



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0098149  
(43) 공개일자 2018년09월03일

(51) 국제특허분류(Int. Cl.) C12Q 1/6895 (2018.01) (52) CPC특허분류 C12Q 1/6895 (2018.05) C12Q 2531/113 (2013.01) (21) 출원번호           10-2018-0020804 (22) 출원일자           2018년02월21일 심사청구일자       2018년02월21일 (30) 우선권주장 1020170025085   2017년02월24일   대한민국(KR)	(71) 출원인 주식회사 엠디헬스케어 서울특별시 마포구 월드컵북로56길 9, 1303호(상 암동, 우리기술빌딩) (72) 발명자 김윤근 경기도 남양주시 호평로 93, 1707동 805호 (호평 동, 동원로얄듀크아파트) (74) 대리인 이명진
--	---

전체 청구항 수 : 총 2 항

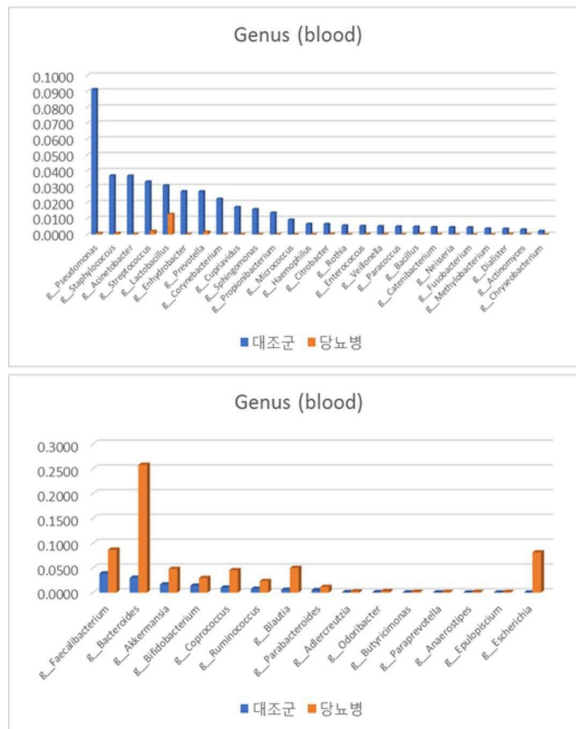
(54) 발명의 명칭 세균 메타게놈 분석을 통한 당뇨병 진단 방법

**(57) 요약**

본 발명은 세균 메타게놈 분석을 통해 당뇨병을 진단하는 방법에 관한 것으로서, 보다 구체적으로는 피검체 유래 샘플을 이용해 세균 메타게놈 분석을 수행하여 특정 세균 유래 세포막 소포의 함량 증감을 분석함으로써 당뇨병의 원인인자, 질병 위험도, 및 경과를 예측하는 방법에 관한 것이다.

(뒷면에 계속)

**대표도** - 도6



환경에 존재하는 세균에서 분비되는 세포막 소포는 체내에 흡수되어 인슐린에 반응하는 장기에 분포하여 인슐린 저항성 등의 대사기능에 영향을 주어 당뇨병을 유도 혹은 억제할 수 있는데, 당뇨병은 증상이 나타나기 전 발생을 예측하는 것이 어려워 효율적인 치료가 어려운 실정이므로, 본 발명에 따른 인체 유래 샘플을 이용한 세균 유래 세포막 소포의 메타게놈 분석을 통해 당뇨병의 원인인자를 진단하고, 발병의 위험도를 미리 진단함으로써 당뇨병의 위험군을 조기에 진단 및 예측하여 적절한 관리를 통해 발병 시기를 늦추거나 발병을 예방할 수 있으며, 발병 후에도 원인인자를 진단 할 수 있어 당뇨병의 발병률을 낮추고 치료효과를 높일 수 있다.

이 발명을 지원한 국가연구개발사업

과제고유번호 NRF-2016M3A9B6901516  
 부처명 미래창조과학부  
 연구관리전문기관 한국연구재단  
 연구사업명 바이오의료기술개발사업 차세대응용오믹스사업  
 연구과제명 인체 공생세균 유래소포 기반 진단기술 개발 및 제어 후보물질 발굴  
 기여율 1/2  
 주관기관 주식회사 엠디헬스케어  
 연구기간 2016.05.01 ~ 2020.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호 HI16C1997  
 부처명 보건복지부  
 연구관리전문기관 한국보건산업진흥원  
 연구사업명 포스트게놈다부처유전체사업  
 연구과제명 소변 엑소좀 microRNA 분석 플랫폼 확립  
 기여율 1/2  
 주관기관 주식회사 엠디헬스케어  
 연구기간 2016.08.04 ~ 2019.03.03

---

명세서

청구범위

청구항 1

- (a) 피검체 샘플에서 분리한 세포밖 소포로부터 DNA를 추출하는 단계;
- (b) 상기 추출한 DNA에 대하여 서열번호 1 및 서열번호 2의 프라이머 쌍을 이용하여 PCR을 수행하는 단계; 및
- (c) 상기 PCR 산물의 서열분석을 통하여 정상인 유래 샘플과 세균 유래 세포밖 소포의 함량 증감을 비교하는 단계를 포함하고,

상기 피검체 샘플은 혈액 또는 소변이고,

상기 (c) 단계에서, 피검체 혈액 샘플에서 분리한 써미(Thermi), 푸조박테리아(Fusobacteria), 클로로플렉시(Chloroflexi), 시아노박테리아(Cyanobacteria), 유리고세균(Euryarchaeota), 프로테오박테리아(Proteobacteria), 방선균문(Actinobacteria), 우미균문(Verrucomicrobia), 및 의간균문(Bacteroidetes), 피검체 소변 샘플에서 분리한 테네리쿠테스(Tenericutes)로 이루어진 균으로부터 선택되는 1종 이상의 문(phylum) 세균 유래 세포밖 소포,

피검체 혈액 샘플에서 분리한 사이토파지아(Cytophagia), 데이노코키(Deinococci), 푸조박테리아(Fusobacteriia), 스펅고박테리아(Sphingobacteriia), 플라보박테리아(Flavobacteriia), 알파프로테오박테리아(Alphaproteobacteria), 베타프로테오박테리아(Betaproteobacteria), 간균강(Bacilli), 악티노박테리아(Actinobacteria), 감마프로테오박테리아(Gammaproteobacteria), 클로스트리디아(Clostridia), 우미균강(Verrucomicrobiae), 및 박테로이디아(Bacteroidia), 피검체 소변 샘플에서 분리한 몰리쿠테스(Mollicutes), 코리오박테리아(Coriobacteriia), 델타프로테오박테리아(Deltaproteobacteria), 및 엡실론프로테오박테리아(Epsilonproteobacteria)로 이루어진 균으로부터 선택되는 1종 이상의 강(class) 세균 유래 세포밖 소포,

피검체 혈액 샘플에서 분리한 아에로모나달레스(Aeromonadales), 데이노코카레스(Deinococcales), 사이토파잘레스(Cytophagales), 리조비움목(Rhizobiales), 나이세리아레스(Neisseriales), 오세아노스피릴라레스(Oceanospirillales), 푸조박테리움균목(Fusobacteriales), 스펅고박테리알레스(Sphingobacteriales), 스펅고모나달레스(Sphingomonadales), 슈도모나달레스(Pseudomonadales), 로도피릴라레스(Rhodospirillales), 플라보박테리아레스(Flavobacteriales), 로도사이클러스(Rhodocyclales), 로도박테라레스(Rhodobacterales), 게멜라레스(Gemellales), 카울로박테라레스(Caulobacterales), 악티노마이세탈레스(Actinomycetales), 산토모나다레스(Xanthomonadales), 알테로모나달레스(Alteromonadales), 파스테우렐라레스(Pasteurellales), 바실라레스(Bacillales), 버크홀데리아레스(Burkholderiales), 유산균목(Lactobacillales), 클로스트리디알레스(Clostridiales), 베루코미크로비알레스(Verrucomicrobiales), 및 박테로이테스목(Bacteroidales), 피검체 소변 샘플에서 분리한 스트라메노필레스(Stramenopiles), 슈도모나달레스(Pseudomonadales), 코리오박테리움목(Coriobacteriales), 데설펠로비리온알레스(Desulfovibrionales), 및 캄필로박테라레스(Campylobacterales)로 이루어진 균으로부터 선택되는 1종 이상의 목(order) 세균 유래 세포밖 소포,

피검체 혈액 샘플에서 분리한 아에로모나다시예(Aeromonadaceae), 메틸로박테리아시예(Methylobacteriaceae), 리조비움과(Rhizobiaceae), 브라디리조비아시예(Bradyrhizobiaceae), 할로모나다시예(Halomonadaceae), 사이토파자시예(Cytophagaceae), 나이세리아시예(Neisseriaceae), 푸조박테리아시예(Fusobacteriaceae), 스펅고모나다시예(Sphingomonadaceae), 위크셀라시예(Weeksellaceae), 모락셀라시예(Moraxellaceae), 아에로코카시예(Aerococcaceae), 슈도모나다시예(Pseudomonadaceae), 마이크로코카시예(Micrococcaceae), 프로피오니박테리아시예(Propionibacteriaceae), 인트라스포란지아시예(Intrasporangiaceae), 제멜라시예(Gemellaceae), 플라보박테리아시예(Flavobacteriaceae), 브레비박테리아시예(Brevibacteriaceae), 로도사이클라시예(Rhodocyclaceae), 코리네박테리아시예(Corynebacteriaceae), 버크홀데리아시예(Burkholderiaceae), 로도박테라시예(Rhodobacteraceae), 티시에렐라시예(Tissierellaceae), 카울로박테라시예(Caulobacteraceae), 산토모나다시예(Xanthomonadaceae), 옥살로박테라시예(Oxalobacteraceae), 스태필로코카시예(Staphylococcaceae), 코마모나다시예(Comamonadaceae), 플라노코카시예(Planococcaceae), 파스테우렐라시예(Pasteurellaceae), 액티노마이세타시예(Actinomycetaceae), 엔테로코카시예(Enterococcaceae), 바실라시예(Bacillaceae), 프레보텔라과

