

# Environmental Genomics: Application of DNA Microarray Technology to Environmental Microbiology

조 재 창

NSF Center for Microbial Ecology, Michigan State University

## Genomics, Bioinformatics and DNA Microarray

최근 미생물 genome sequencing 프로그램의 결과로 방대한 양의 미생물 genome sequence 정보가 수집되어 가고 있다. 약 20종의 미생물 genome 분석이 이미 완성되었고, 현재 진행 중인 미생물 genome 프로젝트도 대략 100종이 더 되는 것으로 추정된다. Microbial genomics 시대의 시작을 알리는 이러한 연구 성과들을 바탕으로 다음 단계의 과제라면 방대한 양의 자료들로부터 유용한 정보를 추출해 내는 것이라고 할 수 있겠다. 그리고 그 정보들을 생태학적 관점에서 해석하고 genomics technology를 이용하여 환경미생물학 및 다양한 응용 분야들의 많은 질문들을 풀어 나갈 수 있을 것이다. 본 총설에서 DNA microarray를 환경미생물학과 미생물생태학 분야에 소개하고, 저자의 연구 결과를 토대로 몇 가지 적용 가능한 예와 현시점에서의 방법론적 한계에 대해 기술함으로써 DNA microarray 기법 및 genomics technology가 환경미생물학 분야에 유용한 도구의 하나로 도입되기를 기대한다.

DNA microarray는 많은 종류의 DNA sequence들을 고정화 기저(microscope slide 또는 wafer)상에 규칙적으로 고정화(immobilization)하여 배열시킨 것으로서, DNA probe들의 현미경적 크기의 정렬(array)이라고 할 수 있겠다. 기본적으로 Southern hybridization의 원리를 이용하며, 시료 내에 probe들에 상응하는 DNA/RNA sequence들이 있는지를 감지하게 된다. 크게 cDNA microarray와 oligonucleotide microarray 두 종류가 있다. cDNA array는 partial(expressed sequence tag; EST) 또는 full-length complementary DNA(cDNA) sequence들로 구성되며 probe들은 대개 PCR법으로 제작된다. Oligonucleotide array는 15-40 mer 정도로 구성되며 특정한 coding region에 hybridization되도록 설계된다. Oligonucleotide array의 경우에는 한 유전자당 10~20개의 probe가 사용되며 mismatch probe들이 함께 이용될 경우도 있다. Microarray를 사용함으로써 얻을 수 있는 장점들은 다음과 같다. Probe들을 고밀도로 배열할 수 있기 때문에 수 만개 이상의 hybridization

을 한번의 실험으로 가능하게 하고, hybridization solution의 양을 줄이고 시료의 농도를 높일 수 있으며, reaction kinetics를 증진시킬 수 있다. Genome이 sequencing 되어 있는 경우, genome에 대한 global information이나 global gene expression pattern이 분석될 수 있다. 자동화된 설비를 이용함으로써 대량 생산이 가능하고 data 수집 software나 system도 함께 개발 가능하다. Multiple-color fluorescence dye를 이용한 hybridization이 가능하기 때문에 독립적인 실험 결과들을 함께 비교할 수 있다.

이제 microarray hybridization은 PCR이나 DNA sequencing이 생물학에 기여했던 만큼이나 큰 역할을 하고 있다. 개발 초기 단계에는 유전자의 기능을 밝히기 위해 사용되었으며[1, 2], *Saccharomyces cerevisiae*의 full sequence를 사용한 연구가 완성된 적이 있다[3, 4]. 그 후 특정한 유전자의 mutation[5], 유전질환의 원인 유전자 탐색[6], 바이러스 유전자의 발현이나 변이[7] 등을 연구하는데도 이용되었다. Oligonucleotide array를 이용한 DNA sequencing이나 genotyping 등도 보고되었다[8]. Affymetrix(Santa Clara, California)가 photolithography 방법으로 제작한 oligonucleotide array인 GeneChips를 상업화하였고, 현재 인간이나 yeast의 open reading frame(ORF)들을 연구하기 위한 GeneChip을 판매하고 있으며, 이들을 이용하여 gene expression, polymorphism, genotyping 등의 연구가 가능하다. Prokaryote의 genome 연구를 위한 microarray는 아직 초기 단계이며, 미생물학, 환경미생물학 그리고 미생물생태학 분야에서는 새로운 mutation이나 항생제의 metabolic effect, 특정 유전자 검색, 환경시료에서 DNA sequence나 genome의 탐색 및 검출, pathogenicity에 관련된 gene regulation 연구, bioremediation이나 biogeochemical process에 관련된 대사경로나 regulatory network의 연구, evolutionary divergence 연구를 위한 natural population의 screening, community structure 분석 등에 응용될 수 있을 것으로 사료된다. 하지만, 시장규모 때문에 soil bacteria 같은 environmental isolate들의 연구를 위한 상품화된 DNA microarray를 기대하기는 힘들고 연구자들마다 각각의 연구 목적에 맞는 array를 설계하고 제작

해야 될 것이다.

