Plant Pathol. J. 21(2) : 93-98 (2005)

Mini-Review

The Plant Pathology Journal
©The Korean Society of Plant Pathology

Metagenome, the Untapped Microbial Genome, toward Discovery of Novel Microbial Resources and Application into the Plant Pathology

Seon-Woo Lee*
Division of Applied Biology, College of Natural Resources and Life Science, Dong-A University, Pusan 604-714, Korea
(Received April 11, 2005; Accepted on May 4, 2005)

Molecular ecological studies of microbial communities revealed that only tiny fraction of total microorganisms in nature have been identified and characterized, because the majority of them have not been cultivated. A concept, metagenome, represents the total microbial genome in natural ecosystem consisting of genomes from both culturable microorganisms and viable but non-culturable bacteria. The construction and screening of metagenomic libraries in culturable bacteria constitute a valuable resource for obtaining novel microbial genes and products. Several novel enzymes and antibiotics have been identified from the metagenomic approaches in many different microbial communities. Phenotypic analysis of the introduced unknown genes in culturable bacteria could be an important way for functional genomics of unculturable bacteria. However, estimation of the number of clones required to uncover the microbial diversity from various environments has been almost impossible due to the enormous microbial diversity and various microbial population structure. Massive construction of metagenomic libraries and development of high throughput screening technology should be necessary to obtain valuable microbial resources. This paper presents the recent progress in metagenomic studies including our results and potential of metagenomics in plant pathology and agriculture.

Keywords : metagenome, microbial diversity, microbial resources, unculturable microorganisms

---

Since first observation of bacteria using a microscope by Antonie van Leeuwenhoek and development of pure culture technique by Robert Koch, microbiologists have been dependent on pure culture of bacteria to study microorganisms. However, microbiologists has recognized for generations that the majority of microscopically visualized cells are viable but do not form visible colonies on culture plates, so called the phenomena "great plate-count anomaly" (Amann et al., 1995). Introduction of molecular tools in the area of microbial ecology greatly increased the understanding of microbial community structure and the reality of microbial diversity. The molecular ecological approaches, the culture-independent method, included the extraction of microbial nucleic acids *in situ* and subsequent amplification of microbial small subunit rRNA genes, the 16S rDNA (Pace, 1997). The downstream tools include DNA sequence analysis, *in situ* hybridization with rDNA probe, quantitative dot blot hybridization, microbial community structure analysis and development of community rDNA database. The culture-independent studies of various microbial communities demonstrated that portion of the viable but non-culturable bacteria in natural ecosystem is over 99%. In addition, the pioneering work by Carl Woese introduced the new taxon concept of "Domain" based on the phylogenetic analysis of 16S rRNA genes, that classify total organisms into three different Domains; *Bacteria, Archaea,* and *Eucarya* (Woese and Fox 1977; Woese et al., 1990).

This advancement of microbial ecology opened new era of microbiology, which could de described as "renaissance of microbiology". Based on the "great plate-count anomaly", one may assume that most of microbial resources have not been really explored and potential to discover novel microbial products from the unculturable bacteria may be huge. Metagenomic approach is a functional genomics approach of unculturable micro-organisms to discover novel microbial resources from unculturable bacteria (Rondon et al., 2000). In this paper, emergence of metagenomics from microbial ecology, recent progress, technological limitation, and future direction are described. In addition, potential application of metagenomic approach to the plant pathology is briefly discussed.

