

특집: 식품유래 미생물의 산업적 활용

NGS 분석에 기반한 전통발효식품 미생물 연구 및 산업적 활용

남 영 도

한국식품연구원 헬스케어연구단

Identification and Industrial Application of Fermented Food Microorganisms Analyzed by NGS Technology

Young-Do Nam

Research Group of Healthcare, Korea Food Research Institute, Wanju, Jeonbuk 55365, Korea

서 론

한국의 대표적인 발효식품으로는 김치, 젓갈, 장류 및 주류 등을 들 수 있으며 이들 발효식품은 대부분 전통적인 방법으로 제조되어 자연에서 유래된 다양한 미생물에 의해 발효되는 특성을 가진다. 현재까지의 많은 연구 결과들은 한국 전통발효식품이 영양적 가치 이외에도 항산화, 항당뇨, 항비만 등의 다양한 건강 기능성을 가지고 있는 것으로 보고하고 있다. 전통발효식품은 주재료로 사용되는 야채류와 곡물류 혹은 젓갈의 소재로 사용되는 다양한 해양 생물의 표면과 생체 내부에 포함하고 있는 미생물에 의해서나 부재료로 사용되는 소금이나 물에 포함되어 있는 미생물에 의해 발효가 진행되며, 심지어는 제조되는 작업 환경이나 메주, 누룩을 띄우는 제국실에서 유래된 미생물들도 발효식품의 특성에 영향을 주는 것으로 알려져 있다.

현재까지 배양법에 기반한 전통적인 미생물 분석법으로 확인한 미생물은 전체 미생물의 단 1%에 불과하며, 비교적 연구가 많이 된 식품 미생물의 경우에도 상당히 많은 미생물의 정보가 미지로 남아 있다. 하지만 최근의 분자생물학적 분석 기술의 발달과 함께, NGS(Next generation sequencing)로 대변되는 대규모 염기서열 분석 방법의 적용을 통해 현재 우리는 발효식품에 존재하는 대부분 미생물에 대한 정보를 알 수 있는 정도에 이르렀다.

앞서 언급한 것처럼 전통발효식품은 주변 환경에서 유래한 미생물에 의해 자연적으로 발효가 진행되는 식품으로 항상 일관된 제품 특성을 보장하지는 않기에 최근에는 일부 발효식품 제조에 종균 미생물을 활용하고 있다. 하지만 보다 과학적이고 영양성분이나 건강 기능성분 강화 등의 기능이 보장되는 종균 미생물의 활용을 위해서는 각 미생물의 유전적 특성을 확인하고 보강하고자 하는 유효 성분의 생산 능력을 가지는지 여부 등을 확인하여야 하며, 이를 위해서는 NGS를 이용한 게놈 해독 및 생물정

보학을 통한 유전자 확인이 필요한 시점이다.

따라서 본 기고에서는 발효식품 미생물 분석에 사용된 여러 분자생물학적 방법과 특히 NGS 기법을 활용한 예를 소개하고 전통발효식품으로부터 확보한 미생물 자원의 게놈 해독을 통해 유용 유전자를 포함하는 기능성 종균의 산업적 활용 방안에 대한 정보를 제공하고자 한다.

NGS 방법에 기반한 발효식품 미생물 분석

미생물은 우리 지구의 대부분을 담당하고 있으며, 식품에서도 발효, 숙성, 오염 및 부패에 이르기까지 다양한 부분에서 미생물이 주된 역할을 하고 있지만, 현재까지 배양법에 기반하여 발견된 미생물은 전체 미생물의 약 1% 정도로 매우 적은 수준이다(1). 이는 기존의 배양법에 기초한 미생물 분석방법이 자연 상태에서 존재하는 미생물의 생장 조건을 재현하는 데 어려움이 있고, 인간이 미생물 간의 신호 물질 교환이나 물질 이용에 대한 상호 작용을 완전히 조절하는 것이 불가능하기 때문이며, 이로 인해 배양법에 기반한 분석 방법은 점차 분자생물학적인 비배양적 분석방법으로 대체되어 왔다(2).