## Microarray Basics

DNA microarray 설계와 연구에는 다양한 scheme들이 가능하기 때문에, 여기서는 환경미생물학이나 미생물생태학 분야의 연구를 위한 microarray의 설계 및 제작에 필요한 내용만을 기술한다. 세부적인 내용은 cDNA microarray를 이용한 gene expression 연구를 다룬 Eisen과 Brown의 논문[9]을 참조하기 바란다. Microarray project의 기본적인 흐름도를 Fig. 1에 제시하였다. 어떤 실험 요소들이 필요한가에 대한 perspective를 위해서 실험 설계의 일반적인 내용들만을 나열하였다. Array 설계와 수천 또는 그 이상의 probe들을 수집(oligonucleotide synthesis, PCR) 하기 위해선 많은 시간과 비용이 소요되기 때문에, 먼저 research question을 충분히 고려한 다음, 그 실험 목적에 부합하는 array format을 결정해야 할 것이다. 실험 목적은 앞서 언급한 gene expression analysis, genotype analysis, community composition(structure) analysis 등이나 또는 이들을 응용시킨 것일 수 있겠다. 다음 단계에서는 어떤 probe들을 이용할 것인지를 결정한다. Oligonucleotide array인 경우 주로 database(i.e., RDP, GeneBank)에서 얻어진 정보를 이용하게 된다. 그 외 cDNA library(ESTs)나 특정 PCR primer를 사용해서 얻어진 environmental library가 이용될 수 있다. 아직 모든 bacteria의 genome이 sequencing되어 있지

않으므로 environmental isolates의 ORF들을 probe로 사용하기는 힘들다. 저자의 연구에서는 genome sequence를 모르는 bacteria의 경우, ORF에 상응하는 random genome fragments를 array probe로 사용한 적이 있다[10]. 이와 같이 shotgun cloning 등의 방법으로 얻어진 genome fragment들은 sequencing 되지 않은 bacteria나 environmental isolate들로부터 probe를 수집하는데 유용하며, 특히 환경 시료일 경우 각각의 실험 목적이나 대상 환경에 맞는 array를 제작 할 수 있다는 장점이 있다. Probe들에 대한 정보가 수집되고 probe들의 종류가 결정되면 probe를 제작하게 되는데, oligonucleotide probe의 경우 직접 합성법을 이용하고 그 외의 경우는 대부분 PCR법을 이용하게 된다. 후자의 경우, PCR product의 quality control이 매우 중요하며 product들의 농도와 순도 모두 모든 microarray 실험 결과에 중요한 영향을 주게 된다.

Array fabrication은 대부분 robot을 이용한 자동화된 설비를 이용한다. Bowtell이 array printer들의 제작에 필요한 다양한 구성요소들의 list를 정리하였다[11]. Oligonucleotide array인 경우 실리콘 wafer상에서 photolithography법으로 합성시키거나 polyacrylamide로 코팅된 microchip이 사용된다. 그 외의 경우는 다양한 물질들로 표면 처리된 glass slide가 사용된다.

다음은 분석할 시료의 준비 단계이다. Bacterial cell들의 수집이나 환경시료 수집 후, 시료들로부터 DNA/RNA를 추출하게 된다. 환경 시료의 경우 DNA/RNA 농도가 낮고 추출 후에도 순도가 낮으므로, 충분한 labelling reaction이 가능하도록 microarray hybridization 기준에 적합한 DNA/RNA를 얻는 방법이 필요하다. 일반적으로 PCR법의 주형 DNA로 사용될 만한 수준의 순도가 요구된다. 토양을 포함한 다양한 환경 시료들로부터 고순도의 DNA/RNA를 추출하는 방법들이 보고되었으며[12, 13], 이러한 방법들을 microarray hybridization용 target DNA/RNA 준비에 사용할 수 있다. DNA/RNA 시료가 준비되면 이들을 fluorescent dye를 이용하여 labelling하는데, reverse transcriptase, Klenow DNA polymerase, 또는 direct chemical labelling 등의 방법들이 사용된다. 또한 두 개의 시료가 각각 다른 색깔의 fluorescent dye로 labelling 된 다음 함께 분석될 수도 있으며, 세 개 이상의 dye를 사용하는 방법들도 개발 중에 있다. 현재 두 개의 dye를 이용하여 두 개의 다른 배양(또는 환경) 조건하에서의 gene expression profile 분석이나, 대조군-시료 쌍의 분석 등이 가능하나, 앞으로 multiple color labelling이 가능하게 되면 보다 다양한 환경 조건 등을 대상으로 연구하는 것이 가능할 것으로 기대된다.