Microbial diversity. A study on microbial biomass estimation has shown that prokaryotes are the most dominant organisms (Whitman et al., 1998). Majority of bacteria present in nature are not culturable (Amann et al., 1995; Hugenholtz and Pace, 1996). Estimation of the portion of culturable bacteria on standard culture media is variable from 0.01% to 1% of total microorganisms, dependent on the different microbial communities. Microbial diversity has been revealed by culture-independent studies

*Corresponding author.
Phone) +82-51-200-7551, FAX) +82-51-200-7505
E-mail) seonlee@dau.ac.kr

with the analysis of the complexity of 16S rRNA genes from various microbial communities. These studies indicated that microbial community structure is highly complicated and microbial diversity is much higher than ever imagined (Bintrim et al., 1997; Borneman et al., 1996; Liesack and Stackebrandt, 1992). Culture-independent studies showed that there are several thousand microbial species per gram of forest soil (Torsvik et al., 1990, 1998). Although the microbial diversity in nature has been estimated by means of culture-independent methods, it is uncertain whether it would be possible to find a universal way to cultivate and characterize most bacteria present in natural ecosystems (Rondon et al., 1999a).

Recently, Torsvik et al. (2002a) summarized the reality of microbial diversity in various soil environments and sediments. Microbial DNA diversity in forest soil and pasture soil appeared to carry more than 6,000 and 3,500 different microbial genome representatives, respectively. Therefore, we focused on forest soil environments to explore the soil microbial genes and products by means of a metagenomic approach (Lee et al., 2004). We analyzed microbial community of one of forest soils by culture independent method and found that the *Acidobacteria* group was the most numerically abundant phylotype (manuscript in preparation).

Metagenome. Discovery of microbial products has been dependent on cultivation of microbes and subsequent screening for desirable activity. This classical approach indicated that rediscovery rate of known microbial products is increasing and the probability of obtaining novel resources keeps decreasing (Handelsman et al., 1998). A recently developed metagenomic approach employs cloning of the total microbial genome, the so-called 'metagenome', directly isolated from natural environments in culturable bacteria such as *Escherichia coli* (Beja et al., 2000; Handelsman et al., 1998; Rondon et al., 1999a, 2000) and discovering novel microbial resources (Handelsman, 2004). The basis of metagenomic approach originated from the molecular ecological studies of microbial diversity, indicating that majority of micro-organisms in nature was not cultivated by standard cultivation techniques.

Difficulties in microorganism cultivation exclude the majority of the microbial diversity in soils from the utilization of the microbial products (Amann et al., 1995). Therefore, there might be a high probability of finding novel microbial products, such as antibiotics and enzymes, from soil metagenome. Initial studies have demonstrated that the foreign gene expression in a heterologous host is technically feasible for exploring novel microbial resources (Rondon et al., 1999b, 2000).

In addition, the combination of phylogenetic marker screening of metagenomic libraries and genomics could reveal the physiology of as-yet-uncultured microorganisms, only identified by culture-independent studies (Liles et al., 2003; Quaiser et al., 2002). Ultimately, isolating previously uncultured microorganisms in soil (Janssen et al., 2002; Joseph et al., 2003) can be advanced by genetic and physiological information of the uncultured bacteria revealed by the analysis of metagenome (Handelsman, 2004).

