

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2016년 12월 29일 (29.12.2016)



(10) 국제공개번호
WO 2016/209037 A1

- (51) 국제특허분류:
C12Q 1/68 (2006.01)
- (21) 국제출원번호: PCT/KR2016/006792
- (22) 국제출원일: 2016년 6월 24일 (24.06.2016)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보:
10-2015-0089772 2015년 6월 24일 (24.06.2015) KR
10-2016-0079254 2016년 6월 24일 (24.06.2016) KR
- (71) 출원인: 경북대학교 산학협력단 (KYUNGPOOK NATIONAL UNIVERSITY INDUSTRY-ACADEMIC COOPERATION FOUNDATION) [KR/KR]; 41566 대구시 북구 대학로 80 (산격동), Daegu (KR).
- (72) 발명자: 이동우 (LEE, Dong Woo); 41573 대구시 북구 대현로 10길 82, 104동 303호, Daegu (KR). 이용직 (LEE, Yong Jik); 42156 대구시 수성구 수성로 165, B동 1509호, Daegu (KR). 이상재 (LEE, Sang Jae); 46957 부산시 사상구 백양대로 700번길 23, 104동 401호, Busan (KR). 이한승 (LEE, Han Seung); 48516 부산시 남구 분포로 111, 121동 402호, Busan (KR).
- (74) 대리인: 김순웅 (KIM, Soon Woong); 08380 서울시 구로구 디지털로 31길 20 (구로동) 601호, Seoul (KR).

(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

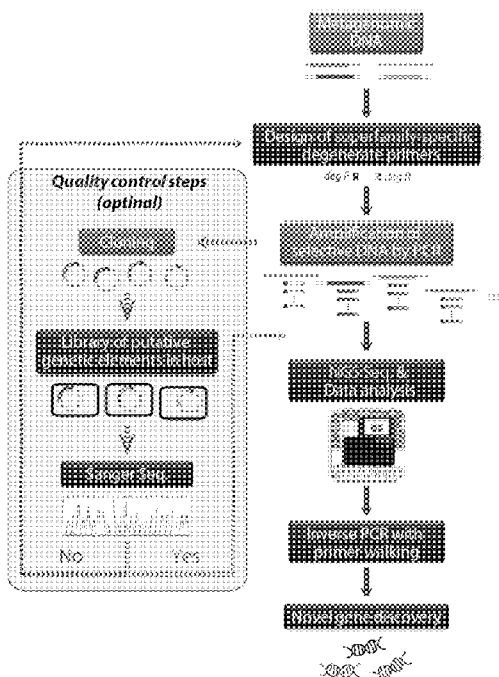
(84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

공개:

- 국제조사보고서와 함께 (조약 제 21 조(3))
- 청구범위 보정 기한 만료 전의 공개이며, 보정서를 접수하는 경우 그에 관하여 별도 공개함 (규칙 48.2(h))
- 명세서의 서열목록 부분과 함께 (규칙 5.2(a))

(54) Title: METHOD FOR EXPLORING USEFUL GENETIC RESOURCES THROUGH BULK METAGENOME ANALYSIS AND USE THEREOF

(54) 발명의 명칭 : 대용량 메타지놈 분석을 통한 유용 유전자원 탐색 방법 및 이의 이용



(57) Abstract: The present invention relates to a degenerate primer for analyzing bulk metagenome information, and a method and a kit for analyzing metagenome information using the same. The method for analyzing metagenome information, according to the present invention, and a primer set used therefor can be applied to the present invention in order to rapidly determine the utility value of extensive metagenome samples. In particular, a superfamily-specific degenerate primer of the present invention can be used to rapidly determine the presence or absence of genetic information of target peptides in a metagenome by a simple method, and therefore can be used to collect large amounts of useful peptide resource information at a high speed from various metagenome samples. In addition, based on the method of the present invention, it can be used to explore a new peptide gene through design and manufacturing of a superfamily-specific degenerate primer of new target peptides. In addition, the method of the present invention can be applied to research relating to polypeptides, oligopeptides, antibiotic resistant genes, antimicrobial peptides, anti-fungal peptides, oligopeptides, markers, or Single-Nucleotide Polymorphism (SNP), as well as enzymes.

(57) 요약서: 본 발명은 대용량 메타지놈 정보 분석을 위한 디제너레이트 프라이머 및 이를 이용하여 메타지놈 정보를 분석하는 방법 및 키트에 관한 것이다. 본 발명에 따른 메타지놈의

[다음 쪽 계속]



정보 분석 방법, 및 이에 이용되는 프라이머 세트는 방대한 메타지놈 샘플 효용 가치의 빠른 관별을 위하여 본 발명이 적용될 수 있다. 특히, 본 발명의 슈퍼패밀리-특이적 (superfamily-specific) 디제너레이트 프라이머를 이용하여 메타지놈에서 목적 펩타이드들의 유전정보의 유무를 간단한 방법에 의해 빠르게 탐색할 수 있으므로 다양한 메타지놈 샘플로부터 고속으로 대량의 유용 펩타이드 자원 정보를 수집하는데 사용이 가능하다. 또한, 이러한 본 발명의 방법을 기반으로 하여 목적하는 새로운 펩타이드들의 슈퍼패밀리-특이적 디제너레이트 프라이머 설계 및 제작을 통하여 새로운 펩타이드 유전자를 탐색하는데 사용이 가능하다. 또한, 이러한 방법은 효소뿐 만 아니라, 폴리펩타이드, 올리고펩타이드, 항생제 내성 유전자, 항균펩타이드, 항진균펩타이드, 올리고펩타이드, 마커 또는 SNP (Single-Nucleotide Polymorphism) 관련 연구에도 본 발명의 방법이 적용될 수 있다.

명세서

발명의 명칭: 대용량 메타지놈 분석을 통한 유용 유전자원 탐색 방법 및 이의 이용

기술분야

- [1] 본 발명은 대용량 메타지놈 분석을 통한 초고속 유용 유전자원 탐색 방법 및 이의 이용에 관한 것이다.

배경기술

- [2] 메타지놈(metagenome)이란 특정 자연 환경에 존재하는 모든 미생물의 유전체 집합으로 정의된다. 미생물은 전체 생물 중 가장 많은 수를 차지하고 있지만, 미생물들이 존재하는 원래의 자연환경과 유사한 조건을 제공해 주기가 어렵기 때문에 환경에 존재하고 있는 미생물들의 99% 이상은 실험실에서 배양할 수 없다. 메타지놈을 이용하면 실험실에서의 배양 가능 여부와 관계없이 자연 환경에서 존재하는 다양한 미생물 유래의 DNA를 직접 추출하여 유전자의 탐색, 동정 및 기능 분석이 가능하다.
- [3] 이러한 메타지놈 시료의 분석을 통해 미생물들의 특정 환경 내에서의 역할을 규명함으로써 미생물을 이용한 폐수처리, 화학적 화합물 생산, 의약품 생산, 바이오 에너지 생산, 산업공정에서의 이용 등 산업, 환경, 에너지, 농업, 의료 등 인간의 삶과 관련된 전반적인 모든 분야에 도움을 줄 수 있다. 최근 메타지놈 연구가 각광을 받는 이유는 앞에서 언급한 대로 그 응용분야인 신물질 또는 신규한 효소의 탐색에 유용하기 때문이다. 특히, 효소는 생물체가 만들어 내는 생촉매(biocatalyst)로써 화학 합성 공정에서 필요한 촉매반응에 미생물 및 메타지놈에서 유래한 효소를 이용하려는 시도가 이루어지고 있다. 효소는 화학촉매에 비해 고도의 조정밀성, 특이성, 선택성 및 고효율성의 특성을 지니기에 다양한 산업분야로의 이용성은 매우 우수하다. 효소의 촉매 효율성은 비슷한 반응 조건 하에서의 비효소적 반응에 비해 $10^8 \sim 10^{14}$ 배나 되며 화학 촉매반응에 비해 반응의 선택성이 매우 높다. 이뿐만 아니라 고온, 고압, 유기용매하의 반응 등 특수한 상황에서도 반응이 가능한 효소가 존재할 경우, 미래 산업용 제제로 활용가치가 매우 높은 품목이다.
- [4] 초기에는 효소의 가격 면에서 경제적 타당성이 매우 낮아 현실적으로 사용이 어려웠으나, 근래는 유전자조작기술, 생물공학기술 및 분자진화기술을 포함한 신기술의 눈부신 발전으로 말미암아 새로운 기능효소개발을 위한 고속탐색(High-throughput screening) 및 개량이 가능하여 신응용 분야로의 확대 적용이 가능하다.
- [5] 이러한 효소의 탐색을 위해 효소의 존재유무 및 활성을 측정할 수 있는 방법에는 미생물을 활용한 탐색법과 배양이 불가(non-culturable)한 메타지놈을 활용한 탐색법이 있다. 전자의 경우, 다양한 균주 컬렉션(Strain collection)이

필수이며, 이들 균주의 직접적인 배양 및 활성 비교를 통한 후보 균주의 선정이 선행되어야 한다. 또한, 많은 노동력 및 연구비가 필요하고 오랜 시간이 소요되는 단점이 있다. 이를 극복하고자 최근에 초고속 탐색 시스템(High-throughput screening system)이 개발 및 활용되고 있으나, 목적하는 유용 효소의 확인을 위해서는 반드시 후보 균주로부터 게놈디엔에이(gDNA)의 정제 및 DNA library 구축, 발현 시스템을 활용한 효소 활성의 재확인 및 선정 library균주의 염기서열 재분석을 통해서만이 목적 유용 효소 자원을 확보할 수 있다. 후자의 경우에는 환경시료로부터 바로 gDNA를 정제한 메타지놈을 활용하여 DNA library를 구축하는 방법이며, 이 또한 발현시스템을 활용한 활성 비교 (Function-based enzyme screening) 또는 디제너러시 프라이머 (degeneracy primers)를 활용한 목적 유용 효소 자원의 DNA를 확보하는 방법 (DNA sequence-based enzyme screening)으로 구분된다. 활성에 기초한 효소 탐색법의 경우 메타지놈의 직접적인 클로닝에 의해 라이브러리(Library)를 제작한 뒤 발현을 통해 기질과 반응하여 원하는 물질로 변환시키는 효소의 존재를 직접적으로 확인하는 방법이다. 이 방식은 유전자원의 존재를 확실하게 확인할 수 있다는 장점이 있지만, 라이브러리의 제작과 발현과정에 상당한 노동력과 시간이 투자되며, 발현 시스템으로 주로 대장균 시스템을 주로 이용하므로 이중 유전자 발현 시스템(Heterologous gene expression system)의 한계를 가진다. 프로모터(promoter)나 코돈(codon)의 구성 또는 사용이 대장균과 맞지 않는 진화적인 유연관계가 먼 세균의 유전자의 경우 그 발현이 영향을 받을 수도 있으며 이러한 노력과 투자에 비해서 탐색의 효율이 낮다는 단점이 있다. 염기서열에 기초한 효소 탐색법의 경우 목적 효소의 특징적인 motif에 근거하여 제작된 프라이머를 이용하여 중합효소연쇄반응을 통해 목적하는 효소의 염기서열을 확인하는 방법이다. 하지만 이 방법 또한 특정 효소의 motif를 근거로 제작한 프라이머를 사용하므로 확보한 염기서열이 어느 한 효소에 치우쳐서 증폭이 되는 문제가 있으며 목적하는 효소가 들어가 있는 DNA library를 직접적으로 시퀀싱을 통한 염기서열의 확인을 해야한다는 문제가 있다.