Hybridization 과정은 기본적으로 Southern type hybridization과 유사하며, hybridization 반응의 specificity, normalization, sensitivity는 microarray와 hybridization buffer에 포함되어 있는 internal standard들을 이용해서 control된다. Hybridization

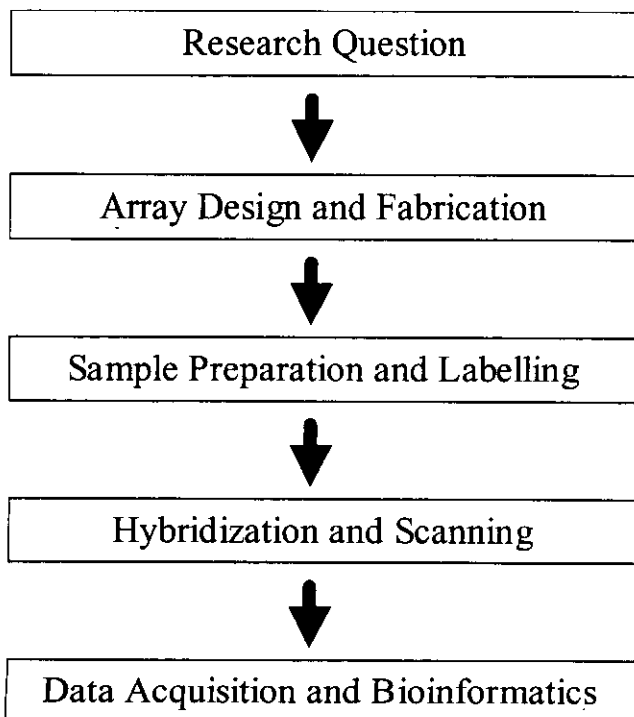


Fig. 1. DNA microarray experimental strategy.

결과는 confocal scanning laser나 CCD camera 등으로 구성된 microarray scanner로 검색되며, 화상분석 software로 분석된다. 화상분석은 화소(pixel) 단위로 이루어지며 background subtraction 후, 각 pixel들의 hybridization intensity가 계산된다. Microarray hybridization 결과는 항상 두 개의 다른 fluorescent dye로 표지된 DNA/RNA간의 상대적 hybridization 비율(ratio)로만 표시된다. Array fabrication시에 실제 immobilization되는 probe의 양은 probe 마다 조금씩 다르며, hybridization 역시 소량의(주로 microliter scale) hybridization buffer로 mixing없이 component들의 diffusion에만 의존하여 수행되기 때문에, hybridization rate도 array의 부분마다 조금씩 다르다. 하지만, 해당 probe에 반응하는 두 개의 다른 target DNA/RNA의 competitive hybridization 정도는 probe의 양이나 hybridization rate에 비교적 영향을 덜 받기 때문에, microarray hybridization 결과는 단위가 없는 competitive hybridization ratio들로 표시된다. 화소 단위들로부터 ratio를 얻는 여러 가지 방법이 있다. 일반적으로 화소들의 각 laser channel 별로 평균값이나 중심값(median)을 얻은 다음, 이들로부터 ratio들을 계산하는 방법이 사용된다. 저자의 연구결과에 따르면 pixel들의 전체 평균값이나 중심값을 사용하지 않고, pixel 별로 channel간 signal intensity ratio를 계산한 다음, 그 ratio들의 중심값을 이용하는 것이 background signal이나 nonspecific signal의 영향을 덜 받는 것으로 나타났다[10]. 평균값보다는 중심값의 이용이 선호되는 이유는 dust particle 등이나 기타 nonspecific signal의 영향을 덜 받게 하기 위함이다. 최종적으로 ratio들로 구성된 data matrix를 얻은 후, 다양한 분석에 들어간다. Microarray data의 많은 경우 ratio data의 분포(distribution)가 편중(skewed)된 경우가 많기 때문에, 통계 분석 전에 log transformation같은 normalization 과정을 수행하기도 한다.

Microarray 실험의 일반적인 내용들을 실험 설계부터 결과를 얻는 과정까지 살펴보았다. 대부분의 microarray 연구가 Fig. 2a와 같은 scheme을 사용한다. 이를 이용한 대표적인 예는 DeRisi 등이 행한 gene expression analysis이다[3]. DeRisi 등은 starved cell과 nutrient-rich condition에서 배양된 cell로부터 추출된 mRNA population들을 비교하였는데, 두 개의 mRNA population을 fluorochrome Cy3와 Cy5로 표지하고 microarray와 hybridization한 후, 상대적인 발현 정도를 hybridization ratio 값으로부터 분석하였다. 동일한 scheme이 환경미생물학이나 미생물생태학 연구에도 이용될 수 있다. 가능한 적용분야들에 대한 내용은 다음절에 기술한다. 한편 competitive hybridization을 이용하지 않고, direct binding을 이용한 방법이 저자의 연구에서 시도된 바 있다[14]. Fig. 2b에 direct binding version을 소개하였다. 이는 환경 시료로부터 target DNA/RNA를 정량적으로 검출하기 위한 목적으로 생물산업