Library construction. Library construction with soil

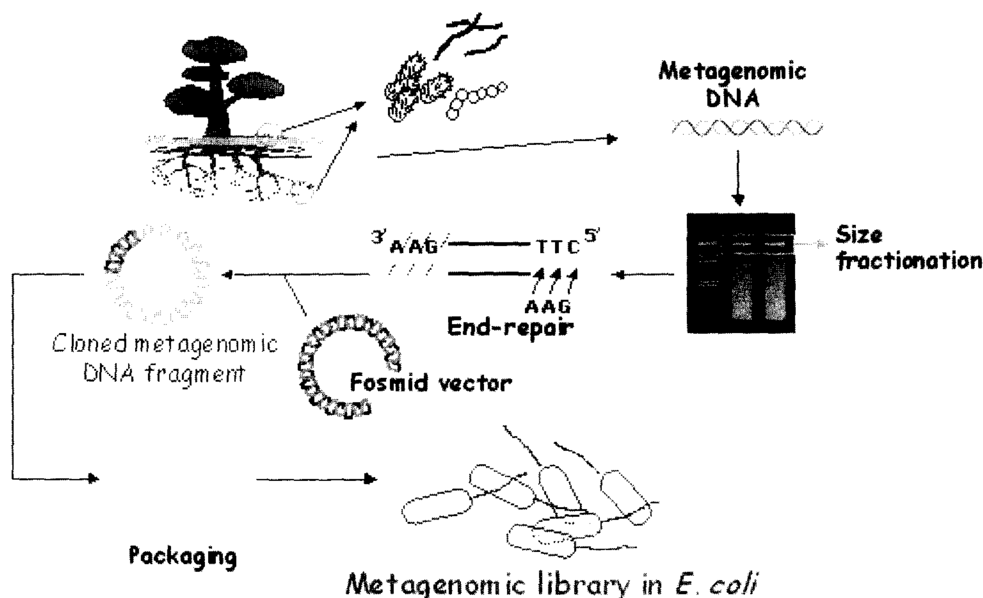

Fig. 1. The library construction process in a fosmid with metagenomic DNA directly isolated from forest topsoil.

metagenome is the conventional process to generate genomic DNA library (Fig. 1). However, technical difficulties are frequently confronted during library construction with the metagenomic DNA directly isolated from soils. There are several protocols described to extract intact pure metagenomic DNA from soils (Berry et al., 2003; Gabor et al., 2003; Rondon et al., 2000; Zhou et al., 1996). The major problem to manipulate the DNA is due to the frequent contamination of humic substances from soils. The removal of the humic substance is still essential step to increase the efficiency of library construction. Because of the low frequency of finding desirable genes from a metagenomic library of diverse microbial genomes, cloning efficiency is an important factor in constructing a large clone library, which should include most of the microbial DNA in the soil (Lorenz and Schleper, 2002).

Fosmids are good vectors for constructing metagenomic libraries because of their high cloning efficiency, improved stability in *E. coli*, and optimum (40 kb) insert size (Kim et al., 1992). Cloning of intact large DNA fragments still carry the discrepancy of population effect of microorganisms in a specific soil. The genomes of many rare microorganisms were probably not included in the library. An rRNA gene analysis from the our forest soil revealed *Acidobacteria* to be the most numerically abundant clone type, therefore our metagenomic library may contain large portion of genes from this group (manuscript in preparation). Several metagenomic libraries were constructed from forest topsoils to explore forest soil microbial genetic resources. Our library, which was constructed using 1 μg of forest topsoil DNA, contained from 30,000 to 150,000 clones, with an approximate average insert size over 35 kb (Lee et al., 2004). The libraries represented almost several hundreds to a few thousands microbial genome equivalents. An alternative strategy for improving the probability of finding desirable genes in the metagenome would be to construct the library after certain enrichment processes (Schloss and Handelsman, 2003).

Exploiting metagenomic library to find useful resources is mostly based on the functional expression of the cloned environmental DNA in the cloning host (Rondon et al., 2000; Wang et al., 2001). However, it is also be possible to explore environmental microbial genome based on DNA homology dependent screening (Quaiser et al., 2002).

Biocatalysts from metagenome. Several studies have exhibited that the metagenomic approach allows for searches of various enzymes such as lipase/esterase, protease, oxidoreductase, and nitrilase in various soils and microbial habitats (Lorenz et al., 2002). Some enzymes are important biocatalysts for various industrial applications. Especially, lipases have been recognized as good biocatalysts due to their useful features such as having no requirements for cofactors, remarkable stability in organic solvents, broad substrate specificity, stereoselectivity, and positional selectivity (Jaeger and Eggert, 2002).

