994).** Cold adaptation of proteins. Purification, characterization, and sequence of the heat-labile subtilisin from the antarctic psychrophile *Bacillus TA41*. *J. Biol. Chem.*, 269: 17448-17453.
- 20) **Garsoux, G., Lamotte, J., Gerday, C. and Feller, G. (2004).** Kinetic and structural optimization to catalysis at low temperatures in a psychrophilic cellulase from the Antarctic bacterium *Pseudoalteromonas haloplanktis*. *Biochem. J.*, 384: 247-253.
- 21) **Daniel, R. (2005).** The metagenomics of soil. *Nat Rev. Microbiol.*, 3: 470–478.
- 22) **Rusch, D. B., Halpern, A. L., Sutton, G., Heidelberg, K. B., Williamson, S., Yooseph, S., Wu, D., Eisen, J. A., Hoffman, J. M., Remington, K., Beeson, K., Tran, B., Smith, H., Baden-Tillson, H., Stewart, C., Thorpe, J., Freeman, J., Andrews-Pfannkoch, C., Venter, J. E., Li, K., Kravitz, S., Heidelberg, J. F., Utterback, T., Rogers, Y. H.,**

- Falcón, L. I., Souza, V., Bonilla- Rosso, G., Eguiarte, L. E., Karl, D. M., Sathyendranath, S., Platt, T., Bermingham, E., Gallardo V., Tamayo-Castillo, G., Ferrari, M. R., Strausberg, R. L., Nealson, K., Friedman, R., Frazier, M. and Venter, J.C. (2007). The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol.* 5: e77. doi:10.1371/journal.pbio.0050077
- 23) Simon, C. and Daniel, R. (2011). Metagenomic analyses: past and future trends. *Appl. Environ. Microbiol.*, 77: 1153-1161.
- 24) Benton, W. D. and Davis, R. W. (1977). Screening  $\lambda$ gt recombinant clones by hybridization to single plaques in situ. *Science*, 196: 180-182.
- 25) Grunstein, M. and Hogness, D. S. (1975). Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. *Proc. Natl. Acad. Sci.*, 72: 3961-3965.
- 26) Yu, L. and Bloem, L. J. (1993). Use of the polymerase chain reaction to screen phage libraries. *Methods Mol. Biol.*, 15: 211-215.
- 27) Mendez, M. J., Klapholz, S., Brownstein, B. H. and Gemmill, R. M. (1991). Rapid screening of a YAC library by pulsed-field gel southern blot analysis of pooled YAC clones. *Genomics*, 10: 661-665.
- 28) William, S., Helene Feil, A. and Copeland, A. (2012). Bacterial genomic DNA isolation using CTAB. Available: <http://jgidoegov/collaborate-with-jgi/pmo-overview/protocols-sample-preparation-information/>.
- 29) Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R. and Leunissen, J. A. (2007). Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Research* 35: W71-W74; doi:10.1093/nar/gkm306.
- 30) Pushpam, P. L., Rajesh, T. and Gunasekaran, P. (2011). Identification and characterization of alkaline serine protease from goat skin surface metagenome. *AMB Express*, 1: 3.
- 31) Schmeisser, C., Steele, H. and Streit, W. R. (2007). Metagenomics, biotechnology with non-culturable microbes. *Appl. Microbiol. Biotechnol.*, 75: 955-962.
- 32) Sharma, P., Kumari, H., Kumar, M., Verma, M., Kumari, K., Malhotra, S., Khurana, J. and Lal, R. (2008). From bacterial genomics to metagenomics: concept, tools and recent advances. *Indian J. Microbiol.*, 48: 173-194.