(Prevotellaceae), 스트렙토코카시에(Streptococcaceae), 베일로넬라시에(Veillonellaceae), 유산균과(Lactobacillaceae), 루미노코카시에(Ruminococcaceae), 모지박테리아시에(Mogibacteriaceae), 베루코미크로비아시에(Verrucomicrobiaceae), 라치노스피라시에(Lachnospiraceae), 오도리박테라시에(Odoribacteraceae), 박테로이다시에(Bacteroidaceae), 및 바르네시엘라시에(Barnesiellaceae), 피검체 소변 샘플에서 분리한 브라디리조비아시에(Bradyrhizobiaceae), 셀룰로모나다시에(Cellulomonadaceae), 슈도모나다시에(Pseudomonadaceae), 모락셀라시에(Moraxellaceae), 코마모나다시에(Comamonadaceae), 엔테로코카시에(Enterococcaceae), 클로스트리디움과(Clostridiaceae), 코리오박테리움과(Coriobacteriaceae), 리케넬라시에(Rikenellaceae), 데설포비브리오나시에(Desulfovibrionaceae), 및 헬리코박테라시에(Helicobacteraceae)로 이루어진 군으로부터 선택되는 1종 이상의 과(family) 세균 유래 세포막 소포, 또는

피검체 혈액 샘플에서 분리한 할로모나스(Halomonas), 메틸로박테리움(Methylobacterium), 나이세리아(Neisseria), 푸조박테리움(Fusobacterium), 카이스토박터(Kaistobacter), 아그로박테리움(Agrobacterium), 포르피로모나스(Porphyrmonas), 쿠프리아비두스(Cupriavidus), 아시네토박터(Acinetobacter), 슈도모나스(Pseudomonas), 크리세오박테리움(Chryseobacterium), 스펅고모나스(Sphingomonas), 로티아(Rothia), 마이크로코커스(Micrococcus), (Enhydrobacter), 프로피오니박테리움(Propionibacterium), 브레비박테리움(Brevibacterium), 코리네박테리움(Corynebacterium), 라우트로피아(Lautropia), 파라코쿠스(Paracoccus), (Staphylococcus), 헤모필루스(Haemophilus), 카테니박테리움(Catenibacterium), 아나에로코커스(Anaerococcus), 프레보텔라(Prevotella), 액티노마이세스(Actinomyces), 베일로넬라(Veillonella), 시트로박터(Citrobacter), 엔테로코커스(Enterococcus), 스트렙토코커스(Streptococcus), 디알리스터(Dialister), 바실러스(Bacillus), 유산균속(Lactobacillus), 비피도박테리움(Bifidobacterium), 페칼리박테리움(Faecalibacterium), 파라박테로이데스(Parabacteroides), 파라프레보텔라(Paraprevotella), 아커만시아(Akkermansia), 루미노코커스(Ruminococcus), 아들러크레우치아(Adlercreutzia), 부티리시모나스(Butyricimonas), 오도리박터(Odoribacter), 코프로코커스(Coprococcus), 아나로스티페스(Anaerostipes), 블라우티아(Blautia), 박테로이데스(Bacteroides), 및 에풀로피스키움(Epulopiscium), 피검체 소변 샘플에서 분리한 리조비움(Rhizobium), 쿠프리아비두스(Cupriavidus), 아시네토박터(Acinetobacter), 슈도모나스(Pseudomonas), 유산균속(Lactobacillus), 시트로박터(Citrobacter), 엔테로코커스(Enterococcus), 파라코쿠스(Paracoccus), 클렙시엘라(Klebsiella), 알로바쿨럼(Allobaculum), 데설포비브리오(Desulfovibrio), 및 플렉시스피라(Flexispira)로 이루어진 군으로부터 선택되는 1종 이상의 속(genus) 세균 유래 세포막 소포의 함량을 비교하는 것을 특징으로 하는, 당뇨병 진단을 위한 정보제공방법.

**청구항 2**

제1항에 있어서,

상기 혈액은 전혈, 혈청, 혈장, 또는 혈액 단핵구인 것을 특징으로 하는, 정보제공방법.

**발명의 설명**

**기술 분야**

[0001] 본 발명은 세균 메타게놈 분석을 통해 당뇨병을 진단하는 방법에 관한 것으로서, 보다 구체적으로는 피검체 유래 샘플을 이용해 세균 메타게놈 분석을 수행하여 특정 세균 유래 세포막 소포의 함량 증감을 분석함으로써 당뇨병의 원인인자, 발병 위험도, 및 질병경과를 진단하는 방법에 관한 것이다.

**배경 기술**

[0003] 21세기 고령화 사회에서 식이습관 등의 생활환경의 변화로 최근 50년 사이에 질병패턴에 커다란 변화가 생겼다. 특히, 과거 50년 전까지만 하더라도 영양결핍이 문제가 되었지만, 최근 경제발전으로 인해 먹거리가 풍성해지면서 포화지방의 과섭취로 인한 당뇨, 비만 등의 대사질환의 유병률이 급증하고 있다. 당뇨병의 발생과 관련해서 오랫동안 대사의 문제로 접근했지만, 최근 고지방식이에 따른 염증반응이 당뇨, 비만의 발생과 관련이 있다는 연

구결과가 주목을 받고 있다. 대사증후군 (metabolic syndrome)은 심혈관질환의 주요 위험인자로서 인구의 20-30% 유병률을 보이고, 인슐린저항성에 의한 제2형 당뇨병, 고혈압, 콜레스테롤 이상, 지혈 이상 등의 특징을 보인다. 대사증후군의 가장 중요한 병태생리는 간, 근육, 지방조직과 같이 인슐린이 작용하는 장기에 발생하는 인슐린저항성 (insulin resistance)을 특징으로 하는 당뇨병이다. 당뇨병의 진단은 혈액에서 포도당 수치를 측정함으로써 이루어지는데, 대부분 당뇨병은 질환이 이미 진행된 경우에 치료가 상당히 어려운 바, 당뇨병의 발생 및 원인인자를 미리 예측하여, 고위험군에서 당뇨병 발생을 예방하는 방법을 제공하는 것이 효율적인 방법이다.

[0004] 한편, 인체에 공생하는 미생물은 100조에 이르러 인간 세포보다 10배 많으며, 미생물의 유전자수는 인간 유전자수의 100배가 넘는 것으로 알려지고 있다. 미생물총(microbiota 혹은 microbiome)은 주어진 거주지에 존재하는 세균(bacteria), 고세균(archaea), 진핵생물(eukarya)을 포함한 미생물 군집(microbial community)을 말하고, 장내 미생물총은 사람의 생리현상에 중요한 역할을 하며, 인체 세포와 상호작용을 통해 인간의 건강과 질병에 큰 영향을 미치는 것으로 알려져 있다. 우리 몸에 공생하는 세균은 다른 세포로의 유전자, 단백질 등의 정보를 교환하기 위하여 나노미터 크기의 소포(vesicle)를 분비한다. 점막은 200 나노미터(nm) 크기 이상의 입자는 통과할 수 없는 물리적인 방어막을 형성하여 점막에 공생하는 세균인 경우에는 점막을 통과하지 못하지만, 세균 유래 소포는 크기가 대개 100 나노미터 크기 이하라서 비교적 자유롭게 점막을 통과하여 우리 몸에 흡수된다.

[0005] 환경 유전체학이라고도 불리는 메타게놈학은 환경에서 채취한 샘플에서 얻은 메타게놈 자료에 대한 분석학이라고 할 수 있다(국내공개특허 제2011-0073049호). 최근 16s 리보솜 RNA(16s rRNA) 염기서열을 기반으로 한 방법으로 인간의 미생물총의 세균 구성을 목록화하는 것이 가능해졌으며, 16s 리보솜 RNA의 유전자인 16s rDNA 염기서열을 차세대 염기서열분석 (next generation sequencing, NGS) platform을 이용하여 분석한다. 그러나 당뇨병 발병에 있어서, 혈액 또는 소변 등의 인체 유래물에서 세균 유래 소포에 존재하는 유전자 메타게놈 분석을 통해 당뇨병의 원인인자를 동정하고 당뇨병을 예측하는 방법에 대해서는 보고된 바가 없다.

### 발명의 내용

#### 해결하려는 과제

[0007] 본 발명자들은 당뇨병의 원인인자, 발병 위험도, 및 질병경과를 진단하기 위하여, 피검체 유래 샘플인 혈액 및 소변을 이용해 세균 유래 세포밖 소포로부터 유전자를 추출하고 이에 대하여 메타게놈 분석을 수행하였으며, 그 결과 당뇨병의 원인인자로 작용할 수 있는 세균 유래 세포밖 소포를 동정하였는바, 이에 기초하여 본 발명을 완성하였다.

[0008] 이에, 본 발명은 세균 유래 세포밖 소포에 대한 메타게놈 분석을 통해 당뇨병을 진단하기 위한 정보제공방법을 제공하는 것을 목적으로 한다.

[0009] 그러나 본 발명이 이루고자 하는 기술적 과제는 이상에서 언급한 과제에 제한되지 않으며, 언급되지 않은 또 다른 과제들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

#### 과제의 해결 수단

[0011] 상기와 같은 본 발명의 목적을 달성하기 위하여, 본 발명은 하기의 단계를 포함하는, 당뇨병 진단을 위한 정보 제공방법을 제공한다 :

[0012] (a) 피검체 샘플에서 분리한 세포밖 소포로부터 DNA를 추출하는 단계;

[0013] (b) 상기 추출한 DNA에 대하여 서열번호 1 및 서열번호 2의 프라이머 쌍을 이용하여 PCR을 수행하는 단계; 및

[0014] (c) 상기 PCR 산물의 서열분석을 통하여 정상인 유래 샘플과 세균 유래 세포밖 소포의 함량 증감을 비교하는 단계.