In our study, lipolytic activity was screened from our metagenomic libraries. A total of fifty-five lipolytic active clones were isolated from approximately 210,000-member library based on tributyrin hydrolysis. Fifteen of them were subcloned in a high copy number plasmid and lipolytic activities of the corresponding gene products were confirmed by colorimetric analysis. Six of them were most similar to hormone-sensitive lipase homologs from cultured bacteria and archaea (Fig. 2), although the identity of the deduced amino acid sequences were in a range of 35-49% to the known hormone-sensitive lipases. Enzymatic analysis and sequence analysis indicated that the hormone sensitive lipases from forest soil metagenome were more likely carboxyesterases, which hydrolyze the short chain

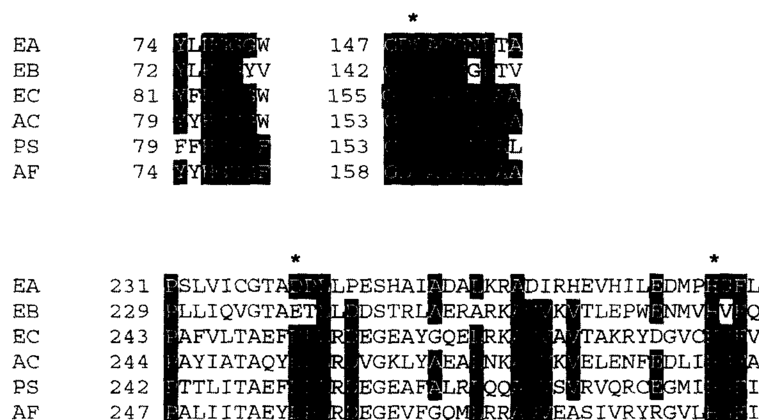

Fig. 2. Alignment of the conserved regions containing enzyme active sites of lipases from the forest topsoil metagenome (EA, EB, EC:), from *Alicyclobacillus acidocaldarius* (AC), from cold-adapted *Pseudomonas* sp. B11-1 (PS), and from hyperthermophilic *Archaeoglobus fulgidus* (AF). Asterisks indicate the catalytic triads (serine, aspartate or glutamate, histidine) of the hormone-sensitive lipase family.

acyl-ester (Lee et al., 2004). Further analysis showed that one of the enzymes could be developed as a useful biocatalyst because of its specific resolution of racemic mixture of organic compounds (unpublished).

Metabolites from metagenome. A number of recent metagenomic studies also showed that the approach provides a unique opportunity to search for novel microbial bioactivities and the anticancer drugs (Handelsman, 2004; Pettit, 2004). Several antibiotics and novel secondary metabolites were identified from metagenomic studies together with their responsible genes, including long chain acyl amino acid antibiotics and turbomycin antibiotics, and genes encoding polyketide synthases (Handelsman, 2004). The bioactive compounds from metagenome were obtained from various soils and microbial communities including unculturable bacterial symbiont of invertebrates (Piel, 2002). Therefore, diverse microbial communities could be subject to search for the novel bioactivities using metagenomic approach.

In a course of obtaining novel antibiotics from soil metagenome, we have adopted a bioassay-based library screening method using a yeast or *Bacillus subtilis* as target organisms to select antimicrobial active clones (Rondon et al., 2000). Several clones with biological activity were isolated and their genes were identified. One of them showed that a gene cluster in the clone consisted of more than 15 ORFs to express antifungal activity (unpublished). Among several metagenomic clones, one clone was characterized as producing indole-derived pigment antibiotics. The corresponding genes were identified and characterized (manuscript in preparation). Our results and results from other researchers suggested that the metagenomic library construction with intact large DNA inserts is an essential procedure to search for novel bioactivities and it is technically feasible.

Potential in plant pathology and agriculture. Discovery of novel antimicrobial compounds is a conventional but plausible strategy to develop agrochemicals with novel modes of action to manage plant diseases. The original compounds and their derivatives are good chemical candidates for lead generation and further development of agrochemicals (Corran et al., 1998). Previous studies demonstrated that metagenome from various microbial communities could be potential reservoir to obtain candidate antimicrobial compounds with novel mode of action. Since the selection of bioactive clones indicated that the responsible genes are cloned in a culturable bacterium, further manipulation of the responsible genes to improve the activity or production could be possible.