분자생물학적으로 전통발효식품의 미생물을 분석하는 방법으로는 DGGE(Denaturant gradient gel electrophoresis)로 김치 발효 기간 중의 세균과 효모의 변화를 분석한 예가 있으며(3), Microarray 방법으로 전체 전사체의 분석을 통해 김치 발효에 실제로 작용하는 미생물의 변화를 확인한 예가 있다(4). 배양법에 기반한 미생물 분석방법과는 달리 DNA에 기반한 16S rDNA와 같은 마커 유전자의 cloning/sequencing, DGGE, Microarray 등의 방법은 전통발효식품에서 배양할 수 없는 미생물의 존재 여부를 확인할 수 있다는 장점이 있으나, 여전히 시간과 노력이 많이 들고 주요 미생물만을 확인할 수 있다는 단점이 있었다.

2000년대 후반부터는 NGS 분석 방법 중에 하나인 454 pyrosequencing을 이용한 미생물 분석 방법이 점차 보

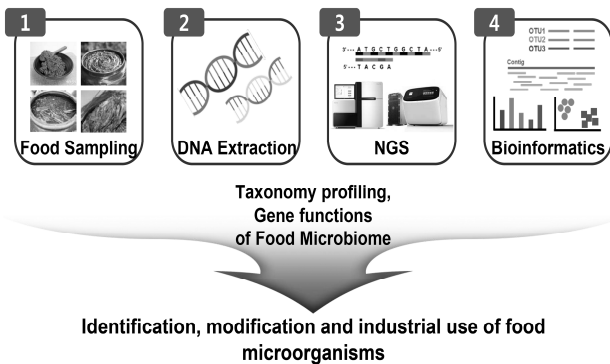


Fig. 1. Overall analytical procedures for food microorganisms based on NGS technology.

편화되기 시작하여 토양, 해수, 생체뿐만 아니라 다양한 식품 미생물의 분석에 적용되어 왔다. 일반적으로 사용되는 NGS 분석 방법은 샘플로부터 DNA를 추출하여 정제하고 적당한 길이의 절편으로 DNA를 나누거나(150~600 bp) 특이적인 프라이머를 이용하여 증폭하는 과정을 거치며, sequencing을 위한 라이브러리 제작 및 NGS 기기 운용을 통해 식품 미생물의 DNA 염기서열을 확보하게 된다(5). NGS를 이용하여 확보한 발효식품의 DNA 염기서열로부터 실제로 식품 미생물의 조성을 확인하는 과정은 생물정보학적 해석을 통해 완성되게 되는데, 각 DNA 염기서열의 QC(quality control) 과정을 통해 sequencing error를 제거하고 유사한 sequence read별로 클러스터링을 진행하여 OTU(operational taxon unit)를 만들고 각 OTU의 동정을 통해 식품 미생물의 특성을 규명하게 된다(Fig. 1)(6).