- [6] 중합효소연쇄반응법(PCR)은 유전자 검출 방법 중 하나로 특정 디옥시리보핵산 영역을 시험관 내에서 중합효소에 의해 증폭하는 방법이다. 중합효소연쇄반응법은 수 시간의 효소 반응으로 유전자의 폭발적인 대량 생산을 가능하게 하였으며, 그 응용성이 매우 높은 디옥시리보핵산 검출의 중심 기법이 되었고, 분자 생물학적인 접근을 필요로 하는 모든 생명과학 분야에서 필수적인 기술이 되었다. 한편, "디제너러시(degeneracy)"는 한 종류의 아미노산이 복수의 코돈(codon)에 대응하는 현상을 말한다. 디제너러시로 인하여 어떤 특이한 아미노산 서열을 코딩하는 유전자의 염기서열은 결정할 수 없지만, 가능성이 있는 염기서열 모두를 포함하는 올리고뉴클레오타이드(oligonucleotide)를 합성할 수 있다. 디제너레이트 올리고뉴클레오타이드를 프라이머로 사용하여 중합효소연쇄반응을 하면

아미노산 서열정보를 바탕으로 그 단백질을 코딩하는 유전자의 DNA 단편이 얻어질 가능성이 있다. 이와 같은 목적으로 디자인된 프라이머를 디제너레이트 프라이머, 중합효소연쇄반응을 디제너레이트 중합효소연쇄반응법이라고 한다. 이 방법은 기지의 아미노산 서열로부터 그 유전자를 클로닝하는 기술이지만 현재는 풍부하게 축적된 핵산의 일차구조(염기서열) 정보를 기초로 유전자군을 검색할 때 널리 이용된다. 효소의 실질적인 존재 유무나 활성은 알 수 없어서 최종적으로는 발현을 통해 확인을 거쳐야 한다는 단점이 있지만, 노동 집약적인 방법에서 벗어나 메타지놈으로부터 목적하는 DNA를 선별할 수 있다는 장점이 있다.

[7]

[8] [선행기술문헌]

[9] 대한민국 출원번호 : 1020090130204

[10]

발명의 상세한 설명

기술적 과제

[11] 이에, 본 발명자들은 산업적 유용 효소 등의 펩타이드 탐색에 있어서 메타지놈 시료의 효용 가치를 신속하고 정확하게 판단하기 위한 방법으로 다양한 효소 등의 펩타이드의 유전자를 대량 증폭하고 증폭된 단편들의 대용량 염기서열의 효율적이고 신속한 분석을 위해 바코드 시퀀스 (bar-code sequence)로 활용할 수 있는 목적 펩타이드에 특이적인 슈퍼패밀리(Superfamily) 디제너레이트 프라이머(degenerate primer)를 설계하고 제작하여 중합연쇄반응에 이용하여 메타지놈 시료로부터 대량으로 증폭된 DNA 단편들을 차세대 염기서열 분석방법을 통하여 유용 펩타이드의 대용량 염기서열을 확인함으로써 본 발명을 완성하였다.

[12] 이에 따라, 본 발명의 일 양상은 목적 유전자의 슈퍼 패밀리 특이적 (superfamily-specific) 디제너레이트 (degenerate) 프라이머 세트를 이용한 메타지놈 정보의 분석 방법에 관한 것이다.

[13] 또한, 본 발명의 다른 양상은 상기 목적 유전자의 슈퍼 패밀리 특이적 (superfamily-specific) 디제너레이트 (degenerate) 프라이머 세트에 관한 것이다.

[14]

과제 해결 수단

[15] 본 발명의 일 양상은

[16] (1) 목적 유전자의 슈퍼 패밀리 특이적 (superfamily-specific) 디제너레이트 (degenerate) 프라이머 세트로 주형인 메타지놈 시료의 유전자를 증폭하는 단계; 및

[17] (2) 상기 유전자의 증폭 산물의 서열을 분석하는 단계; 를 포함하는, 메타지놈 정보의 분석 방법을 제공한다.

- [18] 상기 목적 유전자의 슈퍼패밀리 특이적 디제너레이트 프라이머 세트는
- [19] i) 35% 이상의 아미노산 상동성을 나타내는 목적 단백질을 코딩하는 유전자에 특이적이고;
- [20] ii) 프라이머와 목적 유전자가 결합하는 부위의 서열이 80% 이상의 상동성을 나타내고;
- [21] iii) 디제너러시 비율 (degeneracy ratio)이 1,500 이하이며;
- [22] iv) 어닐링의 온도가 70°C 이하이고;
- [23] v) 프라이머의 길이가 30 염기 이하이며;
- [24] vi) 프라이머 결합부위 안쪽의 상동성이 40% 이하이며;
- [25] vii) 증폭 단편의 크기가 1,000 bp 이하일 수 있다.
- [26] 더욱 바람직하게 상기 프라이머 세트는 상기 디제너러시 비율이 800 이하;
- [27] 상기 프라이머 길이가 25 염기 이하; 및
- [28] 상기 증폭 단편의 크기가 500 bp 이하일 수 있다.
- [29] 상기 메타지놈 시료는 메타지놈 라이브러리, 상기 메타지놈 라이브러리가 도입된 벡터, 또는 상기 벡터로 형질전환된 세포일 수 있다.
- [30] 또한, 상기 메타지놈 시료는 토양, 늪지대, 화산, 갯벌, 염전, 담수, 해수, 채액, 소변, 농산물, 수산물 또는 분변일 수 있다.
- [31] 또한, 상기 목적 유전자는 폴리펩타이드, 올리고펩타이드, 효소, 항생제 내성 유전자, 항균펩타이드, 항진균펩타이드, 올리고펩타이드, 또는 마커를 코딩하는 것 일 수 있다.
- [32] 또한, 상기 (2)의 서열 분석 단계는
- [33] 단일-분자 실시간 시퀀싱 (Single-molecule real-time sequencing), 이온 반도체 시퀀싱 (Ion semiconductor), 파이로시퀀싱 (Pyrosequencing), SBS (Sequencing by synthesis) (Illumina), SBL (Sequencing by ligation), 및 연쇄 정지반응 (Chain termination)으로 이루어진 군으로부터 선택된 1종 이상의 방법을 이용하는 것일 수 있다.
- [34]
- [35] 또한, 본 발명의 다른 양상은 메타지놈 정보 분석을 위하여 목적 유전자의 슈퍼패밀리 특이적 디제너레이트 프라이머 세트로서,
- [36] i) 35% 이상의 아미노산 상동성을 나타내는 목적 단백질을 코딩하는 유전자에 특이적이고;
- [37] ii) 프라이머와 목적 유전자가 결합하는 부위의 서열이 80% 이상의 상동성을 나타내고;
- [38] iii) 디제너러시 비율 (degeneracy ratio)이 1,500 이하이며;
- [39] iv) 어닐링의 온도가 70°C 이하이고;
- [40] v) 프라이머의 길이가 30 염기 이하이며;
- [41] vi) 프라이머 결합부위 안쪽의 상동성이 40% 이하이며;
- [42] vii) 증폭 단편의 크기가 1,000 bp 이하인 것인, 프라이머 세트를 제공한다.

- [43] 상기 프라이머 세트는 바람직하게는 상기 디제너러시 비율이 800 이하;
- [44] 상기 프라이머 길이가 25 염기 이하; 및
- [45] 상기 증폭 단편의 크기가 500 bp 이하인 것일 수 있다.

발명의 효과

- [46] 본 발명에 따른 메타지놈의 정보 분석 방법, 및 이에 이용되는 프라이머 세트는 방대한 메타지놈 샘플 효용 가치의 빠른 판별을 위하여 본 발명이 적용될 수 있다. 특히, 본 발명의 슈퍼패밀리-특이적 (superfamily-specific) 디제너레이트 프라이머를 이용하여 메타지놈에서 목적 펩타이드들의 유전정보의 유무를 간단한 방법에 의해 빠르게 탐색할 수 있으므로 다양한 메타지놈 샘플로부터 고속으로 대량의 유용 펩타이드 자원 정보를 수집하는데 사용이 가능하다. 또한, 이러한 본 발명의 방법을 기반으로 하여 목적하는 새로운 펩타이드들의 슈퍼패밀리-특이적 디제너레이트 프라이머 설계 및 제작을 통하여 새로운 펩타이드 유전자를 탐색하는데 사용이 가능하다. 또한, 이러한 방법은 효소뿐만 아니라, 폴리펩타이드, 올리고펩타이드, 항생제 내성 유전자, 항균펩타이드, 항진균펩타이드, 올리고펩타이드, 마커 또는 SNP (Single-Nucleotide Polymorphism) 관련 연구에도 본 발명의 방법이 적용 될 수 있다.

도면의 간단한 설명

- [47] 도 1은 본 발명의 전체모식도인 대용량 메타지놈 정보분석을 통한 초고속 대용량 효소 탐색방법을 나타낸 것이다.
- [48] 도 2는 본 발명에 사용된 슈퍼패밀리-특이적 (superfamily-specific) 디제너레이트 프라이머 제작 방법을 나타낸 것이다.
- [49] 도 3은 본 발명에 사용된 슈퍼패밀리-특이적 (superfamily-specific) 디제너레이트 프라이머를 이용하여 타겟유전자 증폭을 위한 PCR 최적 조건(A)과 이에 따른 결과(B) 및 음성 대조군(negative control) PCR 결과(C)를 나타낸 것이다.
- [50] 도 4는 본 발명의 방법으로 분석한 메타지놈 샘플 중 일 예로 인도네시아 화산지역 메타지놈 샘플의 미생물 커뮤니티(microbial community) 확인을 위한 문(phylum) 및 속(genus) 분석을 나타낸 것이다.
- [51] 도 5는 본 발명의 방법으로 각각의 슈퍼패밀리 특이적 디제너레이트 프라이머 및 메타지놈 샘플을 사용하여 PCR 수행 후 목적 유전자 확인을 위한 전기영동 분석 결과를 나타낸 것이다.
- [52] 도 6은 본 발명의 방법으로 분석한 유전자 중 일 예로 피루브산 키나아제 (Pyruvate kinase) 유전자의 메타지놈 샘플별 미생물 phylum 분석을 나타낸 것이다.

