

Molecular Methods for Studying the Human Microbiota

Yoon Hee Choi¹, Jin Chung² and Hee Sam Na^{2*}

¹Department of Hematology-Oncology, Dongnam Institute of Radiological and Medical Sciences; ²Department of Oral Microbiology, School of Dentistry, Pusan National University, Busan, Korea

Vast array of microbes colonize to each anatomical environment of human body. Culture based methods are important in investigating the microbial structure, but they are extremely biased in their evaluation of microbial diversity by selecting particular population of microbiota. Recent advance in molecular technology has allowed sophisticated analysis of complex human microbiota by culture-independent methods. Here, we will discuss features of tools for human microbiota studies including Roche-454 and Illumina platform. We will also briefly discuss features of some strategies that are commonly applied to these platforms including 16S rRNA targeting and shotgun sequencing. New platforms such as PacBio and Oxford Nanopore are also introduced.

Key Words: Human microbiota, Next-Generation sequencing, Roche-454, Illumina

서 론

저자는 human microbiota 연구 방법에 대한 review article인 'Genomic approaches to studying the human microbiota. Nature 2012: 489; 250-256'과 'Experimental and analytical tools for studying the human microbiome. Nature Reviews Genetics 2012: 13; 47-58'를 읽고 이를 중심으로 이 분야에 대한 요약과 견해를 아래와 같이 전달하고자 합니다 (1, 2).

요 약

인체에 존재하는 다양한 미생물에 대한 연구는 최근에 발달한 next-generation sequencing 기법의 도입과 함께 눈부시게 발전하고 있다. 가장 대표적인 분석 기반으로는 Roche-454와 Illumina가 있다. Roche-454는 400 bps 내외

의 비교적 긴 염기 서열을 단 시간 이내 읽어 들일 수 있는 장점이 있고, Illumina는 100~150 bps의 짧은 염기 서열을 분석하는 반면 방대한 양의 정보를 읽어 들일 수 있는 장점을 갖는다. 이러한 분석의 특징과 장·단점을 비교하고 더불어 최근에 소개되고 있는 PacBio, Oxford Nanopore 등에 대한 소개도 곁들이고자 한다.

해 설

연구 배경

인체에는 다양한 미생물이 피부 및 점막에 존재한다. 피부, 장, 구강, 여성 생식기 등에 존재하는 다양한 미생물들은 각 장소에 따라 그 분포가 다르고 인체의 성장, 영양, 건강, 위생 상태 등 다양한 변수에 따라 상당한 차이와 변화를 보인다. 또한 연구 기법의 개발에 따라 이전에는 알려지지 않았던 미생물이 알려지고 있으며 질병과의 연관 관계도 새롭게 밝혀지고 있다. 다양한 미생물은

Received: January 21, 2013 / Revised: January 23, 2013 / Accepted: January 30, 2013

*Corresponding author: Hee Sam Na. Department of Oral Microbiology, School of Dentistry, Pusan National University, Mool Keum, Beomoe Ri, Yangsan, 626-810, Korea.

Phone: +82-51-510-8252, Fax: +82-51-510-8246, e-mail: heesamy@pusan.ac.kr

**This work was supported by a 2-Year Research Grant of Pusan National University.

© This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>).

각각의 특성을 가질 뿐만 아니라 서식하는 인체의 환경에 따른 특징을 갖게 된다. 인체의 미생물을 연구하는 목적은 이러한 미생물 구성체의 구조와 변화를 측정하고 각 미생물 간의 상호 관계를 분석하고 어떤 물질이 생산되고 소모되는지를 밝힐 뿐만 아니라 인체와의 상호 작용과 질환에 따른 차이를 알아내는데 있다.

인간과 미생물의 상호 관계에 대한 많은 연구가 진행되고 있음에도 인체에 존재하는 미생물에 대한 것은 많은 것이 아직도 잘 알려지지 않았다. 제2의 유전체로 불리기도 하는 인체 내 미생물은 인간의 유전자보다 100배 이상의 유전자를 갖고 있으며 장내 미생물에서만도 300만개 이상의 유전자가 보고되었다 (3~5). 복잡한 미생물 생태계를 정의하고 이를 연구할 수 있는 방법을 개발하는 것은 21세기 미생물학의 큰 과제 중에 하나가 되었다.