NGS 분석 방법에 기반한 미생물 군집 분석은 단일 미생물의 확보뿐만 아니라 식품에 실제 존재하는 미생물의

구성을 파악하고 현재까지 발굴되지 않은 신규의 미생물을 특정 분리하는 방법에도 적용되고 있다. 현재까지 우리나라 전통발효식품의 발효 및 숙성 과정에 관여하는 미생물의 군집 변화 연구는 김치, 장류, 막걸리, 젓갈류 등 다양한 분야에 적용되어 왔으며, 유전체학, 대사체학 및 단백질체학과의 연계를 통해 발효식품에서의 미생물 기능을 규명하는 데 적용되어 오고 있다. NGS 분석 기술이 실제로 발효식품 미생물에 사용된 예로서는 amplicon sequencing 기법을 이용하여 국내에서 생산되는 젓갈의 세균과 고세균 구성을 밝히는데(7), 한국의 각 지역에서 생산된 된장(8), 청국장(9), 고추장(10) 등의 전통 콩 발효식품 내에 존재하는 미생물의 특성을 확인하는 데 사용되어 왔다(Fig. 2). 또한 발효식품 내에 존재하는 전체 유전체를 분석하여 미생물의 종류와 양뿐만 아니라 기능 유전체의 구성을 함께 분석할 수 있는 방법으로서 metagenome sequencing을 이용한 김치 발효에 관련한 미생물을 NGS 기법으로 분석한 예가 있다(11). 전통적인 방법으로 생산되는 막걸리는 사용되는 누룩의 종류에 따라 알코올 생산성과 관능적 특성이 달라지며 막걸리의 특성은 누룩에 포함되어 발효를 주도하는 미생물에 의해 결정될 것이라고 예상되어 왔으나, 실제로 세균과 곰팡이류가 발효 과정 전체에 어떤 영향을 끼치는지는 NGS 기법을 활용한 분석이 이루어지기 전까지는 확실히 밝혀지지 않았었다(12). 한국의 대표적인 전통발효식품의 NGS 분석을 통해 우리는 이제까지 알지 못했던 다양한 미생물이 발효식품의 발효와 숙성에 관여한다는 사실을 알게 되었으며, 발효식품이 제조되는 지역과 제조 원료 및 방법에 따라 각기 독특한 기능적, 관능적 특성을 가지게 된다는 사실을 알게 되었다.

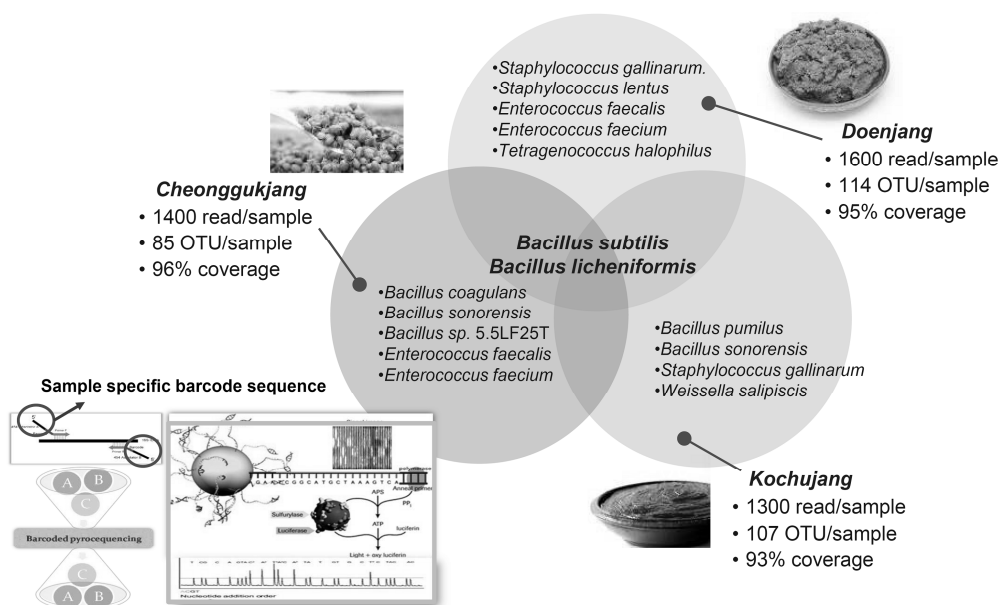


Fig. 2. Representative bacterial species of Korean traditional fermented soy product identified by NGS technology.

발효식품 유래 유용 미생물의 발굴

과거에 식물, 동물, 미생물 등의 유전자원은 인류의 공동자산으로서 여겨져 왔으나 유전자원 및 이와 관련한 지식의 경제적 가치에 대한 인식이 변화함에 따라 현재는 자원 이용국가와 보유국 간의 개발 이익 공유의 논의로 이동하는 추세이다. 나고야의정서 발효 이후 우리나라 산업계에서는 매년 최소 136억 원에서 최대 639억 원의 추가 부담 비용이 발생할 것으로 예측되며, 우리나라는 주로 유전자원을 이용하는 국가로 최근 국회의 나고야의정서 비준 및 국내 발효로 유전자원에 대한 대응방안이 필요한 시점이다(13).