[0015] 그리고, 본 발명은 하기의 단계를 포함하는, 당뇨병 진단방법을 제공한다 :

[0016] (a) 피검체 샘플에서 분리한 세포밖 소포로부터 DNA를 추출하는 단계;

[0017] (b) 상기 추출한 DNA에 대하여 서열번호 1 및 서열번호 2의 프라이머 쌍을 이용하여 PCR을 수행하는 단계; 및

- [0018] (c) 상기 PCR 산물의 서열분석을 통하여 정상인 유래 샘플과 세균 유래 세포막 소포의 함량 증감을 비교하는 단계.
- [0019] 또한, 본 발명은 하기의 단계를 포함하는, 당뇨병의 발병 위험도 예측방법을 제공한다 :
- [0020] (a) 피검체 샘플에서 분리한 세포막 소포로부터 DNA를 추출하는 단계;
- [0021] (b) 상기 추출한 DNA에 대하여 서열번호 1 및 서열번호 2의 프라이머 쌍을 이용하여 PCR을 수행하는 단계; 및
- [0022] (c) 상기 PCR 산물의 서열분석을 통하여 정상인 유래 샘플과 세균 유래 세포막 소포의 함량 증감을 비교하는 단계.
- [0023] 본 발명의 일구현예로, 상기 피검체 샘플은 혈액 또는 소변이고,
- [0024] 상기 (c) 단계에서, 피검체 혈액 샘플에서 분리한 써미(Thermi), 푸조박테리아(Fusobacteria), 클로로플렉시(Chloroflexi), 시아노박테리아(Cyanobacteria), TM7, 유리고세균(Euryarchaeota), 프로테오박테리아(Proteobacteria), 방선균문(Actinobacteria), 우미균문(Verrucomicrobia), 및 의간균문(Bacteroidetes), 피검체 소변 샘플에서 분리한 테네리쿠테스(Tenericutes)로 이루어진 군으로부터 선택되는 1종 이상의 문(phylum) 세균 유래 세포막 소포,
- [0025] 피검체 혈액 샘플에서 분리한 사이토파지아(Cytophagia), 데이노코키(Deinococci), 푸조박테리아(Fusobacteriia), 스펡고박테리아(Sphingobacteriia), 플라보박테리아(Flavobacteriia), 알파프로테오박테리아(Alphaproteobacteria), 베타프로테오박테리아(Betaproteobacteria), TM7-3, 간균강(Bacilli), 악티노박테리아(Actinobacteria), 감마프로테오박테리아(Gammaproteobacteria), 클로스트리디아(Clostridia), 우미균강(Verrucomicrobiae), 및 박테로이디아(Bacteroidia), 피검체 소변 샘플에서 분리한 몰리쿠테스(Mollicutes), 코리오박테리아(Coriobacteriia), 델타프로테오박테리아(Deltaproteobacteria), 및 입실론프로테오박테리아(Epsilonproteobacteria)로 이루어진 군으로부터 선택되는 1종 이상의 강(class) 세균 유래 세포막 소포,
- [0026] 피검체 혈액 샘플에서 분리한 아에로모나달레스(Aeromonadales), 데이노코카레스(Deinococcales), 사이토파잘레스(Cytophagales), 리조비움목(Rhizobiales), 나이세리아레스(Neisseriales), 오세아노스피릴라레스(Oceanospirillales), 푸조박테리움균목(Fusobacteriales), 스펡고박테리알레스(Sphingobacteriales), 스펡고모나달레스(Sphingomonadales), 슈도모나달레스(Pseudomonadales), 로도피릴라레스(Rhodospirillales), 플라보박테리아레스(Flavobacteriales), 로도사이클러스(Rhodocyclales), 로도박테라레스(Rhodobacterales), 게멜라레스(Gemellales), 카울로박테라레스(Caulobacterales), 악티노마이세탈레스(Actinomycetales), 산토모나다레스(Xanthomonadales), 알테로모나달레스(Alteromonadales), 파스테우렐라레스(Pasteurellales), 바실라레스(Bacillales), 버크홀데리알레스(Burkholderiales), 유산균목(Lactobacillales), 클로스트리디알레스(Clostridiales), RF32, 베루코미크로비알레스(Verrucomicrobiales), 및 박테로이데스목(Bacteroidales), 피검체 소변 샘플에서 분리한 스트라메노필레스(Stramenopiles), 슈도모나달레스(Pseudomonadales), 코리오박테리움목(Coriobacteriales), 데설펠로브리오날레스(Desulfovibrionales), 및 캄필로박테라레스(Campylobacterales)로 이루어진 군으로부터 선택되는 1종 이상의 목(order) 세균 유래 세포막 소포,
- [0027] 피검체 혈액 샘플에서 분리한 아에로모나다시예(Aeromonadaceae), 메틸로박테리아시예(Methylobacteriaceae), 리조비움과(Rhizobiaceae), 브라디리조비아시예(Bradyrhizobiaceae), 할로모나다시예(Halomonadaceae), 사이토파자시예(Cytophagaceae), 나이세리아시예(Neisseriaceae), 푸조박테리아시예(Fusobacteriaceae), 스펡고모나다시예(Sphingomonadaceae), 위크셀라시예(Weeksellaceae), 모락셀라시예(Moraxellaceae), 아에로코카시예(Aerococcaceae), 슈도모나다시예(Pseudomonadaceae), 마이크로코카시예(Micrococcaceae), 프로피오니박테리아시예(Propionibacteriaceae), 인트라스포란지아시예(Intrasporangiaceae), 제멜라시예(Gemellaceae), 플라보박테리아시예(Flavobacteriaceae), 브레비박테리아시예(Brevibacteriaceae), 로도사이클라시예(Rhodocyclaceae), 코리네박테리아시예(Corynebacteriaceae), 버크홀데리아시예(Burkholderiaceae), 로도박테라시예(Rhodobacteraceae), 티시에렐라시예(Tissierellaceae), 카울로박테라시예(Caulobacteraceae), 산토모나다시예(Xanthomonadaceae), 옥살로박테라시예(Oxalobacteraceae), 스태필로코카시예(Staphylococcaceae), 코마모나다시예(Comamonadaceae), 플라노코카시예(Planococcaceae), 파스테우렐라시예(Pasteurellaceae), 액티노마이세타시예(Actinomycetaceae), S24-7, 엔테로코카시예(Enterococcaceae), 바실라시예(Bacillaceae), 프레보텔라과(Prevotellaceae), 스트렙토코카시예(Streptococcaceae), 베일로넬라시예(Veillonellaceae), 유산균과(Lactobacillaceae), 루미노코카시예(Ruminococcaceae), 모지박테리아시예(Mogibacteriaceae), 베루코미크로비아시예(Verrucomicrobiaceae), 라치노스피라시예(Lachnospiraceae), 오도리박테라시예(Odoribacteraceae), 박

테로이다시에(Bacteroidaceae), 및 바르네시엘라시에(Barnesiellaceae), 피검체 소변 샘플에서 분리한 브라디리조비아시에(Bradyrhizobiaceae), 셀룰로모나다시에(Cellulomonadaceae), 슈도모나다시에(Pseudomonadaceae), 모락셀라시에(Moraxellaceae), 코마모나다시에(Comamonadaceae), 엔테로코카시에(Enterococcaceae), 클로스트리디움과(Clostridiaceae), 코리오박테리움과(Coriobacteriaceae), 리케넬라시에(Rikenellaceae), 데설펜브리오나시에(Desulfovibrionaceae), 및 헬리코박테라시에(Helicobacteraceae)로 이루어진 군으로부터 선택되는 1종 이상의 과(family) 세균 유래 세포막 소포, 또는

[0028] 피검체 혈액 샘플에서 분리한 할로모나스(Halomonas), 메틸로박테리움(Methylobacterium), 나이세리아(Neisseria), 푸조박테리움(Fusobacterium), 카이스토박터(Kaistobacter), 아그로박테리움(Agrobacterium), 포르피로모나스(Porphyrromonas), 쿠프리아비두스(Cupriavidus), 아시네토박터(Acinetobacter), 슈도모나스(Pseudomonas), 크리세오박테리움(Chryseobacterium), 스펅고모나스(Sphingomonas), 로티아(Rothia), 마이크로코커스(Micrococcus), (Enhydrobacter), 프로피오니박테리움(Propionibacterium), 브레비박테리움(Brevibacterium), 코리네박테리움(Corynebacterium), 라우트로피아(Lautropia), 파라코쿠스(Paracoccus), (Staphylococcus), 헤모필루스(Haemophilus), 카테니박테리움(Catenibacterium), 아나에로코커스(Anaerococcus), 프레보텔라(Prevotella), 액티노마이세스(Actinomyces), 베일로넬라(Veillonella), 시트로박터(Citrobacter), 엔테로코커스(Enterococcus), 스트렙토코커스(Streptococcus), 디알리스터(Dialister), 바실러스(Bacillus), 유산균속(Lactobacillus), 비피도박테리움(Bifidobacterium), 페칼리박테리움(Faecalibacterium), 파라박테로이데스(Parabacteroides), 파라프레보텔라(Paraprevotella), 아커만시아(Akkermansia), 루미노코커스(Ruminococcus), 아들리크레우치아(Adlercreutzia), 부티리시모나스(Butyricimonas), 오도리박터(Odoribacter), 코프로코커스(Coprococcus), 아나로스티페스(Anaerostipes), 블라우티아(Blautia), 박테로이데스(Bacteroides), 및 에풀로피스키움(Epulpiscium), 피검체 소변 샘플에서 분리한 리조비움(Rhizobium), 쿠프리아비두스(Cupriavidus), 아시네토박터(Acinetobacter), 슈도모나스(Pseudomonas), 유산균속(Lactobacillus), 시트로박터(Citrobacter), 엔테로코커스(Enterococcus), 파라코쿠스(Paracoccus), 클렙시엘라(Klebsiella), SMB53, 알로바쿨럼(Allobaculum), 데설펜브리오(Desulfovibrio), AF12, 및 플렉시스피라(Flexispira)로 이루어진 군으로부터 선택되는 1종 이상의 속(genus) 세균 유래 세포막 소포의 함량 증감을 비교하는 단계를 포함할 수 있다.