미생물 생태계를 연구함에 있어서 그 복잡함은 연구를 어렵게 한다. 수백 종의 미생물이 있고 이들을 고전적인 방법으로 정량 한다는 것은 다수의 미생물이 지금까지 배양되지 않았기 때문에 불가능하다. 미생물의 분포 또한 그 폭이 너무 넓어서 소수로 존재하는 미생물을 확인하기 위해서는 깊이 있는 측정이 필요하다. 미생물의 수를 측정하는 연구 방법으로 배양에 비의존적인 기법은 약 25년 정도의 역사를 가지며 다양한 분자방법론적 접근법이 제시되었다 (6). 초기에는 5S 또는 16S rRNA 유전자의 염기 서열을 분석하여 종을 확인하였다 (7). 연구가 보다 활발해짐에 따라 next-generation sequencing (NGS) 기법이 개발되어 16S rRNA 염기 서열 분석이나 whole-genome shotgun sequencing을 통해 보다 심도 깊은 연구가 진행되고 있다 (8).

16S rRNA 염기 서열 분석에 의한 미생물의 분류

미생물 생태계의 분석은 표적(예, 16S rRNA) 유전자나 shotgun sequencing을 통해 유전자를 분류함으로 이루어 진다. 미생물의 생태에 대한 정보를 얻는 방법으로는 각 미생물 종 별로 갖는 독특한 유전자를 분석하는 방법[oligonucleotide에 대한 hybridization을 통한 array 기법 등 (9)], polymerase chain reaction (PCR) 산물의 다양한 fingerprinting 방법(예로 single-strand conformation polymorphism 또는 terminal restriction fragment length polymorphism)에 의한 분석, 또는 표적 PCR 산물의 DNA 염기 서열 분석법 등이 있다.

16S rRNA는 각 세균 종 간에 상당한 차이를 보인다.

16S rRNA는 약 1.5 kb 가량 되는 유전자로 9개의 짧은 초가변영역(hypervariable region)을 가지고 있으며 세균을 분류하는데 사용된다. 세균 종을 정의 하는 것은 상당히 어렵지만 16S rRNA 유전자 염기 서열이 97% 이상 동일한 경우 같은 종으로 여겨지고 있다.

NGS 기법 도입 이전에는 다양한 종에서 유전자 증폭이 가능한 PCR primer를 사용하여 전체 16S rRNA를 cloning하고 Sanger 기법으로 cloning된 유전자의 염기 서열을 분석하여 이루어 졌다. Sanger 기법에 의한 분석은 상당한 비용과 노력을 요구하기에 연구의 깊이의 제한 요소가 되어 실험 당 100여 개의 염기 서열만을 분석하였다. 이 연구법은 대표적인 세균의 분포를 확인하는 데는 유용하였지만 소수의 세균을 확인하는 데는 제한적이었다.

16S rRNA 유전자를 NGS 기법으로 분석함에 따라 비용과 노력이 획기적으로 개선되었다. 가장 대표적인 연구 기반으로는 Roche의 454 기반과 Illumina 기반 두 가지가 있다. Roche-454는 세균총 분석의 대표 주자였다 (10). 454 pyrosequencing은 약 400개의 염기를 읽을 수 있다. 미국의 Human Microbiome Project (HMP)는 16S rRNA의 V3에서 V5 영역을 표적으로 454를 이용한 분석을 수행하여 각 시료당 평균 7,000여 개의 염기 서열 결과를 얻었다 (11). 실험을 진행함에 있어서 개인 간의 차이를 16S rRNA 유전자 분석을 통해 진행할 때는 어떤 세균들이 존재하는지를 아는 것 보다는 16S rRNA 유전자를 분석하는데 primer가 적당한지와 준비된 시료 간의 차이 정도가 유사한지가 중요하다. 예로, 16S rRNA를 증폭하는데 널리 사용되는 F27-R338 primer는 세균에 대한 특이성이 매우 높지만 장내 세균 가운데 *Bifidobacterium*을 잘 증폭하지 못하는 단점을 갖는다 (12). 또한, 그람 양성 세균은 음성 세균에 비해 DNA를 추출하기가 어렵기 때문에 시료에서 그람 음성 세균의 DNA가 더 많이 뽑힐 가능성도 높다. 이러한 제한점에도 불구하고 세균의 종에 대한 민감도의 감소는 감내할 만 하며 NGS는 비용 대비 효과가 우수하여 대규모 연구를 가능하게 한다.