우리나라에서 식품 유래의 미생물을 주요 대상으로 하는 미생물은행은 한국식품연구원의 식품유전자은행과 세계김치연구소의 김치미생물은행이 있다. 식품유전자는 은행의 경우 보관된 미생물 자원이 발효식품의 스타터 개발에 용이하게 사용될 수 있으나, 해당 미생물 자원의 기본 정보만 구축되어 있을 뿐 유전적 또는 기능적 특성에 대한 정보가 연계되어 있지 않아 실제 식품산업에의 활용은 미진한 실정이다. 현재 국내에서 사용되고 있는 스타터균주의 90% 이상은 수입에 의존하고 있는 상황이며, 전통식품 미생물의 발굴과 이의 유전 및 기능적 특성 정보 구축은 바이오식품산업의 기반 강화에 큰 의미를 가진다(14).

현재 *Aspergillus oryzae*, *Bacillus subtilis*, *Bacillus licheniformis* 등이 일부 장류 발효의 종균으로서 사용되고 있으며, *Leuconostoc mesenteroides* 균종은 김치 발효의 종균으로 사용되고 있으나 현재 사용되는 종균 미생물은 발효식품의 안정적인 생산과 기호성 증진을 위해 개발되어 왔을 뿐 특정 기능성을 강화하기 위한 종균으로서의 활용은 거의 이루어지지 않고 있다. 하지만 전통 장

류로부터 분리한 *Bacillus amyloliquefaciens*로부터 항당뇨 소재인 1-deoxynojirimycin의 생합성 유전자의 존재를 확인하고 이를 실제 기능성 발효식품 제조에 적용한 예가 있으며(15), *Bacillus thermoglucosidasius* 균주로부터 L-arabinose isomerase의 유전자를 발굴하고 재조합 단백질로 발현하여 저칼로리 대체 감미료인 tagatose의 생산을 확인한 예가 있다(16). 또한 *Bacillus thuringiensis* 유래의 항당뇨 소재인 4-hydroxy-L-isoleucine 생산 유전자를 재조합하여 이의 생산성을 검증함으로써 한국 전통식품 유래의 미생물이 가지고 있는 유용 유전자를 활용하고 이를 이용한 발효식품 종균 개발이 점차 확대되고 있는 추세이다(Fig. 3)(17).

발효식품 유래 미생물의 게놈 해독

미생물 유전체는 미생물 내에 존재하는 모든 유전정보를 의미하는 것으로 대부분 chromosomal DNA에 저장되어 있으며 일부는 plasmid의 형태로 존재한다. 미생물 유전체의 크기는 대략 0.5~10 Mb의 size를 가지며 원핵생물에 비하여 그 크기가 1/100~1/1000밖에 되지 않으나 상대적으로 높은 유전자 밀도를 가지는 특징이 있어서 비교적 적은 노력으로 미생물의 전체 게놈 상에 존재하는 유전자의 특성이나 유용 기능성 유전자의 발굴 및 활용에 좋은 소재가 될 수 있다(18).

인류가 1995년에 처음으로 *Haemophilus influenzae* 유전체 완전 해독을 통해 생물체의 유전정보를 확보한 이후 급속도로 발전한 대단위 염기서열 분석 기술과 생물정보학의 혁신적 진보를 바탕으로 근래에는 비교적 적은 비용과 시간으로 미생물의 유전체를 해독하고 그 정보를 확보할 수 있게 되었다. GOLD(Genome online database)에 등록된 미생물의 게놈 분석 현황을 보면 2013년