[0029] 본 발명의 다른 구현예로, 상기 혈액은 전혈, 혈청, 혈장, 또는 혈액 단핵구일 수 있다.

**발명의 효과**

[0031] 환경에 존재하는 세균에서 분비되는 세포막 소포는 체내에 흡수되어 당뇨병 발생에 직접적인 영향을 미칠 수 있으며, 당뇨병은 증상이 나타나기 전 발생을 예측하는 것이 어려워 효율적인 치료가 어려운 실정이므로, 본 발명에 따른 인체 유래 샘플을 이용한 세균 유래 세포막 소포의 메타게놈 분석을 통해 당뇨병의 원인인자 및 발병의 위험도를 미리 예측함으로써 당뇨병의 위험군을 조기에 진단 및 예측하여 적절한 관리를 통해 발병 시기를 늦추거나 발병을 예방할 수 있으며, 발병 후에도 원인인자를 진단 할 수 있어 당뇨병의 경과를 좋게 하여 치료 효과를 높일 수 있다.

**도면의 간단한 설명**

[0033] 도 1a은, 마우스에 장내 세균과 세균유래 소포 (EV)를 구강으로 투여한 후, 시간별로 세균과 소포의 분포양상을 촬영한 사진이고, 도 1b는 구강으로 투여한 후 12시간째에, 혈액, 간, 및 여러 장기를 적출하여, 세균과 소포의 체내 분포양상을 평가한 그림이다.

도 2는 당뇨병환자 및 정상인 혈액에서 세균 유래 소포를 분리한 후, 메타게놈 분석을 수행하여 문(phylum) 수준에서 진단적 성능이 유의한 세균 유래 소포(EVs)의 분포를 나타낸 결과이다.

도 3은 당뇨병환자 및 정상인 혈액에서 세균 유래 소포를 분리한 후, 메타게놈 분석을 수행하여 강(class) 수준에서 진단적 성능이 유의한 세균 유래 소포(EVs)의 분포를 나타낸 결과이다.

도 4는 당뇨병환자 및 정상인 혈액에서 세균 유래 소포를 분리한 후, 메타게놈 분석을 수행하여 목(order) 수준에서 진단적 성능이 유의한 세균 유래 소포(EVs)의 분포를 나타낸 결과이다.

도 5는 당뇨병환자 및 정상인 혈액에서 세균 유래 소포를 분리한 후, 메타게놈 분석을 수행하여 과(family) 수준에서 진단적 성능이 유의한 세균 유래 소포(EVs)의 분포를 나타낸 결과이다.

도 6은 당뇨병환자 및 정상인 혈액에서 세균 유래 소포를 분리한 후, 메타게놈 분석을 수행하여 속(genus) 수준에서 진단적 성능이 유의한 세균 유래 소포(EVs)의 분포를 나타낸 결과이다.

도 7은 당뇨병환자 및 정상인 소변에서 세균 유래 소포를 분리한 후, 메타게놈 분석을 수행하여 문(phylum) 수준에서 진단적 성능이 유의한 세균 유래 소포(EVs)의 분포를 나타낸 결과이다.

도 8은 당뇨병환자 및 정상인 소변에서 세균 유래 소포를 분리한 후, 메타게놈 분석을 수행하여 강(class) 수준에서 진단적 성능이 유의한 세균 유래 소포(EVs)의 분포를 나타낸 결과이다.

도 9는 당뇨병환자 및 정상인 소변에서 세균 유래 소포를 분리한 후, 메타게놈 분석을 수행하여 목(order) 수준에서 진단적 성능이 유의한 세균 유래 소포(EVs)의 분포를 나타낸 결과이다.

도 10은 당뇨병환자 및 정상인 소변에서 세균 유래 소포를 분리한 후, 메타게놈 분석을 수행하여 과(family) 수준에서 진단적 성능이 유의한 세균 유래 소포(EVs)의 분포를 나타낸 결과이다.

도 11은 당뇨병환자 및 정상인 소변에서 세균 유래 소포를 분리한 후, 메타게놈 분석을 수행하여 속(genus) 수준에서 진단적 성능이 유의한 세균 유래 소포(EVs)의 분포를 나타낸 결과이다.

**발명을 실시하기 위한 구체적인 내용**

- [0034] 본 발명은 세균 메타게놈 분석을 통해 당뇨병을 진단하는 방법에 관한 것으로서, 본 발명자들은 피검체 유래 샘플을 이용해 세균 유래 세포밖 소포로부터 유전자를 추출하고 이에 대하여 메타게놈 분석을 수행하였으며, 당뇨병의 원인인자로 작용할 수 있는 세균 유래 세포밖 소포를 동정하였다.
- [0035] 이에, 본 발명은 (a) 피검체 샘플에서 분리한 세포밖 소포로부터 DNA를 추출하는 단계;
- [0036] (b) 상기 추출한 DNA에 대하여 서열번호 1 및 서열번호 2의 프라이머 쌍을 이용하여 PCR을 수행하는 단계; 및
- [0037] (c) 상기 PCR 산물의 서열분석을 통하여 정상인 유래 샘플과 세균 유래 세포밖 소포의 함량 증감을 비교하는 단계를 포함하는 당뇨병 진단을 위한 정보제공방법을 제공한다.
- [0038] 본 발명에서 사용되는 용어, "당뇨병 진단" 이란 환자에 대하여 당뇨병이 발병할 가능성이 있는지, 당뇨병이 발병할 가능성이 상대적으로 높은지, 당뇨병의 원인인자가 무엇인지, 또는 당뇨병이 이미 발병하였는지 여부를 판별하는 것을 의미한다. 본 발명의 방법은 임의의 특정 환자에 대한 당뇨병 발병 위험도가 높은 환자로써 특별하고 적절한 관리를 통하여 발병 시기를 늦추거나 발병하지 않도록 하는데 사용할 수 있다. 또한, 본 발명의 방법은 당뇨병을 조기에 진단하여 가장 적절한 치료방식을 선택함으로써 치료를 결정하기 위해 임상적으로 사용될 수 있다.
- [0039] 본 발명에서 사용되는 용어, "메타게놈(metagenome)" 이란 "군유전체" 라고도 하며, 흙, 동물의 장 등 고립된 지역 내의 모든 바이러스, 세균, 곰팡이 등을 포함하는 유전체의 총합을 의미하는 것으로, 주로 배양이 되지 않는 미생물을 분석하기 위해서 서열분석기를 사용하여 한꺼번에 많은 미생물을 동정하는 것을 설명하는 유전체의 개념으로 쓰인다. 특히, 메타게놈은 한 종의 게놈 또는 유전체를 말하는 것이 아니라, 한 환경단위의 모든 종의 유전체로서 일종의 혼합유전체를 말한다. 이는 오믹스적으로 생물학이 발전하는 과정에서 한 종을 정의할 때 기능적으로 기존의 한 종뿐만 아니라, 다양한 종이 서로 상호작용하여 완전한 종을 만든다는 관점에서 나온 용어이다. 기술적으로는 빠른 서열분석법을 이용해서, 종에 관계없이 모든 DNA, RNA를 분석하여, 한 환경 내에서의 모든 종을 동정하고, 상호작용, 대사작용을 규명하는 기법의 대상이다. 본 발명에서는 바람직하게 혈액에서 분리한 세균 유래 세포밖 소포를 이용하여 메타게놈 분석을 실시하였다.
- [0040] 본 발명에서 사용되는 용어, "세균 유래 소포" 란 세균 및 고세균이 분비하는 막으로 형성된 나노크기의 물질로서, 소포에 세균에서 유래하는 유전자를 갖고 있는 물질을 총칭한다.
- [0041] 본 발명에 있어서, 상기 피검체 샘플은 혈액, 또는 소변일 수 있고, 상기 혈액은 바람직하게 전혈, 혈청, 혈장, 또는 혈액 단핵구일 수 있으나, 이것으로 제한되는 것은 아니다.
- [0042] 본 발명의 실시예에서는 상기 세균 유래 세포밖 소포에 대한 메타게놈 분석을 실시하였으며, 문(phylum), 강(class), 목(order), 과(family), 및 속(genus) 수준에서 각각 분석하여 실제로 당뇨병 발생의 원인으로 작용할

수 있는 세균 유래 소포를 동정하였다.