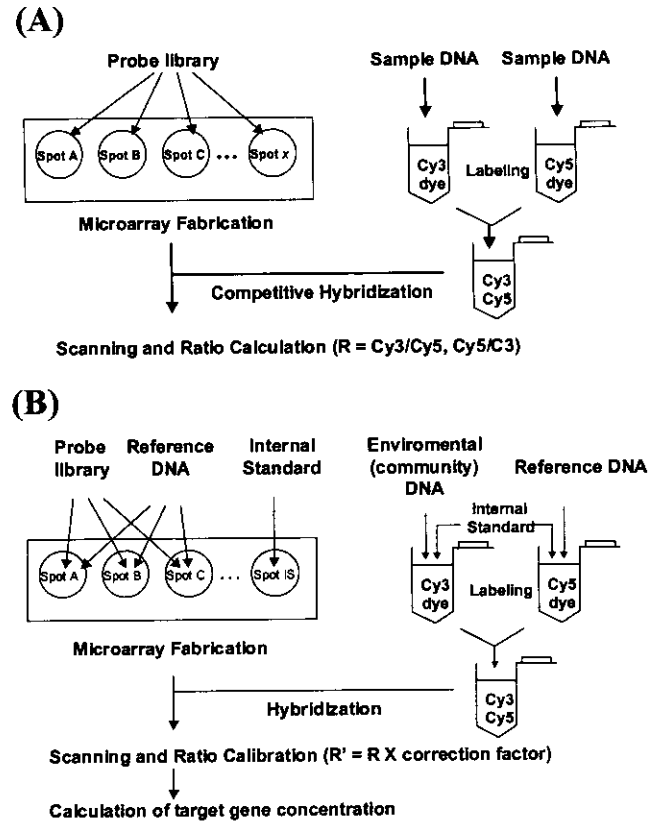


Fig. 2. DNA microarray formats and hybridization schemes using competitive hybridization(a) and direct binding(b).

개발되었으며, competitive hybridization version에서와 같은 pairwise hybridization이 필요하지 않다. Reference DNA는 probe들의 signal을 표준화(standardization)할 수 있는 어떠한 DNA라도 가능하며, reference DNA와 probe를 혼합한 다음 array로 fabrication 시키고 시료와 reference DNA/RNA의 mixture를 target DNA/RNA로 사용한다. Fig. 2b와 같은 direct binding version의 또 다른 장점은 두 개 이상의 독립적인 실험 결과들이 두 개 이상의 fluorochrome을 쓰지 않고도 함께 분석 가능하다는 점이다. 현존하는 대부분의 microarray scanner들이 두 개의 channel을 scanning하도록 고안 되어있고 실험 방법들도 두 개의 fluorochrome을 사용하는 것으로 되어있다는 것을 고려할 때, direct binding version이나 이의 변형된 형태들이 다양한 조건이나 환경의 분석이 요구되는 환경미생물학과 미생물생태학 분야에 유용할 것으로 사료된다.

## Applications to Environmental Microbiology

Microbial genomics로부터 환경미생물학과 미생물생태학 분야가 혜택을 볼 수 있는 여러 가지 부분이 있다. 앞서 언급한 바와 같이 microarray는 주로 gene expression pattern의 분석, 새로운 metabolic pathway나 regulatory network들을

탐색하는 microbial functional genomics 연구에 주로 사용되고 있으나, 이들 연구로부터 얻어진 기본적인 지식과 발견들이 환경미생물학 분야에서도 유용한 정보가 될 것이다. Microbial genome sequencing의 결과들이 몇 종류 미생물들의(*Bacillus subtilis*, *Escherichia coli*, *Haemophilus influenzae*, *Helicobacter pylori*, *Mycobacterium tuberculosis*, *Pseudomonas aeruginosa*, *Shewanella putrefaciens*) 모든 open reading frame(ORF)을 포함하는 array의 설계에 사용되었다. 상당한 연구가 진행되었고, *E. coli*의 경우 약 60% 이하의 genome만이 이미 알려진 gene들과 homology를 보이고 있으며, 그 외의 종들 역시 40~60% 이하가 될 것으로 추정되고 있다. 이러한 functional genomics의 결과들과 자연계에 존재하는 미생물들의 방대한 diversity를 고려한다면 미래의 생물 자원의 하나로서 미생물들의 가치가 얼마나 중요한가를 가늠케 한다. Torsvik 등의 연구 결과에 따르면 토양 1 gram내에는 약 3000개 이상 다른 종류의 genome이 존재한다고 하였으며[15], 이는 Bergey's manual에 수록된 모든 세균 종의 종류에 상응하는 양이다. 저자의 연구에서도 동일한 토양시료에서 다른 종류의 bioremediation treatments만을 수행하더라도 해석하기 힘들만큼의 복잡하고 다양한 community structure변화가 일어남을 보고하였었고[16, 17], 잘 알려진 세균의 하나인 *Pseudomonas*의 경우만도 방대한 양의 genotype들이 존재하며 그 genotype들은 DNA homology에 관계없이 모두 다른 genome structure를 갖는 것으로 밝혀졌다[18]. Functional genomics 역시 environmental genomics와 함께 연구될 때 그 진정한 의미와 가치를 부각시킬 수 있을 것으로 생각된다.