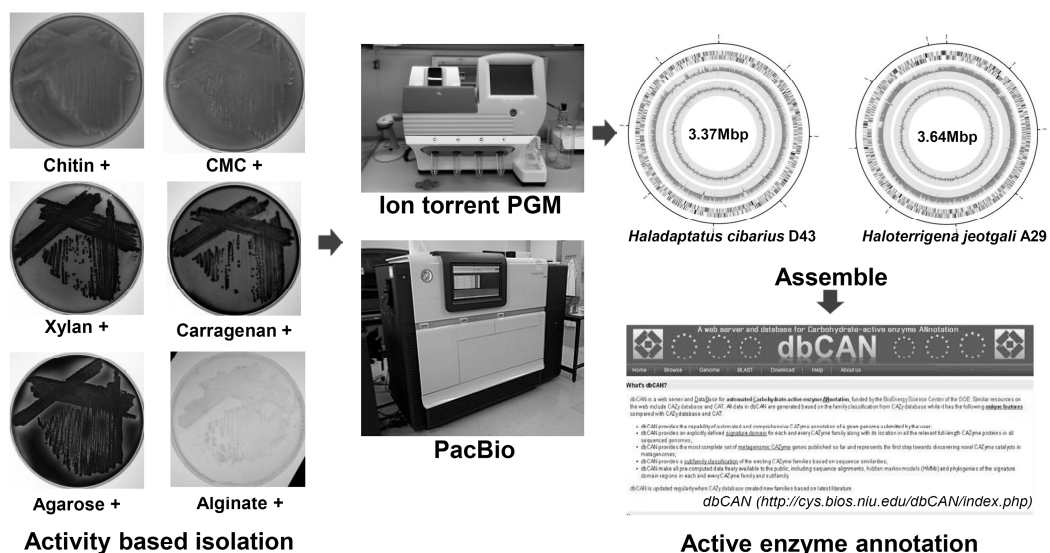


Fig. 3. Function based bacterial screening and functional gene annotation process.

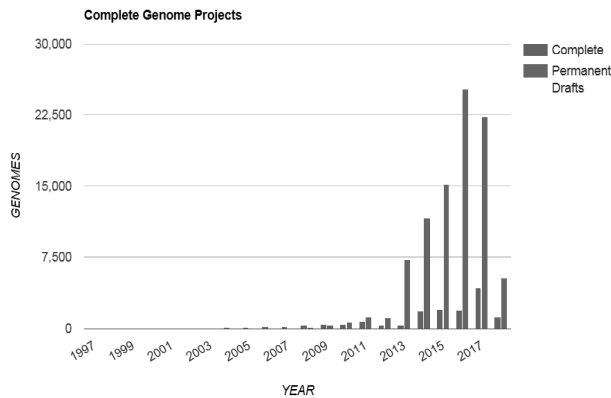


Fig. 4. Complete and permanent draft genome totals in GOLD (by year and status).

부터 급격한 미생물 유전체의 분석 증가를 확인할 수 있으며, 현재까지 전 세계적으로 진행된 미생물의 유전체는 약 70,000개이고 이중 식품이나 농업에 관련된 미생물 유전체는 약 4,068개이다(Fig. 4)(19).

1977년에 생어(Sanger)에 의해 개발된 생어 시퀀싱 방법은 dideoxynucleotides(dNTPs)를 이용한 연쇄 종결 반응을 통해 DNA의 염기서열 정보를 확인하는 방법이며, 캐필러리 장비를 통한 자동화로 현재는 한 번에 약 600~800 b 이상의 DNA 정보를 확인할 수 있다. 초기에 진행된 미생물의 게놈 해독은 생어 시퀀싱 방법으로 진행되어 왔지만 한 번에 분석할 수 있는 샘플과 시간 그리고 노동력의 한계로 인하여 최근의 미생물 유전체 해독에는 차세대 유전체 분석 기술(NGS)이 주로 사용되어진다.

NGS 기술 기반은 하나의 유전체를 무수히 많은 조각으로 나누고 각 조각을 증폭하는 과정을 거쳐 동시에 많은 수의 염기서열을 해독한 뒤 생물정보학적 방법을 통해 조립하는 과정으로 이루어지며, 이를 통해 기존의 생어 시퀀싱 방법에 비하여 시간과 노동력을 비약적으로 단축시킬 수 있었다. 현재는 2세대 NGS 기기인 454 GS FLX, SOLiD 등은 거의 사용되고 있지 않으며, Ion Torrent, Illumina 장비와 long read sequencing 장비인 Pacbio RS II 장비가 미생물 게놈 분석에 주로 사용되고, 단일 가닥의 DNA를 long read로 증폭과정 없이 해독할 수 있는 MinION과 같은 3세대 NGS 장비의 활용이 점차 늘어가고 있다(Table 1).

