

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국



(43) 국제공개일
2011년 1월 6일 (06.01.2011)

PCT

(10) 국제공개번호
WO 2011/002237 A2

- (51) 국제특허분류:
C12N 15/10 (2006.01) C12Q 1/02 (2006.01)
C12Q 1/68 (2006.01)
- (21) 국제출원번호: PCT/KR2010/004273
- (22) 국제출원일: 2010년 7월 1일 (01.07.2010)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보:
10-2009-0059787 2009년 7월 1일 (01.07.2009) KR
- (71) 출원인 (US 을(를) 제외한 모든 지정국에 대하여): 한국생명공학연구원 (KOREA RESEARCH INSTITUTE OF BIOSCIENCE AND BIOTECHNOLOGY) [KR/KR]; 대전광역시 유성구 어은동 52, 305-333 Daejeon (KR).
- (72) 발명자: 겸
- (75) 발명자/출원인 (US 에 한하여): 송재준 (SONG, Jae Jun) [KR/KR]; 대전광역시 서구 둔산 2 동 수정타운아파트 18 동 403 호, 302-120 Daejeon (KR). 최중현

(CHOI, Jong Hyun) [KR/KR]; 대전광역시 유성구 지족동 가나파로스빌 2 차 605 호, 305-330 Daejeon (KR). 한윤전 (HAN, Yun Jon) [KR/KR]; 전라북도 정읍시 신정동 1404 번지, 580-185 Jeollabuk-do (KR). 김중수 (KIM, Joong Su) [KR/KR]; 대전광역시 서구 만년동 상아아파트 101-907, 302-739 Daejeon (KR).

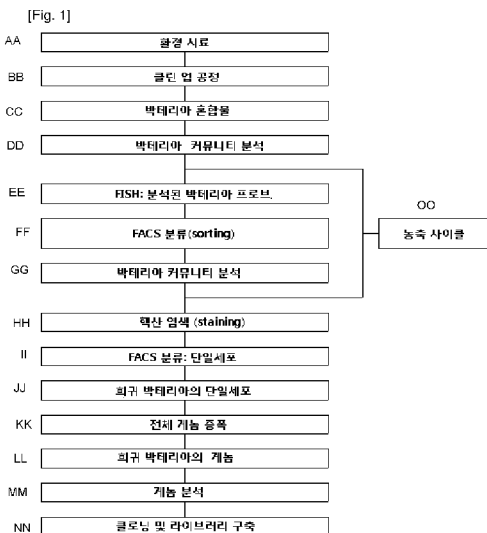
(74) 대리인: 이처영 (LEE, Cheo Young); 서울특별시 강남구 역삼동 648-23 여삼빌딩 11 층 5T 국제특허법률사무소, 135-080 Seoul (KR).

(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, 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, IS, JP, KE, KG, KM, KN, KP, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, 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.

[다음 쪽 계속]

(54) Title: METHOD FOR PREPARING GENOMIC LIBRARY DERIVED FROM SINGLE CELL OF RARE MICROORGANISMS OR NOVEL MICROORGANISMS BASED ON DOMINANT SPECIES REMOVAL TECHNOLOGY

(54) 발명의 명칭 : 우점종 제거 기술에 기반한 희소 미생물 또는 신규 미생물의 단일 세포 유래 게놈 라이브러리의 제조방법



AA ... Environmental sample
BB ... Clean-up process
CC ... Bacterial mixture
DD ... Bacterial community analysis
EE ... FISH: analyzed bacterial probe
FF ... FACS sorting
GG ... Bacterial community analysis
HH ... Nucleic acid staining
II ... FACS sorting: single cell
JJ ... Single cells from rare microorganisms
KK ... Whole genome amplification
LL ... Genome of rare microorganisms
MM ... Genome analysis
NN ... Cloning and construction of library
OO ... Concentration cycle

(57) Abstract: The present invention relates to a method for preparing a genomic library derived from a single cell of rare microorganisms or novel microorganisms. More specifically, the present invention is characterized by: analyzing a community of a microorganism group from collected samples; performing FISH and first FACS process for removing a dominant microorganism group based on the analysis and performing second FACS process to sort the microorganisms into single cells; amplifying non-culturable rare microorganisms or novel microorganisms which rarely exist in an environment into a genome derived from a single cell from performing multiple displacement amplification; and constructing a gene library. The method of the present invention can obtain and provide a genome library of non-culturable microorganisms and novel microorganisms existing in a real environment, which are difficult to detect by a conventional method and can obtain large quantities of genes of rare and novel microorganisms. In addition, the method is useful since the method enables the identification of the origin of the obtained gene unlike the conventional method for preparing a metagenome library and a library based on diversity can be manufactured by removing a dominant microorganism existing in an environment.

(57) 요약서:

[다음 쪽 계속]

WO 2011/002237 A2



(84) **지정국** (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, MD, 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, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

공개:

- 국제조사보고서 없이 공개하며 보고서 접수 후 이를 별도 공개함 (규칙 48.2(g))
- 명세서와 별도로 규칙 13 의 2 에 의하여 제출한 기탁된 생물학적 물질에 관한 표시와 함께 (규칙 13 의 2.4(d)(i) 및 48.2(a)(viii))
- 명세서의 서열목록 부분과 함께 (규칙 5.2(a))

.본 발명은 희소 미생물 또는 신규 미생물의 단일 세포 유래 게놈 라이브러리의 제조방법에 관한 것으로, 보다 상세하게는 수집한 시료의 미생물군에 대하여 커뮤니티 분석, 이에 근거한 우점종 미생물군 제거를 위한 FISH 와 1 차 FACS 공정의 수행 및 단일세포로 분리하기 위한 2 차 FACS 공정을 수행한 다음, 다중분리증폭법을 수행하여, 환경에 소수 존재하는 난배양성 희소 미생물 또는 신규 미생물을 단일 세포 유래의 게놈으로 증폭하여 유전자 라이브러리로 제조하는 것을 특징으로 한다. 본 발명에 따른 라이브러리의 제조방법은 기존의 방법으로 검출이 어려웠던 실제 환경에 존재하는 난배양성 희소미생물 및 신규 미생물의 게놈 라이브러리를 획득, 제공하며, 희소 미생물 및 신규 미생물에 대한 대량의 유전자를 확보할 수 있게 한다. 또한, 기존의 메타게놈 라이브러리 제조방법과 달리 확보된 유전자의 유래를 확인할 수 있게 하며 환경에 존재하는 우점의 미생물을 제거함으로써 다양성에 기반을 둔 라이브러리의 제조가 가능하게 하므로 유용하다.

명세서

발명의 명칭: 우점종 제거 기술에 기반한 희소 미생물 또는 신규 미생물의 단일 세포 유래 게놈 라이브러리의 제조방법

기술분야

- [1] 본 발명은 환경 내의 희소 미생물 또는 신규 미생물을 농축하는 것을 특징으로 하는, 희소 미생물 또는 신규 미생물의 단일 세포 유래 게놈 라이브러리의 제조방법에 관한 것이다.

[2]

배경기술

- [3] 미생물의 SSU rRNA(small subunit ribosomal RNA)의 유전자인 16S rDNA의 염기서열을 기준으로 한 계통분류학은, 환경 미생물에 대한 다양성을 분석하는 방법으로서 배양에 의존하지 않고 기존의 분류체계에 비해 미생물 다양성을 부각시키는데 결정적인 역할을 하여 왔고 (Woese *et al.*, *Proc. Natl. Acad. Sci. USA.*, 87(12):4576, 1990), 이를 통하여 특정 분류군의 미생물이 자연계에 존재하는지를 미생물의 배양 없이도 알아볼 수 있게 되었다.
- [4] 더 나아가, DGGE (Denaturing Gradient Gel Electrophoresis), T-RFLP (Terminal-Restriction Fragment Length Polymorphism), ARISA (Amplified Ribosomal Intergenic Spacer Analysis), 클로닝(Cloning)과 같은 분자 수준의 미생물 생태 분석 기법이 개발되면서, 이제까지 알 수 없었던 미생물의 다양성을 밝혀낼 수 있게 되었다. 그러나, 각각의 분석방법의 발전에도 불구하고 환경시료에 희소하게 존재하는 미생물에 대한 분석 및 추출을 위해서는 많은 양의 시료를 통한 광범위한 분석이 필요한 실정이다. 또한, 현재까지 인간이 순수 분리하여 배양 가능한 미생물의 수는 자연계 전체 미생물의 1% 정도에 불과하다고 알려져 있다 (Pace, *Science*, 276(5313): 734, 1997).
- [5] 이렇게 배양이 불가능한 미생물 또는 환경에 소수로 존재하는 미생물로부터 공업적으로 유용한 효소의 유전자나, 미생물이 생산하는 유용한 물질을 생합성하는 유전자를 얻어 산업적으로 이용하고자, 메타게놈 유전자 라이브러리를 통한 방법을 대표적으로 하여 여러 가지 방법이 연구되어 활용되고 있다. 그러나 각각의 방법에는 실제 유전자의 원천 미생물이 무엇인지 알 수 없다는 단점과 환경에 존재하는 희소 미생물의 경우 완전한 형태의 유전자를 얻을 확률이 낮다는 단점 등이 제기되고 있다.
- [6] 최근까지 미배양 미생물의 검출과 분리를 위한 여러 가지 연구가 이루어져 오고 있는데, 이러한 연구 중에는 미생물의 커뮤니티(community)를 분석한 후 미배양 미생물에 대한 FISH를 수행하고 형광 표지된 미생물을 단일세포로 분리한 후, 이 단일세포의 게놈에 대한 다중 분리 증폭법을 이용하여 유전자를 확보하는 연구(T. Kvist *et al. Appl. Microbiol. Biotechnol.*, 74:926, 2007)나, 환경시료를

충분히 희석하고 핵산을 형광 염색하여 FACS를 통한 단일세포 형태로 분리한 후 이 단일세포의 게놈에 대한 다중 분리 증폭법을 이용하여 유전자를 확보하는 연구 (R. Stepanauskas *et al.*, *Proc. Natl. Acad. Sci.*, 104(21):9052, 2007) 등이 있다.

- [7] 그러나 Kvist 등의 방법은 현재까지 밝혀진 유전자 염기서열을 기반으로 목적 미생물을 표지 선별하는 방법으로 실험을 통해 얻을 수 있는 유전체의 다양성이 목적 유전자에 편향되어 나타난다는 단점을 지니고 있다. 또한, Stepanauskas 등의 연구는 실험에 적용한 환경시료에 존재하는 미생물 군집을 생태학적인 측면에서 정확하게 분석할 수 있다는 장점이 있으나 우점을 차지하고 있는 미생물로 인해 희소하게 존재하는 미생물의 검출 및 분석이 어렵다는 단점을 지니고 있다.

- [8] 이에 본 발명자들은 자연 환경에 존재하는 희소미생물을 배양에 의존하지 않고 농축하여 기존에 알려진 방법으로 검출하지 못했던 희소 미생물에 대한 검출 한계를 높임과 더불어 높은 효율로 신규 유전자를 확보하고자 예의 노력한 결과, 미생물의 커뮤니티 분석, 2차래의 FACS 기법 및 다중분리증폭법(MDA: multiple displacement amplification)를 수행하여 난배양성 희소 미생물 및 신규 미생물에 대한 단일세포 유래의 게놈 라이브러리를 제작할 수 있음을 발견하고, 본 발명을 완성하였다.

[9]

[10] 발명의 요약

- [11] 본 발명의 목적은 기존의 방법으로 검출하지 못했던 희소 미생물 또는 신규 미생물의 게놈 라이브러리를 제조하는 방법을 제공하는 것이다.

- [12] 상기 목적을 달성하기 위하여, 본 발명은 다음 단계를 포함하는, 희소 미생물 또는 신규 미생물의 단일 세포 유래 게놈 라이브러리의 제조방법을 제공한다:

- [13] (a) 수집된 시료의 미생물군 커뮤니티 분석을 수행하는 단계;

- [14] (b) 상기 커뮤니티 분석에서 우점종 또는 공지의 미생물로 확인된 미생물을 표지하는 단계;

- [15] (c) 상기 표지된 미생물을 1차 FACS (Fluorescence Associated Cell Sorter) 기법을 이용하여 제거하는 단계;

- [16] (d) 상기 표지된 미생물이 제거된 미생물군 전체를 표지한 다음, 2차 FACS 기법을 이용하여 상기 미생물군 전체를 단일 세포로 분리하는 단계;

- [17] (e) 상기 분리된 단일세포의 게놈을 다중분리증폭법 (multiple displacement amplification, MDA)를 이용하여 증폭하는 단계; 및

- [18] (f) 상기 증폭된 단일세포의 게놈을 유전자 라이브러리로 제조하는 단계.

- [19] 본 발명은 또한, 다음 단계를 포함하는, 희소 미생물 또는 신규 미생물의 단일 세포 유래 게놈 라이브러리를 이용한 활성효소 또는 이를 코딩하는 유전자의 스크리닝 방법을 제공한다:

- [20] (a) 수집된 시료의 미생물군 커뮤니티 분석을 수행하는 단계;

- [21] (b) 상기 커뮤니티 분석에서 우점종 또는 공지의 미생물로 확인된 미생물을

표지하는 단계;

- [22] (c) 상기 표지된 미생물을 1차 FACS (Fluorescence Associated Cell Sorter) 기법을 이용하여 제거하는 단계;
- [23] (d) 상기 표지된 미생물이 제거된 미생물군 전체를 표지한 다음, 2차 FACS 기법을 이용하여 상기 미생물군 전체를 단일 세포로 분리하는 단계;
- [24] (e) 상기 분리된 단일세포의 게놈을 다중분리증폭법 (multiple displacement amplification, MDA)를 이용하여 증폭하는 단계;
- [25] (f) 상기 증폭된 단일세포의 게놈을 유전자 라이브러리로 제조하는 단계;
- [26] (g) 상기 제조된 게놈 라이브러리에 타겟 효소 기질을 처리하여 기질반응이 나타난 클론을 선별하는 단계; 및
- [27] (h) 상기 선별된 클론으로부터 타겟 효소 또는 이를 코딩하는 유전자를 수득하는 단계.
- [28] 본 발명은 또한, 다음 단계를 포함하는, 미생물 다양성이 증가된 메타게노믹(metagenomic) 라이브러리의 제조방법을 제공한다:
- [29] (a) 수집된 환경 시료의 미생물군 커뮤니티 분석을 수행하는 단계;
- [30] (b) 상기 커뮤니티 분석에서 우점종 또는 공지의 미생물로 확인된 미생물을 표지하는 단계;
- [31] (c) 상기 표지된 미생물을 FACS (Fluorescence Associated Cell Sorter) 기법을 이용하여 제거하는 단계; 및
- [32] (d) 상기 표지된 미생물이 제거된 미생물군으로부터 DNA를 분리하고 증폭하여 메타게노믹 라이브러리를 제조하는 단계.
- [33] 본 발명의 다른 특징 및 구현에는 다음의 상세한 설명 및 첨부한 특허청구범위로부터 더욱 명백해 질 것이다.

