장내미생물 균총 연구를 위한 차세대 염기서열 분석 기법의 응용

Application of next generation sequencing technology in gut microbiome analysis

임예슬 · 김근배*

Yea-Seul Lim and Geun-Bae Kim*

중앙대학교 동물자원과학과

Department of Animal Science and Technology, Chung-Ang University

I. 서론

전통적으로 특정 미생물의 동정에는 배지법(culture method)이 이용되었으나, 지난 20여년간 배지를 이용하지 않는 동정이 가능한 분자생물학적인 기법들(culture-independent molecular biology methods)이 많이 개발되었으며, 이 기술들은 미생물의 동정뿐만 아니라 다양한 환경에 서식하는 미생물의 다양성을 연구하는데 크게 기여하고 있다(Siqueira *et al.*, 2012).

특히 Sanger 등(1977)이 개발한 염기서열분석법에 의하여 미생물 균총에 대한 초기 연구가 가능하게 되었다. Sanger법은 세균의 16S rRNA를 증폭한 후 이것을 클로닝하고 염기서열을 분석함으로써 미생물 균총을 확인할 수 있는 기술이다. 이 기술은 DNA의 염기서열을 밝히고 분자생물학의 발전에 크게 기여한 기술이지만, 다양한 시료에서 얻은 많은 수의 유전자를 동시에 분석하기에는 많은 분석비용 및 시간 등 기술적 한계점을 가지고 있기 때문에 미생물 생태학을 연구하기에 어려움을 가지고 있었다.

최근 들어, 대용량 염기서열 분석이 가능한 새로운 기술이 개발되었는데, Sanger법이 "1세대 (first-generation)" 기술로 여겨지고 있기 때문에 새롭게 개발된 기술들을 "차세대 염기서열 분석 기술(Next-Generation DNA Sequencing, NGS technology)" 이라고 부르게 되었다. NGS 기술 중에서도 pyrosequencing(파이로시퀀싱) 기술은 한 번의 분석으로 엄청나게 많은 염기서열을 분석할 수 있는 장점으로 인하여 특정 시료 중에 포함되어 있는 다양한 종류의 미생물균총(microbiome)을 분석하는데 혁신적인 수단을 제공하게 되었다(Kunin *et al.*, 2010).

NGS 기술은 미생물 생태학 연구에 유용하게 이용되고 있을 뿐만 아니라 다양한 게놈프로젝

Corresponding authors: Geun-Bae Kim
Department of Animal Science and Technology, Chung-Ang University, Anseong 456-756, Korea
Tel: 82-31-670-3027
Fax: 82-31-676-5986
E-mail: kimgeun@cau.ac.kr

Table 1. Summary and comparison of next-generation sequencing technologies

Platform	Library preparation	Chemistry	Read length (bp)	Bases per run	Run time
454 GS FLX Titanium	Emulsion PCR	Pyrosequencing	400	500 Mb	10 h
Illumina/solexa	Bridge PCR	Reversible terminators	100	18-35 Gb	4-9 d
SOLiD	Emulsion PCR	Sequencing by ligation	50	30-50 Gb	7-14 d
Helicos	Single molecule	Reversible terminators	32	37 Gb	8 d
Sanger	PCR and cloning	Dye terminators	800	800 bp	3 h

(Adopted from Siqueira *et al.*, 2012)

트에서도 그 성능을 발휘하고 있다. 예를 들면, Sanger법을 이용한 첫 번째 human genome project는 13년 동안 약 27억 달러를 사용하여 2003년 완성된 반면, NGS를 사용한 두 번째 HGP는 2008년에 약 150만 달러를 사용하여 불과 5개월만에 완성되었다(Wheeler *et al.*, 2008).

본 논문에서는 차세대 염기서열 분석 기술 중에서 파이로시퀀싱에 초점을 맞추어 미생물 다양성 분석의 기본 원리를 알아보고, 이 기술을 활용한 최근의 연구동향 중에서 장내세균총이 건강에 미치는 영향에 관한 다양한 응용 예를 알아보고자 한다.

Ⅱ. 본론

1. 차세대 염기서열 분석 기술(Next-Generation Sequencing, NGS Technology)

미생물을 배지에 접종하여 배양한 다음 미생물을 동정하는 전통적인 방법으로는 한계가 있기 때문에, 최근 들어 16S rRNA 유전자 분석을 통한 분자생물학적 동정방법이 일반적으로 사용되고 있다. 그러나, DNA 추출, 16S rRNA 유전자의 증폭, 대장균 숙주 내로 클로닝, 클로닝된 증폭 유전자의 염기서열 분석을 하는 기존의 Sanger 방법은 많은 시간이 소요되고 동시에 분석할 수 있는 시료의 수에도 제한이 있다. 이를 대체하기 위해 차세대 염기서열 분석 기술이 연구 되었고 대표적으로 454 FLX(Roche, USA), HiSeq(Illumina, USA), SOLiD(Life technologies, USA) 등의 방법이 있다. 각각의 분석 방법은 분석된 염기서열의 길이, 처리량 등에 차이가 있으며 다양한 생명과학 분야에 널리 사용되고 있다(Table 1).

가. Roche (454) GS FLX Sequencer

NGS 중 가장 처음 출시된 454 sequencer는 "파이로시퀀싱" 이라는 원리에 기반한 기술이다. 파이로시퀀싱이란 DNA polymerase가 nucleotide monomer를 붙일 때 생성되는 pyrophosphate를 탐지함으로써 염기서열을 읽어내는 기술이다. Pyrophosphate가 방출되면서 sulfurylase에 의해 adenosine-5'-phosphosulfate(APS)를 adenosine triphosphate(ATP)로 바뀌는 반응이 일어나고, 이때 생성된 ATP는 luciferase에 의해 luciferin이 oxyluciferin이 되면서 빛이 발생한다. 따라서 이 반응에서 방출되는 light signal은 하나의 nucleotide의 extension이 이루어졌다는 것을 알려 준다.

Roche 454 sequencer는 clonal amplification을 위해 oil과 물의 친수성과 소수성의 원리를 이용한 emulsion PCR을 사용하며, 이 과정을 통해 증폭된 microbead에 붙어있는 DNA 단편들에서 생성되는 phosphate의 signal 강도를 laser로 확인하여 그 결과를 수치화한다(Fig. 1.).