현재까지 유전체 정보가 보고된 한국 발효식품 유래 미생물로는 김치 및 젓갈 유래의 *Lactobacillus*, *Weissella*, *Leuconostoc*, *Pediococcus*, *Carnobacterium*, 장류 유래의 *Bacillus*, *Staphylococcus*, *Enterococcus* 속을 포함하여 다양한 미생물이 있으며, 향후 NGS 방법으로 분석하고 발굴한 미생물 유전체의 유전자 기능을 각 유전자를 제거하여 어떤 기능이 상실되는지를 확인하거나, 특정 환경 또는 조건에서의 유전자 발현 양상을 조사하거나 또는 각 유전자를 발현시켜 생산되는 단백질의 기능을 분석하는 등의 방법을 이용한 실험적 검증을 통해 산업적인 기능성 발효식품 종균으로 활용이 가능하다.

결론

나고야의정서 발효 이후, 최근의 유전자원 및 이와 관

Table 1. Comparison of different sequencing platform

Platform	Reads per run	Reads length	Read type	Throughput	Advantages	Disadvantages	Applications
454 GS FLX	1 M	600	SE, PE	450 Mb	Long read length Short run time	Expensive Low throughput High error rate (1%) No longer available	16S amplicon
SOLiD	3 B	75	SE	160 Gb	High accuracy	Expensive Long run time	16S amplicon
Ion Torrent PGM/ Proton	4~5.5 M/ 60~80 M	200, 400/ 200	SE	1.2 to 2 Gb/ 10 Gb	Short run time Low cost Different chips available	High error rate (1%)	16S amplicon Shotgun
Illumina Miseq/ Hiseq	250 M/ 5 B	300/ 150	SE, PE	15 Gb/ 1.8 Tb	High throughput Low sequencing cost	Short read length Long run time	16S amplicon Targeted gene Shotgun Metatranscriptome
Pacbio RS II (P6-C4)/ Sequel	50 K/ 500 K	10~ 15 K	—	2 Gb/ 20 Gb	Single-molecule-real-time sequencer Short run time Long read length	High error rate (10~15%) Lack of application to meta-analysis	16S amplicon Shotgun Transcriptome
MinION	Variable	Variable	—	5 Gb	Single-molecule-real-time sequencer Small size instrument Low capital cost	High error rate (5~15%) Short read length	Targeted gene Shotgun Transcriptome

련한 지식의 경제적 가치에 대한 인식 변화는 우리가 가지고 있는 미생물 유전자원에 대하여 다시 한번 생각해 보기를 요구하고 있다. 한국의 전통발효식품은 영양적, 건강 기능적으로 우수성이 높이 평가되고 있으며 유산균과 바실러스 균종을 대표로 하는 다양한 미생물을 포함하는 유전자원 pool로써도 고려될 수 있다. 따라서 건강 기능성이 우수한 전통발효식품으로부터 안정적인 품질과 기능성의 향상을 기대할 수 있는 종균의 발굴 및 활용은 현재의 국가 간 유전자원 이용에 대한 마찰을 줄이고 감소하고 있는 전통발효식품 시장을 다시 한번 활성화할 수 있는 하나의 방안이 될 수 있겠다.

NGS로 대변되는 최신 대용량 염기서열 분석방법은 전통발효식품에 존재하는 미생물의 특성을 규명하는데 유용하게 사용되고 있으며, 생물정보학과의 결합을 통해 기능 유전자의 발굴 및 활용에 유용한 도구로써 사용될 수 있다. 대단위 염기서열 분석방법을 이용하여 전통발효식품 내의 미생물의 구성을 확인하고, 발효식품 내의 유용 미생물을 선택적으로 발굴, 유전체 해독을 통한 유용 유전자의 기능을 검증하는 등의 프로세스는 향후 한국의 전통발효식품 유래 유용 미생물 자원이 산업적으로 활용되는 데 도움을 줄 수 있을 것이라 생각한다.