- [0043] 보다 구체적으로 본 발명의 일실시예에서는, 피검자 유래 혈액 샘플에 존재하는 소포에 대하여 세균 메타게놈을 문 수준에서 분석한 결과, *Thermi*, *Fusobacteria*, *Chloroflexi*, *Cyanobacteria*, *TM7*, *Euryarchaeota*, *Proteobacteria*, *Actinobacteria*, *Verrucomicrobia*, 및 *Bacteroidetes* 문 세균 유래 세포밖 소포의 함량이 당뇨병환자와 정상인에 사이에 유의한 차이가 있었다(실시예 4 참조).
- [0044] 보다 구체적으로 본 발명의 일실시예에서는, 피검자 유래 혈액 샘플에 존재하는 소포에 대하여 세균 메타게놈을 강 수준에서 분석한 결과, *Cytophagia*, *Deinococci*, *Fusobacteriia*, *Sphingobacteriia*, *Flavobacteriia*, *Alphaproteobacteria*, *Betaproteobacteria*, *TM7-3*, *Bacilli*, *Actinobacteria*, *Gammaproteobacteria*, *Clostridia*, *Verrucomicrobiae*, 및 *Bacteroidia* 강 세균 유래 세포밖 소포의 함량이 당뇨병환자와 정상인에 사이에 유의한 차이가 있었다(실시예 4 참조).
- [0045] 보다 구체적으로 본 발명의 일실시예에서는, 피검자 유래 혈액 샘플에 존재하는 소포에 대하여 세균 메타게놈을 목 수준에서 분석한 결과, *Aeromonadales*, *Deinococcales*, *Cytophagales*, *Rhizobiales*, *Neisseriales*, *Oceanospirillales*, *Fusobacteriales*, *Sphingobacteriales*, *Sphingomonadales*, *Pseudomonadales*, *Rhodospirillales*, *Flavobacteriales*, *Rhodocyclales*, *Rhodobacterales*, *Gemellales*, *Caulobacterales*, *Actinomycetales*, *Xanthomonadales*, *Alteromonadales*, *Pasteurellales*, *Bacillales*, *Burkholderiales*, *Lactobacillales*, *Clostridiales*, *RF32*, *Verrucomicrobiales*, 및 *Bacteroidales* 목 세균 유래 세포밖 소포의 함량이 당뇨병환자와 정상인에 사이에 유의한 차이가 있었다(실시예 4 참조).
- [0046] 보다 구체적으로 본 발명의 일실시예에서는, 피검자 유래 혈액 샘플에 존재하는 소포에 대하여 세균 메타게놈을 과 수준에서 분석한 결과, *Aeromonadaceae*, *Methylobacteriaceae*, *Rhizobiaceae*, *Bradyrhizobiaceae*, *Halomonadaceae*, *Cytophagaceae*, *Neisseriaceae*, *Fusobacteriaceae*, *Sphingomonadaceae*, *Weeksellaceae*, *Moraxellaceae*, *Aerococcaceae*, *Pseudomonadaceae*, *Micrococcaceae*, *Propionibacteriaceae*, *Intrasporangiaceae*, *Gemellaceae*, *Flavobacteriaceae*, *Brevibacteriaceae*, *Rhodocyclaceae*, *Corynebacteriaceae*, *Burkholderiaceae*, *Rhodobacteraceae*, *Tissierellaceae*, *Caulobacteraceae*, *Xanthomonadaceae*, *Oxalobacteraceae*, *Staphylococcaceae*, *Comamonadaceae*, *Planococcaceae*, *Pasteurellaceae*, *Actinomycetaceae*, *S24-7*, *Enterococcaceae*, *Bacillaceae*, *Prevotellaceae*, *Streptococcaceae*, *Veillonellaceae*, *Lactobacillaceae*, *Ruminococcaceae*, *Mogibacteriaceae*, *Verrucomicrobiaceae*, *Lachnospiraceae*, *Odoribacteraceae*, *Bacteroidaceae*, 및 *Barnesiellaceae* 과 세균 유래 세포밖 소포의 함량이 당뇨병환자와 정상인에 사이에 유의한 차이가 있었다(실시예 4 참조).
- [0047] 보다 구체적으로 본 발명의 일실시예에서는, 피검자 유래 혈액 샘플에 존재하는 소포에 대하여 세균 메타게놈을 속 수준에서 분석한 결과, *Halomonas*, *Methylobacterium*, *Neisseria*, *Fusobacterium*, *Kaistobacter*, *Agrobacterium*, *Porphyromonas*, *Cupriavidus*, *Acinetobacter*, *Pseudomonas*, *Chryseobacterium*, *Sphingomonas*, *Rothia*, *Micrococcus*, *Enhydrobacter*, *Propionibacterium*, *Brevibacterium*, *Corynebacterium*, *Lautropia*, *Paracoccus*, *Staphylococcus*, *Haemophilus*, *Catenibacterium*, *Anaerococcus*, *Prevotella*, *Actinomyces*, *Veillonella*, *Citrobacter*, *Enterococcus*, *Prevotella*, *Streptococcus*, *Dialister*, *Bacillus*, *Lactobacillus*, *Bifidobacterium*, *Faecalibacterium*, *Parabacteroides*, *Paraprevotella*, *Akkermansia*, *Ruminococcus*, *Adlercreutzia*, *Butyricimonas*, *Odoribacter*, *Coprococcus*, *Anaerostipes*, *Blautia*, *Bacteroides*, 및 *Epulopiscium* 속 세균 유래 세포밖 소포의 함량이 당뇨병환자와 정상인에 사이에 유의한 차이가 있었다(실시예 4 참조).
- [0048] 보다 구체적으로 본 발명의 일실시예에서는, 피검자 유래 소변 샘플에 존재하는 소포에 대하여 세균 메타게놈을 문 수준에서 분석한 결과, *Tenericutes* 문 세균 유래 세포밖 소포의 함량이 당뇨병환자와 정상인에 사이에 유의한 차이가 있었다(실시예 5 참조).
- [0049] 보다 구체적으로 본 발명의 일실시예에서는, 피검자 유래 소변 샘플에 존재하는 소포에 대하여 세균 메타게놈을 강 수준에서 분석한 결과, *Mollicutes*, *Coriobacteriia*, *Deltaproteobacteria*, 및 *Epsilonproteobacteria* 강 세균 유래 세포밖 소포의 함량이 당뇨병환자와 정상인에 사이에 유의한 차이가 있었다(실시예 5 참조).
- [0050] 보다 구체적으로 본 발명의 일실시예에서는, 피검자 유래 소변 샘플에 존재하는 소포에 대하여 세균 메타게놈을 목 수준에서 분석한 결과, *Verrucomicrobia*, 및 *Cyanobacteria* 목 세균 유래 세포밖 소포의 함량이 당뇨병환자와 정상인에 사이에 유의한 차이가 있었다(실시예 5 참조).

- [0051] 보다 구체적으로 본 발명의 실시예에서는, 피검자 유래 소변 샘플에 존재하는 소포에 대하여 세균 메타게놈을 과 수준에서 분석한 결과, Verrucomicrobia, 및 Cyanobacteria 과 세균 유래 세포밖 소포의 함량이 당뇨병환자와 정상인에 사이에 유의한 차이가 있었다(실시예 5 참조).
- [0052] 보다 구체적으로 본 발명의 실시예에서는, 피검자 유래 소변 샘플에 존재하는 소포에 대하여 세균 메타게놈을 속 수준에서 분석한 결과, Verrucomicrobia, 및 Cyanobacteria 속 세균 유래 세포밖 소포의 함량이 당뇨병환자와 정상인에 사이에 유의한 차이가 있었다(실시예 5 참조).
- [0053] 본 발명은 상기와 같은 실시예 결과를 통해, 혈액 및 소변으로부터 분리한 세균 유래 세포밖 소포에 대하여 메타게놈 분석을 실시함으로써 정상인과 비교하여 당뇨병환자에서 함량이 유의하게 변화한 세균 유래 소포들을 동정하였으며, 메타게놈 분석을 통해 상기 각 수준에서 세균 유래 소포들의 함량 증감을 분석함으로써 당뇨병을 진단할 수 있음을 확인하였다.
- [0054] 이하, 본 발명의 이해를 돕기 위하여 바람직한 실시예를 제시한다. 그러나 하기의 실시예는 본 발명을 보다 쉽게 이해하기 위하여 제공되는 것일 뿐, 하기 실시예에 의해 본 발명의 내용이 한정되는 것은 아니다.
- [0056] **[실시예]**
- [0057] **실시예 1. 장내 세균 및 세균 유래 소포의 체내 흡수, 분포, 및 배설 양상 분석**
- [0058] 장내 세균과 세균 유래 소포가 위장관을 통해 전신적으로 흡수되는 지를 평가하기 위하여 다음과 같은 방법으로 실험을 수행하였다. 마우스의 위장에 형광으로 표지한 장내세균과 장내 세균 유래 소포를 각각 50  $\mu$ g의 용량으로 위장관으로 투여하고 0분, 5분, 3시간, 6시간, 12시간 후에 형광을 측정하였다. 마우스 전체 이미지를 관찰한 결과, 도 1a에 나타난 바와 같이, 상기 세균(Bacteria)인 경우에는 전신적으로 흡수되지 않았지만, 세균 유래 소포(EV)인 경우에는, 투여 후 5분에 전신적으로 흡수되었고, 투여 3시간 후에는 방광에 형광이 진하게 관찰되어, 소포가 비뇨기계로 배설됨을 알 수 있었다. 또한, 소포는 투여 12시간까지 체내에 존재함을 알 수 있었다.
- [0059] 장내세균과 장내 세균유래 소포가 전신적으로 흡수된 후, 여러 장기로 침윤된 양상을 평가하기 위하여, 형광으로 표지한 50  $\mu$ g의 세균과 세균유래 소포를 상기의 방법과 같이 투여한 다음 12시간째에 마우스로부터 혈액(Blood), 심장(Heart), 폐(Lung), 간(Liver), 신장(Kidney), 비장(Spleen), 지방조직(Adipose tissue), 및 근육(Muscle)을 적출하였다. 상기 적출한 조직들에서 형광을 관찰한 결과, 도 1b에 나타난 바와 같이, 상기 장내 세균(Bacteria)은 각 장기에 흡수되지 않은 반면, 상기 장내 세균 유래 세포밖 소포(EV)는 혈액, 심장, 폐, 간, 신장, 비장, 지방조직, 및 근육에 분포하는 것을 확인하였다.
- [0061] **실시예 2. 혈액 및 소변으로부터 소포 분리 및 DNA 추출**
- [0062] 혈액 및 소변으로부터 소포를 분리하고 DNA를 추출하기 위해, 먼저 10 ml 튜브에 혈액 또는 소변을 넣고 원심분리(3,500 x g, 10min, 4 $^{\circ}$ C)를 실시하여 부유물을 가라앉혀 상등액을 회수한 후 새로운 10 ml 튜브에 옮겼다. 0.22  $\mu$ m 필터를 사용하여 상기 회수한 상등액으로부터 세균 및 이물질을 제거한 후, 센트리프렐 튜브(centrifuge filters 50 kD)에 옮기고 1500 x g, 4 $^{\circ}$ C에서 15분간 원심분리하여 50 kD 보다 작은 물질은 버리고 10 ml까지 농축 시켰다. 다시 한 번 0.22  $\mu$ m 필터를 사용하여 박테리아 및 이물질을 제거한 후, Type 90ti 로터로 150,000 x g, 4 $^{\circ}$ C에서 3시간 동안 초고속원심분리방법을 사용하여 상등액을 버리고 덩어리진 pellet을 생리식염수(PBS)로 녹여 소포를 수득하였다.
- [0063] 상기 방법에 따라 혈액 및 소변으로부터 분리한 소포 100  $\mu$ l를 100 $^{\circ}$ C에서 끓여서 내부의 DNA를 지질 밖으로 나오게 한 후 얼음에 5분 동안 식혔다. 다음으로 남은 부유물을 제거하기 위하여 10,000 x g, 4 $^{\circ}$ C에서 30분간 원심분리하고 상등액 만을 모은 후 Nanodrop을 이용하여 DNA 양을 정량하였다. 이후 상기 추출된 DNA에 세균 유래 DNA가 존재하는지 확인하기 위하여 하기 표 1에 나타난 16s rDNA primer로 PCR을 수행하여 상기 추출된 유전자에 세균 유래 유전자가 존재하는 것을 확인하였다.