미세가공기술을 이용해 한 개의 microbead 만이 들어갈 수 있는 reactor에 sulfurylase와 luciferase가 포함된 packing bead가 들어가 sequencing 반응을 하게 된다. 이 platform은 4

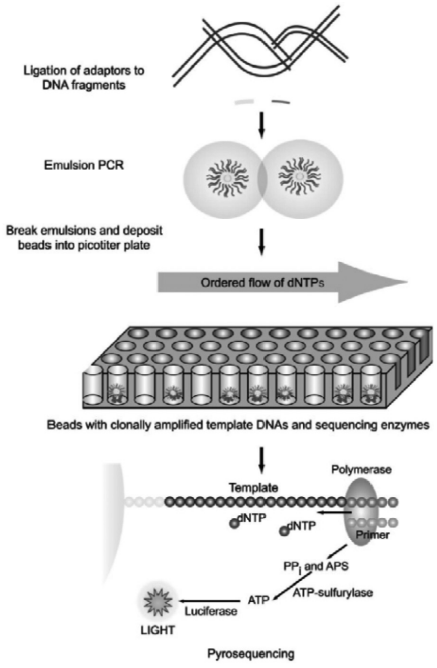

Fig. 1. Roche 454 Workflow: library construction ligates 454-specific adapters to DNA fragments and couples amplification beads with DNA in an emulsion PCR to amplify fragments before sequencing (Adopted from Voelkerding *et al.*, 2009).

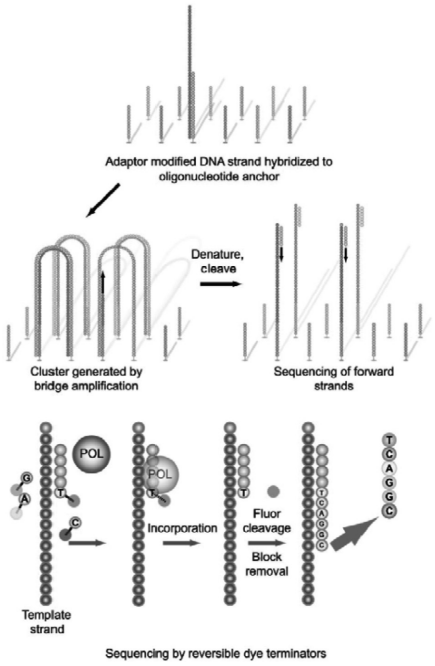

Fig. 2. Illumina Genome Analyzer sequencing (Adopted from Voelkerding *et al.*, 2009).

천만~5천만개의 sequence를 동시에 측정할 수 있으며, 한번의 run을 통해 100 Mbp를 읽어낼 수 있다. 분석된 염기서열이 다른 platform에 비해 긴 200 bp을 분석하기 때문에 시퀀싱 방법으로 가장 많이 사용되지만, 전체 처리량은 다른 두 개의 platform에 비해 떨어지는 단점이 있다.

나. Illumina Genome Analyzer

Illumina 사에서 출시한 Genome Analyzer는 원리는 Roche 454와 비슷하지만 clonal amplification에 bridge amplification 방법을 이용하는 것이 다르다(Fig. 2.).

DNA 단편들의 양 끝에 adaptor oligonucleotide를 부착시킨 후 이를 glass flow cell의 표면에 흘려주게 되면, 표면에 고정되어 있는 primer에 hybridization 하게 된다. 이 상태에서 PCR을 실시 하여 single DNA fragment로부터 유래한 cluster를 형성하고 nucleotide 마다 각각의 다른 형광 tag을 읽어 염기서열을 확인한다. 이 방법으로 한번의 run을 통해 총 130 Mbp의 sequence를 읽어낼 수 있지만, 읽어낸 염기서열의 길이는 Roche 454보다 짧은 32~40 bp으로 미생물 분류를 위한 정보량이 상대적으로 부족하다는 단점이 있다.

다. Applied Biosystems SOLiD Sequencer

가장 늦게 출시된 ABI 사의 SOLiD sequencer는 clonal amplification 방법으로는 Roche 454와 같은 emulsion PCR을 이용하지만, DNA ligase를 이용한 독특한 sequencing 방법을 사용한다. Microbead를 glass slide의 표면에 공유결합으로 붙인 뒤, 4가지 색의 형광 dye이 달린 1, 2 또는 4, 5 bp 만 염기가 정해져 있는 8 bp의 oligonucleotide가 순차적으로 ligation 되고 각각

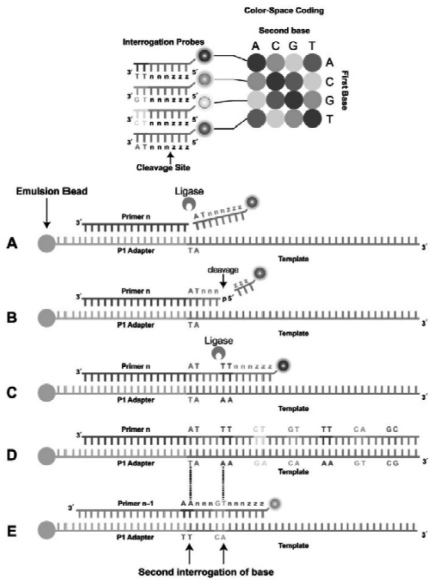

Fig. 3. Applied Biosystems SOLiD sequencing by ligation (Adopted from Voelkerding *et al.*, 2009).

의 경우 형광을 감지하여 sequencing 한다(Fig. 3.).

Signal을 검출한 후 6~8번째 base 부분과 형광 tag은 제거되고 이러한 과정을 반복하여 universal primer의 부착을 shifting하면서 보정하면 전체의 염기서열을 완성할 수 있다. 특히 각 형광 signal이 두 개의 sequence 정보를 알려주는 2-base encoding system이 이 platform의 최대 장점이며 정확한 sequencing readout을 주는 기반이 된다. ABI SOLiD의 평균 read length는 25~35 bp이며, 한번의 run을 통해 3~4 Gbp를 읽어낼 수 있다. 보통 한번의 run은 7일 정도가 소요된다고 한다.