또 다른 분석 장비로 Illumina 기반의 염기 서열 분석이 유전체 사업에 적용되고 있다 (13, 14). Illumina 기반의 염기 서열 분석은 100~150 bps의 비교적 짧은 길이만을 분석할 수 있기 때문에 단일 초가변영역을 대상으로 분석을 진행할 수 있다. 그러나 이러한 민감도의 감소에도 불구하고 Illumina 기반의 연구도 현재 활발히 진행되고

있다. Illumina의 가장 큰 장점은 상대적으로 낮은 분석 가격과 많은 수의 염기 분석에 있고 이러한 장점으로 인해 16S rRNA 유전자 분석이나 다른 유전체 분석에 많이 이용되고 있다 (15).

Shotgun sequencing을 통한 미생물의 분류

특정 유전자의 염기 서열 분석을 통한 분류는 상당히 강력한 분석 도구로 현재 사용되고 있으나 기능이나 유전적인 정보와 관련해서는 그 내용이 제한적이다. 유전자에 대한 정보가 알려진 미생물의 경우 이를 참고로 생태계에서 발견되는 유사한 유전자의 기능의 유추가 가능하다. 기능적인 추론의 단계를 넘어서는 경우 유전자 기반 통계가 요구되고 이러한 유전자의 분류는 추출된 미생물계의 DNA를 shotgun sequencing을 통해 분석함으로 얻어질 수 있다. 수 백 종 이상의 미생물이 공존하는 생태계에서 중요한 기능을 담당하는 소수의 세균을 분석하기 위해서는 심도 깊은 분석이 요구된다. 장내 세균의 경우 1 ml에 10^{11} 가량의 미생물이 존재하고, 10^6 가운데 1로 존재하는 소수라도 1 ml당 10^5 이 존재하게 되고, 이 정도면 특정 대사물이나 독소를 만들기에 충분한 양으로 전체 생태계와 숙주에 영향을 미칠 수 있다 (16).

Illumina 염기 서열 분석은 대변 시료의 경우 시료당 Metagenomics of the Human Intestinal Tract (MetaHIT)의 경우 약 4 Gb (5)와 HMP project의 경우 약 10 Gb (3)의 자료를 생산하였다. 이런 심도 깊은 분석에서는 시료 내 미생물계 전체에서 1% 미만을 차지하는 *E. coli*라고 하더라도 거의 전체 유전체가 분석되고 이 보다 적게 분포하는 미생물이라 하더라도 일부 유전체의 분석이 이루어지게 된다.

Shotgun sequencing은 16S rRNA를 분석과 더불어 전체 미생물계의 구성에 대한 정보를 제공한다. 16S rRNA를 표적으로 하는 분석법은 PCR primer에 따라 특정 미생물의 유전자가 증폭되지 않아 정보의 치우침이 나타날 수 있는 반면 shotgun sequencing은 이러한 가능성은 없다. 반면, shotgun sequencing은 16S rRNA 분석보다 민감도가 떨어지는데 이는 일부 유전 정보만이 16S rRNA에서 유래하기 때문이다. 또 다른 접근 방법으로는 미생물의 참고 유전체(reference genome)에 shotgun 염기 서열을 배열하는 방법으로 (17, 18), 각 참고 유전체와 비교하여 상대적으로 발견되는 유전자를 비교함으로 미생물의 상대적 빈도를 알아낼 수 있다.

16S rRNA를 통해 미생물 종을 확인하거나 shotgun sequencing을 통해 미생물을 비교하는데 참고 유전체가 필요하다는 것은 자명하다. 현재 방법으로는 주로 Illumina shotgun sequencing에 의존하여 유전 정보를 획득하고 도구를 사용하여 유전체에 배열함으로 참고 유전체를 제작하고 있다. 결과물은 완벽한 유전체라기 보다는 대부분의 미생물의 유전자를 포함하는 질 높은 기초 자료이다. HMP (11)와 Genomic Encyclopedia of Bacteria and Archaea (GEBA)와 같은 프로그램을 통해 참고 유전체가 만들어지고 있다.