표 1

pr imer		서열	서열 번호
16S rDNA	16S_V3_F	5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG-3'	1
	16S_V4_R	5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC-3	2

**실시예 3. 혈액 및 소변에서 추출한 DNA를 이용한 메타게놈 분석**

상기 실시예 2의 방법으로 유전자를 추출한 후, 상기 1에 나타낸 16S rDNA 프라이머를 사용하여 PCR을 실시하여 유전자를 증폭시키고 시퀀싱(Illumina MiSeq sequencer)을 수행하였다. 결과를 Standard Flowgram Format(SFF) 파일로 출력하고 GS FLX software(v2.9)를 이용하여 SFF 파일을 sequence 파일(.fasta)과 nucleotide quality score 파일로 변환한 다음 리드의 신용도 평가를 확인하고, window(20 bps) 평균 base call accuracy가 99% 미만(Phred score <20)인 부분을 제거하였다. 질이 낮은 부분을 제거한 후, 리드의 길이가 300 bps 이상인 것만 이용하였으며(Sickle version 1.33), 결과 분석을 위해 Operational Taxonomy Unit(OTU)은 UCLUST와 USEARCH를 이용하여 시퀀스 유사도에 따라 클러스터링을 수행하였다. 구체적으로 속(genus)은 94%, 과(family)는 90%, 목(order)은 85%, 강(class)은 80%, 문(phylum)은 75% 시퀀스 유사도를 기준으로 클러스터링을 하고 각 OTU의 문, 강, 목, 과, 속 레벨의 분류를 수행하고, BLASTN와 GreenGenes의 16S DNA 시퀀스 데이터베이스(108,453 시퀀스)를 이용하여 97% 이상의 시퀀스 유사도 갖는 박테리아를 분석하였다(QIIME).

**실시예 4. 혈액에서 분리한 세균유래 소포 메타게놈 분석 기반 당뇨병 진단모형**

상기 실시예 3의 방법으로, 당뇨병환자 61명과 나이와 성별을 매칭한 정상인 122명의 혈액에서 소포를 분리한 후 메타게놈 시퀀싱을 수행하였다. 진단모형 개발은 먼저 t-test에서 두 군 사이의 p값이 0.05 이하이고, 두 군 사이에 2배 이상 차이가 나는 군주를 선정하고 난 후, logistic regression analysis 방법으로 진단적 성능 지표인 AUC(area under curve), 민감도, 및 특이도를 산출하였다.

혈액 내 세균유래 소포를 문(phylum) 수준에서 분석한 결과, Thermi, Fusobacteria, Chloroflexi, Cyanobacteria, TM7, Euryarchaeota, Proteobacteria, Actinobacteria, Verrucomicrobia, 및 Bacteroidetes 문 세균에서 선택되는 하나 이상의 바이오마커로 진단모형을 개발하였을 때, 당뇨병에 대한 진단적 성능이 유의하게 나타났다 (표 2 및 도 2 참조).

표 2

Taxon	대조군		당뇨병		t-test p-value	Ratio	Training Set			Test Set		
	Mean	SD	Mean	SD			AUC	Sensitivity	Specificity	AUC	Sensitivity	Sp ec if ic it y
p__[Thermi]	0.0022	0.0050	0.0000	0.0000	0.0000	0.01	0.71	0.98	0.00	0.70	1.00	0.04
p__Fusobact eria	0.0059	0.0110	0.0000	0.0000	0.0000	0.01	0.77	0.84	0.26	0.71	0.81	0.39
p__Chlorofl exi	0.0007	0.0022	0.0000	0.0000	0.0015	0.01	0.67	0.99	0.00	0.50	1.00	0.00
p__Cyanobac teria	0.0167	0.0646	0.0001	0.0001	0.0055	0.01	0.91	0.84	0.76	0.87	0.88	0.65
p__TM7	0.0041	0.0084	0.0001	0.0001	0.0000	0.03	0.76	0.86	0.24	0.66	0.78	0.30
p__Euryarch aeota	0.0013	0.0034	0.0001	0.0001	0.0001	0.06	0.66	0.99	0.00	0.60	1.00	0.04
p__Proteoba cteria	0.3781	0.1300	0.1119	0.1119	0.0000	0.30	0.99	0.96	0.97	0.94	0.94	1.00

p_Actinobacteria	0.1072	0.0686	0.0357	0.0357	0.0000	0.33	0.90	0.83	0.95	0.89	0.84	1.00
p_Verrucomicrobia	0.0171	0.0185	0.0481	0.0481	0.0000	2.81	0.93	0.90	0.68	0.95	0.94	0.70
p_Bacteroidetes	0.0857	0.0653	0.2915	0.2915	0.0000	3.40	1.00	1.00	1.00	0.98	0.91	1.00

[0074] 혈액 내 세균유래 소포를 강(class) 수준에서 분석한 결과, Cytophagia, Deinococci, Fusobacteriia, Sphingobacteriia, Flavobacteriia, Alphaproteobacteria, Betaproteobacteria, TM7-3, Bacilli, Actinobacteria, Gammaproteobacteria, Clostridia, Verrucomicrobiae, 및 Bacteroidia 강 세균에서 선택되는 하나 이상의 바이오마커로 진단모형을 개발하였을 때, 당뇨병에 대한 진단적 성능이 유의하게 나타났다 (표 3 및 도 3 참조).

표 3

[0075]

Taxon	대조군		당뇨병		t-test		Training Set			Test Set		
	Mean	SD	Mean	SD	p-value	Ratio	AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity
c_Cytophagia	0.0011	0.0028	0.0000	0.0000	0.0000	0.01	0.70	0.94	0.08	0.60	0.97	0.09
c_Deinococci	0.0022	0.0050	0.0000	0.0000	0.0000	0.01	0.71	0.98	0.00	0.70	1.00	0.04
c_Fusobacteriia	0.0059	0.0110	0.0000	0.0000	0.0000	0.01	0.77	0.84	0.26	0.71	0.81	0.39
c_Sphingobacteriia	0.0021	0.0048	0.0000	0.0000	0.0000	0.01	0.72	0.97	0.08	0.72	0.97	0.09
c_Flavobacteriia	0.0048	0.0070	0.0000	0.0000	0.0000	0.01	0.88	0.83	0.71	0.81	0.75	0.61
c_Alphaproteobacteria	0.0644	0.0444	0.0010	0.0010	0.0000	0.02	1.00	0.99	1.00	1.00	1.00	0.96
c_Betaproteobacteria	0.0486	0.0428	0.0014	0.0014	0.0000	0.03	0.97	0.94	1.00	0.97	0.88	1.00
c_TM7-3	0.0040	0.0084	0.0001	0.0001	0.0000	0.03	0.76	0.86	0.24	0.63	0.75	0.30
c_Bacilli	0.1423	0.0723	0.0196	0.0196	0.0000	0.14	0.98	0.97	1.00	0.95	0.91	1.00
c_Actinobacteriia	0.0993	0.0686	0.0310	0.0310	0.0000	0.31	0.88	0.83	0.95	0.88	0.88	0.96
c_Gammaproteobacteria	0.2630	0.1112	0.1093	0.1093	0.0000	0.42	0.96	0.94	0.97	0.91	0.91	1.00
c_Clostridia	0.1904	0.1130	0.4011	0.4011	0.0000	2.11	0.97	0.93	1.00	0.97	0.97	1.00
c_Verrucomicrobiae	0.0168	0.0183	0.0481	0.0481	0.0000	2.86	0.93	0.90	0.71	0.95	0.94	0.70
c_Bacteroidia	0.0769	0.0670	0.2914	0.2914	0.0000	3.79	1.00	1.00	1.00	0.98	0.91	1.00

[0077] 혈액 내 세균유래 소포를 목(order) 수준에서 분석한 결과, Aeromonadales, Deinococcales, Cytophagales, Rhizobiales, Neisseriales, Oceanospirillales, Fusobacteriales, Sphingobacteriales, Sphingomonadales, Pseudomonadales, Rhodospirillales, Flavobacteriales, Rhodocyclales, Rhodobacteriales, Gemellales, Caulobacteriales, Actinomycetales, Xanthomonadales, Alteromonadales, Pasteurellales, Bacillales, Burkholderiales, Lactobacillales, Clostridiales, RF32, Verrucomicrobiales, 및 Bacteroidales 목 세균에

서 선택되는 하나 이상의 바이오마커로 진단모형을 개발하였을 때, 당뇨병에 대한 진단적 성능이 유의하게 나타났다 (표 4 및 도 4 참조).