[53]

발명의 실시를 위한 최선의 형태

- [54] 본 발명의 일 양상은

- [55] (1) 목적 유전자의 슈퍼 패밀리 특이적 (superfamily-specific) 디제너레이트 (degenerate) 프라이머 세트로 주형인 메타지놈 시료의 유전자를 증폭하는 단계; 및
- [56] (2) 상기 유전자의 증폭 산물의 서열을 분석하는 단계; 를 포함하는, 메타지놈 정보의 분석 방법을 제공한다.
- [57] 상기 목적 유전자의 슈퍼패밀리 특이적 디제너레이트 프라이머 세트는
- [58] i) 35% 이상의 아미노산 상동성을 나타내는 목적 단백질을 코딩하는 유전자에 특이적이고;
- [59] ii) 프라이머와 목적 유전자가 결합하는 부위의 서열이 80% 이상의 상동성을 나타내고;
- [60] iii) 디제너러시 비율 (degeneracy ratio)이 1,500 이하이며;
- [61] iv) 어닐링의 온도가 70°C 이하이고;
- [62] v) 프라이머의 길이가 30 염기 이하이며;
- [63] vi) 프라이머 결합부위 안쪽의 상동성이 40% 이하이며;
- [64] vii) 증폭 단편의 크기가 1,000 bp 이하일 수 있다.
- [65] 바람직하게는 상기 프라이머는
- [66] i) 35% 이상 100% 이하의 아미노산 상동성을 나타내는 목적 단백질을 코딩하는 유전자에 특이적이고;
- [67] ii) 프라이머와 유전자가 결합하는 부위의 서열이 80% 이상 100% 이하의 상동성을 나타내고;
- [68] iii) 디제너러시 비율(degeneracy ratio)가 0 이상 1,500 이하이며;
- [69] iv) 어닐링의 온도가 30°C 이상 70°C 이하이고;
- [70] v) 프라이머의 길이가 4 염기 이상 30 염기 이하이며;
- [71] vi) 프라이머 결합부위 안쪽의 상동성이 0% 이상 40% 이하이며;
- [72] vii) 증폭 단편의 크기가 100 bp 이상 1,000 bp 이하일 수 있다.
- [73]
- [74] 더욱 바람직하게 상기 프라이머 세트는 상기 디제너러시 비율이 0 이상 800 이하;
- [75] 상기 프라이머 길이가 4 이상 25 염기 이하; 및
- [76] 상기 증폭 단편의 크기가 100 bp 이상 500 bp 이하일 수 있다.
- [77] 상기 메타지놈 시료는 메타지놈 라이브러리, 상기 메타지놈 라이브러리가 도입된 벡터, 또는 상기 벡터로 형질전환된 세포일 수 있다.
- [78] 또한, 상기 메타지놈 시료는 토양, 늪지대, 화산, 갯벌, 염전, 담수, 해수, 채액, 소변, 농산물, 수산물 또는 분변일 수 있다.
- [79] 상기 메타게놈 라이브러리는 자연 또는 특정 지역에 있는 미생물 군집을 채취하고 유전체를 직접 추출하여 벡터에 도입시킴으로써 이루어질 수 있다. 벡터는 사용자의 목적에 따라 플라스미드, 포스미드(fosmid), 코스미드(cosmid), BAC(bacterial artificial chromosome), YAC(yeast artificial chromosome) 등이

사용될 수 있다.

- [80] 또한, 상기 목적 유전자는 폴리펩타이드, 올리고펩타이드, 효소, 항생제 내성 유전자, 항균펩타이드, 항진균펩타이드, 올리고펩타이드, 또는 마커를 코딩하는 것일 수 있다.
- [81] 상기 폴리펩타이드, 펩티드 또는 올리고펩타이드는 펩타이드 결합을 통한 아미노산의 중합체로, 상기 펩타이드는 효소, 항균 펩타이드 또는 마커일 수 있다.
- [82] 상기 항균 펩타이드는 박테리아, 진균, 프로토조아, 바이러스 등에 대하여 항생 효과를 나타내는 것으로서, 박테리오신, 디펜신, 인돌리시딘, 카텔리시딘, 락소페리신, 라이소자임 등일 수 있으며, 그 종류를 제한하지 않는다.
- [83] 상기 효소는 카탈라아제 하이드로페록시다아제 II(Catalase HydroperoxidaseII, KatE), 카탈라아제 하이드로페록시다아제 I (Catalase HydroperoxidaseI, CatA), 4-아미노부티레이트 트랜스아미나아제 (4-aminobutyrate transaminase, GabT), UDP-N-아세틸글루코사민-2-에피머라아제 (UDP-N-acetylglucosamine-2-epimerase, WegB), L-아라비노스 아이소머라아제 (L-arabinose isomerase, AraA), 피타아제 (Phytase, Phy), 서브틸리신 (Subtilisin, AprE), 셀룰라아제-1 (Cellulase-1, YtoP), 셀룰라아제-2 (Cellulase-2, EglS), 트리오세포스페이트 아이소머라아제 (Triosephosphate isomerase, TpiA), 피루브산 키나아제 (Pyruvate kinase, PyrK) 등일 수 있으며, 그 종류를 제한하지 않는다.
- [84] 상기 마커는 질병의 발병, 예후, 유전형, 약제에 대한 민감성 등을 예측 및 진단하기 위한 표지자로서, 그 질환의 종류 등에 대한 제한이 없다.
- [85] 또한 상기 목적 유전자는 항생제 내성 유전자일 수 있다.
- [86] 상기 항생제 내성은 미생물이 항생제에 노출되어도 생존할 수 있는 약제 내성으로, 이는 다양한 유전적 변형에 의하여 유발될 수 있다. 상기 항생제로는 암피실린, 아미노글리코시드, 베타락탐, 에리스로마이신, 메시실린, 반코마이신, 퀴놀론계 항생제 등일 수 있으며 그 종류에 한정되지 않는다.
- [87]
- [88] 또한, 상기 (2)의 서열 분석 단계는
- [89] 단일-분자 실시간 시퀀싱 (Single-molecule real-time sequencing), 이온 반도체 시퀀싱 (Ion semiconductor), 파이로시퀀싱 (Pyrosequencing), SBS (Sequencing by synthesis) (Illumina), SBL (Sequencing by ligation), 및 연쇄 정지반응 (Chain termination)으로 이루어진 군으로부터 선택된 1종 이상의 방법을 이용하는 것일 수 있다.
- [90] 또한, 본 발명의 다른 양상은 메타지놈 정보 분석을 위하여 목적 유전자의 슈퍼패밀리 특이적 디제너레이트 프라이머 세트로서,
- [91] i) 35% 이상의 아미노산 상동성을 나타내는 목적 단백질을 코딩하는 유전자에 특이적이고;

- [92] ii) 프라이머와 목적 유전자가 결합하는 부위의 서열이 80% 이상의 상동성을 나타내고;
- [93] iii) 디제너러시 비율 (degeneracy ratio)이 1,500 이하이며;
- [94] iv) 어닐링의 온도가 70°C 이하이고;
- [95] v) 프라이머의 길이가 30 염기 이하이며;
- [96] vi) 프라이머 결합부위 안쪽의 상동성이 40% 이하이며;
- [97] vii) 증폭 단편의 크기가 1,000 bp 이하인 것인, 프라이머 세트를 제공한다.
- [98] 상기 프라이머 세트는 바람직하게는
- [99] i) 35% 이상 100% 이하의 아미노산 상동성을 나타내는 목적 단백질을 코딩하는 유전자에 특이적이고;
- [100] ii) 프라이머와 유전자가 결합하는 부위의 서열이 80% 이상 100% 이하의 상동성을 나타내고;
- [101] iii) 디제너러시 비율(degeneracy ratio)가 0 이상 1,500 이하이며;
- [102] iv) 어닐링의 온도가 30°C 이상 70°C 이하이고;
- [103] v) 프라이머의 길이가 4 염기 이상 30 염기 이하이며;
- [104] vi) 프라이머 결합부위 안쪽의 상동성이 0% 이상 40% 이하이며;
- [105] vii) 증폭 단편의 크기가 100 bp 이상 1,000 bp 이하일 수 있다.
- [106] 상기 프라이머 세트는 더욱 바람직하게는
- [107] 상기 디제너러시 비율이 0 이상 800 이하;
- [108] 상기 프라이머 길이가 4 이상 25 염기 이하; 및
- [109] 상기 증폭 단편의 크기가 100 bp 이상 500 bp 이하일 수 있다.
- [110] 본 명세서에서 사용되는 용어 프라이머는 합성 또는 자연의 올리고뉴클레오타이드를 의미한다. 프라이머는 핵산쇄(주형)에 상보적인 프라이머 연장 산물의 합성이 유도되는 조건, 즉, 뉴클레오타이드와 DNA 중합효소와 같은 중합체의 존재, 그리고 적합한 온도와 pH의 조건에서 합성의 개시점으로 작용한다. 증폭의 최대 효율을 위하여, 바람직하게는 프라이머는 단일쇄이다. 바람직하게는, 프라이머는 올리고디옥시리보뉴클레오타이드이다. 본 발명의 프라이머는 자연(naturally occurring) dNMP(즉, dAMP, dGMP, dCMP 및 dTMP), 변형 뉴클레오타이드 또는 비-자연 뉴클레오타이드를 포함할 수 있다.
- [111] 바람직하게는, 본 발명의 목적 유전자의 슈퍼패밀리 특이적 디제너레이트 프라이머 세트는 서열번호 1에 나타난 염기서열로 이루어진 포워드 디제너레이트 프라이머 및 서열번호 2에 나타난 염기서열로 이루어진 리버스 디제너레이트 프라이머 세트; 서열번호 3에 나타난 염기서열로 이루어진 포워드 디제너레이트 프라이머 및 서열번호 4에 나타난 염기서열로 이루어진 리버스 디제너레이트 프라이머 세트; 서열번호 5에 나타난 염기서열로 이루어진 포워드 디제너레이트 프라이머 및 서열번호 6에 나타난 염기서열로 이루어진 리버스 디제너레이트 프라이머 세트; 서열번호 7에 나타난 염기서열로 이루어진 포워드