[34]

도면의 간단한 설명

- [35] 도 1은 본 발명에 따른 라이브러리 제조방법을 나타낸 개략도이다.
- [36] 도 2는 본 발명에 따른 라이브러리 제조방법 및 이를 이용한 효소와 유전자의 스크리닝 방법을 도식화한 그림이다.
- [37] 도 3은 본 발명의 일 실시예의 인공미생물 군집을 이용한 검증실험을 도식화한 그림이다.
- [38] 도 4는 FACS를 이용한 희소 미생물 농축 결과를 나타낸다.
- [39] 도 5는 Eco3 probe를 이용 농축된 시료를 FACS를 이용 단일세포 단위로 분류하고 분류된 시료에 대한 게놈 증폭과 16S rDNA, 셀룰레아제 유전자의 증폭을 확인한 결과를 나타낸다.
- [40] 도 6은 단일세포로 분리된 셀룰레아제 1,2를 포함한 게놈 증폭 시료의 16S rDNA PCR 산물을 클로닝하고, 시퀀싱 하여 *J. naejangsensis* 균주의 16S rDNA 시퀀스와 비교한 결과이다 (*J.naejangsanensis* 16S, SEQ ID No: 13; Single Cell MDA

16S #1, SEQ ID No.: 14; Single Cell MDA 16S #2, SEQ ID No.: 15; Single Cell MDA 16S #3, SEQ ID No.: 16).

- [41] 도 7는 다중분리 증폭법 통해 증폭된 게놈에 S1 뉴클레아제를 처리하여 단일 가닥의 DNA를 제거하고, 이중가닥만의 DNA에 대한 크기를 확인한 결과이다.
- [42] 도 8은 다중분리 증폭법 통해 증폭된 게놈을 이용하여 fosmid 라이브러리를 제작한 후 셀룰레아제 활성을 지니는 클론만을 선별한 결과이다.
- [43] 도 9는 셀룰레아제 활성을 지니는 fosmid 클론의 삽입 단편의 크기를 *Bam*HI 제한효소 처리를 통해 확인한 실험 결과이다.
- [44] 도 10는 셀룰레아제 활성을 지니는 fosmid 클론의 fosmid 플라스미드를 주형으로 사용하여 각각의 셀룰레아제 프라이머를 사용하여 PCR을 실시 PCR산물을 확인한 결과이다.
- [45] 도 11은 단일세포로 분리된 셀룰레아제 1 유전자를 포함한 게놈 증폭 시료와 셀룰레아제 활성을 지니는 fosmid 클론에 대한 셀룰레아제 1 유전자 시퀀싱 결과를 *J. naejangsanesis* 균주의 셀룰레아제 1 시퀀스와 비교한 결과로서, Cell #1, #2는 *J. naejangsanesis* 균주의 genomic DNA를 주형으로 사용하였고, MDA #1, #2는 단일 세포 유래 MDA 산물을 주형으로 사용하였으며, Fosmid #1, #2는 셀룰레아제 활성을 나타내는 Fosmid 클론인 MDC #11 클론을 주형으로 사용하여 각각 PCR 증폭하였고 증폭된 산물을 클로닝 하여 시퀀싱한 결과를 나타낸다(Cell #1, SEQ ID No: 17; Cell #2, SEQ ID No.: 18; MDA #1, SEQ ID No.: 19; MDA #2, SEQ ID No.: 20; Fosmid #1, SEQ ID No.: 21; Fosmid #2, SEQ ID No.: 22).
- [46] 도 12은 단일세포로 분리된 셀룰레아제 2 유전자를 포함한 게놈 증폭 시료와 셀룰레아제 활성을 지니는 fosmid 클론에 대한 셀룰레아제 2 유전자 시퀀싱 결과를 *J. naejangsanesis* 균주의 셀룰레아제 2 시퀀스와 비교한 결과로서, Cell #1, #2는 *J. naejangsanesis* 균주의 genomic DNA를 주형으로 사용하였고, MDA #1, #2는 단일 세포 유래 MDA 산물을 주형으로 사용하였으며, Fosmid #1, #2는 셀룰레아제 활성을 나타내는 Fosmid 클론인 MDC #11 클론을 주형으로 사용하여 각각 PCR 증폭하였고 증폭된 산물을 클로닝 하여 시퀀싱한 결과를 나타낸다(Cell #1, SEQ ID No: 23; Cell #2, SEQ ID No.: 24; MDA #1, SEQ ID No.: 25; MDA #2, SEQ ID No.: 26; Fosmid #1, SEQ ID No.: 27; Fosmid #2, SEQ ID No.: 28).

[47]

[48] 발명의 상세한 설명 및 구체적인 구현예

[49] 다른 식으로 정의되지 않는 한, 본 명세서에서 사용된 모든 기술적 및 과학적 용어들은 본 발명이 속하는 기술분야에서 숙련된 전문가에 의해서 통상적으로 이해되는 것과 동일한 의미를 갖는다. 일반적으로, 본 명세서에서 사용된 명명법 및 이하에 기술하는 실험 방법은 본 기술분야에서 잘 알려져 있고 통상적으로 사용되는 것이다.

- [50] 본 발명의 상세한 설명 등에서 사용되는 주요 용어의 정의는 다음과 같다.
- [51] 본원에서, "미생물군"이란 시료의 미생물 집합체, 즉 미생물 군집을 의미하며, 이의 "커뮤니티 분석"이란 시료의 미생물 군집의 유전적 다양성을 분석하는 방법으로, ARISA (Amplified riosomal intergenic spacer analysis), T-RFLP (Terminal restriction fragment length polymorphism), DGGE (Denaturing gradient gel electrophoresis), TGGE (Temperature Gradient Gel Electrophoresis), ARDRA (Amplified Ribosomal DNA Restriction Analysis) 및 DNA microarray 등의 방법에 의하여 이루어질 수 있다.
- [52] 여기서 "ARISA"란, 배양기반 접근이나, ss rRNA gene 클론 라이브러리 구성하는 실험 없이, 미생물 다양성 및 커뮤니티 조성의 대략값을 제공하는 미생물 분석방법으로, rRNA 오페론에서 small(16S) 및 large(23S) 서브유닛 rRNA gene 사이의 유전자간 부위(intergenic region)를 전체 박테리아 커뮤니티 DNA에 대하여 PCR 증폭시키는 것을 포함한다.
- [53] "T-RFLP"란, 증폭된 유전자의 라벨된 말단에 가장 가까운 제한효소의 사이트의 위치에 기반한 미생물 커뮤니티의 프로파일링을 위한 분자생물학적 기법을 말한다. 이는 하나 이상의 제한효소를 사용하여 단일 유전자의 PCR 증폭 산물의 혼합물을 자르고, DNA 시퀀서를 사용하여 개개의 말단 단편 각각의 사이즈를 탐지하는 것에 기반한다.
- [54] "DGGE"란, PCR 기반 DNA 산물을 분리하기 위하여 사용되는 유전적 핑거프린팅 기법으로, PCR 산물이 폴리아크릴아마이드 겔을 통해 이동하면서 점점 증가하는 화학적 변성제(chemical denaturant)의 농도에 부딪혀, 약한 분리(melting) 도메인을 가지는 이중-가닥 PCR 산물일수록 먼저 분리(denaturation)되기 시작하고 그에 따라 이동속도가 크게 느려지고, 이에 다른 박테리아 유래의 다른 DNA 서열은 다른 변성 농도(denaturant concentration)에서 분리되는 원리를 이용한다.
- [55] "TGGE"란, 상기 DGGE와 유사한 유전적 핑거프린팅 기법으로 화학적 변성제 대신에 온도를 점점 증가시켜서 PCR 산물을 분리시키는 원리를 사용한다.
- [56] "ARDRA"란, 미생물의 16S rDNA를 증폭한 다음, 제한효소로 증폭된 PCR 산물을 자른 후, 아가로스 겔에서 제한효소 처리된 산물을 전기영동하여 밴드 패턴을 분석하는 분자생물학적 기법이다.
- [57] 한편, "FISH (Fluorescence *in situ* hybridization)"는 대표적으로 rRNA 올리고뉴클레오타이드 프로브(rRNA oligonucleotide probe)를 이용하여 자연 환경 및 폐수 처리장에 존재하는 미생물 군집구조 분석 및 정량화, 그리고 미생물 동정 및 분포를 이해하는데 배양에 따른 오차 없이 빠르고 믿을만한 정보를 제공하는 방법으로 알려져 있다. 특히 FISH 결과를 co focal laser scanning microscope (CLSM)를 사용하여 관찰함으로써 미생물의 특성을 공간적으로 분석할 수 있는 방법이다. 본 발명에서는 상기의 방법을 환경시료에 존재하는 우점 미생물을 표지하는 방법으로 사용하였다.

- [58] "FACS (Fluorescence Associated Cell Sorter)"는 유액상태의 입자나 세포가 일정 감지지역을 통과할 때 각각의 입자나 세포를 신속하게 측정하여, 한 세포가 갖는 여러 특징을 동시에 측정하고 경우에 따라 특징한 부분을 선택하여 분석하고 분류(sorting)할 수 있는 장비이다. 본 발명에서는 상기의 장비를 희소 미생물의 농축과 농축된 미생물을 단일세포로 분류하는 방법에 이용하였다.
- [59] "다중분리증폭법 (multiple displacement amplification, MDA)"은 매우 적은 양의 DNA를 시퀀싱 반응에서 기질로서 사용될 정도의 충분한 양으로 증폭하는데 사용되는 기법으로, 기존에 널리 알려진 원형 DNA 염기서열 증폭용으로 개발된 박테리오파지 $\phi 29$ 의 폴리머레이즈를 이용하여 단일 세포로부터 직접 전체 게놈을 증폭하는데 사용할 수 있다.
- [60] 본원에서, "메타게놈(metagenome)"이란, 자연계에 존재하는 모든 미생물의 유전체를 통칭하는 것으로 정의된다.
- [61] 본 발명은 일 관점에서, 희소 미생물 또는 신규 미생물의 단일 세포 유래 게놈 라이브러리의 제조방법에 관한 것으로, 다음 단계를 포함하는 것을 특징으로 한다:
- [62] (a) 수집된 시료의 미생물군 커뮤니티 분석을 수행하는 단계;
- [63] (b) 상기 커뮤니티 분석에서 우점종 또는 공지의 미생물로 확인된 미생물을 표지하는 단계;
- [64] (c) 상기 표지된 미생물을 1차 FACS (Fluorescence Associated Cell Sorter) 기법을 이용하여 제거하는 단계;
- [65] (d) 상기 표지된 미생물이 제거된 미생물군 전체를 표지한 다음, 2차 FACS 기법을 이용하여 상기 미생물군 전체를 단일 세포로 분리하는 단계;
- [66] (e) 상기 분리된 단일세포의 게놈을 다중분리증폭법 (multiple displacement amplification, MDA)를 이용하여 증폭하는 단계; 및
- [67] (f) 상기 증폭된 단일세포의 게놈을 유전자 라이브러리로 제조하는 단계.
- [68] 도 1 및 도 2는 본 발명에 따른 희소 미생물 또는 신규 미생물의 단일 세포 유래 게놈 라이브러리의 제조방법의 개략도를 나타낸다.
- [69] 본 발명에 있어서, 상기 시료는 물, 토양, 공기, 식품, 폐기물, 동식물 장내 및 동식물 조직 중 어느 하나 이상에서 수집된 것임을 특징으로 할 수 있으나, 이에 한정되는 것은 아니다. 이때, 상기 물은 강수(江水), 해수(海水), 호수(湖水) 및 우수(雨水) 등을 포함하고, 폐기물은 하수, 폐수 등을 포함하며, 상기 동식물은 인체를 포함한다.
- [70] 채집된 시료는 바람직하게는 비미생물적 요소(non microbial particles)를 제거하고, 미생물 커뮤니티 분석 시 명확한 결과를 얻기 위하여 클린 업(clean up) 공정을 추가로 수행하여, 박테리아 혼합물(bacterial cocktail)의 형태로 제공될 수 있다.
- [71] 본 발명에서, 수집된 미생물 커뮤니티의 분석은 ARISA (automated riosomal intergenic spacer analysis), T-RFLP (Terminal restriction fragment length

polymorphism), DGGE (Denaturing gradient gel electrophoresis), TGGE (Temperature Gradient Gel Electrophoresis), ARDRA (Amplified Ribosomal DNA Restriction Analysis) 및 DNA microarray 등의 방법에 의하여 이루어질 수 있으며, 미생물의 커뮤니티를 분석하는 방법이면 이에 제한되지 않는다. 다만, 상기 방법들의 경우 PCR 반응의 제한성으로 인하여 정확한 커뮤니티의 분석이 가능하지 않은 경우가 많이 있으며 우점종으로 인지된 미생물에 대한 제한적인 정보만을 얻을 수 있다는 단점이 있다. 이에, 초기 환경 시료의 미생물 커뮤니티를 분석하는 방법으로 16S rDNA를 PCR 반응을 통해 증폭한 후 증폭된 16S rDNA에 대한 clone library를 제작하여 일정 수 이상의 16S rDNA clone을 시퀀싱하고 이를 분석하여 시료의 우점종인 미생물을 밝혀내는 방법을 사용하는 것이 바람직하다. 커뮤니티의 분석은 상기 방법을 둘 이상 조합하여 수행할 수 있으며, 바람직하게는 상기 16S rDNA를 타겟으로 한 방법에 의하여 1차 분석한 후, ARISA, T-RFLP, DGGE, TGGE, ARDRA 및 DNA microarray 중 어느 하나 이상의 방법을 수행함으로써 이루어질 수 있다.