2. 파이로시퀀싱을 이용한 다양성 분석

가. 파이로시퀀싱

여러 NGS 중에서 가장 많이 이용되고 있는 파이로시퀀싱은 앞서서 설명한 그대로 DNA 중합효소가 뉴클레오티드 한 개를 합성할 때 생성되는 pyrophosphate를 검출하여 염기서열을 해독하는 기술이다. 파이로시퀀싱은 PCR 프라이머에 독특한 염기서열(바코드)을 삽입함으로써, 시퀀싱 반응을 할 때, 동시에 많은 종류의 시료를 혼합하여 분석할 수 있다. 이 때 분석에 사용되는 유전자는 세균과 고세균에서 리보솜 RNA(rRNA)의 소 단위체(16S rRNA) 영역의 염기서열이 이용된다. 16S rRNA 유전자는 다른 유전자에 비해, 첫째, 변이가 적은 영역은 PCR 프라이머 디자인에 사용되며, 둘째, 변이가 많은 영역은 군집 구성원의 정확한 분류학적, 계통학적 동정을 할 수 있게 하고, 셋째, 분류군 사이의 수평적 유전자전이(horizontal gene transfer)가 드물며, 넷째, 현재 광범위한 16S rRNA 염기서열 데이터가 데이터베이스로 축적되어 있어, 정확한 분류가 가능하므로 다양성분석에 있어 많은 이점을 갖고 있다. 진균의 경우는 18S rRNA 유전자나 ITS 영역(Internal Transcribed Spacer region)이 이용된다.

나. 파이로시퀀스 분석 과정

파이로시퀀싱을 통한 분석 방법은 일반적으로 한번에 수 천개에서 수 만개의 염기서열을 얻을 수 있다. 여기서 얻은 데이터를 해석하기 위해 다양한 컴퓨터 프로그램과 알고리즘이 개발되었고, 이를 분석하는 방법도 많이 연구되고 있다.

1) 염기서열 선별(오류발생 염기서열 제거)

파이로시퀀싱을 통해 얻어진 염기서열은 먼저 선별 과정을 통해 PCR 과정의 오류로 발생한 오류 염기서열(artifact)을 제거한다. 보통 오류의 비율은 0.5~1.1% 인 것으로 알려져 있으나 오류를 0.5% 이하로 낮추기 위해 GS20 시스템을 이용하여 분석하기도 한다(Gilles *et al.*, 2011; Huse *et al.*, 2007). 또한 PCR 과정에서 두 종 이상의 염기서열이 합쳐져 새로운 염기서열이 생성되는 경우(이를 '키메라'라고 부름) 이

를 제거하는 과정이 필요하다. 키메라는 파이로시퀀싱으로 얻은 염기서열의 길이가 200 bp 이하였을 때는 거의 고려되지 않다가 기술의 발전으로 더 긴 염기서열까지 분석이 가능하게 되었을 때부터 중요하게 여겨졌다(Huse *et al.*, 2008). 이들로 인해 실제 미생물 다양성이 과대평가될 수 있기 때문이다. 키메라는 16S rRNA 유전자를 두 부분으로 나눈 뒤 다른 염기서열 간 유사도를 비교하여 판정하고 제거한다. 이를 제거하기 위해 많은 프로그램이 개발되었으나 아직 키메라의 완벽한 제거는 어렵다.

2) 염기서열 정렬(alignment)

다음으로 염기서열을 정렬(alignment)하는 과정을 거친다. 각 개체간의 유전자는 오랜 시간 진화를 겪으면서 염기의 종류가 변하기도 하고, 새로운 염기가 삽입되거나 삭제되기도 하며 서로 조금씩 달라졌다. 이를 비교하기 위해 염기서열들을 비교하고 gap을 삽입하여 각각의 염기가 본래의 위치 (homologous position)에 오도록 한다. 방법에는 pairwise alignment, multiple alignment, profile-based alignment가 있다(Schloss, 2009).

3) 미생물 군집구조 분석

세 번째 단계로 정렬된 시료에 들어있는 미생물들의 군집구조를 분석하는 과정을 수행한다. 분석 과정은 크게 두 가지를 기반으로 한다(Schloss and Westcott, 2011). 하나는 미생물 분류학을 기반으로 현재 데이터베이스로 가지고 있는 미생물의 유사도에 근거하여 염기서열을 동정하는 것이고, 다른 하나는 시료에서 얻은 염기서열 간의 유사도에 근거하여 조작상분류단위(Operational Taxonomic Unit; OTU)를 만들어 이들끼리 비교하여 미생물 다양성을 평가하는 것이다. 첫 번째 방법은 현재 존재하는 미생물 유사도를 이용하기 때문에 좀더 빠르게 분석할 수 있으나 아직 제대로 분석되지 않은 미생물의 경우 정확한 데이터를 얻기 어렵다는

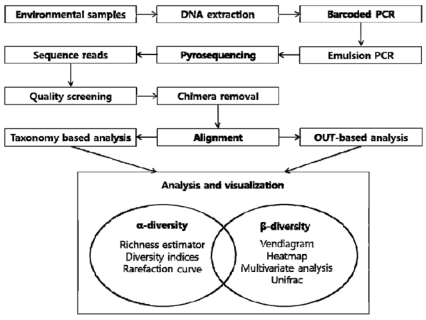

Fig. 4. General scheme of the pyrosequencing analysis of microbial community(안 등, 2012)

단점이 있다. 또한 프로그램의 분류 체계에 따라 다르게 나타날 수 있어 하나의 데이터베이스를 선택하여 진행하는 것이 유리하다. 현재 사용 가능한 데이터베이스는 RDP, GenBank, Greengenes, EzTaxon-e이 있다(안 등, 2012). 두 번째 방법인 OTU에 근거하여 미생물의 군집구조를 분석하는 과정에 사용하는 유사도는 연구자가 임의로 정의할 수 있으나, 각 염기서열 간의 거리를 계산하여 OUT를 설정해야 하기 때문에 컴퓨터에 부하가 많이 걸리게 된다. 분석과정은 각 염기서열 간의 거리를 계산하는 과정과 염기서열들을 OTU로 나누는 과정으로 되어있으며, 거리를 계산하기 전에 gap을 넣어 염기서열을 정리한다. 그 뒤 서로 간의 거리를 계산 한 뒤 그 결과를 가지고 각 염기서열들을 원하는 알고리즘을 사용하여 OTU로 묶는다. 세균 군집분석의 경우 OUT 간의 최소 거리는 일반적으로 3%가 권장된다(Kunin *et al.*, 2010).