환경미생물학은 미생물생태학적 지식들의 응용이라고 할 수 있으며, 미생물생태학의 궁극적인 목표는 환경에서의 미생물 군집구조(structure)와 기능(function)의 파악이라고 할 수 있겠다. 다시 말하자면 어떠한 미생물들이 얼마의 양만큼, 어떤 조합으로, 어떻게 존재하고, 각각 어떤 기능들을 수행하고 있으며, 그 군집구조와 기능이 어떻게 유지되고 있는가를 파악하는 것이다. 환경 시료로부터 병원성 세균이나 분해균주, 유용한 물질들을 생산하는 미생물들을 검출하고 탐색하는 작업이나, bioremediation 또는 bioreactor내에 있는 오염물질의 분해에 관련된 미생물들을 추적하는 작업은 미생물 군집 중에서 특정 종류의 미생물을 정량적으로 관찰하는 것이라고 할 수 있겠다.

다양한 종류의 microarray들이 환경미생물학과 미생물 생태학 분야에 이용될 수 있을 것으로 사료된다. 크게 다섯 가지로 구분하여 정리하여 보았다. 첫째, type strain들이나 environmental isolate들의 genome들로 구성된 community genome array가 가능할 것이다. 이는 토양, sediment, water, bioreactor 등의 미생물 군집 composition이나 dynamics를 연구하는데 이용 가능할 것으로 사료된다. 각 probe는 whole genome DNA-

DNA hybridization을 수행하며 reverse sample genome probing(RSGP)[19]의 microarray version이라고 할 수 있겠다. Microarray상에 genome들을 probe로 사용한 예는 아직 보고된바 없지만, ORF array와 hybridization kinetics만 크게 다르지 않다면 가능할 것으로 사료된다. 또한 유용성 여부는 대상 환경의 microbial diversity와 biomass의 양에 달려있으며, 단점으로는 배양 가능한 세균들만이 연구 대상이 될 수 있다는 것이다. 자연계에 존재하는 대부분의 세균들이 아직 배양 불가능하나, 앞으로 배양 기술들이 발전하게 되면 이 문제는 쉽게 해결될 것으로 기대된다. 현 시점에서 community genome array는 RSGP의 대응으로 biomass양이 크고 diversity가 낮은 환경에는 적용 가능할 것으로 생각되며, 특히 bioreactor등에 적용할 경우 biodegrader들의 composition과 dynamics를 연구하는데 유용할 것으로 사료된다. 약간 다른 접근 방법이지만, 저자의 경우 앞서 언급한 genome fragment로 구성된 shotgun array를 이용하여 bacterial species 동정법과 genome similarity index에 관한 연구를 수행한 바 있는데 [10, 20], 유사한 접근 방법을 이용한다면 whole genome에 기초한 RSGP에서 제기되었던 cross-hybridization 등에 의한 문제점들을 해결할 수 있을 것이며 specificity를 향상시킬 수 있을 것이다.

둘째, 미생물 군집구조 분석을 위한 SSU rDNA array가 가능할 것이다. 약 12,000개 이상의 prokaryotic sequence가 수록되어 있는 RDP database(<http://rdp.msu.edu/RDP>) 등으로부터 얻은 sequence정보를 바탕으로 oligonucleotide array를 제작할 수 있으며, Affymetrix 등에서 제공하는 GeneChip기법 등이 사용될 수 있다. 우선 ribosomal DNA sequence들의 phylogenetic framework가 선행되어야 할 것이다. Sequence conservation정도에 따라 highly conserved sequence들은 넓은 범위의 taxonomic group들을, 그리고 hypervariable sequence 들은 genus나 species수준의 taxonomic group들을 대상으로 하는 oligonucleotide probe들로 이용될 수 있을 것이다. 장점으로는 community genome array와는 달리 대상 생태계로부터 균주들을 분리하지 않아도 된다는 점이고, 한번의 설계와 제작으로 만들어진 array를 다른 종류의 생태계와 환경에도 이용할 수 있다는 점이다. 단점으로는 rDNA를 이용한 방법의 resolution 문제이다. 많은 연구자들이 16S rDNA나 기타 다른 gene들의 sequence 정보를 미생물 동정 등에 이용하고 있지만, 대부분의 경우 그 resolution이 community genome array에서 사용하는 whole genome DNA-DNA hybridization에 미치지 못하고 있다. 예를 들어 *Pseudomonas*(sensu stricto)의 경우 모든 species들의 16S rDNA 유사도는 >93%이며, 몇 종의 species들은 99% 이상의 유사도를 보이고[21], 16S rDNA sequence similarity와 DNA homology와의 관계는 log-function이라고 보고되었다[22, 23]. 오염물질 분해, 병원성,

생리활성물질과 항생제 생산 등과 같은 미생물들의 많은 중요한 특성들이 subspecies 이하의 수준에서 결정된다는 것을 고려하면, single gene sequence의 phylogeny를 바탕으로 미생물의 기능을 유추하기 위해선 많은 주의가 필요하다. SSU rDNA array 역시 이러한 문제점들을 모두 가지게 될 것으로 생각되며, 그 유용성과 versatility에 반해 community genome array 보다는 낮은 resolution을 보일 것으로 사료된다.