Two fundamental questions could be addressed in the plant pathology area related with viable but unculturable bacteria. Are there many more unculturable plant pathogenic bacteria, yet to be identified? There have been several cases previously reported as a viral disease or disease caused by unknown pathogen, but later proven as bacterial diseases (da Graca, 1991). Due to unculturable nature of the pathogenic agent, the metagenomic approach could be an alternative strategy to identify the causative unculturable organism. The major technical limitation to perform the metagenomic approach would be the separation of the microbial DNA from plant-originated nucleic acids. Another potential application includes the plant disease suppressiveness, which could be related with unculturable dominant microorganisms. There have been a number of reports about suppressive soils and the responsible organisms identified based on the culture-dependent method (Weller et al., 2002). However, it is not clear if nonculturable microorganisms have a role in suppressiveness (Weller et al., 2002). Metagenomic approach would be an alternative way to identify the responsible microorganisms and their traits by constructing and screening the metagenomic library from suppressive soils.

Molecular breeding of various crops and other plants engineering with new functions are an important issue in future biotechnology and agricultural industry (Dunwell, 2005). Plant molecular breeding with pest resistance function is also one of measures to increase agricultural production (Campbell et al., 2002). In spite of continued debates of the cultivation of genetically modified crops, the cultivation areas of the engineered crops are increasing in other countries including the US. The majority of the genes introduced into plants are originated from microorganisms. Metagenome could also supply the beneficial genes into the plant biotechnology to engineer the plant with novel functions. Since the origin of the metagenomic genes to be used in the plant biotechnology is not known in the stage of initial screening to search for the gene, the gene cannot be directly used for the introduction into other organisms. Therefore, either the homologs of the target metagenomic genes from the identified bacteria should be used or the original microorganism carrying the target metagenomic genes should be identified.

Future perspectives. Screening the metagenomic library is dependent on either phenotypic expression of the metagenome in culturable bacteria or DNA sequence homology. Both have advantages and disadvantages. Since the phenotypic expression dependent screening is based on the gene expression in culturable bacteria, tapping genetic resources is limited in some fractions of genetic resources. The efforts to develop new host-vector system are necessary to explore diverse microbial genes from various microbial communi-

ties including the Archaea dominated one. Despite of several technical limitations of metagenomic approach to search for microbial resources, the libraries could be used for broad applications in the field of biotechnology, as previously described (Schloss and Handelsman, 2003).

Several strategies have been described for linking phylogenetic groups to their activity and function in soil (Torsvik and Øvreås, 2002b), but defining the origin of a specific metagenomic clone is a great challenge. It is important to define the origin of the clones in terms of the utility of the genes and functional genomics of the uncultured microorganisms to really solve the enigma of culture problem of microorganisms.

Acknowledgements

This research was supported by a grant (MG 05-0103-1-0) from the Microbial Genomics and Application Center of the 21st Century Frontier Research Program funded by the Ministry of Science and Technology of the Republic of Korea.

References

Amann, R. I., Ludwig, W. and Schleifer, K. H. 1995. Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol. Rev.* 59:143-169.

Béjà, O., Suzuki, M. T., Koonin, E. V., Aravind, L., Hadd, A., Nguyen, L. P., Villacorta, R., Amjadi, M., Garrigues, C., Javanovich, S. B., Feldman, R. A. and Delong, E. F. 2000. Construction and analysis of bacterial artificial chromosome libraries from a marine microbial assemblage. *Environ. Microbiol.* 2:516-529.

Berry, A. E., Chiocchini, C., Selby, T., Sosio, M. and Wellington, E. M. H. 2003. Isolation of high molecular weight DNA from soil for cloning into BAC vectors. *FEMS Microbiol. Lett.* 223: 15-20.