4) 미생물 다양성 평가

마지막 단계는 위에서 얻은 결과를 바탕으로 시료 내의 미생물 다양성을 평가하고 서로 다른 시료의 미생물 군집구조를 비교한다. 파이로시퀀싱의 전체적인 과정을 요약하면 Fig. 4에서 보는 바와 같다.

3. 파이로시퀀스 분석 시 고려 사항

가. 시료의 보존

미생물 DNA는 시료 채취 즉시 추출하고 어느 정도 시간이 걸리게 되면 시료는 동결시키는 것이 좋다(Rochelle *et al.*, 1994). 하지만 한 연구에서 시간과 온도를 다르게 보관한 토양, 사람의 피부, 인분 시료를 대상으로 파이로시퀀싱 방법을 이용하여 계통 및 분류학적 군집분석을 수행하였다. 그 결과, 시료 출처와 종류가 DNA 추출 전 보존방법(온도, 시간)에 비해 미생물 군집구조에 더 큰 영향을 줄 수 있다고 보고하였다(Lauber *et al.*, 2010). Wu 등(2010)은 분석 전 시료 보존 방법과 DNA 추출방법의 영향을 조사한 결과, 1) 군집조성변이에 보존방법이 미치는 영향이 시료 간 변이에 비해 작기 때문에, 보존방법은 시험 성격에 맞게 선택될 수 있고, 2) DNA 추출 방법은 군집구조 해석에 큰 영향을 줄 수 있다는 결론에 도달하였다.

나. DNA 추출

명확한 미생물 데이터를 얻기 위해서는 시료량, 혼합, 반복 등을 고려해야 한다. 보통 미생물 DNA는 소량의 시료에서도 추출의 가능하기 때문에 추출 전 시료의 균질화 및 충분한 반복을 두어 시료가 대표성을 가지도록 해야 한다. 또한 시료는 채취에서 DNA 추출까지 사이에 시간이 짧은 것이 좋으며 다른 시료와 섞인다거나 이송, 분석 도중에 다른 시료나 공기 중에 노출로 인해 오염되지 않아야 한다. 시료가 채취된 조건에 따라 서로 다른 물리적, 화학적, 효소학적 방법 등을 사용하며 DNA을 추출하며 보통 이 세 가지 방식이 단독 또는 조합하여 사용되고 있다(안 등, 2012).

다. 증폭 조건

파이로시퀀싱은 기본적으로 염기서열 유전자 자체가 아닌 PCR을 사용하여 특정 유전자만을 증폭시킨 새로운 염기서열을 사용한다. 때문에 이 때 키메라 염기서열이 생기거나 DNA 중합효소의 오류로 염기서열이 변하고, PCR 과정 중에 서로 다른 염기서열 간의 상대적인 풍부도가 달라질 수 있다. PCR 과정 중에 상대적인 풍부도는 프라이머와의 결합력 차이와 PCR 과정에서의 포화 정도의 차이에서 나타날 수 있다(안 등, 2012). 프라이머와의 결합력이 약한 경우는 강한 경우에 비해 증폭된 양이 적어 상대적으로 풍부도가 달라지며, PCR 과정에서 증폭이 포화단계에 진행 될 때 초기 농도가 낮은 것이 농도가 높은 것 보다 오랫동안 증폭이 가능하여 종간 풍부도 차이가 감소하게 된다. 이러한 문제들은 알맞은 프라이머를 선택하고(Watanabe *et al.*, 2001), annealing 온도를 낮추거나 PCR반복 횟수를 감소시켜 해결될 수 있다(Acinas *et al.*, 2005; Ishii and Fukui, 2001). 하지만 아직까지 표준화된 방법은 없으며, 연구자가 각자의 실험목적, 신뢰성, 비용, 효율성을 판단하여 적절한 PCR 조건을 선택해야 할 것이다.

라. 분석 경로

파이로시퀀싱을 실시한 뒤 나오는 대량의 데이터는 연구자가 원하는 방향으로의 분석이 필요하다. 일반적으로 한 종류의 프로그램이 아닌 여러 종류의 프로그램들이 사용되며, 이에 따라 여러 연구자들이 다양하게 응용 프로그램들을 개발하고 있다. Mothur, RDP, QIIME, CLcommunity™ 등의 프로그램이 존재하며 그 중 Mothur 프로그램(Schloss *et al.*, 2009)은 현재 가장 많은 명령어와 옵션을 포함하고 있기 때문에 유연성이 가장 높다고 할 수 있다. 이 외에도 다양한 분석 패키지들이 있지만 표준화된

방법이 없고, 분석결과가 어떤 프로그램, 어떤 세부 옵션을 선택했느냐에 따라 달라질 수 있기 때문에, 연구자는 자신이 원하는 방향으로 분석 경로와 옵션을 맞추고, 비교하고자 하는 시료들을 동일한 조건으로 사용하는 것이 필요하다(안 등, 2012).

4. 파이로시퀀싱에 의한 장내 미생물균총의 분석

사람을 포함한 동물의 신체 각 부위에서 숙주와 함께 살아가고 있는 세균은 숙주의 건강에 중요한 역할을 하는 것으로 알려져 있다. 이들 세균은 소화 작용, 비타민의 합성, 병원성 세균 억제, 대사작용 조절, 그리고 면역조절 작용을 통하여 숙주 동물의 건강에 도움을 주고 있다(Dethlefsen and Relman, 2011). 따라서, 건강과 질병 상태에서 인체내 세균 분포에 대한 연구는 이들 미생물의 인체내 역할을 밝히는데 있어서 중요할 뿐만 아니라 다양한 질병에 대한 진단, 예방, 치료의 수단을 확립하는데 중요한 역할을 할 수 있다.