데이터 분석 도구 및 전략

미생물 생태계에 대한 실험 데이터는 유전자에 대한 풍부한 정보를 포함하고 있으며 이를 분석할 수 있는 많은 방법들이 개발되고 있다 (19, 20). 데이터의 분석은 크게 세 단계로 이루어진다. 첫째, 일차 데이터가 적용하고자 하는 목적에 맞게 가공되고 여과된다. 16S rRNA 유전자 정보의 경우 미생물을 잘 분류하기 위해 분석의 질이 중요하다. 초기 가공 단계에서는 low-quality base와 이와 관련된 정보를 제거한 후 염기 서열의 질과 읽어 들인 염기 서열의 길이가 중요하다 (21). Shotgun sequencing의 경우 중복된 정보나 인체 유래 유전자를 제거하는 것이 중요하다 (5).

둘째, 일차 여과된 데이터로부터 유도된 다양한 데이터가 생산된다. 16S rRNA 분석의 경우 16S rRNA 데이터 베이스와 비교하여 종의 분류와 빈도에 관한 자료가 생성된다 (22). Shotgun sequencing의 결과는 Basic Local Alignment Search Tool (BLAST)를 사용하여 GenBank나 Kyoto Encyclopedia of Genes and Genomes (KEGG)로부터 유전자 정보를 비교하여 유전자 명단이나 일치하는 결과의 수를 얻을 수 있다 (3, 23). 참고 유전체에 대한 유전자의 비교는 각 유전체에 대한 깊이 있고 폭 넓은 정보를 제공한다 (17).

셋째, 이렇게 생성된 정보는 가계도나 다른 미생물계와 비교하여 유사성을 표현하거나 생태계의 구조를 표현 할 수 있는 빈도 곡선, 다양성 표 등 다양한 결과물을 산출한다. BLAST에서 일치한 결과는 생태계의 재구성을 연구하는데 필요한 대사 경로를 구성할 수 있다 (24). 참고 유전체에 대한 비교는 변이에 대한 분석과 사회 구성원의 인구 유전자에 대한 정보를 제공할 수 있다. 분석은 또한 어떤 미생물이 함께 나타나거나 함께 거의 나타나

지 않는지 등을 밝힘으로 공생 또는 경쟁 관계에 대한 근거를 제공할 뿐만 아니라 시간에 따른 생태계 구성의 역동성을 분석할 수도 있다 (25).

앞으로의 전망

미생물 유전체 연구는 많은 문제들을 해결하였지만 아직도 풀리지 않은 것이 많다. 현재 연구는 보다 정밀해지고 분석 기법은 날로 발전하고 있다(Table 1). Illumina 기반은 보다 작은 유전체를 sequencing하는데 적합하고 454 기반은 신속하고 보다 긴 염기 서열을 분석하기에 metagenomic 분석에 조금 더 적합하다. PacBio 기반의 긴 염기 서열 분석은 shotgun metagenomics의 단점을 보완하여 서론 다른 종을 구별하는데 도움을 줄 수 있다. Oxford Nanopore는 긴 염기 서열 분석과 짧은 운영 시간으로 미생물에 적용하기에 적합할 것으로 기대된다. 필요한 DNA 량의 감소는 국소 조직에서 미생물계의 구조 분석을 가능하게 할 것으로 기대된다. 짧은 운영 시간과 필요한 시료량의 감소는 임상에 적용하기에 보다 용이할 것으로 생각되고 인간과 미생물계의 최종 연구 목적인 진단, 치료 그리고 예방에 응용하는데 한 걸음 더 가까이

갈 수 있게 되었다.