참고문헌

1. Staley JT, Konopka A. 1985. Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Ann Rev Microbiol* 39: 321-346.
2. Hugenholtz P. 2002. Exploring prokaryotic diversity in the genomic era. *Genome Biol* 3: reviews0003.1-0003.8.
3. Chang HW, Kim KH, Nam YD, Roh SW, Kim MS, Jeon CO, Oh HM, Bae JW. 2008. Analysis of yeast and archeal population dynamics in kimchi using denaturing gradient gel electrophoresis. *Int J Food Microbiol* 126: 159-166.
4. Nam YD, Chang HW, Kim KH, Roh SW, Bae JW. 2009. Metatranscriptome analysis of lactic acid bacteria during kimchi fermentation with genome-probing microarrays. *Int J Food Microbiol* 130: 140-146.
5. Ercolini D. 2003. High-throughput sequencing and metagenomics: moving forward in the culture-independent analysis of food microbial ecology. *Appl Environ Microbiol* 79: 3148-3155.
6. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JJ, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7: 335-336.
7. Roh SW, Kim KH, Nam YD, Chang HW, Park EJ, Bae JW. 2010. Investigation of archaeal and bacterial diversity in fermented seafood using barcoded pyrosequencing. *ISME J* 4: 1-16.
8. Nam YD, Lee SY, Lim SI. 2012. Microbial community analysis of Korean soybean pastes by next-generation sequencing. *Int J Food Microbiol* 155: 36-42.
9. Nam YD, Yi SH, Lim SI. 2012. Bacterial diversity of *cheonggukjang*, a traditional Korean fermented food, analyzed by barcoded pyrosequencing. *Food Cont* 28: 135-142.
10. Nam YD, Park SL, Lim SI. 2012. Microbial composition of the Korean traditional food "*kochujang*" analyzed by a massive sequencing technique. *J Food Sci* 77: M250-M256.
11. Lee M, Song JH, Jung MY, Lee SH, Chang JY. 2017. Large-scale targeted metagenomics analysis of bacterial ecological changes in 88 kimchi samples during fermentation. *Food Microbiol* 66: 173-183.
12. Jung MJ, Nam YD, Roh SW, Bae JW. 2012. Unexpected convergence of fungal and bacterial communities during fermentation of traditional Korean alcoholic beverages inoculated with various natural starters. *Food Microbiol* 30: 112-123.
13. 광충목. 2017. 유전자원 보호를 위한 IP규범형성 관련 국제 논의 동향 및 시사점. 한국지식재산연구원. 심층분석 보고서 제2017-07호.
14. Yi SH, Lee SY. 2011. Application of microorganism from traditional foods and microbial gene bank. *Bulletin of Food Technol* 24(1): 74-79.
15. Seo MJ, Nam YD, Lee SY, Park SL, Yi SH, Lim SI. 2013. Isolation of the putative biosynthetic gene cluster of 1-deoxynojirimycin by *Bacillus amyloliquefaciens* 140N, its production and application to the fermentation of soybean paste. *Biosci Biotechnol Biochem* 77: 398-401.
16. Seo MJ. 2013. Characterization of an L-arabinose isomerase from *Bacillus thermoglucosidasius* for D-tagatose production. *Biosci Biotechnol Biochem* 77: 385-388.
17. Smirnov SV, Kodera T, Samsonova NN, Kotlyarov VA, Rushkevich NY, Kivero AD, Sokolov PM, Hibi M, Ogawa J, Shimizu S. 2010. Metabolic engineering of *Escherichia coli* to produce (2S, 3R, 4S)-4-hydroxyisoleucine. *Appl Microbiol Biotechnol* 88: 719-726.
18. McCutcheon JP, Moran NA. 2012. Extreme genome reduction in symbiotic bacteria. *Nat Rev Microbiol* 10: 13-26.
19. Bernal A, Ear U, Kyrpides N. 2001. Genomes OnLine Database (GOLD): a monitor of genome projects world-wide. *Nucleic Acids Res* 29: 126-127.