디제너레이트 프라이머 및 서열번호 8에 나타난 염기서열로 이루어진 리버스 디제너레이트 프라이머 세트; 서열번호 9에 나타난 염기서열로 이루어진 포워드 디제너레이트 프라이머 및 서열번호 10에 나타난 염기서열로 이루어진 리버스 디제너레이트 프라이머 세트; 서열번호 11에 나타난 염기서열로 이루어진 포워드 디제너레이트 프라이머 및 서열번호 12에 나타난 염기서열로 이루어진 리버스 디제너레이트 프라이머 세트; 서열번호 13에 나타난 염기서열로 이루어진 포워드 디제너레이트 프라이머 및 서열번호 14에 나타난 염기서열로 이루어진 리버스 디제너레이트 프라이머 세트; 서열번호 15에 나타난 염기서열로 이루어진 포워드 디제너레이트 프라이머 및 서열번호 16에 나타난 염기서열로 이루어진 리버스 디제너레이트 프라이머 세트; 서열번호 17에 나타난 염기서열로 이루어진 포워드 디제너레이트 프라이머 및 서열번호 18에 나타난 염기서열로 이루어진 리버스 디제너레이트 프라이머 세트; 서열번호 19에 나타난 염기서열로 이루어진 포워드 디제너레이트 프라이머 및 서열번호 20에 나타난 염기서열로 이루어진 리버스 디제너레이트 프라이머 세트; 또는 서열번호 21에 나타난 염기서열로 이루어진 포워드 디제너레이트 프라이머 및 서열번호 22에 나타난 염기서열로 이루어진 리버스 디제너레이트 프라이머 세트를 포함하며, 목적 유전자의 서열에 따라 달라질 수 있는바 이에 제한되지 않는다.

[112] 또한, 상기 프라이머의 서열을 나타냄에 있어 IUB 중의성(ambiguity) 코드(codes)를 사용한다. 따라서 프라이머 서열에서 A는 아데닌(adenine)을, G는 구아닌(guanine)을, C는 시토신(cytosine)을, T는 티민(thymine)을 나타낸다. 또한, R은 G 또는 A, Y는 T 또는 C, M은 A 또는 C, K는 G 또는 T, S는 G 또는 C, W는 A 또는 T, B는 G 또는 C 또는 T, D는 A 또는 G 또는 T, H는 A 또는 C 또는 T, V는 A 또는 G 또는 C, N은 A 또는 G 또는 C 또는 T를 나타낸다. 이는 본 명세서의 모든 서열을 제시하는 데에 있어 동일하게 적용하였다.

[113]

[114] 또한, 본 발명의 또 다른 양상은 상술한 프라이머 세트를 포함하는 메타지놈 분석 키트를 제공한다.

[115] 본 발명의 키트는 PCR 반응에 필요한 시약, 예컨대, 완충액, DNA 중합효소, DNA 중합효소 조인자 및 디옥시리보뉴클레오타이드-5'-트리포스페이트를 포함할 수 있다. 선택적으로, 본 발명의 키트는 다양한 폴리뉴클레오타이드 분자, 역전사 효소, 다양한 완충액 및 시약, DNA 중합효소의 활성을 억제하는 항체를 포함할 수 있다. 특정 반응에서 이용되는 시약의 최적의 양은 본 명세서의 내용을 파악한 당업자에 의해 용이하게 결정될 수 있다. 전형적으로, 본 발명의 키트는 분리된 패키지 또는 구획에 상술한 성분들을 포함한다.

[116]

[117] 또한, 본 발명의 또 다른 양상은 하기 조건을 만족하는 염기서열을 선별하는 단계를 포함하는, 메타지놈 정보 분석을 위한 목적 유전자의 슈퍼패밀리 특이적

디제너레이트 프라이머 세트의 선별방법을 제공한다:

- [118] i) 35% 이상의 아미노산 상동성을 나타내는 목적 단백질을 코딩하는 유전자에 특이적임;
- [119] ii) 목적 유전자에 결합하는 부위의 서열이 80% 이상의 상동성을 나타냄;
- [120] ii) 디제너러시 비율 (degeneracy ratio)이 1,500 이하임;
- [121] iv) 어닐링의 온도가 70°C 이하임;
- [122] v) 프라이머의 길이가 30 염기 이하임;
- [123] vi) 프라이머 결합부위 안쪽의 상동성이 40% 이하임; 및
- [124] vii) 증폭 단편의 크기가 1,000 bp 이하임.
- [125] 상기 메타지놈 정보 분석을 위한 목적 유전자의 슈퍼패밀리 특이적 디제너레이트 프라이머 세트의 선별방법은 바람직하게는
- [126] i) 35% 이상 100% 이하의 아미노산 상동성을 나타내는 목적 단백질을 코딩하는 유전자에 특이적임;
- [127] ii) 프라이머와 유전자가 결합하는 부위의 서열이 80% 이상 100% 이하의 상동성을 나타냄;
- [128] iii) 디제너러시 비율(degeneracy ratio)가 0 이상 1,500 이하임;
- [129] iv) 어닐링의 온도가 30°C 이상 70°C 이하임;
- [130] v) 프라이머의 길이가 4 염기 이상 30 염기 이하임;
- [131] vi) 프라이머 결합부위 안쪽의 상동성이 0% 이상 40% 이하임; 및
- [132] vii) 증폭 단편의 크기가 100 bp 이상 1,000 bp 이하;인 조건을 만족하는 염기서열을 선별하는 단계를 포함한다.
- [133] 상기 메타지놈 정보 분석을 위한 목적 유전자의 슈퍼패밀리 특이적 디제너레이트 프라이머 세트의 선별방법은 보다 바람직하게는
- [134] 상기 디제너러시 비율이 0 이상 800 이하임;
- [135] 상기 프라이머 길이가 4 이상 25 염기 이하임; 및 상기 증폭 단편의 크기가 100 bp 이상 500 bp 이하;인 조건을 만족하는 염기서열을 선별하는 단계를 포함한다.
- [136]
- [137] 이하, 실시예는 오로지 본 발명을 보다 구체적으로 설명하기 위한 것으로서, 본 발명의 요지에 따라 본 발명의 범위가 이들 실시예에 의해 제한되지 않는다는 것은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 있어서 자명할 것이다.

[138]

[139] 제조예 : Superfamily-specific 디제너레이트 프라이머의 제조

[140] 본 발명에서는 목적 펩타이드의 예시로서 Catalase HydroperoxidaseII (KatE), Catalase HydroperoxidaseI (CatA), 4-aminobutyrate transaminase (GabT), UDP-N-acetylglucosamine-2-epimerase (WegB), L-arabinose isomerase (AraA), Phytase (Phy), Subtilisin (AprE), Cellulase1 (YtoP), Cellulase2 (EglS), Triosephosphate isomerase (TpiA), 및 Pyruvate kinase (PyrK) 효소들의 유전자

정보를 사용하였다. 상기 목적 단백질과의 아미노산 상동성이 35%인 것을 조건으로, 각각의 효소 유전자의 아미노산 서열을 얼라인먼트(alignment)하고 상동성이 80% 이상으로 높은 아미노산 서열을 기반으로 하여 이들을 코딩하는 염기 서열을 확인한 후 다시 얼라인먼트 (alignment)를 하였다. 염기서열의 상보성을 기반으로 하여 Superfamily-specific 디제너레이트 프라이머의 디제너러시 비율(1,500 이하, 바람직하게 800이하), 어닐링 온도값(70°C 이하), 프라이머 길이 (30 염기 이하, 바람직하게 25 염기 이하)를 만족하고, 프라이머 결합 부위 안쪽 염기서열의 상동성이 40% 이하인 낮은 부위 및 증폭 단편의 크기가 1,000 bp 이하(또는 선택적으로 500 bp이하)의 조건을 만족하도록 결정된 염기서열을 바탕으로 하여 디제너레이트 프라이머를 제작하였다.

- [141] 본 발명은 상기와 같은 과정으로 제작된 서열번호 1에 나타난 염기서열로 이루어진 포워드 디제너레이트 프라이머 및 서열번호 2에 나타난 염기서열로 이루어진 리버스 디제너레이트 프라이머로 구성된 상기에 Catalase HydroperoxidaseII (KatE) 유전자 증폭용 프라이머를 제공한다. 여기서, 포워드 디제너레이트 프라이머 및 리버스 디제너레이트 프라이머는 공지된 효소들의 유전자를 증폭하는데 사용된 디제너레이트 프라이머 뿐만 아니라 이와 동일한 방식에 의해 공지된 효소들의 유전자를 증폭할 것으로 추정되는 디제너레이트 프라이머를 모두 포함한다.
- [142] 또한, 서열번호 1 및 2의 프라이머 세트 외에 동일한 방법으로 제조된 디제너레이트 프라이머의 구체적인 예시를 하기에 개시하였다:
- [143] catalase Hydroperoxidase II F: 5'- CRC TTY GAY CAY GAR MGB ATY CC-3' (서열번호 1);
- [144] catalase Hydroperoxidase II R: 5'- ATG AAR AAS ACH GGV AWG TTR TTB CC-3' (서열번호 2);
- [145] catalase Hydroperoxidase I F: 5'- TGG CCN RTH AAG MAR AAR TAY GG-3' (서열번호 3);
- [146] catalase Hydroperoxidase I R: 5'- ACR GTY TCY TCR TCR TTC ATN SCC-3' (서열번호 4);
- [147] 4-aminobutyrate transaminase F: 5'- SCV GAR GCV GTY GAR AAC GC-3' (서열번호 5);
- [148] 4-aminobutyrate transaminase R: 5'- GCC HTC DCC BTG HAY CGG YTC-3' (서열번호 6);
- [149] UDP-N-acetylglucosamine-2-epimerase F: 5'- TCA YGT BGA RGC BGG VYT GMG-3' (서열번호 7);
- [150] UDP-N-acetylglucosamine-2-epimerase R: 5'- TG HAY VCC DCC VGA RTC VGT-3' (서열번호 8);
- [151] L-arabinose isomerase F: 5'- GAA CYT NAA CCA RDC NGC NCA CGG-3' (서열번호 9);