Shotgun sequencing의 단점은 많은 수의 미생물의 sequence가 분석되지 않았을 뿐만 아니라 배양되지 않았다는 것이다. 이러한 미생물은 자료에서 제대로 다루어지지 않고 유전자 결과가 알려지지 않음으로 분류된다. Shotgun 데이터로부터 얻은 정보를 사용하여 새로운 미생물에 대한 유전자 염기 서열을 유전체에 입력하는 경우 contig의 크기가 작아서 하위 생물이 많은 것처럼 보일 수 있고 섞여있는 상태에서 재조합은 상당한 도전 과제이다. 긴 염기 서열을 읽을 수 있는 PacBio와 Oxford Nanopore은 이런 점에서 큰 도움이 될 수 있을 것으로 기대된다. 배양되지 않은 미생물에 대한 참고 유전자 분석이 활발히 진행되고 있다. 각 미생물을 cell sorter를 통해 분리하고 single-cell DNA 정제법을 통해 sequencing한 후 기준의 염기 서열과 비교하여 새로운 유전체로 등록하게 된다 (26, 27).

유전 정보와 관련하여 큰 문제는 미생물이 살아 있는지 숙주의 방어나 항생제에 의해 죽었는지 알 수 없다는 것이다. 이러한 것은 살아있는 세포에 의해 일어날 수 있는 transcriptom 분석, proteomic 또는 metabolomic 데이

Table 1. Comparison of sequencing technologies (modified from ref. 1 and 2)

Platform	Methods	Characteristics	16S rRNA	Shotgun	Relative cost factor (per Mb)	Scale of reads per sample	Raw error rate (%)	Comment
Sanger-based or capillary based instrument	Fluorescent, dideoxy terminator	750-base reads High accuracy	Full length sequenced with 2~3 reads	Long reads help with database comparisons	100	10^2	0.001	Most costly method Relatively low throughput, so low coverage of 16S or shotgun
Roche-454	Pyrosequencing light emission	400-base reads	Up to 3 variable regions per read	Long reads help with database comparisons	1	10^3	1	Cost limits shotgun coverage but 16S coverage is good
Illumina	Fluorescent, stepwise sequencing	100~150 base reads	Only 1 variable region per read	Short reads do not seem to limit analysis	0.1	$10^5 \sim 10^6$	<1	Very high coverage owing to high instrument output and very low cost
PacBio	Fluorescent, single-molecule sequencing	Up to 10-kilobase reads Low accuracy	Accuracy an issue for correct taxon identification	Long reads could help assembly	1.5	10^3	15	Attractive for long reads, but lower accuracy limits applications
Ion Torrent	Proton detection	More than 200-base reads	Like other NGS	Like Illumina	0.4	10^3	2	Expect high coverage, but longer reads than Illumina

터와 함께 분석함으로 극복될 수 있다.

인간 유전학과 유전체학의 발달은 인간 유전자형과 미생물계의 표현형과의 상관 관계 연구를 가능하게 하고 있다. 인간의 유전형과 관련하여 미생물계의 특징을 관찰하는 방법이 최근 소개되었다 (28). 숙주와 미생물계 연구의 발달은 면역과의 관계에 대한 연구로 확대되고 있다 (29). 지속적인 통계적인 발전과 분석 기법의 개발은 실험 설계와 분석을 보다 용이하게 할 것으로 기대된다.