- [72] 미생물 커뮤니티를 분석한 다음, 우점종으로 확인된 미생물, 또는 알려진 데이터와 대비 시 공지의 미생물로 확인된 미생물을 타겟으로 하는 프로브를 제작하여 환경시료에 존재하는 우점 미생물 또는 공지 미생물을 표지하였다. 이때 바람직하게는 형광표지 프로브를 제작하여 FISH(Fluorescence *In Situ* Hybridization)를 수행하여 표지할 수 있으며, 상기 형광표지 프로브는 16S rDNA를 타겟으로 하는 것이 바람직하다. 표지되는 형광은 바람직하게는, 사용하는 FACS 기기에서 검출이 용이한 형광시약을 사용한다. 한편, FISH의 수행은 FISH 이후 FACS에 적용하기 위해 수용액 상에서 실시되는 것이 바람직하다. 또한 FISH 수행을 위한 미생물의 고정의 경우 추후 핵산의 다중분리 증폭법을 이용한 증폭 효율을 높이기 위해 100% 에탄올을 이용한 고정방법이 바람직하다.
- [73] 이후, 표지된 미생물 및 비표지된 미생물을 포함하는 미생물 군에 대하여 1차 FACS (Fluorescence Associated Cell Sorter) 기법을 수행하여, 표지된 미생물을 제거하고, 비표지된 미생물만 수집할 수 있다. 즉, FACS 기법을 사용하여 희소미생물 또는 신규 미생물을 농축할 수 있다. 이때, 도 2에 나타난 바와 같이, 희소 미생물 또는 신규 미생물도 별도의 다른 형광표지 프로브로 표지한 다음, 이들을 FACS 기법을 통하여 분리하는 방법으로 희소 미생물 또는 신규 미생물을 농축할 수 있다.
- [74] 이때, 바람직하게는 상기 농축 단계는 표지된 미생물이 제거된 잔류한 미생물 군의 커뮤니티를 재분석하여 우점종 또는 공지의 미생물이 확인되는 경우 표지하여 제거하는 단계를 추가로 반복하여 수행하여 충분한 우점종 또는 공지 미생물의 제거가 이루어지도록 할 수 있다.
- [75] 이후, 상기와 같이 표지된 미생물이 제거된 잔류 미생물군 전체를 표지한 다음, 2차 FACS 기법을 이용하여 상기 미생물군 전체를 단일 세포로 분리하는

단계로서, 이때, 잔류 미생물군 전체의 표지는 DNA에 대해 특이성을 나타내면서 DNA 구조에 미치는 영향이 적고, 사용하는 FACS 기기에서 검출이 용이한 형광시약을 사용하는 것이 바람직하다.

- [76] 단일세포로 분리한 후, 각 단일세포의 게놈을 다중분리증폭법 (multiple displacement amplification, MDA)를 이용하여 증폭한다. 이때, 상기 다중분리증폭법은 Phi29 DNA 폴리머라제를 이용하며, 비 특이적 증폭을 방지하기 위해 청결한 환경을 유지함은 물론 사용 시약에 대해서도 핵산의 오염을 방지하는 전처리 과정을 거치는 것이 바람직하다.
- [77] 상기 공정을 거쳐 증폭된 게놈은 증폭된 게놈의 16S rRNA를 공지 미생물의 16S rRNA와 대비하여 신규한 미생물인지를 판정한 다음, 신규한 미생물로 판정된 게놈만을 라이브러리로 제조하는 것을 특징으로 할 수 있다. 또한, 증폭된 게놈을 분석하여 단일세포로부터 유래된 것인지 확인하는 단계를 추가로 수행하여, 단일세포로부터 유래된 것으로 판정된 게놈만을 라이브러리로 제조하는 것을 특징으로 할 수 있다. 이때, 단일세포 유래인지의 판정은 증폭된 게놈의 16S rDNA의 시퀀싱시 단일 피크의 크로마토그램을 형성하는지 조사함으로써 이루어질 수 있다. 이외에 추가적으로 DNA 핑거프린팅 분석(DNA fingerprinting analysis), DGGE, T-RFLP 등의 방법으로도 이루어질 수 있다.
- [78] 라이브러리로 제작은 shot gun 라이브러리의 경우 chimeric의 형성 비율이 높기 때문에 fosmid나 BAC과 같은 큰 크기의 DNA 단편을 이용하는 라이브러리의 제작이 바람직하다. 본 연구에서는 다중분리증폭법을 통해 증폭된 단일세포의 게놈을 fosmid 라이브러리로 제조하여 라이브러리의 활용성을 확인하였다. 또한 다중분리증폭법을 통해 증폭된 게놈 DNA의 경우 과량의 가지를 형성하여 라이브러리 제작에 여러 가지로 제약을 주기 때문에 단일 가닥의 가지를 제거 해주기 위해 S1 뉴클레아제를 사용하여 형성된 가지 형태를 제거해 주는 과정이 필요하다.
- [79] 본 발명의 일 실시예에서는 본 발명의 방법의 유용성의 검증을 위하여, 인공 미생물 균집을 제작하고 본 발명의 방법을 적용하여 실험을 수행한 결과, 인공 미생물 균집 중 희소종으로 도입한 미생물인 *J.naejangsanensis* 균주(KCTC 11523BP)의 단일 세포 유래 게놈 라이브러리가 제작될 수 있음을 입증하였다.
- [80] 아울러, 본 발명의 다른 실시예에서는 상기 실시예에서 제작된 *J.naejangsanensis* 균주의 단일세포 유래 게놈 라이브러리 클론으로부터 각각 셀룰라아제 1,2를 확인하였으며, 이를 통해 본 발명에 따라 제작된 희소 미생물 또는 신규 미생물의 단일 세포 유래 게놈 라이브러리가 신규 효소 및 그를 코딩하는 유전자의 발견에 유용하게 이용될 수 있음을 입증하였다.
- [81] 이에, 본 발명은 다른 관점에서, 상기 게놈 라이브러리를 이용하는 것을 특징으로 하는 희소 미생물 또는 신규 미생물의 활성효소 및 이를 코딩하는 유전자의 스크리닝 방법을 제공한다.

- [82] 상세하게는, 상기 스크리닝 방법은, 상기 방법에 따라 희소 미생물 또는 신규 미생물의 단일 세포 유래 게놈 라이브러리를 제작한 후, 상기 제조된 게놈 라이브러리에 타겟 효소기질을 처리하여 기질반응이 나타난 클론을 선별하고, 상기 선별된 클론으로부터 타겟 효소 또는 이를 코딩하는 유전자를 수득함으로써 수행될 수 있다. 이때, 선별된 클론으로부터 타겟 효소 및 이를 코딩하는 유전자를 수득하는 방법은 통상의 알려진 유전공학적 방법을 이용하여 수행될 수 있다.
- [83] 또한 단일 세포유래의 라이브러리뿐 만아니라 우점 공지 미생물을 제거한 희소 난배양성 미생물 농축액 자체에 대한 유전자 증폭 및 라이브러리 제작이 가능하며 이를 통한 다양성이 증가된 형태의 메타게노믹(metagenomic) 라이브러리의 제작이 가능하다. 따라서, 본 발명은 또 다른 관점에서, (a) 수집된 환경 시료의 미생물군 커뮤니티 분석을 수행하는 단계; (b) 상기 커뮤니티 분석에서 우점종 또는 공지의 미생물로 확인된 미생물을 표지하는 단계; (c) 상기 표지된 미생물을 FACS (Fluorescence Associated Cell Sorter) 기법을 이용하여 제거하는 단계; 및 (d) 상기 표지된 미생물이 제거된 미생물군으로부터 DNA를 분리하고 증폭하여 메타게노믹 라이브러리를 제조하는 단계를 포함하는 것을 특징으로 하는 미생물 다양성이 증가된 메타게노믹(metagenomic) 라이브러리의 제조방법에 관한 것이다. 우점종 또는 공지의 미생물이 제거된 미생물군으로부터 DNA를 분리, 증폭하여 메타게노믹 라이브러리를 제조하는 방법은 통상의 알려진 메타게노믹 라이브러리 제조방법에 따라 수행될 수 있다.
- [84]
- [85] 실시예
- [86] 이하, 실시예를 통하여 본 발명을 더욱 상세히 설명하고자 한다. 이들 실시예는 오로지 본 발명을 예시하기 위한 것으로서, 본 발명의 범위가 이들 실시예에 의해 제한되는 것으로 해석되지는 않는 것은 당업계에서 통상의 지식을 가진 자에게 있어서 자명할 것이다.
- [87] 특히, 하기 실시예에서는 본 발명의 효과의 검증을 위하여 희소 미생물을 포함한 인공 미생물 군집을 제작하여 본 발명에 따른 방법에 의하여 희소종이 검출될 수 있음을 확인하였으나, 자연 환경 내의 샘플을 수집하여 본 발명의 방법을 적용하여서도 동일한 결과를 얻을 수 있다는 것은 당업계에서 통상의 지식을 가진 자에게 자명한 사항이라 할 것이다.
- [88]
- [89] 실시예 1: 인공 미생물군집 시료의 제작 및 미생물 커뮤니티의 분석
- [90]
- [91] 5종의 공지 미생물균주(*Escherichia coli* EPI300-T1, *Bacillus cereus* KCTC 3624, *Pseudomonas aeruginosa* KCTC 1637, *Proteus mirabilis* KCTC 2510, *Jeongupia naejangsanesis*)를 순수 배양한 후, 각각의 미생물을 다음과 같은 비율로 혼합하여 희소 미생물을 포함한 인공 미생물 군집 시료를 제조하였다(도 3). J.

naejangsanensis 균주의 경우 본 실험실에서 셀룰레아제 활성균주로서 분리, 동정한 균주로서 2009년 7월 1일자로 한국생명공학연구원 생물자원센터(KCTC)에 기탁한 균주이고 (*Jeongeupia naejangsanensis* BIO-TAS 4-1 KCTC 11523BP), *E. coli* EPI300-T1 균주의 경우 Epicentre사(USA)로부터 구입하였다. 상기의 균주를 제외한 미생물들(*B. cerius* KCTC 3624, *P. aeruginosa* KCTC 1637, *P. mirabilis* KCTC 2510)은 한국생명공학연구원 생물자원센터(KCTC)로부터 분양 받아 실험에 이용하였다.

- [92] 먼저 시료의 제조를 위하여 각각의 미생물의 LB 배지에 16시간 배양한 후 흡광광도계를 이용 600nm에서 흡광도를 측정하였다. 이 후 8,000g로 10분간 4°C에서 다시 원심분리여, 상등액을 버리고 얻어진 세포를 앞서 측정한 흡광도를 기준으로 하여 1X PBS 완충액으로 흡광도 1.0이 되도록 재현탁 하였다. 그 다음, 상기의 방법을 통해 얻어진 미생물 현탁액을 각각 *E. coli* 9 ml, *B. cerius* 0.5 ml, *P. aeruginosa* 0.5 ml, *P. mirabilis* 0.1 ml, *J. naejangsanensis* 0.1 ml의 부피로 혼합하여 *P. mirabilis* 및 *J. naejangsanensis* 균주를 희소 미생물로 한 인공 미생물 균집을 제조하였다.
- [93] 본 실시예에서는 본 발명에 따른 방법의 효과를 검증하기 위하여, 희소미생물을 포함한 인공 미생물 균집을 제조하여 확인하였으나, 실제 환경에 적용할 때에는, 바람직하게는 채집된 시료로부터 비 미생물적 요소(non microbial particles)를 제거하고, 미생물 커뮤니티 분석 시 명확한 결과를 얻기 위하여 클린업(clean up) 공정을 추가로 수행하여, 박테리아 혼합물(bacterial cocktail)의 형태로 제공될 수 있다. 아울러, 환경 시료에 대한 미생물 커뮤니티 분석은 16S rDNA를 PCR 반응을 통해 증폭한 후 증폭된 16S rDNA에 대한 clone library를 제작하여 일정 수 이상의 16S rDNA clone을 시퀀싱하고 이를 분석하여 시료의 우점종인 미생물을 밝혀내는 방법을 사용하는 것이 바람직하다.
- [94]
- [95] 실시예 2: FISH 및 1차 FACS의 수행
- [96]
- [97] 2-1: FISH의 수행
- [98] 실시예 1의 인공 미생물 균집에서 우점종이 되도록 혼합한 *E. coli*를 표지하고 FISH를 수행하기 위하여, 다음과 같이 *E. coli*에 대한 16S rRNA를 타겟으로 하는 FITC(Fluorescein isothiocyanate)로 형광표지한 프로브를 2종 제조하였다.
- [99]
- [100] Eco2: 5'
FITC-GTAACGTCAATGAGCAAAGGTATTAACCTTTACTCCCTTCC-3'
(서열번호 1)
- [101] Eco3: 5' FITC-CCGAAGGCACATTCTCATCTCTGAAAACTTCCGTGGATG-3'
(서열번호 2)
- [102]

- [103] 그 다음 실시예 1-1에서 제작한 인공 미생물균집의 미생물 조성의 변화를 막고 16 rRNA의 분해를 막기 위해, 100% 에탄올을 동일 부피로 첨가한 후, 4 °C의 온도에서 16시간 동안 방치하여 고정화를 실시하였다. 이 후 1X PBS 완충액으로 2회 세척한 후 최종 세포의 농도가 $10^8 \sim 10^9$ cell/ml이 되도록 현탁하여 실험에 이용하였다. 상기 시료 1 ml을 8,000 g, 10분간 원심분리한 후, 20% 포름아마이드(formamide) 농도의 혼성화 완충용액(hybridization buffer: 표 1, 최종 프로브 농도 50 ng/ μ l) 1 ml로 재현탁하여 46 °C에서 1시간 30분간 혼성화를 실시하였다.
- [104] 혼성화 시료를 다시 8,000 g, 10분간 원심분리한 후, 세척완충용액(washing buffer: 표 2) 1 ml로 재현탁 하여 48 °C에서 1시간 동안 incubation 하였다. 그 다음 상기 반응 시료를 다시 8,000 g, 10분간 원심분리한 후, 1 X PBS 완충액 1 ml을 이용하여 재현탁 한 후 추후 실험을 위해 4 °C의 상태로 냉장 보관하였다.