세 번째로는 environmental functional gene array가 가능하다. Microarray expression profile은 유전자들의 functional information을 이해하는데 매우 좋은 도구가 된다. 이러한 functional information은 고도로 발현되는 유전자들이나 biogeochemical cycling, bioremediation, 그리고 biocatalysis에서 중요한 기능을 담당하는 유전자들을 찾아내는데 유용할 것이다. Cell regulatory function 역시 carbon supply, energy source, electron acceptor의 양과 종류가 다른, 다양한 환경 조건하에서의 regulation을 분석한다는 측면에서 연구될 수 있을 것이다. 그 외 quorum sensing, chemotaxis, 항생제 생산과 같은 환경 인자들에 의해 조절되는 기능들도 대상 pathway의 regulation과 genetic expression pattern등을 바탕으로 연구될 수 있을 것이다.

두 가지 방식의 array가 가능할 것으로 생각된다. 먼저 gene expression 분석을 위한 oligonucleotide array이다. Bioremediation이나 biogeochemical cycling에 관여하는 enzyme를 coding하고 있는 유전자들에 complementary한 oligonucleotide probe들을 사용하는 것이다. Array의 제작은 oligonucleotide probe 제작에 필요한 sequence information에 달려있다. 최근에 시작한 BSD(<http://bsd.cme.edu>)와 같은 biodegradative strain들의 database와 분해 유전자들의 gene sequence 정보가 그 예가 될 것이다. SSU rDNA database와는 달리 microbial functional gene들의 database는 많이 뒤떨어져 있는 실정이다. 연구자들은 GeneBank와 같은 database나 여러 실험실들로부터 연구대상이 되는 유전자들의 정보를 수집해야 할 것이다. 환경으로부터 추출된 DNA를 위와 같은 functional gene oligonucleotide array의 target DNA로 사용한다면 대상 유전자들의 검출이나 sequence variation 측정에 사용될 수 있을 것이며, mRNA population이 사용된다면 expression pattern을 분석하는데 이용될 수 있을 것이다. 다른 방식으로는 ORF array와 같이 대상 유전자들의 PCR product들을 probe로 사용하는 것이 가능할 것이다. 분리된 균주들을 이용하여 PCR product를 얻을수도 있고, 대상 환경으로부터 연구에 필요한 유전자들의 library를 community DNA의 PCR을 이용하여 직접 얻는 방법도 가능할 것이다. 이 경우 gene expression profile과 함께 대상 유전자들의 diversity 역시 동시에 분석 가능할 것으로 사료된다. ORF를 이용한 environmental functional gene array의 또 다른 예로서, 예를 들면 bioremediation시의

분해균주로 사용되는 미생물의 ORF들로 구성된 array를 이용하여 bioremediation이나 bioreactor 가동중의 expression profile을 monitoring 한다면 오염물질의 분해를 최적화 하거나 접종 균주의 생존을 증진시킬 수 있는 환경조건을 파악하는데 유용할 것이라고 사료된다. 이 경우 대상 균주의 genome이 sequencing되어 있지 않더라도 앞서 언급한 random genome fragment를 이용하여 array를 제작한 다음, expression profile이나 분해 유전자 검출 등의 결과 수집 후에 선별된 genome fragment들을 sequencing함으로써 identity를 알아내는 접근 방법을 사용할 수도 있을 것이다.

Environmental functional gene array의 적용에는 하나의 풀어야 할 과제가 있다. Microarray hybridization을 위해선 비교적 많은 양의 RNA가 필요한데 효과적인 추출방법의 개발이나 hybridization signal detection 분야의 발전이 필요하다. 또 다른 고려 사항은 환경 시료마다 biomass의 양이 다르고 total RNA 및 mRNA의 비율도 다르기 때문에, 이들을 함께 분석하기 위한 data normalization과 같은 통계학적인 기법의 개발도 필요할 것으로 사료된다.

네 번째로는, 특정 population의 genetic polymorphism이나 genetic diversity연구를 위한 population biology array가 가능할 것이다. Oligonucleotide probe들을 이용한 DNA sequencing이나 ORF array의 competitive hybridization을 이용한 접근 방법이 가능하며, 유사한 연구가 *M. tuberculosis*[8]와 *S. cervisiae*[24], human cytomegalovirus[7]을 대상으로 수행된 적이 있다.