미국 국립보건원(NIH)은 2007년부터 연구비 1억 1,500만 달러를 투입한 인간 미생물군집 프로젝트(Human Microbiome Project; HMP)를 통하여 인체의 다양한 기관(구강, 비강, 피부, 소화관, 비뇨기관 등)에 분포하는 세균의 조성을 심도 있게 분석하고 이들 세균 조성의 변화가 건강과 질병에 미치는 영향을 연구해 오고 있다(Turnbaugh *et al.*, 2007). 이 연구를 위하여 여러 연구자들에 의하여 동시에 진행된 파이로시퀀싱 데이터가 이용되었으며, 이들 연구의 첫 번째 관심사는 많은 사람들이 공통적으로 가지고 있는 핵심 미생물균총(core microbiome)은 무엇일까 하는 것이었다(Gonzalez *et al.*, 2011). 핵심 미생물균총이란 모든 사람 또는 대부분의 사람들의 신체 부위에 존재하는 미생물 분류군(또는 미생물 유전자도 포함)으로 이해할 수 있다.

인간 미생물군집 프로젝트에 참여한 많은 연구들은 인체의 특정 부위에만 관심을 가지고 그곳에 존재하고 있는 다양한 미생물 균총을 연구하였다. 그리고 Costello 등(2009)과 같은 연구자들은 multiplex barcode를 이용한 파이로시퀀싱을 통하여 건강한 사람들의 신체 부위 27군데로부터 채취한 시료를 가지고 미생물균총을 분석하여 개인별 비교, 동일한 사람의 신체 부위별 비교 연구를 수행하였다. 이들은 시료당 평균 1,315 염기서열을 분석하여 전체적으로 100만개의 미생물 염기서열 데이터베이스를 얻을 수 있었다. 전체적으로 22 phylum(문: 모든 생물의 분류체계에서 가장 상위 개념에 해당; 문 – 강 – 목 – 과 – 속 – 종)에 속하는 미생물 집단이 확인되었으며, 이들 중 92%는 다음과 같은 네 종류의 주요 phylum에 속하였다; Actinobacteria(37%), Firmicutes(34%), Proteobacteria(12%), 그리고 Bacteriodetes(9.5%). 각각의 신체 부위 마다 서로 다른 미생물균총이 자리 잡고 있었으며, 시간의 변화에 상관없이 우점종으로 존재하는 미생물 분류군이 있다는 사실도 확인하였다.

가. 장내세균의 다양성

소화관은 사람의 신체 부위 중에서 가장 많은 숫자의 미생물을 포함하고 있을 뿐만 아니라 지구상에서 가장 밀도 높은 미생물 군집을 가지고 있는 장소로도 알려져 있다(Egert *et al.*, 2006). Turnbaugh 등(2010)은 유전적으로 동일한 일란성 쌍둥이의 분변 시료에 있는 세균의 다양성 연구를 위한 파이로시퀀싱 분석을 통하여 이들 장내에 존재하는 800여 종의 미생물 군집을 확인할 수 있었으며 대부분의 세균들은 매우 낮은 빈도수로 존재함을 보고하였다.

아프리카와 이탈리아에 살고 있는 어린이들을 대상으로 한 연구 결과에 의하면, 아프리카 어린이들이 Bacteroidetes 문에 속하는 세균을 더 많이 가지고 있었으며, 특히 *Prevotella*와 *Xylanibacter* 속의 미생물이 특이적으로 우점하고 있었다(De Filippo *et al.*, 2010). 아프리카 어린이들에서 발견되는 대표적인 Bacteroidetes

문의 세균들은 공통적으로 cellulose 나 xylan의 가수분해에 필요한 유전자들을 가지고 있었으며, 이는 그들이 주로 섭취하는 식물성 위주의 음식으로부터 에너지를 획득하는데 필요하기 때문에 장내균총이 그러한 방향으로 진화하였음을 설명하고 있다. 이 연구 결과를 통하여 사람의 장내 균총은 지역에 따라 서로 다를 수 있음을 알게 되었으며, 이들이 섭취하는 식단이 그러한 변이를 일으키는 한 가지 요인이 될 것으로 추측하였다.

나. 장내세균과 비만

장내 미생물균총과 비만과의 연관성을 처음으로 보고한 것은 Ley 등(2006)이었다. 이들은 클로닝과 Sanger 시퀀싱 방법에 의하여 장내세균을 분석한 결과, 비만인 사람은 날씬한 사람에 비하여 Bacteroidetes 문에 대한 Firmicutes 문의 비율이 높다는 사실을 발견하였다. 이후에 Armougom 등(2008)은 barcode 파이로시퀀싱을 이용하여 비용을 절감하면서도 훨씬 많은 수의 미생물 균총 분석을 통하여 비만인 사람들의 장내 세균에 대한 Firmicutes와 Bacteroidetes의 비율을 성공적으로 분석할 수 있었다. Zhang 등(2009)도 파이로시퀀싱 기술을 이용하여 정상 체중인 사람, 병적으로 비만인 사람, 그리고 비만치료를 위해 위우회수술(gastric bypass surgery)를 받은 사람을 대상으로 대략 18만개 정도의 염기서열 분석을 통하여 정상 체중과 비만인 사람에서는 Firmicutes 문이 우점하고 있지만 위우회수술을 받은 사람에서는 Firmicutes 문이 현저하게 감소하고 그 대신 Gammaproteobacter 문이 상대적으로 증가한다고 보고하였다. 이 연구의 최종 결론은 개인간의 차이가 있음에도 불구하고 비만인 사람과 위우회수술을 받은 사람은 정상 체중인 사람과 비교하여 장내세균의 조성에 많은 차이점이 있다는 것이었다.

사람의 유전적 요인, 환경 요인, 그리고 비만 여부가 장내균총에 미치는 영향을 알아보기 위하여 Turnbaugh 등(2009)은 정상체중과 비만 상태에 있는 일란성 쌍둥이와 이란성 쌍둥이 154명을 대상으로 대략 200만개의 염기서열을 가지고 장내세균총을 분석하였다. 이 연구의 흥미로운 결과중의 하나는 모든 사람들이 가지고 있는 미생물 균총은 종(species) 수준에서는 공통적으로 존재하는 핵심 미생물균총이 존재하지 않는 다는 사실이었다. 그 대신 여러 사람들에게서 공통적으로 발견되는 미생물 유전자가 존재하여 그 들이 기능적 수준에서의 핵심 미생물 균총을 이루고 있었다. 비만인 사람들은 건강한 사람과 비교하여 phylum(문) 수준에서의 균총 변화가 발견되었으며, 미생물균총의 다양성이 급격하게 감소하는 특징이 있다고 보고하였다.