[105]

[106] 표 1

포름아마이드 농도에 따른 혼성화 완충용액 조성 (final volume, 2ml)

% Formamide	Formamide (μ l)	Milli-Q water (μ l)	5M NaCl (μ l)	1M Tris-HCl pH 7.2(μ l)	10% SDS (μ l)
0	0	1598	360	40	2
5	100	1498	360	40	2
10	200	1398	360	40	2
15	300	1298	360	40	2
20	400	1198	360	40	2
25	500	1098	360	40	2
30	600	998	360	40	2
35	700	898	360	40	2
40	800	798	360	40	2
45	900	698	360	40	2
50	1000	598	360	40	2

[107]

[108] 표 2

혼성화 완충용액에서 포름아마이드 농도에 따른 세척완충용액 조성 (final volume, 50ml)

% Formamide	Milli-Q water (μl)	5M NaCl (μl)	1M Tris-HCl pH 7.2(μl)	10% SDS (μl)
0	39.95	9,000	1,000	50
5	42.65	6,300	1,000	50
10	44.45	4,500	1,000	50
15	45.77	3,180	1,000	50
20	46.8	2,150	1,000	50
25	47.46	1,490	1,000	50
30	47.93	1,020	1,000	50
35	48.25	700	1,000	50
40	48.49	460	1,000	50
45	48.65	300	1,000	50
50	48.77	180	1,000	50

[109]

[110] 2-2: 1차 FACS의 수행

[111] 상기 미생물군에 대하여 FACS Aria 장비 (Becton Dickinson Bioscience, USA)를 이용하여 형광표지된 우점 미생물을 분석하고 제거하였다. 이때, 상세하게는 488 nm 레이저와 530 nm 검출기를 이용하여 형광을 검출하였고, 70 μm 의 노즐과 70 psi의 속도로 시료를 주입하였다. 각각 사용된 sheath 용액은 DNA에 오염되지 않은 3차 증류수를 이용하여 제조하였으며, FACS 기기의 경우는 sheath 용액과 시료가 통과하는 관을 FACS clean 용액(Becton Dickinson Bioscience)과 DNA에 오염되지 않은 3차 증류수를 이용하여 충분히 세척한 후 실험에 사용하였다. 분석과 분류에 사용한 소프트웨어는 FACS Diva (Becton Dickinson Bioscience)를 사용하였으며, 분류 방법은 "single cell mode"방법을 사용하여 우점 미생물과 희소 미생물을 분류하는 작업을 실시하였다.

[112] 실시예 2-1에서 FISH를 수행한 시료를 이용하여 FACS를 통해 분석한 결과, 도 4에 나타난 바와 같이, Eco2 프로브 보다 Eco3 프로브에 의한 FISH 결과가 더 좋은 감도를 나타내었다.

[113]

[114] 실시예 3: 2차 FACS의 수행 - 농축된 희소미생물의 단일세포로의 분리

[115]

[116] 실시예 2에 따라 형광표지된 미생물이 제거된 잔류 미생물군 전체에 대해 핵산을 SYTO-16 (Invitrogen, USA) 염료를 이용하여 염색하였다. 잔류

미생물군에 SYTO-16 염료를 최종농도 5 μM 이 되도록 첨가하여 준 후, 실온에서 1시간 동안 염색하였다.

- [117] 이후 염색된 미생물군은 FACS 장비를 이용하여 단일 세포로 분리하였다. 이때, 상세하게는 488 nm 레이저와 530 nm 검출기를 이용하여 형광을 검출하였고, 70 μm 의 노즐과 70 psi의 속도로 시료를 주입하였다. 각각 사용된 sheath 용액은 DNA에 오염되지 않은 3차 증류수를 이용하여 제조 하였으며, FACS 기기의 경우는 sheath 용액과 시료가 통과하는 관을 FACS clean 용액(Becton Dickinson Bioscience)과 DNA에 오염되지 않은 3차 증류수를 이용하여 충분히 세척한 후 실험에 사용하였다. 분석과 분류에 사용한 소프트웨어는 FACS Diva (Becton Dickinson Bioscience)를 사용하였으며, 분류 방법은 "single cell mode"방법을 사용하여 최종 96 well PCR tube(Axygen, USA)에 단일세포의 형태로 분류하였다. 이때 96 well PCR tube에는 3 μl 의 1X PBS 완충액을 넣어 두어 분류된 미생물이 안착되기 쉽도록 하였다. 96 well PCR tube로의 분류는 세로로 1, 3, 5, 7번째 줄로 각각 8개의 well에 1, 3, 5, 7개의 세포가 들어가도록 분류하였다. 분류가 끝난 96 well PCR tube는 곧바로 -20 $^{\circ}\text{C}$ 로 냉동 보관하여 추후 실험에 사용하였다.

[118]

[119] 실시예 4: 분리된 단일 세포의 게놈의 증폭

[120]

[121] 4-1: MDA를 통한 단일 세포의 게놈의 증폭

- [122] 실시예 3에서 분리한 각 단일세포의 게놈을 다중분리증폭법 (multiple displacement amplification, MDA)를 이용하여 증폭하였다. 이때, 다중분리증폭법은 REPLI-g Mini 키트(QIAGEN, USA)를 사용하였으며, 포함된 매뉴얼의 "Amplification of Genomic DNA from Blood or Cells the REPLI-g Mini Kit" 방법에 따라 수행하였다. 상기의 실험결과 MDA를 통해 genomic DNA가 증폭된 것을 확인하기 위해 0.8 % 아가로스겔 전기영동과, 16S rDNA PCR 방법을 사용하였다. 이때, 16S rDNA PCR을 위한 프라이머로는 다음과 같이 제조하여 사용하였다.

[123]

[124] 27F: 5'-GAGTTTGATYMTGGCTCAG-3' (서열번호 3)

[125] 1492R: 5'-GGTTACCTTGTTACGACTT-3' (서열번호 4)

[126]

- [127] 그 결과, 도 5에 나타난 바와 같이, 32개의 시료 중 모든 시료에서 16S rDNA에 대한 PCR product가 생성됨을 확인하였다. 즉, 32개의 시료에서 모두 genomic DNA가 증폭되었음을 것을 확인하였다.

[128]

[129] 4-2: MDA 증폭 산물에 대한 검토

- [130] 실시예 4-1에 의해 증폭된 MDA 산물에 대해 인공 미생물군집 중 희소하게 존재하는 *J. naejangsensis*균주가 지니고 있는 것으로 이미 알고 있는 두 종류의

셀룰레아제 유전자를 검출하고자 셀룰레아제 유전자의 중간 위치를 증폭할 수 있는 프라이머를 다음과 같이 합성하고 PCR을 실시하였다.

[131]

[132] 셀룰레아제 1

[133] EC010 Cell 1 IF: 5'-TCGATACCGGCAACGGCGGC-3' (서열번호 5)

[134] EC010 Cell 1 IR: 5'-CTCAGCAGGTCGAGCACTGCCG-3' (서열번호 6)

[135]

[136] 셀룰레아제 2

[137] Cell 2 IF: 5'-CCGACCCGCAGGACGCAAGC-3' (서열번호 7)

[138] Cell 2 IR: 5'-GCCACCAGTTGCCGTCGTGCTC-3' (서열번호 8)

[139]

[140] 그 결과, 11개의 시료에서 셀룰레아제 유전자의 PCR 산물이 확인되었으며, 그중 6개의 시료에서는 두 종류의 셀룰레아제 유전자가 모두 증폭되었음을 확인하였다. 그리고 단일 세포로 분류하여 genomic DNA가 증폭된 시료에서는 하나의 시료에서 두 종류의 셀룰레아제 유전자가 PCR을 통해 증폭된 것을 확인하였다. 즉, 인공 미생물 군집에서 희소종이었던 *J. naejangsanesis* 군주가 본 발명의 방법에 따라 농축되어 MDA 증폭산물로 수득되어졌음을 확인할 수 있다(도 5).

[141] 또한 16S rDNA를 PCR을 통해 증폭하고, 증폭된 PCR 산물을 pGEM-T easy 벡터 (Promega, USA)에 클로닝 하여 클론 3개를 선발 시퀀싱 한 후 시퀀싱 결과를 분석한 결과 단일 세포 유래의 *J. naejangsanesis* 군주로 확인 되었고(도 6), 따라서 상기의 시료를 이후 Fosmid 라이브러리를 제작하는 실험에 이용하였다.

[142]

[143] 실시예 5: 희소종의 단일 세포 유래 게놈 라이브러리의 제조

[144]

[145] 5-1: S1 뉴클레아제의 처리

[146] MDA를 단일 세포의 게놈의 증폭반응 중 과도한 가지(hyperbranch)가 발생하므로, 이를 해소하기 위하여 S1 뉴클레아제를 사용하였다. 이를 통해 이중가닥의 DNA만을 얻을 수 있었는데, 이를 Fosmid 라이브러리에 적용하기 위해서는 일정크기(25 Kbp) 이상이 되어야하므로, 아가로스겔 전기영동을 통해 크기를 확인하는 실험을 실시하였다.

[147] 그 결과, 도 7에 나타난 바와 같이, S1 뉴클레이즈를 과량으로 처리한 경우에도 대부분의 증폭 산물은 48 Kbp 이상의 크기를 지니는 것을 확인하였고 이를 통해 Fosmid 라이브러리를 제조할 수 있음을 확인하였다.

[148]

[149] 5-2: 게놈 라이브러리의 제조

[150] 실시예 5-1을 통해 이중가닥을 지니는 단일세포 유래의 genomic DNA를

확보하였고, 이를 이용하여 Fosmid 라이브러리를 제작하였다. Fosmid 라이브러리의 제조는 "CopyControl Fosmid Library Production Kit"(Epicentre Biotechnologies, USA)를 이용하여, 제조사에서 제공한 매뉴얼에 기재되어 있는 방법으로 실시하였고, 최종적으로 50,000여 클론을 확보하였다.

[151]

[152] 실시예 6: 활성효소의 스크리닝 및 신규 셀룰레아제 유전자의 분리

[153]

[154] 실시예 5-2에서 제조한 게놈 라이브러리 클론으로부터 활성 셀룰레아제를 스크리닝하기 위하여, 다음의 실험을 수행하였다. 셀룰레아제 활성 기반 효소탐색은, LB_Agar 고체배지를 CMC(Carboxyl Methyl Cellulose) 1 %, Tryphan Blue 0.005 %, 항생제로서 chloramphenicol 12.5 $\mu\text{g}/\text{mL}$ 로 최종 농도가 되도록 첨가하여 제조한 다음, Fosmid 라이브러리 클론을 도말하여 37 °C에서 배양하면서 투명환이 나타나는 클론을 선별함으로써 수행되었다.

[155] 그 결과, 도 8에 나타난 바와 같이, 15 개의 셀룰레아제 활성을 지니는 Fosmid 클론을 확보할 수 있었다.

[156] 추가적으로, 활성 클론의 Fosmid 벡터를 분리하여, 제한효소 *Bam*HI을 처리하여 삽입된 DNA 단편의 크기를 확인한 결과 평균 32 Kbp의 크기로 27-40 Kbp 사이의 DNA가 삽입되어 있는 것을 확인하였다(도 9). 상기의 셀룰레아제 활성 클론에 대한 셀룰레아제 유전자의 중간 위치를 증폭할 수 있는 프라이머를 합성하여 각각 PCR을 통해 증폭한 결과 모든 클론에서 각각의 셀룰레아제 유전자가 증폭되는 것을 확인할 수 있었다(도 10).

[157] 그리고 상기의 셀룰레아제 1과 2의 전체 유전자를 각각 *J. naejangsensis* genomic DNA, 단일세포 유래 MDA 산물, fosmid library MDC #11 clone을 template로 하여 PCR 증폭한 후 pGEM-T (promega, USA) 벡터에 클로닝하여 각각 두 개의 클론을 시퀀싱 한 결과 모든 셀룰레아제 유전자의 시퀀스가 동일하게 나타남을 확인 할 수 있었다 (도 11, 12). 상기 PCR에 사용한 프라이머의 시퀀스는 다음과 같다.

[158]

[159] 셀룰레이즈 1

[160] Cell 1 FF: 5'-TCGATACCGGCAACGGCGGC-3' (서열번호 9)

[161] Cell 1 FR: 5'-CTCAGCAGGTCGAGCACTGCCG-3' (서열번호 10)

[162] 셀룰레이즈 2

[163] Cell 2 FF: 5'-AACATATGTTGAGAGGAGCGCATCCGGTGT-3' (서열번호 11)

[164] Cell 2 FR: 5'-TTGGATCCTTATGGCTGTTGCCACAACCCC-3' (서열번호 12)

[165]

산업상 이용가능성

[166] 이상 설명한 바와 같이, 본 발명에 따른 라이브러리의 제조방법은 기존의

방법으로 검출이 어려웠던 실제 환경에 존재하는 난배양성 희소미생물 및 신규 미생물의 게놈 라이브러리를 제공하며, 희소 미생물 및 신규 미생물에 대한 대량의 유전자를 확보할 수 있게 한다. 또한, 기존의 메타게놈 라이브러리 제조방법과 달리 확보된 유전자의 유래를 확인할 수 있게 하며 환경에 존재하는 우점의 미생물을 제거함으로써 다양성에 기반을 둔 라이브러리의 제조가 가능하게 하므로 유용하다.

- [167] 이상으로 본 발명 내용의 특정한 부분을 상세히 기술하였는 바, 당업계의 통상의 지식을 가진 자에게 있어서, 이러한 구체적 기술은 단지 바람직한 실시양태일 뿐이며, 이에 의해 본 발명의 범위가 제한되는 것이 아닌 점은 명백할 것이다. 따라서 본 발명의 실질적인 범위는 첨부된 청구항들과 그것들의 등가물에 의하여 정의된다고 할 것이다.

[168]

서열목록 Free Text

- [169] 전자파일 첨부하였음.