Bintrim, S. B., Donohue, T. J., Handelsman, J., Roberts, G. P. and Goodman, R. M. 1997. Molecular phylogeny of archaea from soil. *Proc. Natl. Acad. Sci. USA* 94:277-282.

Borneman, J., Skroch, P. W., O'Sullivan, K. M., Palus, J. A., Rumjanek, N. G., Jansen, J. L., Nienhuis, J. and Triplett, E. W. 1996. Molecular microbial diversity of an agricultural soil in Wisconsin. *Appl. Environ. Microbiol.* 62:1935-1943.

Campbell, M. A., Fitzgerald, H. A. and Ronald, P. C. 2002. Engineering pathogen resistance in crop plants. *Transgenic Res.* 11: 599-613.

Corran, A. J., Renwick, A. and Dunbar, S. J. 1998. Approaches to in-vitro lead generation for fungicide invention. *Pestic. Sci.* 54:338-344.

da Graca, J. V. 1991. Citrus greening disease. *Annu. Rev. Phytopathol.* 29:109-136.

Dunwell, J. M. 2005. Transgenic crops: the current and next generations. *Methods Mol. Biol.* 286:377-398.

Gabor, E. M., de Vries, E. J. and Janssen, D. B. 2003. Efficient recovery of environmental DNA for expression cloning by indirect extraction methods. *FEMS Microbiol. Ecol.* 44:153-163.

Handelsman, J. 2004. Metagenomics: application of genomics to uncultured microorganisms. *Microbiol. Mol. Biol. Rev.* 68: 669-685.

Handelsman, J., Rondon, M. R., Brady, S. P., Clady, J. and Goodman, R. M. 1998. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.* 5:R245-249.

Hugenholtz, P. and Pace, N. R. 1996. Identifying microbial diversity in the natural environment: a molecular phylogenetic approach. *Trends Biotechnol.* 14:190-197.

Jaeger, K-E. and Eggert, T. 2002. Lipases for biotechnology. *Curr. Opin. Biotechnol.* 13:390-397.

Janssen, P. H., Yates, P. S., Grinton, B. E., Taylor, P. M. and Sait, M. 2002. Improved culturability of soil bacteria and isolation in pure culture of novel members of the divisions *Acidobacteria, Actinobacteria, Proteobacteria,* and *Verrucomicrobia. Appl. Environ. Microbiol.* 68:2391-2396.

Joseph, S. J., Hugenholtz, P., Sangwan, P., Osborne, C. A. and Janssen, P. H. 2003. Laboratory cultivation of widespread and previously uncultured soil bacteria. *Appl. Environ. Microbiol.* 69:7210-7215.

Kim, U-J., Shizuya, H., De Jong, P. J., Birren, B. and Simon, M. I. 1992. Stable propagation of cosmid-sized human DNA inserts in an F factor based vector. *Nucleic Acids Res.* 20: 1083-1085.

Lee, S-W., Won, K., Lim, H. K., Kim, J.-C., Choi, G. J. and Cho, K. Y. 2004. Screening for novel lipolytic enzymes from uncultured soil microorganisms. *Appl. Microbiol. Biotechnol.* 65: 720-726.

Liesack, W. and Stackebrandt, E. 1992. Occurrence of novel groups of the domain Bacteria as revealed by analysis of genetic material isolated from an Australian terrestrial environment. *J. Bacteriol.* 174:5072-5078.

Liles, M. R., Manske, B. F., Bintrim, S. B., Handelsman, J. and Goodman, R. M. 2003 A census of rRNA genes and linked genomic sequences within a soil metagenomic library. *Appl. Environ. Microbiol.* 69:2684-2691.

Lorenz, P. and Schleper, C. 2002. Metagenome - a challenging source of enzyme discovery. *J. Mol. Catal. B-Enzym.* 19-20:13-19.