참 고 문 헌

- 1) Kuczynski J, Lauber CL, Walters WA, Parfrey LW, Clemente JC, Gevers D, *et al.* Experimental and analytical tools for studying the human microbiome. *Nat Rev Genet* 2011;13:47-58.
- 2) Weinstock GM. Genomic approaches to studying the human microbiota. *Nature* 2012;489:250-6.
- 3) Consortium HMP. Structure, function and diversity of the healthy human microbiome. *Nature* 2012;486:207-14.
- 4) Brüls T, Weissenbach J. The human metagenome: our other genome? *Hum Mol Genet* 2011;20:R142-8.
- 5) Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 2010;464:59-65.
- 6) Bae JW. Recent Methodological Approaches to Human Microbiome. *J Bacteriol Virol* 2011;41:1-7.
- 7) Olsen GJ, Lane DJ, Giovannoni SJ, Pace NR, Stahl DA. Microbial ecology and evolution: a ribosomal RNA approach. *Annu Rev Microbiol* 1986;40:337-65.
- 8) Kim W. Application of Metagenomic Techniques: Understanding the Unrevealed Human Microbiota and Explaining the in Clinical Infectious Diseases. *J Bacteriol Virol* 2012;42: 263-75.
- 9) Nelson TA, Holmes S, Alekseyenko AV, Shenoy M, Desantis T, Wu CH, *et al.* PhyloChip microarray analysis reveals altered gastrointestinal microbial communities in a rat model of colonic hypersensitivity. *Neurogastroenterol Motil* 2011; 23:169-77.
- 10) Sogin ML, Morrison HG, Huber JA, Mark Welch D, Huse SM, Neal PR, *et al.* Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc Natl Acad Sci U S A* 2006;103:12115-20.
- 11) Peterson J, Garges S, Giovanni M, McInnes P, Wang L, Schloss JA, *et al.* The NIH Human Microbiome Project. *Genome Res* 2009;19:2317-23.
- 12) Hayashi H, Sakamoto M, Benno Y. Evaluation of three different forward primers by terminal restriction fragment length polymorphism analysis for determination of fecal bifidobacterium spp. in healthy subjects. *Microbiol Immunol* 2004;48:1-6.
- 13) Lazarevic V, Whiteson K, Huse S, Hernandez D, Farinelli L, Osterås M, *et al.* Metagenomic study of the oral microbiota by Illumina high-throughput sequencing. *J Microbiol Methods* 2009;79:266-71.
- 14) Gloor GB, Hummelen R, Macklaim JM, Dickson RJ, Fernandes AD, MacPhee R, *et al.* Microbiome profiling by illumina sequencing of combinatorial sequence-tagged PCR products. *PLoS One* 2010;5:e15406.
- 15) Hummelen R, Fernandes AD, Macklaim JM, Dickson RJ, Changalucha J, Gloor GB, *et al.* Deep sequencing of the vaginal microbiota of women with HIV. *PLoS One* 2010; 5:e12078.
- 16) Luckey TD. Introduction to intestinal microecology. *Am J Clin Nutr* 1972;25:1292-4.
- 17) Martin J, Sykes S, Young S, Kota K, Sanka R, Sheth N, *et al.* Optimizing read mapping to reference genomes to determine composition and species prevalence in microbial communities. *PLoS One* 2012;7:e36427.
- 18) Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, *et al.* Enterotypes of the human gut microbiome. *Nature* 2011;473:174-80.
- 19) Raes J, Foerstner KU, Bork P. Get the most out of your metagenome: computational analysis of environmental sequence data. *Curr Opin Microbiol* 2007;10:490-8.
- 20) Wooley JC, Godzik A, Friedberg I. A primer on metagenomics. *PLoS Comput Biol* 2010;6:e1000667.
- 21) Schloss PD, Gevers D, Westcott SL. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLoS One* 2011;6:e27310.
- 22) Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. UniFrac: an effective distance metric for microbial community comparison. *ISME J* 2011;5:169-72.
- 23) Consortium HMP. A framework for human microbiome research. *Nature* 2012;486:215-21.
- 24) Abubucker S, Segata N, Goll J, Schubert AM, Izard J, Cantarel BL, *et al.* Metabolic reconstruction for metagenomic

- data and its application to the human microbiome. PLoS Comput Biol 2012;8:e1002358.
- 25) Caporaso JG, Lauber CL, Costello EK, Berg-Lyons D, Gonzalez A, Stombaugh J, *et al.* Moving pictures of the human microbiome. Genome Biol 2011;12:R50.
- 26) Chitsaz H, Yee-Greenbaum JL, Tesler G, Lombardo MJ, Dupont CL, Badger JH, *et al.* Efficient de novo assembly of single-cell bacterial genomes from short-read data sets. Nat Biotechnol 2011;29:915-21.
- 27) Dichosa AE, Fitzsimons MS, Lo CC, Weston LL, Preteska LG, Snook JP, *et al.* Artificial polyploidy improves bacterial single cell genome recovery. PLoS One 2012;7:e37387.
- 28) Benson AK, Kelly SA, Legge R, Ma F, Low SJ, Kim J, *et al.* Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. Proc Natl Acad Sci U S A 2010;107:18933-8.
- 29) Hooper LV, Littman DR, Macpherson AJ. Interactions between the microbiota and the immune system. Science 2012;336: 1268-73.