[170]

[170]

출원인 또는 대리인 파일 참조번호 PP-B0940	국제출원번호 PCT/KR2010/004273
-----------------------------------	-----------------------------

기탁된 미생물 또는 기타 생물학적 물질의 표시사항

(특허법령조약 규칙 제13조의2)

A. 하기의 표시사항은 명세서 17쪽 첫째 줄에 기재된 기탁된 미생물 또는 기타 생물학적 물질에 관한 것임.	
B. 기탁물의 확인 ☐ 이외의 추가적 기탁물은 다음 별도의 페이지에 기재 ☐	
기탁기관의 명칭 KCTC 한국 생물자원센터	
기탁기관의 주소(우편번호 및 국명 포함) 대한민국 대전광역시 305-333 유성구 어은동 52	
기탁일자 2009년 07월 01일	기탁번호 KCTC 11523BP
C. 추가적 표시사항(기재사항이 없으면 공란으로 제출) ☐ 별도의 추가 페이지 계속 ☐	
D. 표시사항이 지정하고 있는 지정국(표시사항이 모든 지정국을 위한 사항이 아닌 경우)	
E. 표시사항의 별도 제출(기재사항이 없으면 공란으로 제출)	
하기의 표시사항은 후에 국제사무국에 제출될 것임(표시사항의 일반적 성질을 특정할 것)	

수리환경 사용권	국제사무국 사용권
☐ 당해 용지는 국제출원과 함께 접수되었음	☐ 당해 용지는 국제출원과 함께 접수되었음
담당관	담당관

Form PCT/RO/134 (July 1998, reprint January 2004)

청구범위

- [청구항 1] 다음 단계를 포함하는, 희소 미생물 또는 신규 미생물의 단일 세포 유래 게놈 라이브러리의 제조방법:
 (a) 수집된 시료의 미생물군 커뮤니티 분석을 수행하는 단계;
 (b) 상기 커뮤니티 분석에서 우점종 또는 공지의 미생물로 확인된 미생물을 표지하는 단계;
 (c) 상기 표지된 미생물을 1차 FACS (Fluorescence Associated Cell Sorter) 기법을 이용하여 제거하는 단계;
 (d) 상기 표지된 미생물이 제거된 미생물군 전체를 표지한 다음, 2차 FACS 기법을 이용하여 상기 미생물군 전체를 단일 세포로 분리하는 단계;
 (e) 상기 분리된 단일세포의 게놈을 다중분리증폭법 (multiple displacement amplification, MDA)를 이용하여 증폭하는 단계; 및
 (f) 상기 증폭된 단일세포의 게놈을 유전자 라이브러리로 제조하는 단계.
- [청구항 2] 제1항에 있어서, 상기 시료는 물, 토양, 공기, 식품, 폐기물, 동식물 장내 및 동식물 조직 중 어느 하나 이상에서 채취된 것임을 특징으로 하는 방법.
- [청구항 3] 제1항에 있어서, 상기 미생물군의 커뮤니티 분석은 ARISA (amplified riosomal intergenic spacer analysis), T-RFLP (Terminal restriction fragment length polymorphism), DGGE (Denaturing gradient gel electrophoresis), TGGE (Temperature Gradient Gel Electrophoresis), ARDRA (Amplified Ribosomal DNA Restriction Analysis) 및 DNA 마이크로어레이 중 어느 하나 이상의 방법에 의하여 이루어지는 것을 특징으로 하는 방법.
- [청구항 4] 제1항에 있어서, 상기 미생물군의 커뮤니티 분석은 16S rDNA를 증폭하여 제작한 라이브러리 클론을 시퀀싱하여 분석하는 방법으로 수행되는 것을 특징으로 하는 방법.
- [청구항 5] 제1항에 있어서, 상기 (b)단계의 표지는 FISH (Fluoresence *In Situ* Hygridization) 기법에 의해 수행되는 것을 특징으로 하는 방법.
- [청구항 6] 제5항에 있어서, 상기 (b)단계의 표지는 우점종 또는 공지의 미생물로 확인된 미생물의 16S rDNA를 타겟으로 하는 프로브를 이용하는 것을 특징으로 하는 방법.
- [청구항 7] 제1항에 있어서, 상기 (d) 단계는 표지된 미생물이 제거된 미생물군의 커뮤니티를 재분석하여 우점종 또는 공지의 미생물이 확인되는 경우 표지하여 제거하는 단계를 추가로 반복하여 수행한 다음, 남은 미생물군 전체에 대하여 표지하여 2차 FACS 기법을

이용하여 단일세포로 분리하는 단계를 수행하는 것을 특징으로 하는 방법.

[청구항 8] 제1항에 있어서, 상기 (f)단계에서 증폭된 게놈의 16S rRNA를 공지 미생물의 16S rRNA와 대비하여 신규한 미생물로 확인된 게놈만을 라이브러리로 제조하는 것을 특징으로 하는 방법.

[청구항 9] 제1항에 있어서, 상기 (f)단계에서 증폭된 게놈을 분석하여 단일세포로부터 유래된 것으로 확인된 게놈만을 라이브러리로 제조하는 것을 특징으로 하는 방법.

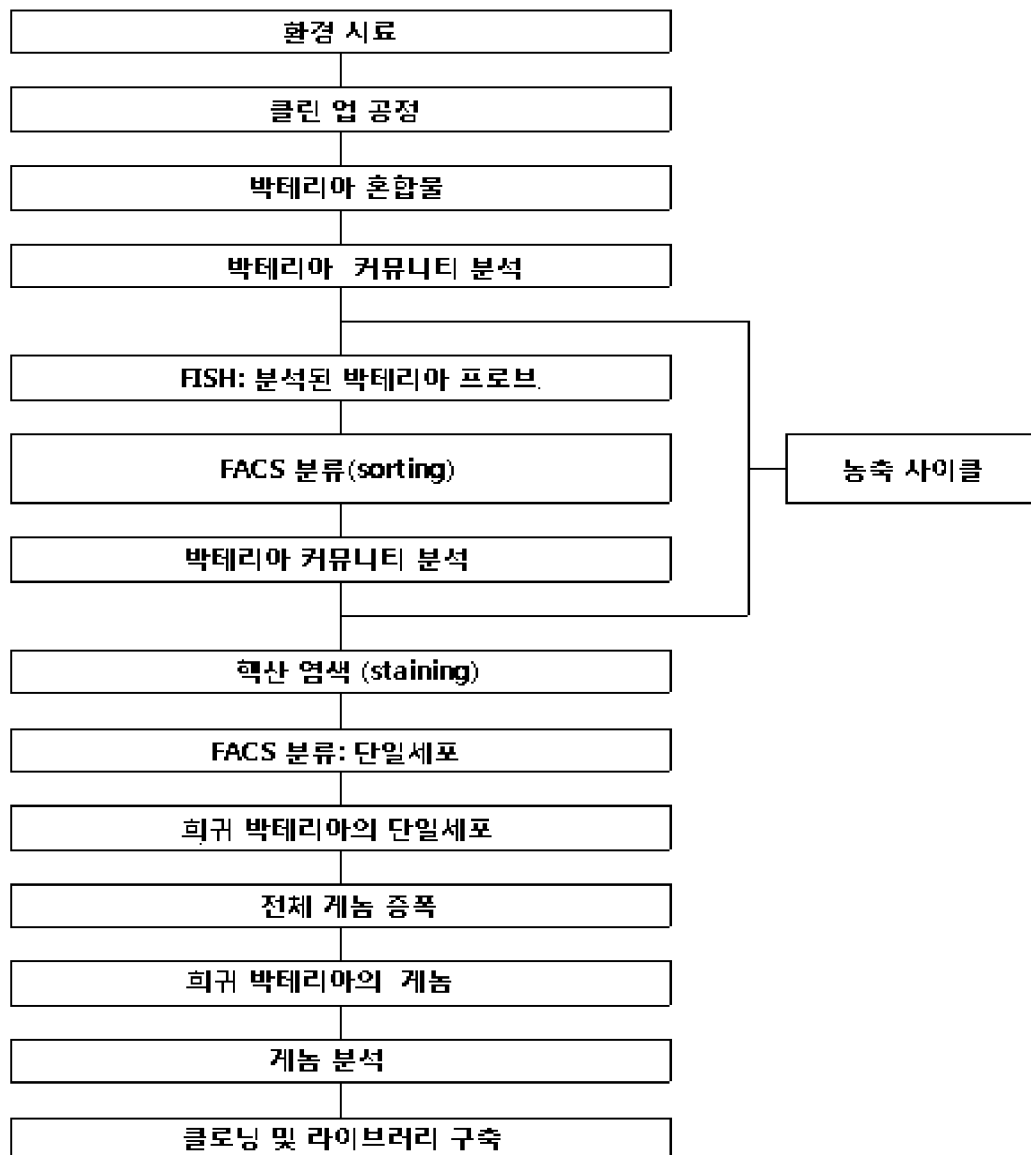
[청구항 10] 다음 단계를 포함하는, 희소 미생물 또는 신규 미생물의 단일 세포 유래 게놈 라이브러리를 이용한 신규 활성효소 또는 이를 코딩하는 유전자의 스크리닝 방법:

- (a) 수집된 시료의 미생물군 커뮤니티 분석을 수행하는 단계;
- (b) 상기 커뮤니티 분석에서 우점종 또는 공지의 미생물로 확인된 미생물을 표지하는 단계;
- (c) 상기 표지된 미생물을 1차 FACS (Fluorescence Associated Cell Sorter) 기법을 이용하여 제거하는 단계;
- (d) 상기 표지된 미생물이 제거된 미생물군 전체를 표지한 다음, 2차 FACS 기법을 이용하여 상기 미생물군 전체를 단일 세포로 분리하는 단계;
- (e) 상기 분리된 단일세포의 게놈을 다중분리증폭법 (multiple displacement amplification, MDA)를 이용하여 증폭하는 단계;
- (f) 상기 증폭된 단일세포의 게놈을 유전자 라이브러리로 제조하는 단계;
- (g) 상기 제조된 게놈 라이브러리에 타겟 효소 기질을 처리하여 기질반응이 나타난 클론을 선별하는 단계; 및
- (h) 상기 선별된 클론으로부터 타겟 효소 또는 이를 코딩하는 유전자를 수득하는 단계.

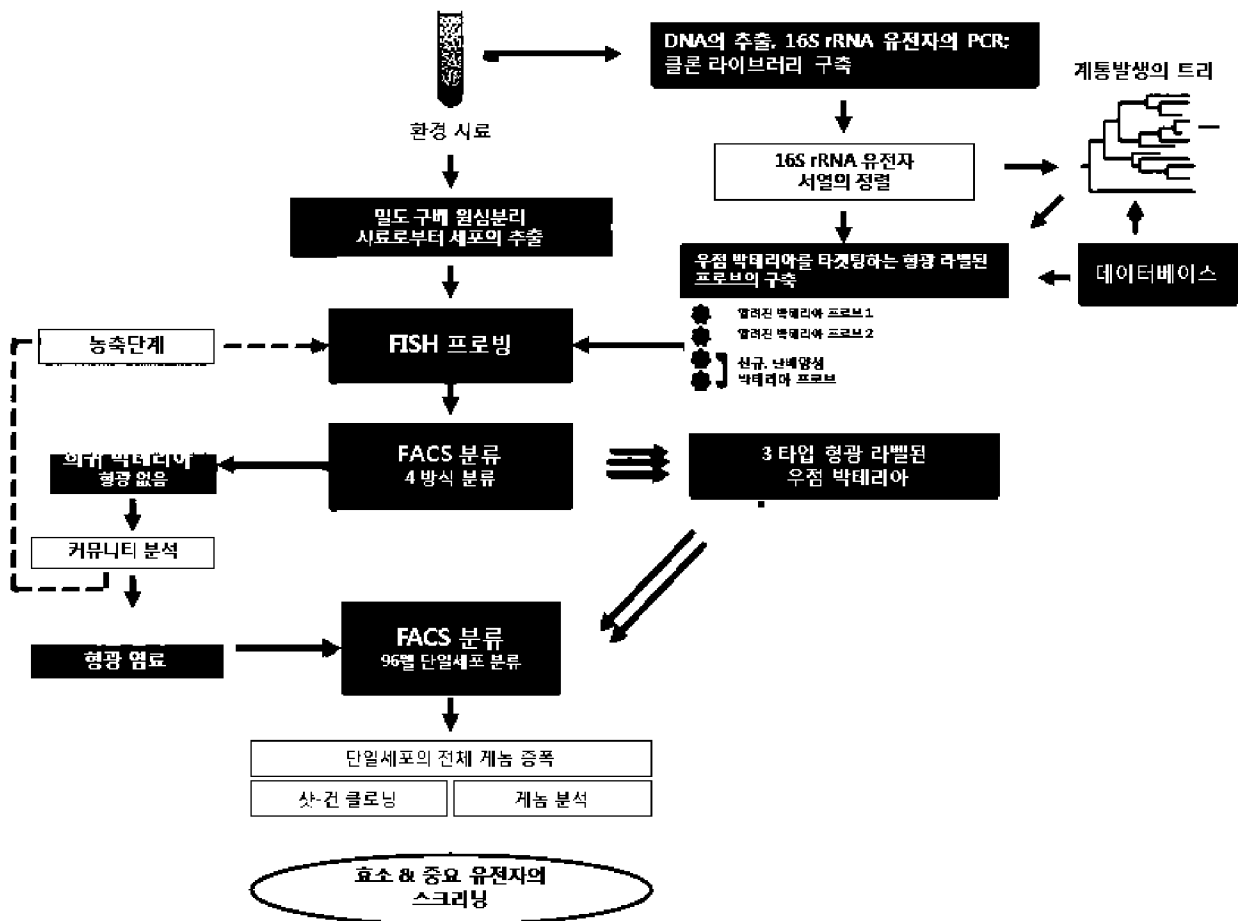
[청구항 11] 다음 단계를 포함하는, 미생물 다양성이 증가된 메타게노믹(metagenomic) 라이브러리의 제조방법:

- (a) 수집된 환경 시료의 미생물군 커뮤니티 분석을 수행하는 단계;
- (b) 상기 커뮤니티 분석에서 우점종 또는 공지의 미생물로 확인된 미생물을 표지하는 단계;
- (c) 상기 표지된 미생물을 FACS (Fluorescence Associated Cell Sorter) 기법을 이용하여 제거하는 단계; 및
- (d) 상기 표지된 미생물이 제거된 미생물군으로부터 DNA를 분리하고 증폭하여 메타게노믹 라이브러리를 제조하는 단계.