Lorenz, P., Liebeton, K., Niehaus, F. and Eck, J. 2002. Screening for novel enzymes for biocatalytic processes: accessing the metagenome as a resource of novel functional sequence space. *Curr. Opin. Biotechnol.* 13:572-577.

Pace, N. R. 1997. A molecular view of microbial diversity and the biosphere. *Science* 276:734-740.

Pettit, R. K. 2004. Soil DNA libraries for anticancer drug discovery. *Cancer Chemother. Pharmacol.* 54:1-6.

Piel, J. 2002. A polyketide synthase-peptide synthase gene cluster from an uncultured bacterial symbiont of *Paederus* beetles. *Proc. Natl. Acad. Sci. USA* 99:14002-14007.

Quaiser, A., Ochsenreiter, T., Klenk, H-P., Kletzin, A., Treusch, A. H., Meurer, G., Eck, J., Sensen, C. W. and Schleper, G.

2002. First insight into the genome of an uncultivated crenarchaeote from soil. *Environ. Microbiol.* 4:603-611.

Rondon, M. R., August, P. R., Bettermann, A. D., Brady, S. F., Grossman, T. H., Liles, M. R., Loiacono, K. A., Lynch, B. A., MacNeil, I. A., Minor, C., Tiong, C. L., Gilman, M., Osburne, M. S., Clardy, J., Handelsman, J. and Goodman, R. M. 2000. Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl. Environ. Microbiol.* 66:2541-2547.

Rondon, M. R., Goodman, R. M. and Handelsman, J. 1999a. The earth's bounty: assessing and accessing soil microbial diversity. *Trends Biotechnol.* 17:403-409.

Rondon, M. R., Raffel, S. J., Goodman, R. M. and Handelsman, J. 1999b. Toward functional genomics in bacteria: analysis of gene expression in *Escherichia coli* from a bacterial artificial chromosome library of *Bacillus cereus*. *Proc. Natl. Acad. Sci. USA* 96:6451-6455.

Schloss, P. D. and Handelsman. J. 2003. Biotechnological prospects from metagenomics. *Curr. Opin. Biotechnol.* 14:303-310.

Torsvik, V., Daae, F. L., Sandaa, R-A. and Øvreås, L. 1998. Novel techniques for analyzing microbial diversity in natural and perturbed environments. *J. Biotechnol.* 64:53-62.

Torsvik, V., Goksoyr, J. and Daae, F. L. 1990. High diversity in DNA of soil bacteria. *Appl. Environ. Microbiol.* 56:782-787.

Torsvik, V., Øvreås, L. and Thingstad, T. F. 2002a. Prokaryotic diversity -magnitude, dynamics, and controlling factors. *Science* 296:1064-1066

Torsvik, V. and Øvreås, L. 2002b. Microbial diversity and function in soil: from genes to ecosystems. *Curr. Opin. Microbiol.* 5:240-245.

Wang, G-Y-S., Graziani, E., Waters, B., Pan, W., Li, X., McDermott, J., Meurer, G., Saxena, G., Anderson, R. J. and Davies, J. 2000. Novel natural products from soil DNA libraries in a streptomycete host. *Org. Lett.* 2:2401-2404.

Weller, D. M., Raaijmakers, J. M., McSpadden Gardener, B. B. and Thomashow, L. S. 2002. Microbial populations responsible for specific soil suppressiveness to plant pathogens. *Annu. Rev. Phytopathol.* 40:309-348.

Whitman, W. B., Coleman, D. C. and Wiebe, W. J. 1998. Prokaryotes: the unseen majority. *Proc. Natl. Acad. Sci. USA* 95:6578-6583.

Woese, C. R. and Fox, G. E. 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. USA* 74:5088-5090.

Woese, C. R., Kandler, O. and Wheelis, M. L. 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. USA* 87:4576-4579.

Zhou, J., Bruns, M. A. and Tiedje, J. M. 1996. DNA recovery from soils of diverse composition. *Appl. Environ. Microbiol.* 62:316-322.