- [152] L-arabinose isomerase R: 5'- CCA GTC DCC TTC NSC NSC RAA NCC-3'
(서열번호 10);
- [153] Phytase F: 5'- ACN CKK AAA ACV GGY GAR GC-3' (서열번호 11);
- [154] Phytase R: 5'- CAR SAG RAA WAT YTC VGY MAG C-3' (서열번호 12);
- [155] Subtilisin F: 5'- CAC GGH ACN CAT KKN GCN GG-3' (서열번호 13);
- [156] Subtilisin R: 5'- GGA GWB GCC ATD SWB GTD CCG-3' (서열번호 14);
- [157] Cellulase-1 F: 5'- ATV TAY GTH ATH RTY GAY TGG C-3' (서열번호 15);
- [158] Cellulase-1 R: 5'- RTG HGT DSC BGH ATA RAA ATG-3' (서열번호 16);
- [159] Cellulase-2 F: 5'- GSN CAY WTG GAY GAR RTS GG-3' (서열번호 17);
- [160] Cellulase-2 R: 5'- CCN AHY TCY TCC TGN ACN GT-3' (서열번호 18);
- [161] Triosephosphate isomerase F: 5'- GGN GCV TMY ACH GGY GAR-3' (서열번호 19);
- [162] Triosephosphate isomerase R: 5'- GC CCA NAY NGG YTC RTA VGC-3'
(서열번호 20);
- [163] Pyruvate kinase F: 5'- CGN HTN AAC TTY TCS CAY GG-3' (서열번호 21);
- [164] Pyruvate kinase R: 5'- CAT BSA STC DAG CAT YTG HGT CGC-3' (서열번호 22).
- [165]
- [166] 따라서, 본 발명은 효소에 존재하는 보존 영역의 염기서열에 상보적인 정방향 디제너레이트 프라이머와 역방향 디제너레이트 프라이머 쌍을 제공하는 것을 목적으로 한다.
- [167] 본 발명의 디제너레이트 프라이머는 특정 효소 superfamily에 속하는 효소 유전자들의 아미노산 서열 분석을 통하여 상동성이 높은 부분(80% 이상)을 선택적으로 증폭하여 검출하기 위한 것이며, 특정 효소 유전자의 유전자의 상보성이 높은 서열에 결합하게 된다.
- [168] 상기 예시한 11 쌍의 디제너레이트 프라이머들을 이용하여 임의의 균주 및 메타지놈에 위에서 언급한 11가지 효소들이 존재하는지 여부를 확인할 수 있다.
- [169]
- [170] 실시예 1: Superfamily-specific 디제너레이트 프라이머의 설계 및 제작
- [171] 기존에 카탈레이즈 하이드로펄옥시데이즈II (KatE)를 생산하는 것으로 알려진 다양한 도메인(Domain)별 대표적인 미생물들의 유전자를 데이터베이스에서 검색한 후 선정된 유전자의 아미노산 서열을 최소 5종류 이상 사용하여 얼라이언트를 수행하였다. 상기 펩타이드에서 상동성이 35% 이상인 것을 1차 선별하였다.
- [172] 다수의 유전자간에 상동성이 80% 이상으로 높은 서열을 2차 선정하여 디제너레이트 프라이머 설계를 위한 구역으로 선정하였다. 선정된 상동성이 높은 아미노산 서열을 다시 염기서열로 변환한 후, 디제너레이트 프라이머 제작시 중요한 요소인 디제너러시 비율(1,500 이하, 바람직하게 800 이하),

어닐링 온도값 (70°C 이하), 프라이머 길이(30 염기이하, 바람직하게 25 염기이하) 조건을 한정하고, 프라이머 결합 부위 안쪽 염기서열의 상동성이 40% 이하로 낮으며, 증폭단편의 크기가 1,000 bp 이하(바람직하게 500 bp이하)가 되도록 결정된 염기서열을 바탕으로 하여 디제너레이트 프라이머를 제작하였다. 다른 공지 효소를 검출하기 위한 나머지 디제너레이트 프라이머들 또한 상기의 방법으로 설계 및 제작하였다.

[173] 디제너레이트 프라이머 설계의 일 예를 도 2에 나타내었다.

[174]

[175] 실시예 2: 중합효소연쇄반응(PCR) 수행 및 증폭된 유전자 단편의 확인

[176] 제작한 카탈레이즈 하이드로필옥시데이즈II (KatE) 탐색용 디제너레이트 프라이머들 (서열번호 1; 5'-CRC TTY GAY CAY GAR MGB ATY CC-3'와 서열번호 2; 5'-ATG AAR AAS ACH GGV AWG TTR TTB CC-3')의 목적 유전자와의 결합 유무를 확인하기 위하여, 본 디제너레이트 프라이머 제작을 위해 KatE 유전자의 아미노산 서열이 사용된 대장균 MG1655의 게놈 디엔에이(genomic DNA)를 주형으로 하여 중합연쇄반응(PCR)을 실시하였다. 나머지 디제너레이트 프라이머들 또한 대장균 MG1655의 게놈 디엔에이(genomic DNA)를 주형으로 하여 중합연쇄반응(PCR)을 실시하였다. 예외적으로 서브틸리신과 셀룰레이스 탐색용 디제너레이트 프라이머들의 목적 유전자와의 결합 유무를 확인하기 위해선, 대장균 MG1655이 아닌 바실러스 서브틸리스 168의 게놈 디엔에이(genomic DNA)를 주형으로 하여 중합연쇄반응(PCR)을 실시하였다.

[177] 이용된 PCR 반응액은 다음과 같다: 100 ng 주형 (수득한 대장균 MG1655 및 바실러스 서브틸리스 168의 게놈 DNA), 12.5 μ l EmeraldAmp GT PCR Master Mix, 0.2 μ M(최종 농도) 각 Forward 및 Reverse 디제너레이트 프라이머, 멸균 증류수를 포함하는 최종 25 μ l의 용액. PCR 반응은 94°C에서 30초의 변성, X°C(54°C 카탈레이즈 하이드로필옥시데이즈II, 카탈레이즈 하이드로필옥시데이즈I; 58°C 가바 트랜스아미네이즈, UDP-N-아세틸글루코사민-2-에피머레이즈, L-아라비노즈 아이소머레이즈; 52°C 파이테이즈, 셀룰레이스2; 56°C 서브틸리신, 파이루베이트 카이네이즈, 51°C 트리오포스페이트 아이소머레이즈, 47°C 셀룰레이스1) 에서 30초의 어닐링 및 72°C에서 30초의 신장으로 구성된 사이클을 30회 수행하였다. PCR 산물을 도 3 과 같이 확인하였다.

[178] 도 3 에 나타난 바와 같이, 증폭된 PCR 산물은 2% 아가로스 겔에서 전기영동을 수행하여 확인하였다. 대장균 MG1655와 바실러스 서브틸리스 168에서 예상했던 크기의 증폭된 유전자 단편(약 280 bp 카탈레이즈 하이드로필옥시데이즈II; 약 350 bp 카탈레이즈 하이드로필옥시데이즈I; 약 300 bp 가바 트랜스아미네이즈; 약 540 bp UDP-N-아세틸글루코사민-2-에피머레이즈; 약 570 bp L-아라비노즈

아이소머레이즈; 약 400 bp 파이테이즈; 약 490bp 서브틸리신; 약 335 bp 셀룰레이스; 약 300 bp 트리오포스페이트 아이소머레이즈, 약 760 bp 파이루베이트 카이나제)을 확인할 수 있었다.

- [179] 또한, 본 발명에 개시된 조건이 최적 조건임을 확인하기 위하여, 음성대조군으로서 PCR 반응 조건 중 어닐링 온도를 70°C 이상으로 설정하여 PCR을 실시하였다. 그 결과, 도 3C에 나타난 바와 같이, 어닐링 온도가 70°C 이상일 때는, 목적하는 단편들이 증폭되지 않는 것을 확인하였다.

[180]

[181] 실시예 3: 증폭된 유전자의 염기서열 분석

- [182] 상기 PCR 산물을 Qiaquick Gel extraction kit (QIAGEN, 독일)를 사용하여 정제한 후, pGEM T-easy 벡터(Promega, 미국)에 라이게이션시켜 대장균 MG1655의 카탈레이즈 하이드로필옥시테이즈 II, 카탈레이즈 하이드로필옥시테이즈 I, 가바 트랜스아미네이즈, UDP-N-아세틸글루코사민 2-에피머화효소, L-아라비노즈 아이소머레이즈, 파이테이즈, 서브틸리신, 셀룰레이스, 트리오포스페이트 아이소머레이즈, 파이루베이트 카이나제를 암호화하는 유전자가 포함된 벡터 pGEM T-katE, catA, gabT, wegB, araA, phy, aprE, ytoP, eglS, tpiA, pyrK를 제작하였다.

- [183] 또한, 상기 재조합 벡터를 형질전환을 위해 준비된 100 μ l의 대장균 DH5 (competent cell)와 혼합한 후, 42°C에서 90초간 열충격 (heat-shock)으로 상기 벡터가 도입된 형질전환 대장균을 제조하고, 암피실린 100 μ g/ml을 포함하는 LB배지에 도말하여 형질전환된 대장균을 선별하였다. 플라스미드 DNA를 회수하기 위해 플라스미드 추출 키트 (Biofact, 한국)를 이용하여 플라스미드 DNA를 회수하고 염기서열 분석을 Solgent 회사(한국)에 의뢰하였다. 염기서열을 분석한 결과, 대장균 MG1655와 바실러스 서브틸리스 168에 포함된 상기 효소들의 유전자와 PCR 산물이 동일한 유전자임을 확인하였다.

[184]

[185] 실시예 4: 메타지놈에 중합효소연쇄반응(PCR) 수행 및 증폭된 유전자 단편의 확인

- [186] 제작한 디제너레이트 프라이머들의 목적 유전자와의 결합능을 확인한 후 메타지놈을 주형으로 활용하여 목적 유전자의 탐색을 위한 중합효소연쇄반응을 수행하였다. 사용된 PCR 반응액은 다음과 같다: 100 ng 주형 (수득한 Soil A 메타지놈 DNA), 12.5 μ l EmeraldAmp GT PCR Master Mix, 0.2 μ M(최종 농도) 각 Forward 및 Reverse 프라이머, 멸균 증류수를 포함하는 최종 25 μ l의 용액. PCR 반응은 94°C에서 30초의 변성, X°C(54°C 카탈레이즈 하이드로필옥시테이즈II, 카탈레이즈 하이드로필옥시테이즈I; 58°C 가바 트랜스아미네이즈, UDP-N-아세틸글루코사민-2-에피머레이즈, L-아라비노즈 아이소머레이즈; 52°C 파이테이즈, 셀룰레이스2; 56°C 서브틸리신, 파이루베이트 카이나제, 51°C 트리오포스페이트 아이소머레이즈, 47°C 셀룰레이스1)에서 30초의 어닐링 및

72°C에서 30초의 신장으로 구성된 사이클을 30회 수행하였다. PCR 산물을 도 5와 같이 확인하였다.

[187] 도 5에 나타난 바와 같이, 증폭된 PCR 산물은 2% 아가로스 겔에서 전기영동을 수행하여 확인하였다. Lane 1이 암호화 유전자의 PCR 증폭산물이다.