다. 장질환 환자의 장내세균총

염증성장질환(Inflammatory Bowel Disease; IBD) 환자에서 장내세균총 조성의 변화 현상은 여러 연구에서 보고된 바 있으며, 이러한 세균총 조성의 변화가 질병을 일으키는 원인으로 작용하는지, 아니면 소화관에서의 염증반응과 장상피조직의 변화로 인해 세균총의 변화가 일어나는 것인지는 확실하지 않다 (Macfarlane *et al.*, 2009). IBD 환자의 장내세균총의 특징은 Proteobacteria 문의 증가와 Firmicutes 문과 Bacteroidetes 문의 감소로 특징지어진다(Frank *et al.*, 2007; Walker *et al.*, 2011). 외인성 미생물 중에서 *Clostridium difficile*의 출현 빈도가 정상인에서 보다 IBD 환자에서 높게 나타났으며, 이 병원균은 IBD 증상이 진정 상태에 있다가 재발을 하는데 기여할 수도 있다고 보고하였다 (Clayton *et al.*, 2009). IBD의 발병 기작에 대해서는 아직까지 완전하게 밝혀지지 않았지만 장점막 면역계에 영향을 미치는 환경적, 유전적 요인들을 포함하는 복잡한 요인에 의한 것으로 이해되고 있다. 따라서, 장내세균총의 변화에 의해서만 IBD 증상이 일어나지는 않지만 장내

세균과 숙주와의 상호작용은 특히 유전적으로 민감한 사람들에서 IBD 증상을 심하게 하는 주요 원인으로 작용할 수는 있다(Melgar and Shanahan, 2010).

장내세균총이 특정 질병에 관여한다는 가장 확실한 증거는 바로 과민성대장증상(Irritable Bowel Syndrome; IBS) 환자로부터 찾아볼 수 있다. IBS는 임상현장에서 가장 흔하게 발견되는 기능성 장기능장애 현상으로 서유럽과 북미에서는 전체 인구의 10~15%에 달하는 사람들이 이 증상을 가지고 있다(Quigley, 2011).

Rajilic-Stojanovic 등(2011)은 IBS 환자 62명을 대상으로 장내세균총 특성을 확인한 결과, IBS 환자는 Bacteriodetes 문에 대한 Firmicutes 문의 비율이 높았으며, *Dorea, Ruminococcus, Clostridium* spp.에 속하는 세균이 많고, *Bifidobacterium* 과 *Faecalibacterium* spp. 세균이 적으며, 메탄생성균의 숫자가 낮은 특징이 있다고 보고하였다.

라. 당뇨환자의 장내세균총

Larsen 등(2010)은 유형-2 당뇨 환자의 장내세균 조성을 건강한 사람의 그것과 비교하기 위하여 미생물의 16S rRNA 유전자에 대한 파이로시퀀싱을 실시하였다. 대략 70만개 정도의 염기서열을 분석한 결과, 당뇨환자에 비하여 건강한 사람의 장내세균총은 Firmicutes 문에 속하는 세균들의 비율이 훨씬 높았다. 또한, 혈당 수치가 높은 사람일수록 Firmicutes 문에 대한 Bacteroidetes 문의 비율은 비례적으로 증가함을 알 수 있었다. Betaproteobacteria 강에 속하는 세균의 비율은 건강한 사람에 비해서 당뇨 환자에서 유의적으로 높았으며, 혈당 수치에 비례하여 높아지는 특징이 있었다. 이들 연구진은 유형-2 당뇨는 장내 세균총의 변화와 연관성을 가지고 있으며, 특히 문과 강 수준에서의 조성 변화가 중요하게 관여하는 것으로 결론지었다.

마. 자폐환자의 장내세균총

Finegold 등(2010)은 다양한 정도의 증상을 가지고 있는 자폐 환자, 증상이 전혀 없는 자폐 환자의 형제자매, 그리고 일반인을 대상으로 분변의 세균총을 분석하기 위하여 파이로시퀀싱 기술을 이용하였다. 문 수준에서 비교해 보면, 심한 자폐증상을 가진 그룹에서는 Bacteroidetes 문에 속하는 세균이 높은 빈도로 발견된 데 반하여 일반인 그룹에서는 Firmicutes 문에 속하는 세균이 많이 발견되었다. 종 수준에서는 자폐 환자들의 경우에는 *Desulfovibrio* 종과 *Bacteroides vulgatus*가 대조군에 비하여 유의적으로 높게 나타났다. 정상인에 비하여 자폐 환자의 분변 세균총에서는 훨씬 높은 세균의 다양성이 발견되었다. 이러한 세균 조성의 차이점은 발견하였으나 저자들은 자폐 증상이 장내 세균총의 변화를 유도하는 것인지 아니면 장내 세균총의 변화가 자폐 질병과 그 증상에 작용을 하는 것인지는 아직 밝혀지지 않은 사실이라는 것을 강조하고 있다.

바. 항생제에 의한 장내세균총의 변화

항생제 치료가 장내세균총에 미치는 영향을 알아보기 위하여 파이로시퀀싱을 이용한 여러 연구들이 시도되었다. Dethlefsen 등(2008)은 건강한 사람에서 항생제 ciprofloxacin의 처방 전후에 장내세균총의 변화를 관찰하기 위하여 90만개 이상의 파이로시퀀싱 reads수를 분석하였다. Ciprofloxacin 치료는 소화관에 존재하는 것으로 분석된 전체 미생물 분류군의 1/3에 해당하는 그룹의 숫자를 감소시키고 결과적으로 세균의 다양성을 감소시켰다고 보고하였다. 항생제에 의한 장내세균총의 파괴가 있었음에도 불구하고, 대부분의 장내세균총은 항생제 처방 중단 4주 이내에 원래의 상태로 회복되었으며, 일부 세균의 균총은 6주가 지나도 회복되지 않는 경우도 있었다.