다섯 번째로는, 환경에서의 특정 유전자 검출/검색용 array가 가능할 것이다. 예를 들어 병원성 세균들의 signature sequence들로 array로 만들게 되면, 음용수나 식품 등에서 병원성 미생물들을 검출하는 방법으로 사용될 수 있을 것이다. Biodegradation에서 주요 기능을 담당하고 있는 유전자들로 제작할 경우에는, 앞서 언급한 environmental functional gene array로서의 기능뿐만 아니라, bioremediation을 수행하기에 앞서서 처리 방법들의 선택에 관한 판단에 큰 도움을 줄 수 있을 것이다. 예를 들면 대상 환경이 분해자 또는 분해 유전자를 포함하고 있는지 여부를 알 수 있으며, 이는 biostimulation이나 bioaugmentation과 같은 bioremediation시의 큰 두 가지 방향을[25] 선택할 때 유용한 정보로서 사용될 수 있을 것이다. Microarray를 이와 같은 목적으로 사용하는 이유는 기존의 분자생물학적인 방법들과는 달리 방대한 양의 정보를 동시에 알아 낼 수 있다는 장점 때문이다. Microarray는 수 천 또는 수 만개 이상의 probe들로 구성될 수 있고, 모든 반응이 한번의 실험으로 완성되기 때문에 중요한 병원성 미생물들의 모든 signature sequence들의 array를 사용할 경우, public health 분야에는 큰 기여를 할 것으로 사료된다. 저자의 연구에서는 *E. coli* O-antigen biosynthesis gene, naphthalene degradation

gene, denitrification gene 등으로 구성된 microarray를 model 생태계에 적용하여 hybridization kinetics 및 detection limit 등에 관한 연구를 수행한 바 있다[14]. 기존의 hybridization kinetics와는 다른 양상이 관찰되었으며 log signal ratio와 log gene mass ratio간의 선형관계가 관찰되었다. 토양 시료의 경우 검출한계는 유전자의 경우 pg, genome의 경우는 ng 수준으로 나타났다. 위 연구는 competitive hybridization이 아닌 Fig. 2b에 제시한 direct binding version을 이용하여 수행되었으며, direct binding version이 다섯 번째 예와 같은 응용 분야에는 유용할 것으로 사료된다.

방대한 양의 정보검색이라는 큰 장점과는 달리 microarray hybridization은 여전히 traditional hybridization법을 이용하고 있다. 때문에 결과를 얻기까지는 보통 12~16 시간의 hybridization 반응이 필요하고, microarray scanning 또한 시간이 요구되며 scanning 장비는 고가이다. 긴 hybridization 시간은 신속한 검출이 요구되는 pathogen detection용 DNA microarray에서는 중요한 문제점으로 지적될 것이다. 저자의 연구에서는 microchip상에서 electronic hybridization법과 electronic hybridization signal detection법을 제시한 바 있다 [26]. 약 1 microliter 정도의 hybridization buffer를 이용하여 2~3분 안에 hybridization 결과를 얻을 수 있고, 얻어진 data는 바로 microarray 분석용 software로 전송됨으로써 모든 과정을 자동화하는 것이 가능할 것으로 기대된다. Nanogene과 Motorola에서도 부분적인 electronic system들의 beta test들을 수행하고 있는데, 이와 같은 electronic hybridization에 대한 연구들이 성공적으로 마무리될 경우, 실험 속도면에서도 큰 혜택을 볼 수 있을 것으로 기대된다. 또한 여러 가지 조건의 실험들을 짧은 시간 안에 반복해서 실행할 수 있기 때문에 실험 결과의 정확도 역시 향상될 수 있을 것으로 생각된다.

DNA microarray와 같은 genomics technology의 환경미생물학과 미생물생태학적 응용에 대한 저자의 의견과 제안을 저자의 연구 결과들을 토대로 간략히 소개하였다. 본 총설에서 제시한 내용 외에도 다양한 접근 방법들이 가능할 것이며, 연구자들마다 연구 목적과 방향에 맞는 변형된 array format과 hybridization scheme이 필요할 것이다. Genomics technology는 생물학의 전반적인 분야에 새로운 연구 도구의 하나로 도입되기 시작했고, 앞으로 그 기술 자체에도 큰 발전이 있을 것으로 사료된다. 환경미생물학과 미생물생태학 연구에서도 genomics technology가 성공적으로 도입되어 유용한 도구로서 사용되기를 기대한다.