표 4

Taxon	대조군		당뇨병		t-test		Training Set			Test Set		
	Mean	SD	Mean	SD	p-value	Ratio	AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity
o_Aeromonadales	0.0007	0.0022	0.0000	0.0000	0.0003	0.00	0.71	0.98	0.11	0.56	0.88	0.09
o_Deinococcales	0.0014	0.0039	0.0000	0.0000	0.0001	0.00	0.68	0.98	0.00	0.51	1.00	0.04
o_Cytophagales	0.0011	0.0028	0.0000	0.0000	0.0000	0.01	0.70	0.94	0.08	0.60	0.97	0.09
o_Rhizobiales	0.0251	0.0233	0.0001	0.0001	0.0000	0.01	0.98	0.96	1.00	0.98	0.97	0.96
o_Neisseriales	0.0073	0.0181	0.0000	0.0000	0.0000	0.01	0.79	0.84	0.39	0.75	0.75	0.35
o_Oceanospirillales	0.0023	0.0060	0.0000	0.0000	0.0001	0.01	0.73	0.93	0.11	0.66	0.91	0.13
o_Fusobacteriales	0.0059	0.0110	0.0000	0.0000	0.0000	0.01	0.77	0.84	0.26	0.71	0.81	0.39
o_Sphingobacteriales	0.0021	0.0048	0.0000	0.0000	0.0000	0.01	0.72	0.97	0.08	0.72	0.97	0.09
o_Sphingomonadales	0.0214	0.0241	0.0002	0.0002	0.0000	0.01	0.99	0.94	0.95	0.95	0.91	0.87
o_Pseudomonadales	0.1607	0.1018	0.0015	0.0015	0.0000	0.01	1.00	0.99	1.00	1.00	0.97	1.00
o_Rhodospirillales	0.0031	0.0063	0.0000	0.0000	0.0000	0.01	0.76	0.90	0.21	0.58	0.78	0.22
o_Flavobacteriales	0.0048	0.0070	0.0000	0.0000	0.0000	0.01	0.88	0.83	0.71	0.81	0.75	0.61
o_Rhodocyclales	0.0017	0.0039	0.0000	0.0000	0.0000	0.01	0.71	0.98	0.03	0.66	1.00	0.04
o_Rhodobacteriales	0.0070	0.0097	0.0001	0.0001	0.0000	0.01	0.88	0.82	0.74	0.80	0.78	0.61
o_Gemellales	0.0012	0.0039	0.0000	0.0000	0.0014	0.01	0.70	0.97	0.16	0.50	0.94	0.13
o_Caulobacteriales	0.0056	0.0081	0.0001	0.0001	0.0000	0.02	0.86	0.80	0.74	0.83	0.81	0.57
o_Actinomycetales	0.0807	0.0680	0.0013	0.0013	0.0000	0.02	0.98	0.98	0.97	1.00	1.00	1.00
o_Xanthomonadales	0.0026	0.0052	0.0000	0.0000	0.0000	0.02	0.76	0.89	0.18	0.78	0.84	0.26
o_Alteromonadales	0.0009	0.0022	0.0000	0.0000	0.0000	0.02	0.70	0.93	0.05	0.56	1.00	0.09
o_Pasteurellales	0.0074	0.0148	0.0002	0.0002	0.0000	0.03	0.83	0.83	0.58	0.82	0.81	0.43
o_Bacillales	0.0535	0.0665	0.0015	0.0015	0.0000	0.03	0.97	0.92	0.92	0.93	0.84	0.96
o_Burkholderiales	0.0393	0.0396	0.0013	0.0013	0.0000	0.03	0.96	0.93	1.00	0.97	0.88	1.00
o_Lactobacillales	0.0866	0.0504	0.0176	0.0176	0.0000	0.20	0.95	0.93	0.97	0.88	0.81	1.00
o_Clostridiales	0.1899	0.1130	0.4007	0.4007	0.0000	2.11	0.97	0.93	1.00	0.97	0.97	1.00
o_RF32	0.0002	0.0009	0.0005	0.0005	0.0010	2.48	0.74	0.97	0.00	0.60	0.94	0.00
o_Verrucomicrobiales	0.0168	0.0183	0.0481	0.0481	0.0000	2.86	0.93	0.90	0.71	0.95	0.94	0.70

[0078]

o__Bacteroidales	0.0769	0.0670	0.2914	0.2914	0.0000	3.79	1.00	1.00	1.00	0.98	0.91	1.00
------------------	--------	--------	--------	--------	--------	------	------	------	------	------	------	------

[0080]

혈액 내 세균유래 소포를 과(family) 수준에서 분석한 결과, Aeromonadaceae, Methylobacteriaceae, Rhizobiaceae, Bradyrhizobiaceae, Halomonadaceae, Cytophagaceae, Neisseriaceae, Fusobacteriaceae, Sphingomonadaceae, Weeksellaceae, Moraxellaceae, Aerococcaceae, Pseudomonadaceae, Micrococcaceae, Propionibacteriaceae, Intrasporangiaceae, Gemellaceae, Flavobacteriaceae, Brevibacteriaceae, Rhodocyclaceae, Corynebacteriaceae, Burkholderiaceae, Rhodobacteraceae, Tissierellaceae, Caulobacteraceae, Xanthomonadaceae, Oxalobacteraceae, Staphylococcaceae, Comamonadaceae, Planococcaceae, Pasteurellaceae, Actinomycetaceae, S24-7, Enterococcaceae, Bacillaceae, Prevotellaceae, Streptococcaceae, Veillonellaceae, Lactobacillaceae, Ruminococcaceae, Mogibacteriaceae, Verrucomicrobiaceae, Lachnospiraceae, Odoribacteraceae, Bacteroidaceae, 및 Barnesiellaceae 과 세균에서 선택되는 하나 이상의 바이오마커로 진단모형을 개발하였을 때, 당뇨병에 대한 진단적 성능이 유의하게 나타났다 (표 5 및 도 5 참조).

표 5

[0081]

Taxon	대조군		당뇨병		t-test		Training Set			Test Set		
	Mean	SD	Mean	SD	p-value	Ratio	AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity
f__Aeromonadaceae	0.0007	0.0022	0.0000	0.0000	0.0003	0.00	0.71	0.98	0.11	0.53	0.88	0.09
f__Methylobacteriaceae	0.0077	0.0098	0.0000	0.0000	0.0000	0.00	0.91	0.82	0.79	0.85	0.75	0.70
f__Rhizobiaceae	0.0115	0.0149	0.0001	0.0001	0.0000	0.00	0.92	0.82	0.89	0.87	0.75	0.91
f__Bradyrhizobiaceae	0.0024	0.0062	0.0000	0.0000	0.0000	0.00	0.74	0.93	0.08	0.78	0.97	0.09
f__Halomonadaceae	0.0022	0.0060	0.0000	0.0000	0.0001	0.01	0.72	0.94	0.08	0.62	0.91	0.13
f__Cytophagaceae	0.0011	0.0028	0.0000	0.0000	0.0000	0.01	0.70	0.94	0.08	0.60	0.97	0.09
f__Neisseriaceae	0.0073	0.0181	0.0000	0.0000	0.0000	0.01	0.79	0.84	0.39	0.75	0.75	0.35
f__Fusobacteriaceae	0.0045	0.0090	0.0000	0.0000	0.0000	0.01	0.74	0.93	0.08	0.64	0.94	0.17
f__Sphingomonadaceae	0.0209	0.0240	0.0002	0.0002	0.0000	0.01	0.98	0.93	0.97	0.95	0.91	0.96
f__[Weeksellaceae]	0.0036	0.0060	0.0000	0.0000	0.0000	0.01	0.85	0.81	0.61	0.78	0.81	0.61
f__Moraxellaceae	0.0655	0.0546	0.0006	0.0006	0.0000	0.01	1.00	0.98	0.97	0.98	0.94	0.96
f__Aerococcaceae	0.0046	0.0082	0.0000	0.0000	0.0000	0.01	0.87	0.80	0.74	0.71	0.63	0.52
f__Pseudomonadaceae	0.0951	0.0802	0.0009	0.0009	0.0000	0.01	1.00	0.98	1.00	1.00	0.97	1.00
f__Micrococcaceae	0.0191	0.0239	0.0002	0.0002	0.0000	0.01	0.97	0.89	0.95	0.89	0.81	0.91
f__Propionibacteriaceae	0.0136	0.0168	0.0001	0.0001	0.0000	0.01	0.96	0.89	0.92	0.92	0.84	0.96
f__Intrasporangiaceae	0.0033	0.0063	0.0000	0.0000	0.0000	0.01	0.83	0.83	0.50	0.66	0.72	0.39
f__Gemellaceae	0.0012	0.0039	0.0000	0.0000	0.0014	0.01	0.71	0.96	0.16	0.51	0.91	0.13
f__Flavobacteriaceae	0.0012	0.0035	0.0000	0.0000	0.0002	0.01	0.70	0.97	0.11	0.61	0.94	0.09