Jakobsson 등(2010)은 항생제 clarithromycin과 metronidazole 치료가 소화관 하부와 목에 존재하는 균총에 미치는 단기적, 장기적인 영향을 알아보는 연구를 진행하였다. 항생제 처방을 받지 않은 대조군의 경우, 미생물 군집이 비교적 안정적으로 유지되는데 반해, 항생제 치료를 하면 즉시 소화관과 목에서 미생물 다양성의 감소가 확인되었다. 특히, 두 부분에서 모두 Actinobacteria 문에 속하는 세균의 감소 현상이 항생제 처방 직후에 나타났다. 대부분의 경우, 미생물균총의 다양성은 항생제 처방 전의 상태로 회복되었지만 어떤 경우에는 미생물 군집의 변화 현상이 4년이 지나도록 회복되지 않았다. 결론적으로 저자들은 흔하게 이용하는 1주간의 항생제 clarithromycin과 metronidazole 처방이 목과 소화관의 미생물균총에 장기적인 변화를 초래한다고 보고하였다.

III. 결론

파이로시퀀싱과 같은 차세대 염기서열 분석 기술의 출현으로 많은 연구자들은 다양한 서식지의 시료를 이용하여 이전에는 상상할 수도 없었던 많은 양의 미생물 염기서열을 해독하고 있으며, 특히 우리의 몸에 존재하고 있는 미생물 균총에 대한 연구의 수준을 한 단계 높여주는데 에도 크게 이바지 하였다. 이 기술을 이용한 연구를 통하여 건강한 사람과 질병에 걸린 사람의 신체(특히, 소화관)에 분포하고 있는 미생물 군집의 구조와 특성에 대한 심도 있는 분석도 이루어지고 있다.

아직은 파이로시퀀싱 기술이 비용이 많이 들고, 많은 양의 DNA가 필요하며, 결과에 대한 분석이 복잡하고, 비교적 짧은 염기서열만을 읽는다는 단점이 있으나, 부가적인 기술의 발전으로 가까운 미래에 파이로시퀀싱 기술은 미생물 생태연구의 분야에서 핵심적인 역할을 수행하게 될 수 있을 것이다. 염기서열의 해독 비용은 지속적으로 감소하고 있으며 대용량 컴퓨터의 기술이 발전으로 데이터의 분석 처리능력은 빠른 속도로 발전하고 있다. 머지않은 미래에, 우리는 개인별 장내세균총의 프로파일을 비싸지 않은 가격으로 확인할 수 있는 날이 올 것으로 예측되며, 특정 질병에 대한 예방이나 진단 그리고 치료의 수단으로 장내세균을 분석하는데 이용되는 핵심적인 기술로 자리잡을 것으로 기대된다.

참고문헌

1. Acinas, S. G., Sarma-Rupavtarm, R., Klepac-Ceraj, V., and Polz, M. F. (2005) PCR-induced sequence artifacts and bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample. *Appl. Environ. Microbiol.* **71**, 8966-8969.
2. Armougom, F. and Raoult, D. (2008) Use of pyrosequencing and DNA barcodes to monitor variations in Firmicutes and Bacteroidetes communities in the gut microbiota of obese humans. *BMC Genomics* **9**, 576.
3. Clayton, E. M., Rea, M. C., Shanahan, F., Quigley, E. M., Kiely, B., Hill, C., and Ross, R. P. (2009) The vexed relationship between *Clostridium difficile* and inflammatory bowel disease: an assessment of carriage in an outpatient setting among patients in remission. *Am. J. Gastoenterol.* **104**, 1162-1169.
4. Costello, E. K., Lauber, C. L., Hamady, M., Fierer, N., Gordon, J. I., and Knight, R. (2009) Bacterial community variation in human body habitats across space and time. *Science* **326**, 1694-1697.
5. De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J. B., Massart, S., Collini, S., Pieraccini, G., and Lionetti, P. (2010) Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc. Natl. Acad. Sci. USA.* **170**, 14691-14696.
6. Dethlefsen, L., Huse, S., Sogin, M. L., and Relman, D. A. (2008) The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol.* **6**: e280.
7. Dethlefsen, L. and Relman, D. A. (2011) Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc. Natl. Acad. Sci. USA.* **108(Suppl. 1)**, 4554-4561.
8. Egert, M., de Graaf, A. A., Smidt, H., de Vos, W. M., and Venema, K. (2006) Beyond diversity: functional microbiomics of the human colon. *Trends Microbiol.* **14**, 86-91.
9. Finegold, S. M., Dowd, S. E., Gontcharova, V., Liu, C., Henley, K. E., Wolcott, R. D., Youn, E., Summanen, P. H., Granpeeshed, D., Dixon, D., Liu, M., Molitoris, D. R., and Green, J. A., 3rd. (2010) Pyrosequencing study of fecal microflora of autistic and control children. *Anaerobe* **16**, 444-453.