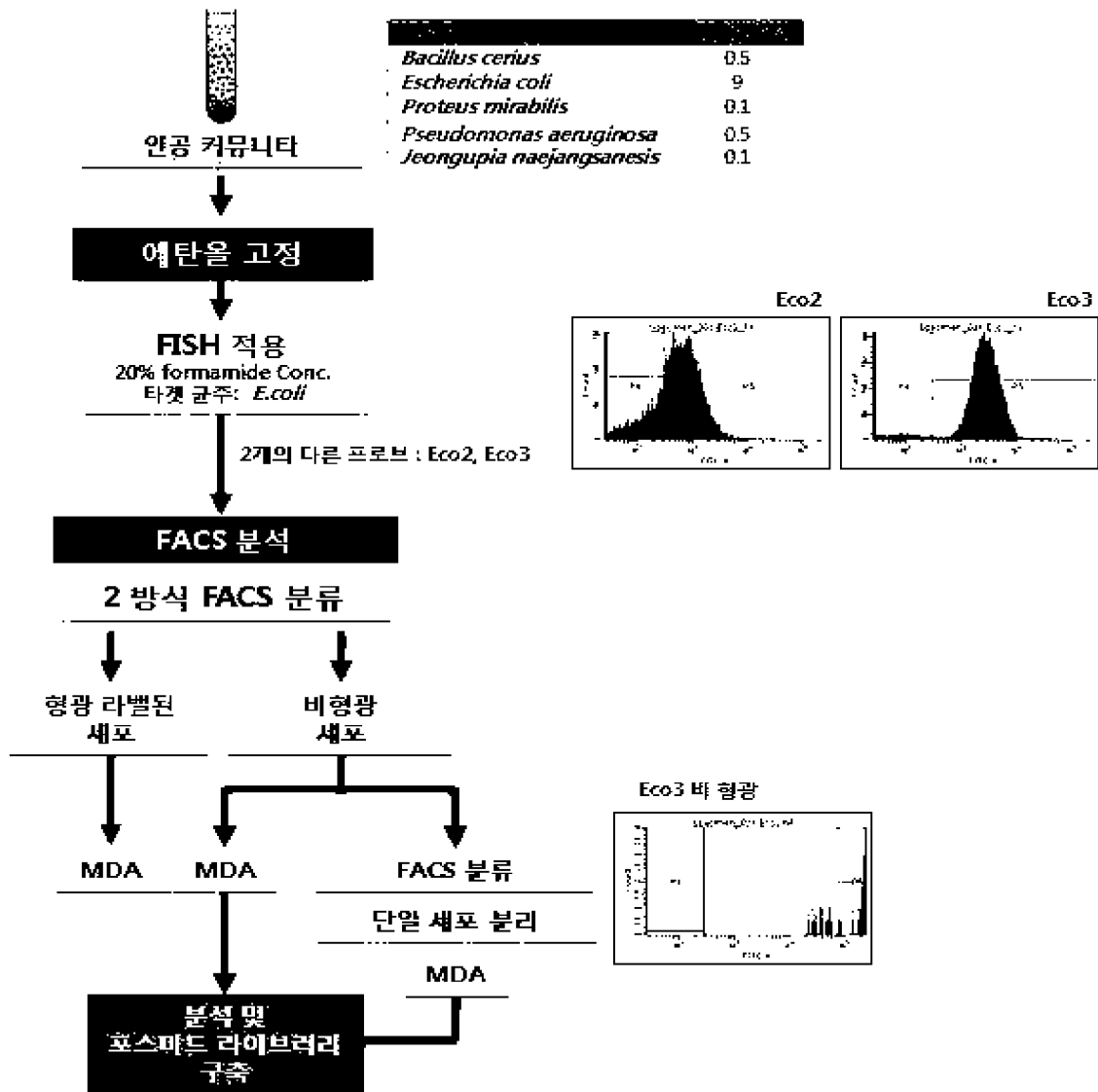
[Fig. 1]



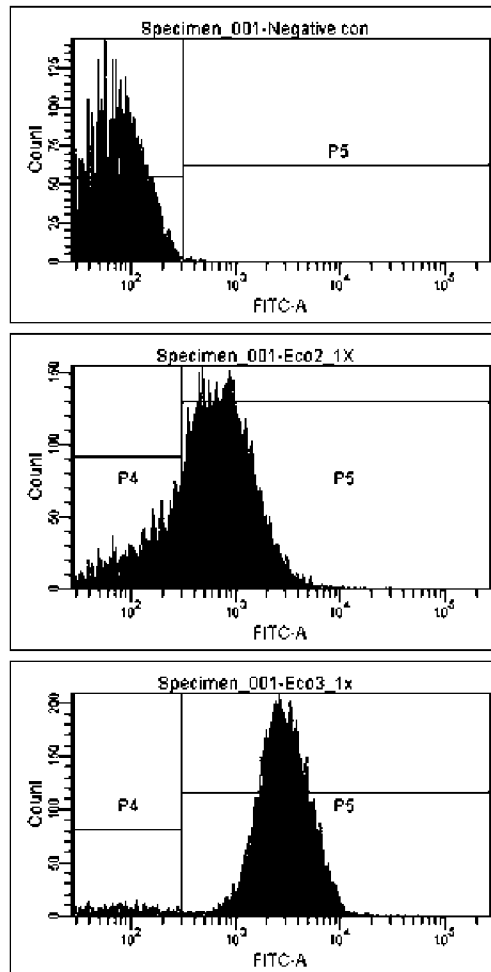
[Fig. 2]



[Fig. 3]



[Fig. 4]



음성 대조군

Tube: Negative con			
Population	#Events	%Parent	%Total
■ All Events	10,000		100.0
■ P1	8,870	88.7	88.7
□ P2	8,413	84.8	84.1
■ P3	8,256	82.6	82.6
⊗ P4	3,983	48.2	39.8
⊗ P5	12	0.1	0.1

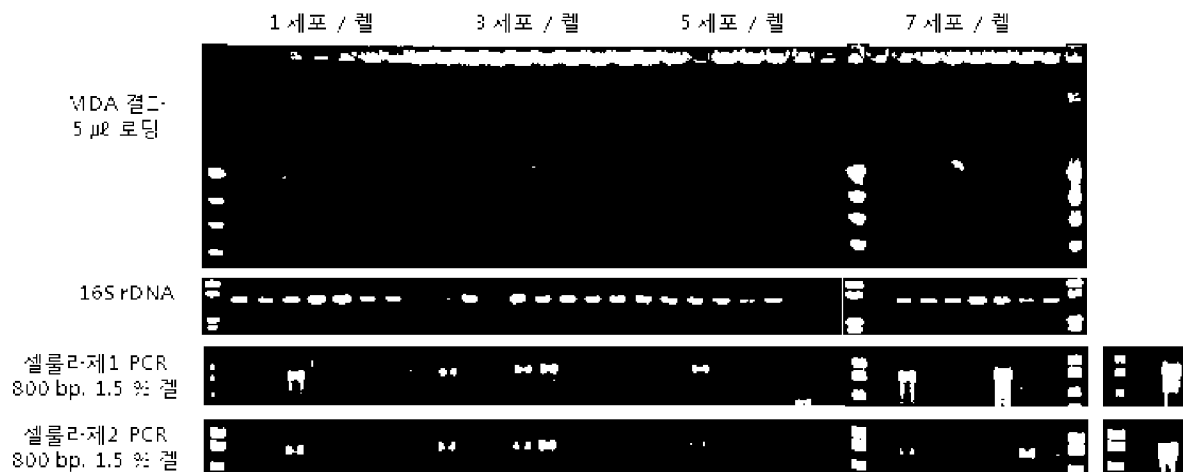
Eco2 프로브

Tube: Eco2_1X			
Population	#Events	%Parent	%Total
■ All Events	10,000		100.0
■ P1	9,169	91.7	91.7
□ P2	8,924	89.2	89.2
■ P3	8,785	87.8	87.8
⊗ P4	1,861	18.6	18.6
⊗ P5	6,338	63.4	63.4

Eco3 프로브

Tube: Eco3_1x			
Population	#Events	%Parent	%Total
■ All Events	10,000		100.0
■ P1	8,870	88.7	88.7
□ P2	8,655	86.6	86.6
■ P3	8,477	84.8	84.8
⊗ P4	414	4.1	4.1
⊗ P5	7,554	75.5	75.5

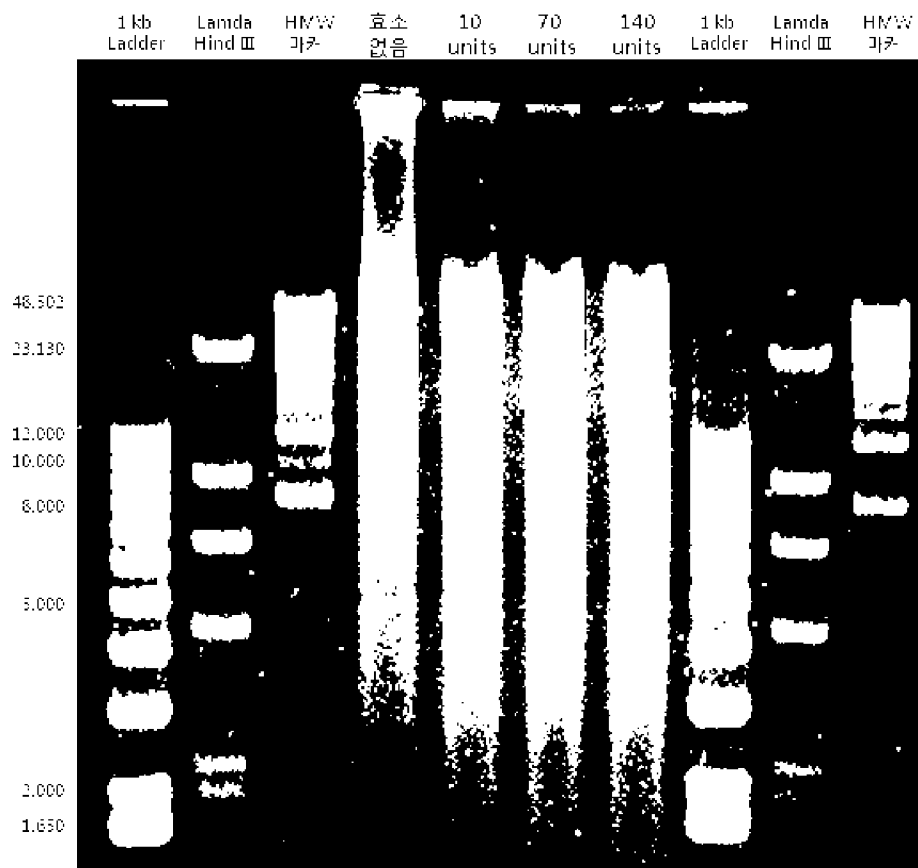
[Fig. 5]



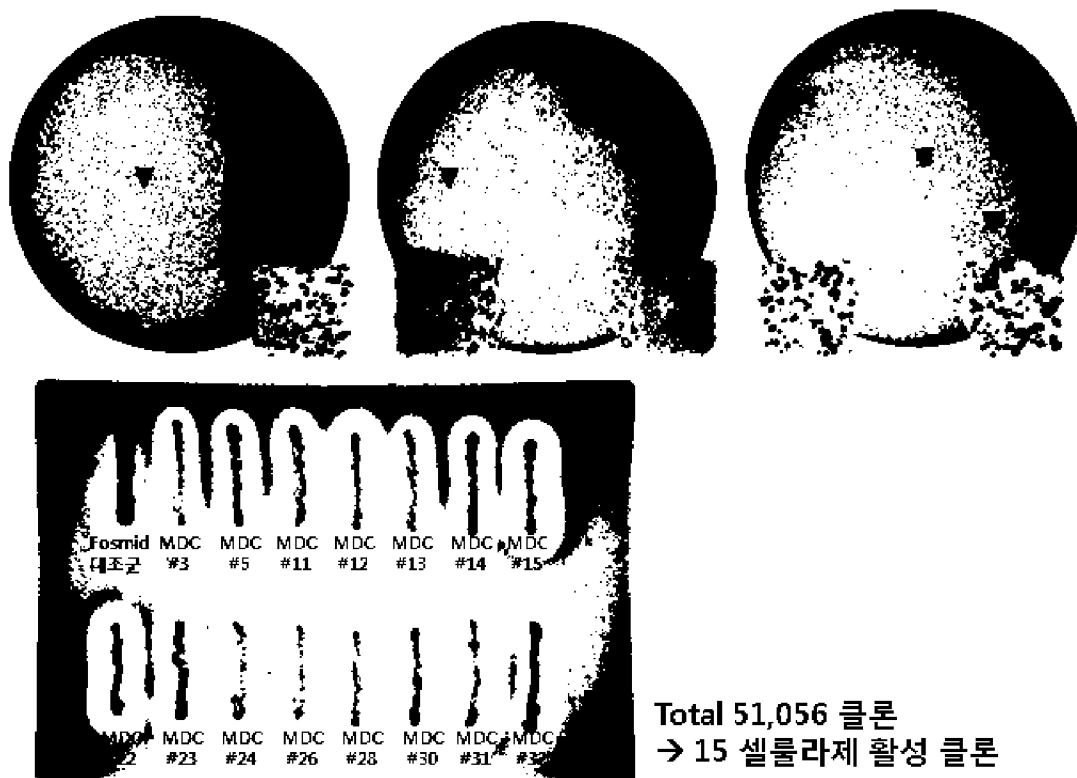
[Fig. 6]

[illegible]

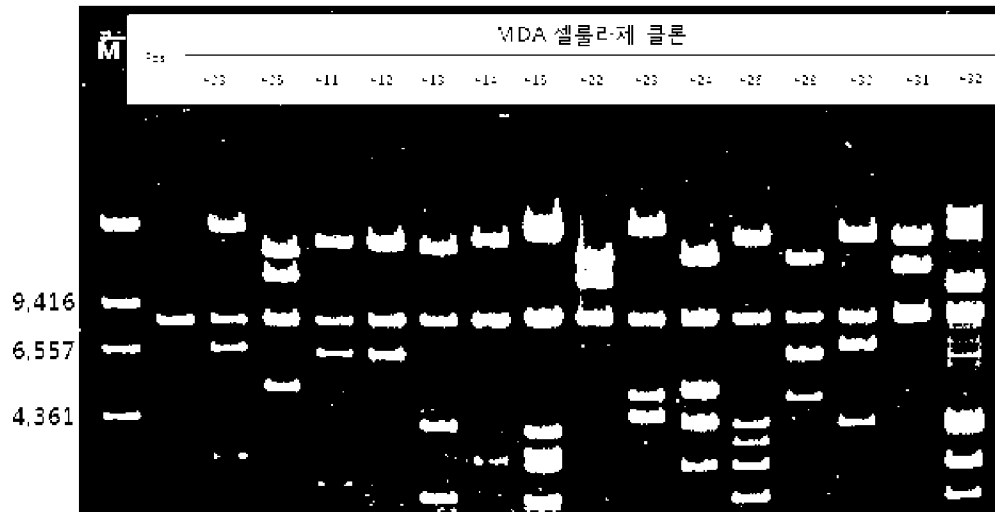
[Fig. 7]