## References

1. Schena, M., D. Shalon, R. W. Davis, and P. O. Brown. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**:

- 467-470.
2. Schena, M., D. Shalon, R. Heller, A. Chai, P. O. Brown, and R. W. Davis. 1996. Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc. Natl. Acad. Sci. USA.* **93**: 10614-10619.
3. DeRisi, J. L., V. R. Lyster, and P. O. Brown. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278**: 680-686.
4. Wodicka, L., H. Dong, M. Mittmann, M. Ho, and D. J. Lockhart. 1997. Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat. Biotechnol.* **15**: 1359-1367.
5. Hacia, J. C., L. C. Brody, M. S. Chee, S. P. Fodor, F. S. Collins. 1996. Detection of heterozygous mutations in BRCA1 using high density oligonucleotide arrays and two-colour fluorescence analysis. *Nat. Genet.* **14**: 441-449.
6. Heller, R. A., M. Schena, A. Chai, D. Shalon, T. Bedilion, J. Gilmore, D. E. Woolley, and R. W. Davis. 1997. Discovery and analysis of inflammatory disease-related genes using cDNA microarrays. *Proc. Natl. Acad. Sci. USA.* **94**: 2150-2155.
7. Chambers, J., A. Angulo, D. Amaratunga, H. Guo, Y. Jiang, J. S. Wan, A. Bittner, K. Frueh, M. R. Jackson, P. A. Peterson, M. G. Erlander, P. Ghazal. 1999. DNA microarrays of the complex human cytomegalovirus genome: profiling kinetic class with drug sensitivity of viral gene expression. *J. Virol.* **73**: 5757-5756.
8. Gingeras, T. R. 1998. Simultaneous genotyping and species identification using hybridization pattern recognition analysis of generic *Mycobacterium* DNA arrays. *Genome Res.* **8**: 435-448.
9. Eisen, M. B., and P. O. Brown. 1999. DNA arrays for analysis of gene expression. *Methods Enzymol.* **303**: 179-205.
10. Cho, J. C., and J. M. Tiedje. 2001. Bacterial species determination from DNA-DNA hybridization using genome fragments and DNA microarrays. *Appl. Environ. Microbiol.* *Accepted.*
11. Bowtell, D. D. L. 1999. Options available - from start to finish - for obtaining expression data by microarray. *Nat. Genet.* **21**: 25-32.
12. Cho, J. C., D. H. Lee, Y. C. Cho, J. C. Cho, and S. J. Kim. 1996. Direct extraction of DNA from soil for amplification of rRNA gene sequences by polymerase chain reaction. *J. Microbiol.* **34**: 229-235.
13. Chandler, D. P., J. R. Stults, S. Cebula, B. L. Schuck, D. W. Weaver, K. K. Anderson, M. Egholm, and F. J. Brockman. 2000. Affinity purification of DNA and RNA from environmental samples with peptide nucleic acid clamps. *Appl. Environ. Microbiol.* **66**: 3438-3445.
14. Cho, J. -C., and J. M. Tiedje. 2001. Detection of

- environmental genes using DNA microarray. Abstract Q-140. 101st General Meeting of the American Society for Microbiology. May 20-24. Orlando, Florida.
15. Torsvik, V., K. Salte, R. Sorheim, and J. Goksoyr. 1991. Comparison of phenotypic diversity and DNA heterogeneity in a population of soil bacteria. *Appl. Environ. Microbiol.* **56**: 776-781.
  16. Cho, J. C., and S. J. Kim. 2000. Computer-assisted PCR-single-strand-conformation polymorphism analysis for assessing shift in soil bacterial community structure during bioremediational treatments. *W. J. Microbiol. Biotechnol.* **16**: 231-235.
  17. Cho, J. C., and J. M. Tiedje. 2000. Biogeography and degree of endemism of fluorescent *Pseudomonas* strains in soil. *Appl. Environ. Microbiol.* **66**: 5448-5456.
  18. Tiedje, J. M., J. C. Cho, A. Murray, D. Treves, B. Xia, and J. Zhou. 2001. Soil teeming with life: New frontiers for soil science. pp.393-412. R. M. Rees, B. C. Ball, C. D. Campbell, and C. A. Watson(eds) Sustainable Management of Soil Organic Matter. CABI publishing.
  19. Greene, E. A., J. G. Kay, K. Jaber, L. G. Stehmeier, G. Voordouw. 2000. Composition of soil microbial communities enriched on a mixture of aromatic hydrocarbons. *Appl. Environ. Microbiol.* **66**: 5282-5289.
  20. Cho, J. C. and J. M. Tiedje. 2001. Bacterial identification chip based on DNA-DNA hybridization. Invention Disclosure. Office of Intellectual Property, Michigan State University.
  21. Cho, J. C., and J. M. Tiedje. 2000. DNA relatedness of world-wide collection of fluorescent *Pseudomonas* genotypes. Abstract N-171. 100th General Meeting of the American Society for Microbiology. May 21-25. Los Angeles, California.
  22. Devereux, R., S. H. He., C. L. Doyle, S. Orkland, D. A. Stahl, J. LeGall, and W. B. Whitman. 1990. Diversity and origin of *Desulfovibrio* species: phylogenetic definition of a family. *J. Bacteriol.* **172**: 3609-3619.
  23. Keswani, J., and B. Whitman. 2001. Relationship of 16S rRNA sequence similarity to DNA hybridization in prokaryotes. *Int. J. Syst. Evol. Microbiol.* **51**: 667-678.
  24. Ferea, T. L., D. Botshein, P. O. Brown, and R. F. Rosenzweig. 1999. Systematic changes in gene expression patterns following adaptive evolution in yeast. *Proc. Natl. Acad. Sci. USA* **96**: 9721-9726.
  25. Cho, J. C., and S. J. Kim. 1997. Biodegradation of phenanthrene in soil microcosms. pp.239-244. In B. C. Alleman, and Leeson, A.(ed.), In situ and on-site bioremediation, vol. 2. Battelle press, Columbus.
  26. Cho, J. C., D. Schooltz, and J. M. Tiedje. 2001. Electronic measurement of DNA hybridization. Invention Disclosure. Office of Intellectual Property, Michigan State University.