[188]

[189] 실시예 5: 메타지놈에서 증폭된 유전자의 염기서열 분석

[190] 8 종류의 메타지놈을 주형으로 사용한 중합효소연쇄반응에서 다양한 목적 유전자의 증폭 양상을 확인하였다. 차세대 염기서열 분석장비의 활용을 하기에 앞서, 메타지놈을 주형으로 한 PCR에서도 제작한 디제너레이트 프라이머들이 목적 유전자들과의 결합능이 있는지를 확인하기 위하여 Soil A 메타지놈을 주형으로 사용하여 PCR을 수행하여 획득한 증폭단편들(약 300 bp 카탈레이즈 하이드로필옥시데이즈II, 약 350 bp 카탈레이즈 하이드로필옥시데이즈I, 약 300 bp 가바 트랜스아미네이즈)을 Qiaquick Gel extraction kit (QIAGEN, 독일)를 사용하여 정제한 후, pGEM T-easy 벡터에 클로닝하여 일반적인 sanger sequencing을 통한 염기서열 분석 결과, 상기의 목적 효소 유전자들이 검출되는 것은 물론이거니와 상기의 효소들의 유전자를 포함하고 있는 다양한 균주들이 검색되는 것을 NCBI 웹사이트의 Blast 검색결과를 통하여 확인할 수 있었다.

[191] 상기의 실험을 통하여 메타지놈으로부터 다양한 목적 유전자의 정보 수집이 가능한 것을 확인한 후 8 종류의 메타지놈을 주형으로 하여 PCR을 통하여 획득한 증폭 단편들을 차세대 염기서열 분석장비 활용에 필요한 충분한 양을 확보하기 위해 다시 한번 획득한 증폭 단편들을 주형으로 사용하고 각각의 효소에 맞는 Superfamily-specific 디제너레이트 프라이머를 사용하여 실시예 4의 PCR 수행 조건으로 PCR을 재수행하였다. 증폭된 PCR산물을 회수한 후 메타지놈별로 증폭된 PCR산물을 모아서 총 8개의 샘플로 차세대 염기서열 분석장비 (Miseq)를 활용하여 염기서열을 분석하였다. 하나의 메타지놈에 여러 종류의 PCR 산물이 혼합되어 있기에 염기서열 결과 분석을 위하여 Superfamily-specific 디제너레이트 프라이머의 염기서열을 barcode sequences로 사용하여 염기서열을 분리, 분석을 수행하였다.

[192] 하기 표 1 및 표 2는 본 발명에 따른 메타지놈 정보 분석방법을 활용하여 다수의 메타지놈 샘플의 유전자 정보를 분석한 결과이다.

[193] [표 1]

[194] 분석 메타지놈 샘플의 microbial community 분석에 따른 phylum 및 genus 개수

[195]

	Phylum	Genus
Ind	6	26
Halo	5	38
HFD	7	25
HFD-M	8	33
Upo	53	1064
Soil A	69	1633
SB	23	380
New-R	32	888

[196]

[197] [표 2]

[198] 분석 메타지놈 샘플에 따른 타겟 효소별 phylum 및 genus 개수 분석결과

[199]

		Ind	Halo	HFD	HFD-M	Upo	Soil A	SB	New-R
PyrK	Phylum	18	12	22	22	31		31	
	Genus	222	252	363	501	961		841	
TpiA	Phylum	22	18	18	20	35	37		
	Genus	352	139	193	238	965	1078		
KatE	Phylum	16	16			30	23		18
	Genus	209	286			792	577		478
CatA	Phylum	19	14	18	22	24	27	27	19
	Genus	178	218	334	384	661	601	525	548
GabT	Phylum	18	12	17			32	22	14
	Genus	113	130	265			882	326	258
WegB	Phylum	17		21	19				
	Genus	145		287	276				
AraA	Phylum	12	11	14	19	29			
	Genus	81	77	199	213	571			
Phy	Phylum				18				
	Genus				275				
AprE	Phylum	15							
	Genus	83							
YtoP	Phylum	17	14	21	17				14
	Genus	162	218	251	269				258
EglS	Phylum	17	11	15	17	34	31		
	Genus	134	124	172	211	825	738		

[200]

[201] 또한, 본 발명에 따른 메타지놈 유전자 정보 분석에 대한 방법을 도 1에 도시화하였다. 도 1은 대용량 메타지놈 정보분석을 통한 초고속 대용량 효소 탐색방법을 나타낸 것이다.

[202]

[203] 이상의 실험결과를 통하여, 본 발명에 따른 메타지놈의 정보 분석 방법, 및 이에 이용되는 프라이머 세트는 방대한 메타지놈 샘플 효용 가치의 빠른 판별을 위하여 본 발명이 적용될 수 있다. 특히, 본 발명의 슈퍼패밀리-특이적 (superfamily-specific) 디제너레이트 프라이머를 이용하여 메타지놈에서 목적 펩타이드들의 유전정보의 유무를 간단한 방법에 의해 빠르게 탐색할 수 있으므로 다양한 메타지놈 샘플로부터 고속으로 대량의 유용 펩타이드 자원 정보를 수집하는데 사용이 가능하다. 또한, 이러한 본 발명의 방법을 기반으로 하여 목적하는 새로운 펩타이드들의 슈퍼패밀리-특이적 디제너레이트 프라이머 설계 및 제작을 통하여 새로운 펩타이드 유전자를 탐색하는데 사용이 가능하다. 또한, 이러한 방법은 효소뿐 만 아니라, 폴리펩타이드, 올리고펩타이드, 항생제 내성 유전자, 항균펩타이드, 항진균펩타이드, 올리고펩타이드, 마커 또는 SNP (Single-Nucleotide Polymorphism) 관련 연구에도 본 발명의 방법이 적용될 수 있다.

청구범위

- [청구항 1] (1) 목적 유전자의 슈퍼 패밀리 특이적 (superfamily-specific) 디제너레이트 (degenerate) 프라이머 세트에 주형인 메타지놈 시료의 유전자를 증폭하는 단계; 및
(2) 상기 유전자의 증폭 산물의 서열을 분석하는 단계; 를 포함하는, 메타지놈 정보의 분석 방법.
- [청구항 2] 청구항 1에 있어서, 목적 유전자의 슈퍼패밀리 특이적 디제너레이트 프라이머 세트는
i) 35% 이상 100%이하의 아미노산 상동성을 나타내는 목적 단백질을 코딩하는 유전자에 특이적이고;
ii) 프라이머와 목적 유전자가 결합하는 부위의 서열이 80% 이상 100% 이하의 상동성을 나타내고;
iii) 디제너러시 비율 (degeneracy ratio)이 0 이상 1,500 이하이며;
iv) 어닐링의 온도가 30°C 이상 70°C 이하이고;
v) 프라이머의 길이가 4 염기 이상 30 염기 이하이며;
vi) 프라이머 결합부위 안쪽의 상동성이 0% 이상 40% 이하이며;
vii) 증폭 단편의 크기가 100 bp 이상 1,000 bp 이하인 것인, 방법.
- [청구항 3] 청구항 2에 있어서, 상기 프라이머 세트는 상기 디제너러시 비율이 0 이상 800 이하; 상기 프라이머 길이가 4 이상 25 염기 이하; 및 상기 증폭 단편의 크기가 100 bp 이상 500 bp 이하인 것인, 방법.
- [청구항 4] 청구항 1에 있어서, 상기 메타지놈 시료는 메타지놈 라이브러리, 상기 메타지놈 라이브러리가 도입된 벡터, 또는 상기 벡터로 형질전환된 세포인 것인, 방법.
- [청구항 5] 청구항 1에 있어서, 상기 메타지놈 시료는 토양, 늪지대, 화산, 갯벌, 염전, 담수, 해수, 체액, 소변, 농산물, 수산물 또는 분변인 것인, 방법.
- [청구항 6] 청구항 1에 있어서, 상기 목적 유전자는 폴리펩타이드, 올리고펩타이드, 효소, 항생제 내성 유전자, 항균펩타이드, 항진균펩타이드, 올리고펩타이드, 또는 마커를 코딩하는 것인, 방법.
- [청구항 7] 청구항 1에 있어서, 상기 (2)의 서열 분석 단계는 단일-분자 실시간 시퀀싱 (Single-molecule real-time sequencing), 이온 반도체 시퀀싱 (Ion semiconductor), 파이로시퀀싱 (Pyrosequencing), SBS (Sequencing by synthesis), SBL (Sequencing by ligation), 및 연쇄 정지반응 (Chain termination)으로 이루어진 군으로부터 선택된 1종 이상의 방법을 이용하는 것인, 방법.
- [청구항 8] 메타지놈 정보 분석을 위한 목적 유전자의 슈퍼패밀리 특이적 디제너레이트 프라이머 세트로서,
i) 35% 이상 100%이하의 아미노산 상동성을 나타내는 목적 단백질을

코딩하는 유전자에 특이적이고;

ii) 프라이머와 목적 유전자가 결합하는 부위의 서열이 80% 이상 100% 이하의 상동성을 나타내고;

iii) 디제너러시 비율 (degeneracy ratio)이 0 이상 1,500 이하이며;

iv) 어닐링의 온도가 30°C 이상 70°C 이하이고;

v) 프라이머의 길이가 4 염기 이상 30 염기 이하이며;

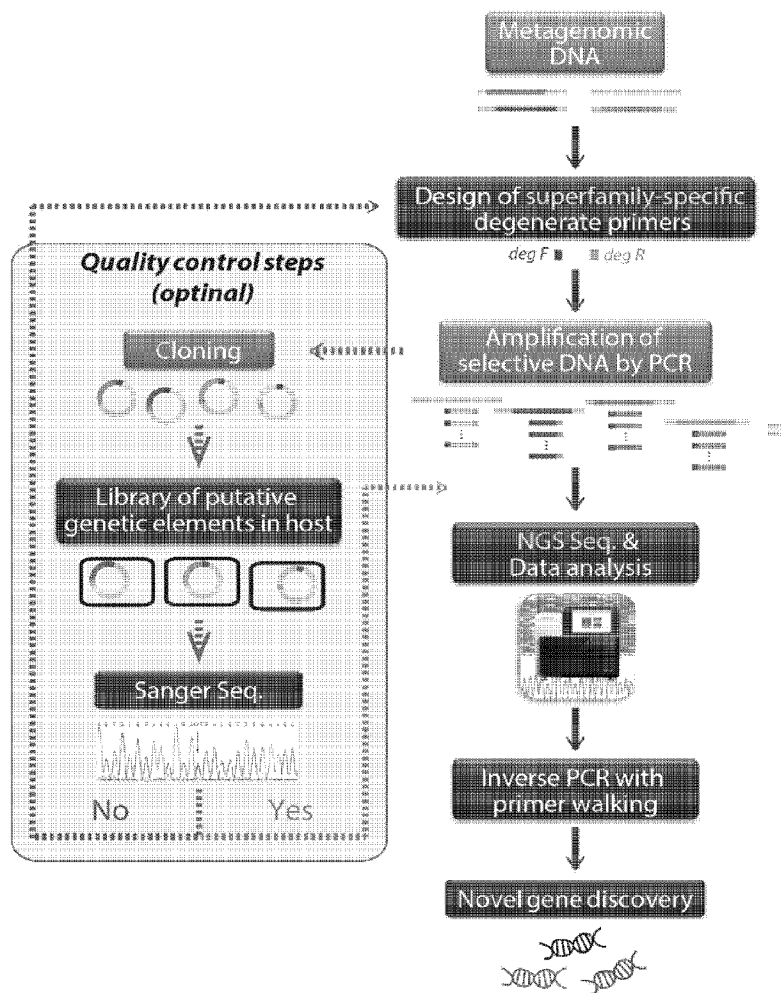
vi) 프라이머 결합부위 안쪽의 상동성이 0% 이상 40% 이하이며;

vii) 증폭 단편의 크기가 100 bp 이상 1,000 bp 이하인 것인, 프라이머 세트.