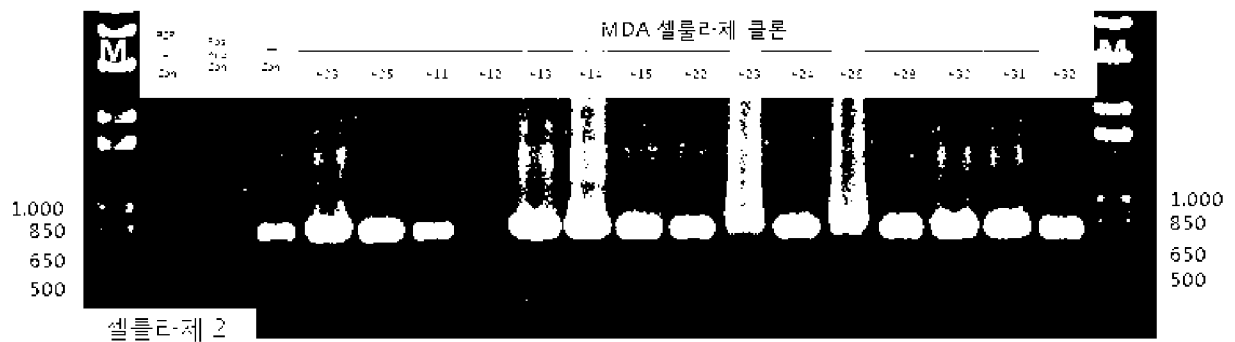
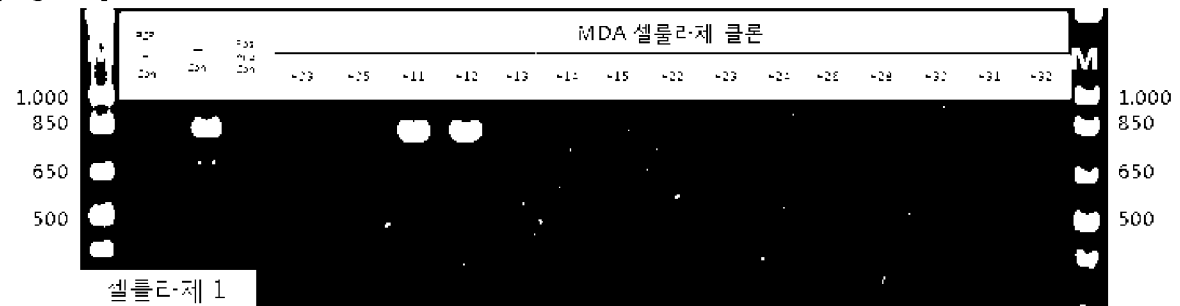
[Fig. 8]



[Fig. 9]



[Fig. 10]



	19	29	39	49	59	69	79	89	99	109	
Cell #1	AACATATGCTCGCACTGGCCCTGCTGCTGCCGTGCGCGGGCCATGCCGCTGACGCGCGCTGGGACGATTTCGTAGCCGTTTCGTGCGCGCGGACGGCGG										
Cell #2	AACATATGCTCGCACTGGCCCTGCTGCTGCCGTGCGCGGGCCATGCCGCTGACGCGCGCTGGGACGATTTCGTAGCCGTTTCGTGCGCGCGGACGGCGG										
MDA #1	AACATATGCTCGCACTGGCCCTGCTGCTGCCGTGCGCGGGCCATGCCGCTGACGCGCGCTGGGACGATTTCGTAGCCGTTTCGTGCGCGCGGACGGCGG										
MDA #2	AACATATGCTCGCACTGGCCCTGCTGCTGCCGTGCGCGGGCCATGCCGCTGACGCGCGCTGGGACGATTTCGTAGCCGTTTCGTGCGCGCGGACGGCGG										
Fosmid #1	AACATATGCTCGCACTGGCCCTGCTGCTGCCGTGCGCGGGCCATGCCGCTGACGCGCGCTGGGACGATTTCGTAGCCGTTTCGTGCGCGCGGACGGCGG										
Fosmid #2	AACATATGCTCGCACTGGCCCTGCTGCTGCCGTGCGCGGGCCATGCCGCTGACGCGCGCTGGGACGATTTCGTAGCCGTTTCGTGCGCGCGGACGGCGG										
	119	129	139	149	159	169	179	189	199	209	
Cell #1	GGTGATCGATACCGGCAACGGCGGCATCAGCCACAGCGAGGGCCAGGGCTACACGCTGCTGCTGGCGGTGGGCCAGAACGACCGTGGCCACCTTCGGGCGC										
Cell #2	GGTGATCGATACCGGCAACGGCGGCATCAGCCACAGCGAGGGCCAGGGCTACACGCTGCTGCTGGCGGTGGGCCAGAACGACCGTGGCCACCTTCGGGCGC										
MDA #1	GGTGATCGATACCGGCAACGGCGGCATCAGCCACAGCGAGGGCCAGGGCTACACGCTGCTGCTGGCGGTGGGCCAGAACGACCGTGGCCACCTTCGGGCGC										
MDA #2	GGTGATCGATACCGGCAACGGCGGCATCAGCCACAGCGAGGGCCAGGGCTACACGCTGCTGCTGGCGGTGGGCCAGAACGACCGTGGCCACCTTCGGGCGC										
Fosmid #1	GGTGATCGATACCGGCAACGGCGGCATCAGCCACAGCGAGGGCCAGGGCTACACGCTGCTGCTGGCGGTGGGCCAGAACGACCGTGGCCACCTTCGGGCGC										
Fosmid #2	GGTGATCGATACCGGCAACGGCGGCATCAGCCACAGCGAGGGCCAGGGCTACACGCTGCTGCTGGCGGTGGGCCAGAACGACCGTGGCCACCTTCGGGCGC										
	219	229	239	249	259	269	279	289	299	309	
Cell #1	GTCCTGGGCATGGACGCGCAGCCGGCTGCAGACGCGCGGGCAGCGCGCTGTTTTTCATGGCAGTGGAAAGCCCGACGCCGACGGCGGCAAGATCGCGGACCGCA										
Cell #2	GTCCTGGGCATGGACGCGCAGCCGGCTGCAGACGCGCGGGCAGCGCGCTGTTTTTCATGGCAGTGGAAAGCCCGACGCCGACGGCGGCAAGATCGCGGACCGCA										
MDA #1	GTCCTGGGCATGGACGCGCAGCCGGCTGCAGACGCGCGGGCAGCGCGCTGTTTTTCATGGCAGTGGAAAGCCCGACGCCGACGGCGGCAAGATCGCGGACCGCA										
MDA #2	GTCCTGGGCATGGACGCGCAGCCGGCTGCAGACGCGCGGGCAGCGCGCTGTTTTTCATGGCAGTGGAAAGCCCGACGCCGACGGCGGCAAGATCGCGGACCGCA										
Fosmid #1	GTCCTGGGCATGGACGCGCAGCCGGCTGCAGACGCGCGGGCAGCGCGCTGTTTTTCATGGCAGTGGAAAGCCCGACGCCGACGGCGGCAAGATCGCGGACCGCA										
Fosmid #2	GTCCTGGGCATGGACGCGCAGCCGGCTGCAGACGCGCGGGCAGCGCGCTGTTTTTCATGGCAGTGGAAAGCCCGACGCCGACGGCGGCAAGATCGCGGACCGCA										
	319	329	339	349	359	369	379	389	399	409	
Cell #1	ACAACGCCAGCGCAGCGCGACCTGCTGATCGCCTGGGCGCTGCTGCGCGGCGCGCAACGCTGGCAGCAGCCGGCCCTACCGCGCCGAGGCACCTGAAAAATCCT										
Cell #2	ACAACGCCAGCGCAGCGCGACCTGCTGATCGCCTGGGCGCTGCTGCGCGGCGCGCAACGCTGGCAGCAGCCGGCCCTACCGCGCCGAGGCACCTGAAAAATCCT										
MDA #1	ACAACGCCAGCGCAGCGCGACCTGCTGATCGCCTGGGCGCTGCTGCGCGGCGCGCAACGCTGGCAGCAGCCGGCCCTACCGCGCCGAGGCACCTGAAAAATCCT										
MDA #2	ACAACGCCAGCGCAGCGCGACCTGCTGATCGCCTGGGCGCTGCTGCGCGGCGCGCAACGCTGGCAGCAGCCGGCCCTACCGCGCCGAGGCACCTGAAAAATCCT										
Fosmid #1	ACAACGCCAGCGCAGCGCGACCTGCTGATCGCCTGGGCGCTGCTGCGCGGCGCGCAACGCTGGCAGCAGCCGGCCCTACCGCGCCGAGGCACCTGAAAAATCCT										
Fosmid #2	ACAACGCCAGCGCAGCGCGACCTGCTGATCGCCTGGGCGCTGCTGCGCGGCGCGCAACGCTGGCAGCAGCCGGCCCTACCGCGCCGAGGCACCTGAAAAATCCT										
	419	429	439	449	459	469	479	489	499	509	
Cell #1	CGCCGACGCTGCGCGGCAAACTGATCAAGCCGTGCGCCTTCGGCCCGCTGCTGATGCCGGCGGAAAAAGGCTTCGACAAGCCTGCCGGCCCCATCGTCAAT										
Cell #2	CGCCGACGCTGCGCGGCAAACTGATCAAGCCGTGCGCCTTCGGCCCGCTGCTGATGCCGGCGGAAAAAGGCTTCGACAAGCCTGCCGGCCCCATCGTCAAT										
MDA #1	CGCCGACGCTGCGCGGCAAACTGATCAAGCCGTGCGCCTTCGGCCCGCTGCTGATGCCGGCGGAAAAAGGCTTCGACAAGCCTGCCGGCCCCATCGTCAAT										
MDA #2	CGCCGACGCTGCGCGGCAAACTGATCAAGCCGTGCGCCTTCGGCCCGCTGCTGATGCCGGCGGAAAAAGGCTTCGACAAGCCTGCCGGCCCCATCGTCAAT										
Fosmid #1	CGCCGACGCTGCGCGGCAAACTGATCAAGCCGTGCGCCTTCGGCCCGCTGCTGATGCCGGCGGAAAAAGGCTTCGACAAGCCTGCCGGCCCCATCGTCAAT										
Fosmid #2	CGCCGACGCTGCGCGGCAAACTGATCAAGCCGTGCGCCTTCGGCCCGCTGCTGATGCCGGCGGAAAAAGGCTTCGACAAGCCTGCCGGCCCCATCGTCAAT										
	519	529	539	549	559	569	579	589	599	609	
Cell #1	CCGTCTGCTACTGGGCTGTTCCCGGGCTTTGCGGAATTGCGCCGCGCGCGACCGTGCGCCCGGTGTGGGCGGAAGTTCGCGCAGAGCGGGCTGGCGCTGCTCGCGG										
Cell #2	CCGTCTGCTACTGGGCTGTTCCCGGGCTTTGCGGAATTGCGCCGCGCGCGACCGTGCGCCCGGTGTGGGCGGAAGTTCGCGCAGAGCGGGCTGGCGCTGCTCGCGG										
MDA #1	CCGTCTGCTACTGGGCTGTTCCCGGGCTTTGCGGAATTGCGCCGCGCGCGACCGTGCGCCCGGTGTGGGCGGAAGTTCGCGCAGAGCGGGCTGGCGCTGCTCGCGG										
MDA #2	CCGTCTGCTACTGGGCTGTTCCCGGGCTTTGCGGAATTGCGCCGCGCGCGACCGTGCGCCCGGTGTGGGCGGAAGTTCGCGCAGAGCGGGCTGGCGCTGCTCGCGG										
Fosmid #1	CCGTCTGCTACTGGGCTGTTCCCGGGCTTTGCGGAATTGCGCCGCGCGCGACCGTGCGCCCGGTGTGGGCGGAAGTTCGCGCAGAGCGGGCTGGCGCTGCTCGCGG										
Fosmid #2	CCGTCTGCTACTGGGCTGTTCCCGGGCTTTGCGGAATTGCGCCGCGCGCGACCGTGCGCCCGGTGTGGGCGGAAGTTCGCGCAGAGCGGGCTGGCGCTGCTCGCGG										
	619	629	639	649	659	669	679	689	699	709	
Cell #1	CCGGCCGCTTTCGGCCAGTGGCAGCTGCCGCCGAGTGGCTGCAGGTGCGCGACACGCTGCGCCCGGGCGATGTCGCACCAAGTACGGCTACAACCGCGCTGGC										
Cell #2	CCGGCCGCTTTCGGCCAGTGGCAGCTGCCGCCGAGTGGCTGCAGGTGCGCGACACGCTGCGCCCGGGCGATGTCGCACCAAGTACGGCTACAACCGCGCTGGC										
MDA #1	CCGGCCGCTTTCGGCCAGTGGCAGCTGCCGCCGAGTGGCTGCAGGTGCGCGACACGCTGCGCCCGGGCGATGTCGCACCAAGTACGGCTACAACCGCGCTGGC										
MDA #2	CCGGCCGCTTTCGGCCAGTGGCAGCTGCCGCCGAGTGGCTGCAGGTGCGCGACACGCTGCGCCCGGGCGATGTCGCACCAAGTACGGCTACAACCGCGCTGGC										
Fosmid #1	CCGGCCGCTTTCGGCCAGTGGCAGCTGCCGCCGAGTGGCTGCAGGTGCGCGACACGCTGCGCCCGGGCGATGTCGCACCAAGTACGGCTACAACCGCGCTGGC										
Fosmid #2	CCGGCCGCTTTCGGCCAGTGGCAGCTGCCGCCGAGTGGCTGCAGGTGCGCGACACGCTGCGCCCGGGCGATGTCGCACCAAGTACGGCTACAACCGCGCTGGC										
	719	729	739	749	759	769	779	789	799	809	
Cell #1	CGTACCGTTGTATCTGGCGTGGGCGGGGCTGCAGACGCGCCG										

Cell #1 A CATATGTT GAGAGGAGCGCATCCGGTGTTCCTCCGCTCGTGAAGAAGACCGTGGTTCGTGTGCTGGCCGCGTGGCCAGCTTGGTTTGTATCCCGCGCATTC
Cell #2 AACATATGTTGAGAGGAGCGCATCCGGTGTTCCTCCGCTCGTGAAGAAGACCGTGGTTCGTGTGCTGGCCGCGTGGCCAGCTTGGTTTGTATCCCGCGCATTC
Posmid #1 AA ATATGTTGAGAGGAGCGCATCCGGTGTTCCTCCGCTCGTGAAGAAGACCGTGGTTCGTGTGCTGGCCGCGTGGCCAGCTTGGTTTGTATCCCGCGCATTC
Posmid #2 AACATATGTTGAGAGGAGCGCATCCGGTGTTCCTCCGCTCGTGAAGAAGACCGTGGTTCGTGTGCTGGCCGCGTGGCCAGCTTGGTTTGTATCCCGCGCATTC
MDA #1 AC ATATGTTGAGAGGAGCGCATCCGGTGTTCCTCCGCTCGTGAAGAAGACCGTGGTTCGTGTGCTGGCCGCGTGGCCAGCTTGGTTTGTATCCCGCGCATTC
MDA #2 GAGGAGCGCGCATCCGGTGTTCCTCCGCTCGTGAAGAAGACCGTGGTTCGTGTGCTGGCCGCGTGGCCAGCTTGGTTTGTATCCCGCGCATTC