10. Frank, D. N., St Amand, A. L., Feldman, R. A., Boedeker, E. C., Harpaz, N., and Pace, N. R. (2007) Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel disease. *Proc. Natl. Acad. Sci. USA.* **104**, 13780-13785.
11. Gilles, A., Meglécz, E., Pech, N., Ferreira, S., Malausa, T., and Martin, J. F. (2011) Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing. *BMC Genomics.* **12**, 1471-2164.
12. Gonzalez, A., Glemente, J. C., Shade, A., Metcalf, J. L., Song, S., Prithiviraj, B., Palmer, B. E., and Knight, R. (2011) Our microbial selves: what ecology can teach us. *EMBO Rep.* **12**, 775-784.
13. Huse, S. M., Huber, J. A., Morrison, H. G., Sogin, M. L., and Welch, D. M. (2007) Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol.* **8**, R143.
14. Huse, S. M., Dethlefsen, L., Huber, J. A., Mark Welch, D., Relman, D. A., and Sogin, M. L. (2008) Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genet.* **4**, e1000255.
15. Ishii, K. and Fukui, M. (2001) Optimization of annealing temperature to reduce bias caused by a primer mismatch in multitemplate PCR. *Appl. Environ. Microbiol.* **67**, 3753-3755.
16. Jakobsson, H. E., Jernberg, C., Andersson, A. F., Sjolund-Karlsson, M., Jansson, J. K., and Engstrand, L. (2010) Short-term antibiotic treatment has differing long-term impacts on the human throat and gut microbiome. *PLoS One* **5**, e9863.
17. Kunin, V., Engelbrektson, A., Ochman, H., and Hugenholtz, P. (2010) Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environ. Microbiol.* **12**, 118-123.
18. Larsen, N., Vogensen, F. K., van den Berg, F. W., Nieseon, D. S., Andreasen, A. S., Pedersen, B., K., Al-Soud, W. A., Sorensen, S. J., Hansen, L. H., and Jakobsen, M. (2010) Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. *PLoS One* **5**, e9085.
19. Lauber, C. L., Zhou, N., Gordon, J. I., Knight, R., and Fierer, N. (2010) Effect of storage conditions on the assessment of bacterial community structure in soil and human-associated samples. *FEMS Microbiol. Lett.* **307**, 80-86.
20. Ley, R. E., Turnbaugh, P. J., Klein, S., and Gordon, J. I. (2006) Microbial ecology: human gut microbes associated with obesity. *Nature* **444**, 1022-1023.
21. Macfarlane, G., Blackett, K., Lakayama, T., Steed, H., and Macfarlane, S. (2009) The gut microbiota in inflammatory bowel disease. *Curr. Pharm. Des.* **15**, 1528-1536.
22. Melgar, S. and Shanahan, F. (2010) Inflammatory bowel disease – from mechanisms to treatment strategies. *Autoimmunity* **43**, 463-477.
23. Quigley, E. M. M. (2011) Antibiotics in irritable bowel syndrome: a novel approach to a challenging disorder. *Clin. Investig. (Lond.)* **1**, 479-482.
24. Rajilic-Stojanovic, M., Biagi, E., Heilig, H. G. H. J., Kajander, K., Kekkonen, R. A., Tims, S., and De Vos, W. M. (2011) Global and deep molecular analysis of microbiota signatures in fecal samples from patients with irritable bowel syndrome. *Gastroenterology* **141**, 1792-1801.
25. Rochelle, P. A., Cragg, B. A., Fry, J. C., Parkes, R. J., and Weightman, A. J. (1994) Effect of sample handling on estimation of bacterial diversity in marine sediments by 16S rRNA gene sequence analysis. *FEMS Microbiol. Ecol.* **15**, 215-224.
26. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* **74**, 5463-5467.
27. Schloss, P. D. (2009) A high-throughput DNA sequence aligner for microbial ecology studies. *PLoS One* **4**, e8230.
28. Schloss, P. D. and Westcott, S. L. (2011) Assessing and improving methods used in operational taxonomic unit-based approaches for 16S rRNA gene sequence analysis. *Appl. Environ. Microbiol.* **77**, 3219-3226.
29. Siqueira, J. F., Jr., Fouad, A. F., and Rocas, I. N. (2012) Pyrosequencing as a tool for better understanding of human microbiomes. *J. Oral Microbiol.* **4**, 10743-10757.
30. Turnbaugh, P. J., Hamady, M., Yatsunenko, T., Cantarel, B. L., Duncan, A., Ley, R. E., Sogin, M. L., Jones, W. J., Roe, B. A., Affourtit, J.P., Egholm, M., Henrissat, B., Heath, A. C., Knight, R., and Gordon, J. I. (2009) A core gut microbiome in obese and lean twins. *Nature* **457**, 480-484.
31. Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R., and Gordon, J. I. (2007) The human microbiome project. *Nature* **449**, 804-810.
32. Turnbaugh, P. J., Quince, C., Faith, J. J., McHardy, A. C., Yatsunenko, T., Niazi, F., Affourtit, J., Egholm, M., Henrissat, B., Knight, R., and Gordon, J. I. (2010) Organismal, genetic, and transcriptional variation in the deeply sequenced microbiomes of identical twins. *Proc. Natl. Acad. Sci. USA.* **107**, 7503-7508.
33. Voelkerding, K. V., Dames, S. A., and Durtschi, J. D. (2009) Next-generation sequencing: from basic research to diagnostics. *Clin. Chem.* **55**, 641-658.
34. Walker, A., Sanderson, J., Churcher, C., Parkes, G., Hudspith, B., Rayment, N., Brostoff, J., Parkhill, J., Dougan, G., and Petrovska, L. (2011) High-throughput clone library analysis of the mucosa-associated microbiota reveals dysbiosis and differences between inflamed and non-inflamed regions of the intestine in inflammatory bowel disease. *BMC Microbiol.* **11**, 7.
35. Watanabe, K., Kodama, Y., and Harayama, S. (2001) Design and evaluation of PCR primers to amplify bacterial 16S ribosomal DNA fragments used for community fingerprinting. *J. Microbiol. Meth.* **44**, 253-262.
36. Wheeler, D. A., Srinivasan, M., Egholm, M., Shen, Y., Chen, L., McGuire, A., *et al.* (2008) The complete genome of an individual by massively parallel DNA sequencing. *Nature* **452**, 872-876.

37. Wu, G. D., Lewis, J. D., Hoffmann, C., Chen, Y. Y., Knight, R., Bittinger, K., Hwang, J., Chen, J., Berkowsky, R., Nessel, L., Li, H., and Bushman, F. D. (2010) Sampling and pyrosequencing methods for characterizing bacterial communities in the human gut using 16S sequence tags. *BMC Microbiol.* **10**, 1471-2180.
38. Zhang, H., DiBaise, J. K., Zuccolo, A., Kudrna, D., Braidotti, M., Yu, Y., Parameswaran, P., Crowell, M. D., Wing, R., Rittmann, B. E., and Krajmalnik-Brown, R. (2009) Human gut microbiota in obesity and after gastric bypass. *Proc. Natl. Acad. Sci. USA.* **106**, 2365-2370.
39. 안재형, 김병용, 김대훈, 송재경, 원항연 (2012) 토양미생물 생태 연구를 위한 증폭 파이로시퀀싱 기법의 응용. 한국토양비료학회지. **45**, 1073-1085.