[청구항 9] 청구항 8에 있어서, 상기 프라이머 세트는 상기 디제너러시 비율이 0 이상 800 이하; 상기 프라이머 길이가 4 이상 25 염기 이하; 및 상기 증폭 단편의 크기가 100 bp 이상 500 bp 이하인 것인, 프라이머 세트.

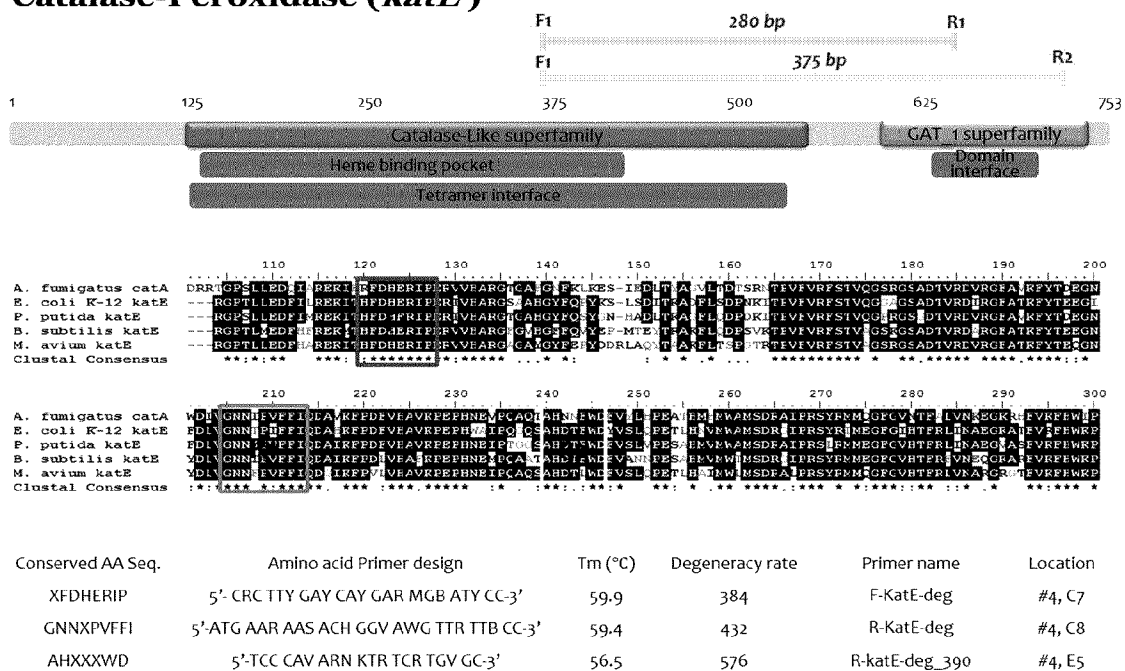
[청구항 10] 제8항의 프라이머 세트를 포함하는 메타지놈 분석 키트.

[도 1]

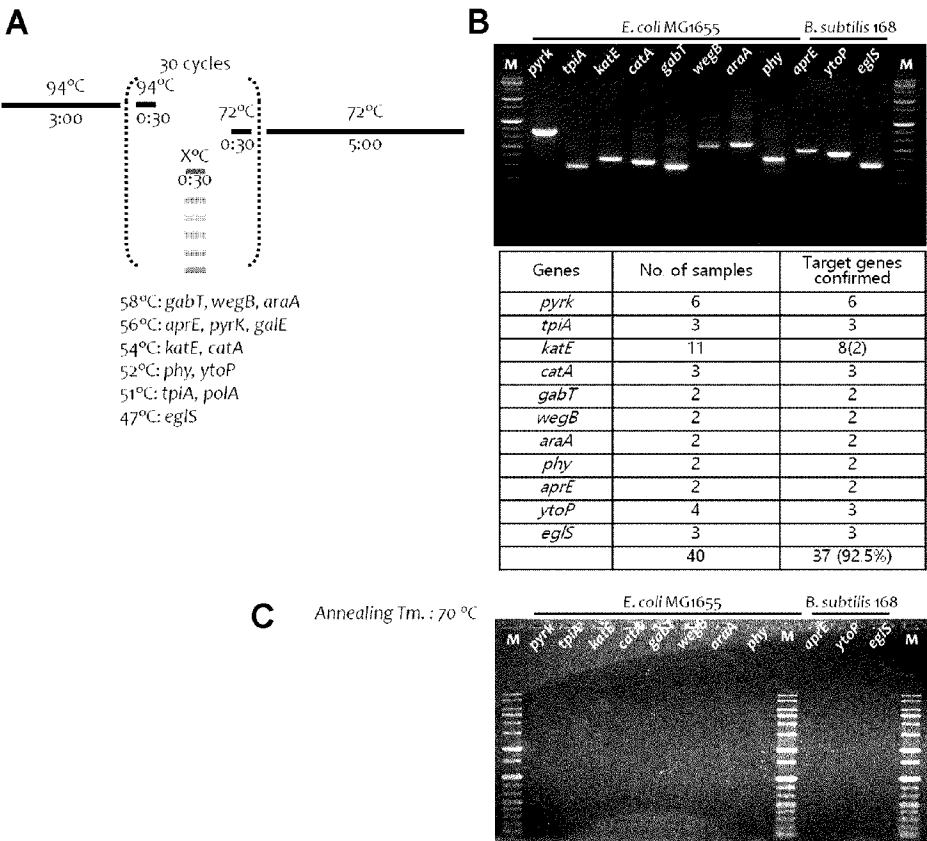


[도 2]

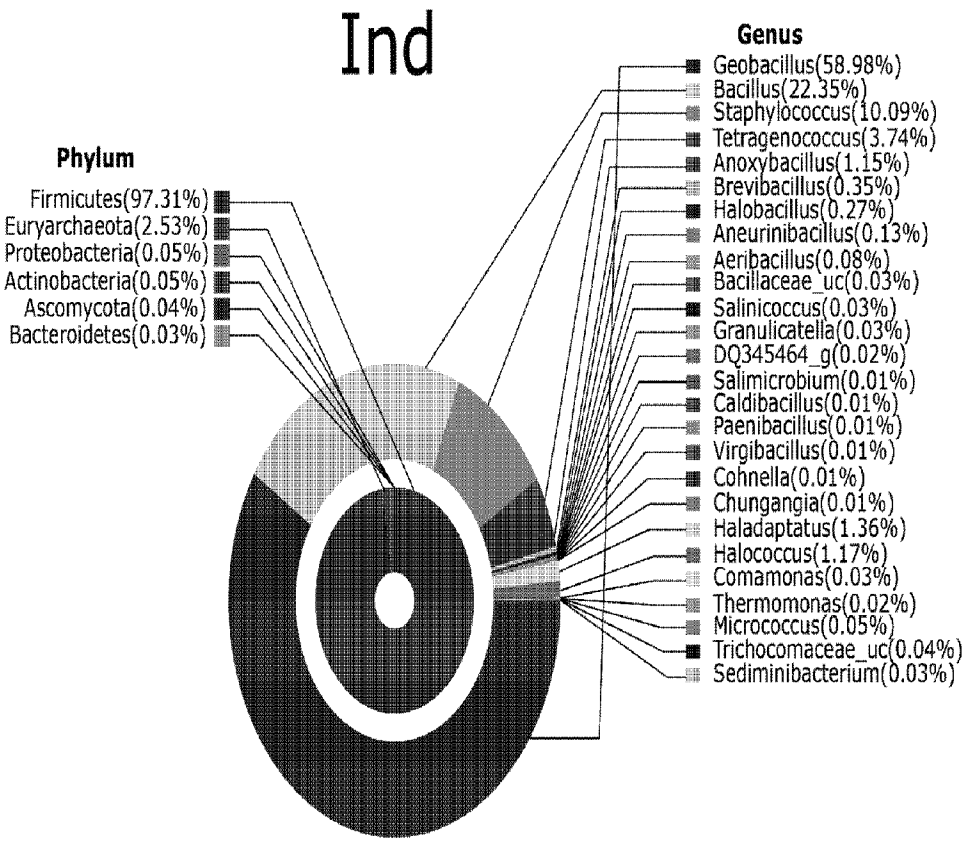
Catalase-Peroxidase (*katE*)