Cell #1 TGCCTTTGGCTACGTCCTTGGCGGATTTACGCCGCCAGTGGTGAATCAGTCCCAAATGCTGAATCAGGTGCAGCCGCCATATGGGGTTGGAAAACCGGT
Cell #2 TGCCCTTTGGCTACGTCCTTGGCGGATTTACGCCGCCAGTGGTGAATCAGTCCCAAATGCTGAATCAGGTGCAGCCGCCATATGGGGTTGGAAAACCGGT
Posmid #1 TGCCCTTTGGCTACGTCCTTGGCGGATTTACGCCGCCAGTGGTGAATCAGTCCCAAATGCTGAATCAGGTGCAGCCGCCATATGGGGTTGGAAAACCGGT
Posmid #2 TGCCCTTTGGCTACGTCCTTGGCGGATTTACGCCGCCAGTGGTGAATCAGTCCCAAATGCTGAATCAGGTGCAGCCGCCATATGGGGTTGGAAAACCGGT
MDA #1 TGCCCTTTGGCTACGTCCTTGGCGGATTTACGCCGCCAGTGGTGAATCAGTCCCAAATGCTGAATCAGGTGCAGCCGCCATATGGGGTTGGAAAACCGGT
MDA #2 TGCCCTTTGGCTACGTCCTTGGCGGATTTACGCCGCCAGTGGTGAATCAGTCCCAAATGCTGAATCAGGTGCAGCCGCCATATGGGGTTGGAAAACCGGT

Cell #1 TATCTGCGCACCGACCGCAGGACGCAAGCCGGCTTTACGTCCTGGCTCCGAGCCGCGCAAGCAGGTGGTGTCCGAAAGCCAGGGCTTCGGCATGC
Cell #2 TATCTGCGCACCGACCGCAGGACGCAAGCCGGCTTTACGTCCTGGCTCCGAGCCGCGCAAGCAGGTGGTGTCCGAAAGCCAGGGCTTCGGCATGC
Posmid #1 TATCTGCGCACCGACCGCAGGACGCAAGCCGGCTTTACGTCCTGGCTCCGAGCCGCGCAAGCAGGTGGTGTCCGAAAGCCAGGGCTTCGGCATGC
Posmid #2 TATCTGCGCACCGACCGCAGGACGCAAGCCGGCTTTACGTCCTGGCTCCGAGCCGCGCAAGCAGGTGGTGTCCGAAAGCCAGGGCTTCGGCATGC
MDA #1 TATCTGCGCACCGACCGCAGGACGCAAGCCGGCTTTACGTCCTGGCTCCGAGCCGCGCAAGCAGGTGGTGTCCGAAAGCCAGGGCTTCGGCATGC
MDA #2 TATCTGCGCACCGACCGCAGGACGCAAGCCGGCTTTACGTCCTGGCTCCGAGCCGCGCAAGCAGGTGGTGTCCGAAAGCCAGGGCTTCGGCATGC

Cell #1 TGGCGGTGGCCGAGATGGCGGCGGTGATCCGACGCCGCCGGGACATTGATGCACTGTTCCGCTACTACCGTGCACCTCCGAGCTCCGACGCCCGTG
Cell #2 TGGCGGTGGCCGAGATGGCGGCGGTGATCCGACGCCGCCGGGACATTGATGCACTGTTCCGCTACTACCGTGCACCTCCGAGCTCCGACGCCCGTG
Posmid #1 TGGCGGTGGCCGAGATGGCGGCGGTGATCCGACGCCGCCGGGACATTGATGCACTGTTCCGCTACTACCGTGCACCTCCGAGCTCCGACGCCCGTG
Posmid #2 TGGCGGTGGCCGAGATGGCGGCGGTGATCCGACGCCGCCGGGACATTGATGCACTGTTCCGCTACTACCGTGCACCTCCGAGCTCCGACGCCCGTG
MDA #1 TGGCGGTGGCCGAGATGGCGGCGGTGATCCGACGCCGCCGGGACATTGATGCACTGTTCCGCTACTACCGTGCACCTCCGAGCTCCGACGCCCGTG
MDA #2 TGGCGGTGGCCGAGATGGCGGCGGTGATCCGACGCCGCCGGGACATTGATGCACTGTTCCGCTACTACCGTGCACCTCCGAGCTCCGACGCCCGTG

Cell #1 GCTGATGGCATGGGCGCAATCGCTGTGGACAAACCGGCTGGTGATACCGAGCGCAGAACTGGGCCACCGAGCGCGATCTCGATGCTGCCTACGGCTTG
Cell #2 GCTGATGGCATGGGCGCAATCGCTGTGGACAAACCGGCTGGTGATACCGAGCGCAGAACTGGGCCACCGAGCGCGATCTCGATGCTGCCTACGGCTTG
Posmid #1 GCTGATGGCATGGGCGCAATCGCTGTGGACAAACCGGCTGGTGATACCGAGCGCAGAACTGGGCCACCGAGCGCGATCTCGATGCTGCCTACGGCTTG
Posmid #2 GCTGATGGCATGGGCGCAATCGCTGTGGACAAACCGGCTGGTGATACCGAGCGCAGAACTGGGCCACCGAGCGCGATCTCGATGCTGCCTACGGCTTG
MDA #1 GCTGATGGCATGGGCGCAATCGCTGTGGACAAACCGGCTGGTGATACCGAGCGCAGAACTGGGCCACCGAGCGCGATCTCGATGCTGCCTACGGCTTG
MDA #2 GCTGATGGCATGGGCGCAATCGCTGTGGACAAACCGGCTGGTGATACCGAGCGCAGAACTGGGCCACCGAGCGCGATCTCGATGCTGCCTACGGCTTG

Cell #1 CTGCTGGCGCATCAGGCTCTGGGGTTTCGACCGGGACGATCAACTACCCGACGAGCGCAGAGACCATCAATCGCTGATGGACAAAGGTGGTGAATCAGT
Cell #2 CTGCTGGCGCATCAGGCTCTGGGGTTTCGACCGGGACGATCAACTACCCGACGAGCGCAGAGACCATCAATCGCTGATGGACAAAGGTGGTGAATCAGT
Posmid #1 CTGCTGGCGCATCAGGCTCTGGGGTTTCGACCGGGACGATCAACTACCCGACGAGCGCAGAGACCATCAATCGCTGATGGACAAAGGTGGTGAATCAGT
Posmid #2 CTGCTGGCGCATCAGGCTCTGGGGTTTCGACCGGGACGATCAACTACCCGACGAGCGCAGAGACCATCAATCGCTGATGGACAAAGGTGGTGAATCAGT
MDA #1 CTGCTGGCGCATCAGGCTCTGGGGTTTCGACCGGGACGATCAACTACCCGACGAGCGCAGAGACCATCAATCGCTGATGGACAAAGGTGGTGAATCAGT
MDA #2 CTGCTGGCGCATCAGGCTCTGGGGTTTCGACCGGGACGATCAACTACCCGACGAGCGCAGAGACCATCAATCGCTGATGGACAAAGGTGGTGAATCAGT

Cell #1 CCGAGTGGACGCTCAAGCTCGGTGATTGGGTCCGGGATGGCGATGCGACCTATGGCAAGGGGCTGCGGGCTCGGACCTGATGTTCAACCACCTTGGCGCG
Cell #2 CCGAGTGGACGCTCAAGCTCGGTGATTGGGTCCGGGATGGCGATGCGACCTATGGCAAGGGGCTGCGGGCTCGGACCTGATGTTCAACCACCTTGGCGCG
Posmid #1 CCGAGTGGACGCTCAAGCTCGGTGATTGGGTCCGGGATGGCGATGCGACCTATGGCAAGGGGCTGCGGGCTCGGACCTGATGTTCAACCACCTTGGCGCG
Posmid #2 CCGAGTGGACGCTCAAGCTCGGTGATTGGGTCCGGGATGGCGATGCGACCTATGGCAAGGGGCTGCGGGCTCGGACCTGATGTTCAACCACCTTGGCGCG
MDA #1 CCGAGTGGACGCTCAAGCTCGGTGATTGGGTCCGGGATGGCGATGCGACCTATGGCAAGGGGCTGCGGGCTCGGACCTGATGTTCAACCACCTTGGCGCG
MDA #2 CCGAGTGGACGCTCAAGCTCGGTGATTGGGTCCGGGATGGCGATGCGACCTATGGCAAGGGGCTGCGGGCTCGGACCTGATGTTCAACCACCTTGGCGCG

Cell #1 GTTTCACGAGCGACCGCGCAGCGACGCTGGCGCGCGGTGCTCGACAAGAGCTACGTGATTGGCGATTTCGGTGTTCGCAACTACGGGCGCAACCCGGT
Cell #2 GTTTCACGAGCGACCGCGCAGCGACGCTGGCGCGCGGTGCTCGACAAGAGCTACGTGATTGGCGATTTCGGTGTTCGCAACTACGGGCGCAACCCGGT
Posmid #1 GTTTCACGAGCGACCGCGCAGCGACGCTGGCGCGCGGTGCTCGACAAGAGCTACGTGATTGGCGATTTCGGTGTTCGCAACTACGGGCGCAACCCGGT
Posmid #2 GTTTCACGAGCGACCGCGCAGCGACGCTGGCGCGCGGTGCTCGACAAGAGCTACGTGATTGGCGATTTCGGTGTTCGCAACTACGGGCGCAACCCGGT
MDA #1 GTTTCACGAGCGACCGCGCAGCGACGCTGGCGCGCGGTGCTCGACAAGAGCTACGTGATTGGCGATTTCGGTGTTCGCAACTACGGGCGCAACCCGGT
MDA #2 GTTTCACGAGCGACCGCGCAGCGACGCTGGCGCGCGGTGCTCGACAAGAGCTACGTGATTGGCGATTTCGGTGTTCGCAACTACGGGCGCAACCCGGT

Cell #1 CTGCTGCGCGATTTCATGCTCAAGCAGGCGACCGATTCCCGCGCGGCTTCGCGCGGGTTTTCGGAAGGTGCGAACGCGGCAACTACAGCTGGAAACAATT
Cell #2 CTGCTGCGCGATTTCATGCTCAAGCAGGCGACCGATTCCCGCGCGGCTTCGCGCGGGTTTTCGGAAGGTGCGAACGCGGCAACTACAGCTGGAAACAATT
Posmid #1 CTGCTGCGCGATTTCATGCTCAAGCAGGCGACCGATTCCCGCGCGGCTTCGCGCGGGTTTTCGGAAGGTGCGAACGCGGCAACTACAGCTGGAAACAATT
Posmid #2 CTGCTGCGCGATTTCATGCTCAAGCAGGCGACCGATTCCCGCGCGGCTTCGCGCGGGTTTTCGGAAGGTGCGAACGCGGCAACTACAGCTGGAAACAATT
MDA #1 CTGCTGCGCGATTTCATGCTCAAGCAGGCGACCGATTCCCGCGCGGCTTCGCGCGGGTTTTCGGAAGGTGCGAACGCGGCAACTACAGCTGGAAACAATT
MDA #2 CTGCTGCGCGATTTCATGCTCAAGCAGGCGACCGATTCCCGCGCGGCTTCGCGCGGGTTTTCGGAAGGTGCGAACGCGGCAACTACAGCTGGAAACAATT

Cell #1 GCCCGTGGCTGGCGGCTGAGTCTGGATTAACTGCTCAATGGCGATACCGCGGCTTGGCGTTGCTGCAGGCGTTGAATCGTGGGTGAGCAGCAGCGC
Cell #2 GCCCGTGGCTGGCGGCTGAGTCTGGATTAACTGCTCAATGGCGATACCGCGGCTTGGCGTTGCTGCAGGCGTTGAATCGTGGGTGAGCAGCAGCGC
Posmid #1 GCCCGTGGCTGGCGGCTGAGTCTGGATTAACTGCTCAATGGCGATACCGCGGCTTGGCGTTGCTGCAGGCGTTGAATCGTGGGTGAGCAGCAGCGC
Posmid #2 GCCCGTGGCTGGCGGCTGAGTCTGGATTAACTGCTCAATGGCGATACCGCGGCTTGGCGTTGCTGCAGGCGTTGAATCGTGGGTGAGCAGCAGCGC
MDA #1 GCCCGTGGCTGGCGGCTGAGTCTGGATTAACTGCTCAATGGCGATACCGCGGCTTGGCGTTGCTGCAGGCGTTGAATCGTGGGTGAGCAGCAGCGC
MDA #2 GCCCGTGGCTGGCGGCTGAGTCTGGATTAACTGCTCAATGGCGATACCGCGGCTTGGCGTTGCTGCAGGCGTTGAATCGTGGGTGAGCAGCAGCGC

Cell #1 AACTGGTGGCGATTTCGGAACCGGATTACGCTGGTGGTTATACGCTGGACGGGAGCGGCTTGAACCGGTACTTCAGATGGCGTTTTCGGGACCGGTTTGGCGCTC
Cell #2 AACTGGTGGCGATTTCGGAACCGGATTACGCTGGTGGTTATACGCTGGACGGGAGCGGCTTGAACCGGTACTTCAGATGGCGTTTTCGGGACCGGTTTGGCGCTC
Posmid #1 AACTGGTGGCGATTTCGGAACCGGATTACGCTGGTGGTTATACGCTGGACGGGAGCGGCTTGAACCGGTACTTCAGATGGCGTTTTCGGGACCGGTTTGGCGCTC
Posmid #2 AACTGGTGGCGATTTCGGAACCGGATTACGCTGGTGGTTATACGCTGGACGGGAGCGGCTTGAACCGGTACTTCAGATGGCGTTTTCGGGACCGGTTTGGCGCTC
MDA #1 AACTGGTGGCGATTTCGGAACCGGATTACGCTGGTGGTTATACGCTGGACGGGAGCGGCTTGAACCGGTACTTCAGATGGCGTTTTCGGGACCGGTTTGGCGCTC
MDA #2 AACTGGTGGCGATTTCGGAACCGGATTACGCTGGTGGTTATACGCTGGACGGGAGCGGCTTGAACCGGTACTTCAGATGGCGTTTTCGGGACCGGTTTGGCGCTC

Cell #1 AGCGCCACCATCGACAGCGGCAACCAAGCTGGCTGAACGCTTTGGAGCGCGCATCGCGGCGCGCGCGTCAACAGGATATTACTCGGACTCGATTTCGCC
Cell #2 AGCGCCACCATCGACAGCGGCAACCAAGCTGGCTGAACGCTTTGGAGCGCGCATCGCGGCGCGCGCGTCAACAGGATATTACTCGGACTCGATTTCGCC
Posmid #1 AGCG