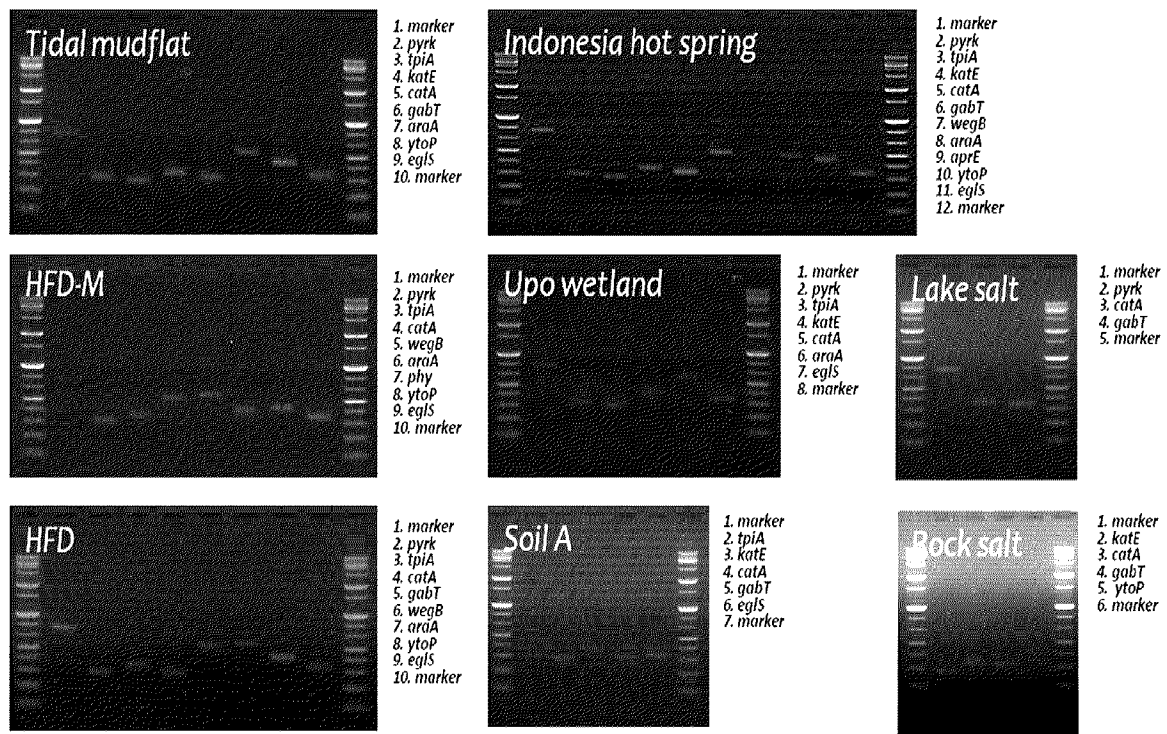
[도3]



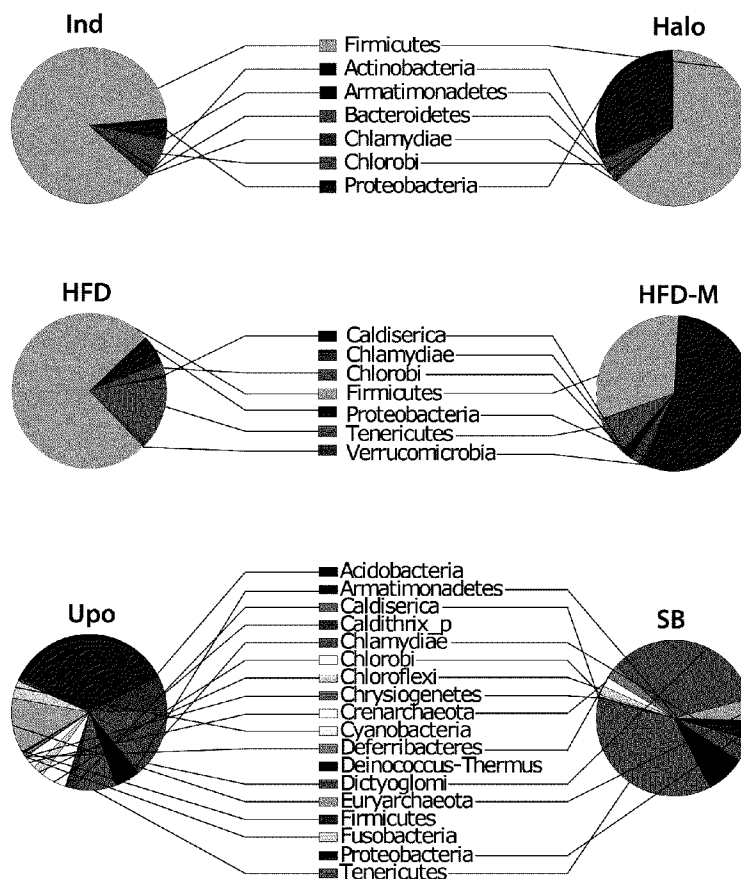
[도4]



[도5]



[도6]

pyrK

INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2016/006792

A. CLASSIFICATION OF SUBJECT MATTER

C12Q 1/68(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12Q 1/68; C12Q 1/34; C12N 15/63; C12N 15/09; C07K 14/00; C07K 17/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean Utility models and applications for Utility models: IPC as above

Japanese Utility models and applications for Utility models: IPC as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS (KIPO internal) & Keywords: superfamily-specific degenerate primer, metagenome, amplification

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ISERTE et al., "Family-specific Degenerate Primer Design: a Tool to Design Consensus Degenerated Oligonucleotides", Biotechnology Research International, vol. 2013, document number 383646, pages 1-9 (21 February 2013) See abstract, figure 1, pages 1-2.	1-10
A	JARIC et al., "Better Primer Design for Metagenomics Applications by Increasing Taxonomic Distinguishability", BMC Proceedings, vol. 7 (supplement 7), document number S4, pages 1-10 (20 December 2013) See abstract; conclusion (pages 6-9)	1-10
A	US 7872100 B2 (VERSECK et al.) 18 January 2011 See abstract; column 3, lines 30-48; column 5, line 51-column 6, line 55.	1-10
A	KR 10-2011-0073049 A (KOREA RESEARCH INSTITUTE OF BIOSCIENCE AND BIOTECHNOLOGY) 29 June 2011 See abstract; claim 1.	1-10
A	JP 2014-103948 A (HIROSHIMA UNIVERSITY) 09 June 2014 See abstract; claims 1-5.	1-10



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

18 OCTOBER 2016 (18.10.2016)

Date of mailing of the international search report

18 OCTOBER 2016 (18.10.2016)

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
Government Complex-Daejeon, 189 Seonsa-ro, Daejeon 302-701,
Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2016/006792

Patent document cited in search report	Publication date	Patent family member	Publication date
US 7872100 B2	18/01/2011	CA 2557476 A1	29/09/2005
		CA 2557476 C	23/02/2016
		DE 102004013842 A1	13/10/2005
		DK 1730300 T3	30/09/2013
		DK 2369009 T3	13/01/2014
		DK 2369010 T3	26/05/2014
		EP 1730300 A2	13/12/2006
		EP 1730300 B1	17/07/2013
		EP 2369009 A2	28/09/2011
		EP 2369009 A3	14/03/2012
		EP 2369009 B1	11/12/2013
		EP 2369010 A2	28/09/2011
		EP 2369010 A3	18/07/2012
		EP 2369010 B1	07/05/2014
		SG 157384 A1	29/12/2009
		US 2008-0171389 A1	17/07/2008
		WO 2005-090595 A2	29/09/2005
		WO 2005-090595 A3	02/03/2006
		ZA 200607776 A	25/06/2008
KR 10-2011-0073049 A	29/06/2011	CN 102459650 A	16/05/2012
		CN 102459650 B	01/10/2014
		EP 2465946 A2	20/06/2012
		KR 10-1222056 B1	14/01/2013
		KR 10-2012-0099202 A	07/09/2012
		US 2012-0238470 A1	20/09/2012
		WO 2010-143871 A2	16/12/2010
		WO 2010-143871 A3	21/04/2011
JP 2014-103948 A	09/06/2014	NONE	

A. 발명이 속하는 기술분류(국제특허분류(IPC))

C12Q 1/68(2006.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

C12Q 1/68; C12Q 1/34; C12N 15/63; C12N 15/09; C07K 14/00; C07K 17/00

조사된 기술분야에 속하는 최소문헌 이외의 문헌

한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우))

eKOMPASS(특허청 내부 검색시스템) & 키워드: 슈퍼 패밀리 특이적 디제너레이트 프라이머, 메타지놈, 증폭

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
X	ISERTE 등, `Family-specific degenerate primer design: a tool to design consensus degenerated oligonucleotides`, Biotechnology Research International, 2013권, 문헌번호 383646, 페이지 1-9 (2013.02.21) 요약, 도면 1, 페이지 1-2 참조.	1-10
A	JARIC 등, `Better primer design for metagenomics applications by increasing taxonomic distinguishability`, BMC Proceedings, 7권 (부록 7), 문헌번호 S4, 페이지 1-10 (2013.12.20) 초록; 결과 (페이지 6-9) 참조.	1-10
A	US 7872100 B2 (VERSECK 등) 2011.01.18 요약; 컬럼 3, 라인 30-48; 컬럼 5, 라인 51 - 컬럼 6, 라인 55 참조.	1-10
A	KR 10-2011-0073049 A (한국생명공학연구원) 2011.06.29 요약; 청구항 1 참조.	1-10
A	JP 2014-103948 A (HIROSHIMA UNIVERSITY) 2014.06.09 요약; 청구항 1-5 참조.	1-10

☐ 추가 문헌이 C(계속)에 기재되어 있습니다.☒ 대응특허에 관한 별지를 참조하십시오.

* 인용된 문헌의 특별 카테고리:

“A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌

“T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌

“E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌

“X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다.

“L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌

“Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다.

“O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌

“&” 동일한 대응특허문헌에 속하는 문헌

“P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌

국제조사의 실제 완료일

2016년 10월 18일 (18.10.2016)

국제조사보고서 발송일

2016년 10월 18일 (18.10.2016)

ISA/KR의 명칭 및 우편주소

대한민국 특허청
(35208) 대전광역시 서구 청사로 189,
4동 (둔산동, 정부대전청사)

팩스 번호 +82-42-481-8578

심사관

김승범

전화번호 +82-42-481-3371



국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
US 7872100 B2	2011/01/18	CA 2557476 A1 CA 2557476 C DE 102004013842 A1 DK 1730300 T3 DK 2369009 T3 DK 2369010 T3 EP 1730300 A2 EP 1730300 B1 EP 2369009 A2 EP 2369009 A3 EP 2369009 B1 EP 2369010 A2 EP 2369010 A3 EP 2369010 B1 SG 157384 A1 US 2008-0171389 A1 WO 2005-090595 A2 WO 2005-090595 A3 ZA 200607776 A	2005/09/29 2016/02/23 2005/10/13 2013/09/30 2014/01/13 2014/05/26 2006/12/13 2013/07/17 2011/09/28 2012/03/14 2013/12/11 2011/09/28 2012/07/18 2014/05/07 2009/12/29 2008/07/17 2005/09/29 2006/03/02 2008/06/25
KR 10-2011-0073049 A	2011/06/29	CN 102459650 A CN 102459650 B EP 2465946 A2 KR 10-1222056 B1 KR 10-2012-0099202 A US 2012-0238470 A1 WO 2010-143871 A2 WO 2010-143871 A3	2012/05/16 2014/10/01 2012/06/20 2013/01/14 2012/09/07 2012/09/20 2010/12/16 2011/04/21
JP 2014-103948 A	2014/06/09	없음	