f__Brevibacteriaceae	0.0014	0.0030	0.0000	0.0000	0.0000	0.01	0.72	0.97	0.08	0.67	0.97	0.04
f__Rhodocyclaceae	0.0017	0.0039	0.0000	0.0000	0.0000	0.01	0.71	0.98	0.03	0.66	1.00	0.04
f__Corynebacteriaceae	0.0224	0.0313	0.0003	0.0003	0.0000	0.01	0.91	0.84	0.84	0.89	0.88	0.70
f__Burkholderiaceae	0.0018	0.0044	0.0000	0.0000	0.0000	0.01	0.78	0.89	0.29	0.65	0.75	0.26
f__Rhodobacteraceae	0.0069	0.0093	0.0001	0.0001	0.0000	0.01	0.88	0.82	0.74	0.80	0.78	0.61
f__[Tissierellaceae]	0.0038	0.0066	0.0001	0.0001	0.0000	0.01	0.82	0.80	0.53	0.70	0.72	0.48
f__Caulobacteraceae	0.0056	0.0081	0.0001	0.0001	0.0000	0.02	0.86	0.80	0.74	0.83	0.81	0.57
f__Xanthomonadaceae	0.0024	0.0052	0.0000	0.0000	0.0000	0.02	0.76	0.87	0.21	0.77	0.84	0.22
f__Oxalobacteraceae	0.0267	0.0394	0.0006	0.0006	0.0000	0.02	0.89	0.87	1.00	0.86	0.81	1.00
f__Staphylococcaceae	0.0376	0.0571	0.0008	0.0008	0.0000	0.02	0.93	0.88	0.89	0.87	0.78	0.83
f__Comamonadaceae	0.0086	0.0102	0.0002	0.0002	0.0000	0.02	0.84	0.78	0.87	0.85	0.78	0.91
f__Planococcaceae	0.0056	0.0074	0.0001	0.0001	0.0000	0.03	0.93	0.87	0.87	0.87	0.84	0.65
f__Pasteurellaceae	0.0074	0.0148	0.0002	0.0002	0.0000	0.03	0.83	0.83	0.58	0.82	0.81	0.43
f__Actinomycetaceae	0.0034	0.0072	0.0001	0.0001	0.0000	0.04	0.76	0.71	0.58	0.64	0.59	0.39
f__S24-7	0.0029	0.0100	0.0001	0.0001	0.0032	0.05	0.73	0.87	0.21	0.51	0.75	0.22
f__Enterococcaceae	0.0059	0.0081	0.0003	0.0003	0.0000	0.06	0.86	0.79	0.84	0.67	0.53	0.78
f__Bacillaceae	0.0083	0.0104	0.0005	0.0005	0.0000	0.06	0.78	0.77	0.37	0.77	0.75	0.57
f__Prevotellaceae	0.0271	0.0306	0.0017	0.0017	0.0000	0.06	0.86	0.82	0.82	0.88	0.84	0.96
f__Streptococcaceae	0.0358	0.0273	0.0023	0.0023	0.0000	0.06	0.92	0.88	0.97	0.87	0.81	1.00
f__Veillonellaceae	0.0121	0.0117	0.0016	0.0016	0.0000	0.13	0.78	0.79	0.39	0.83	0.84	0.43
f__Lactobacillaceae	0.0313	0.0286	0.0129	0.0129	0.0000	0.41	0.81	0.86	0.13	0.61	0.66	0.17
f__Ruminococcaceae	0.0882	0.0769	0.1927	0.1927	0.0000	2.18	0.94	0.91	1.00	0.96	0.94	1.00
f__[Mogibacteriaceae]	0.0003	0.0011	0.0008	0.0008	0.0000	2.35	0.92	0.92	0.53	0.76	0.78	0.57
f__Verrucomicrobiaceae	0.0168	0.0183	0.0481	0.0481	0.0000	2.86	0.93	0.90	0.71	0.95	0.94	0.70
f__Lachnospiraceae	0.0427	0.0326	0.1669	0.1669	0.0000	3.91	1.00	1.00	1.00	1.00	1.00	1.00
f__[Odoribacteraceae]	0.0012	0.0028	0.0046	0.0046	0.0000	4.00	0.92	0.91	0.87	0.91	0.91	0.83
f__Bacteroidaceae	0.0300	0.0240	0.2583	0.2583	0.0000	8.60	1.00	1.00	1.00	1.00	1.00	1.00
f__[Barnesiellaceae]	0.0006	0.0019	0.0057	0.0057	0.0000	8.87	0.99	0.97	1.00	0.96	0.91	1.00

[0083] 혈액 내 세균유래 소포를 속(genus) 수준에서 분석한 결과, Halomonas, Methylobacterium, Neisseria, Fusobacterium, Kaistobacter, Agrobacterium, Porphyromonas, Cupriavidus, Acinetobacter, Pseudomonas, Chryseobacterium, Sphingomonas, Rothia, Micrococcus, Enhydrobacter, Propionibacterium, Brevibacterium,

Corynebacterium, Lautropia, Paracoccus, Staphylococcus, Haemophilus, Catenibacterium, Anaerococcus, Prevotella, Actinomyces, Veillonella, Citrobacter, Enterococcus, Prevotella, Streptococcus, Dialister, Bacillus, Lactobacillus, Bifidobacterium, Faecalibacterium, Parabacteroides, Paraprevotella, Akkermansia, Ruminococcus, Adlercreutzia, Butyricimonas, Odoribacter, Coprococcus, Anaerostipes, Blautia, Bacteroides, Epulopiscium, 및 Escherichia 속 세균에서 선택되는 하나 이상의 바이오마커로 진단모형을 개발하였을 때, 당뇨병에 대한 진단적 성능이 유의하게 나타났다 (표 6 및 도 6 참조).

표 6

Taxon	대조군		당뇨병		t-test		Training Set			Test Set		
	Mean	SD	Mean	SD	p-value	Ratio	AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity
g_Halomonas	0.0020	0.0058	0.0000	0.0000	0.0002	0.00	0.71	0.94	0.05	0.61	0.94	0.13
g_Methylobacterium	0.0037	0.0065	0.0000	0.0000	0.0000	0.01	0.85	0.88	0.61	0.65	0.59	0.48
g_Neisseria	0.0045	0.0115	0.0000	0.0000	0.0000	0.01	0.75	0.92	0.11	0.64	0.94	0.13
g_Fusobacterium	0.0045	0.0090	0.0000	0.0000	0.0000	0.01	0.74	0.93	0.08	0.64	0.94	0.17
g_Kaistobacter	0.0009	0.0030	0.0000	0.0000	0.0020	0.01	0.72	0.97	0.11	0.54	0.88	0.09
g_Agrobacterium	0.0018	0.0053	0.0000	0.0000	0.0003	0.01	0.72	0.96	0.11	0.65	0.94	0.09
g_Porphyromonas	0.0015	0.0037	0.0000	0.0000	0.0000	0.01	0.72	0.93	0.13	0.59	0.91	0.09
g_Cupriavidus	0.0172	0.0367	0.0001	0.0001	0.0000	0.01	0.83	0.73	0.89	0.76	0.63	0.87
g_Acinetobacter	0.0369	0.0442	0.0003	0.0003	0.0000	0.01	0.99	0.96	0.95	0.98	0.94	1.00
g_Pseudomonas	0.0912	0.0794	0.0008	0.0008	0.0000	0.01	1.00	0.99	1.00	1.00	0.97	1.00
g_Chryseobacterium	0.0022	0.0035	0.0000	0.0000	0.0000	0.01	0.83	0.83	0.53	0.75	0.78	0.43
g_Sphingomonas	0.0159	0.0185	0.0001	0.0001	0.0000	0.01	0.97	0.90	0.97	0.92	0.84	0.96
g_Rothia	0.0056	0.0091	0.0000	0.0000	0.0000	0.01	0.88	0.87	0.68	0.73	0.69	0.52
g_Micrococcus	0.0093	0.0154	0.0001	0.0001	0.0000	0.01	0.83	0.76	0.71	0.84	0.84	0.52
g_Enhydrobacter	0.0271	0.0286	0.0002	0.0002	0.0000	0.01	0.96	0.89	0.92	0.97	0.91	0.91
g_Propionibacterium	0.0136	0.0168	0.0001	0.0001	0.0000	0.01	0.96	0.89	0.92	0.92	0.84	0.96
g_Brevibacterium	0.0014	0.0030	0.0000	0.0000	0.0000	0.01	0.72	0.97	0.08	0.67	0.97	0.04
g_Corynebacterium	0.0224	0.0313	0.0003	0.0003	0.0000	0.01	0.91	0.84	0.84	0.89	0.88	0.70
g_Lautropia	0.0013	0.0041	0.0000	0.0000	0.0007	0.02	0.70	0.97	0.11	0.53	0.91	0.09
g_Paracoccus	0.0050	0.0070	0.0001	0.0001	0.0000	0.02	0.86	0.83	0.63	0.73	0.75	0.52
g_Staphylococcus	0.0370	0.0566	0.0008	0.0008	0.0000	0.02	0.93	0.87	0.89	0.87	0.75	0.83
g_Haemophilus	0.0067	0.0140	0.0002	0.0002	0.0000	0.02	0.83	0.84	0.47	0.82	0.84	0.43
g_Catenibacterium	0.0046	0.0098	0.0001	0.0001	0.0000	0.03	0.78	0.83	0.32	0.71	0.78	0.39
g_Anaerococcus	0.0008	0.0019	0.0000	0.0000	0.0000	0.03	0.70	0.97	0.03	0.58	0.94	0.09

[0084]

g_[Prevotella]	0.0020	0.0045	0.0001	0.0001	0.0000	0.04	0.71	0.97	0.03	0.61	0.97	0.09
g_Actinomyces	0.0031	0.0069	0.0001	0.0001	0.0000	0.04	0.75	0.72	0.50	0.61	0.59	0.39
g_Veillonella	0.0052	0.0088	0.0002	0.0002	0.0000	0.04	0.73	0.86	0.21	0.70	0.88	0.22
g_Citrobacter	0.0065	0.0090	0.0004	0.0004	0.0000	0.06	0.80	0.77	0.63	0.66	0.59	0.52
g_Enterococcus	0.0053	0.0076	0.0003	0.0003	0.0000	0.06	0.85	0.81	0.76	0.67	0.53	0.65
g_Prevotella	0.0271	0.0306	0.0017	0.0017	0.0000	0.06	0.86	0.82	0.82	0.88	0.84	0.96
g_Streptococcus	0.0332	0.0272	0.0022	0.0022	0.0000	0.07	0.90	0.87	0.97	0.85	0.81	1.00
g_Dialister	0.0036	0.0060	0.0003	0.0003	0.0000	0.08	0.72	0.94	0.05	0.69	0.94	0.09
g_Bacillus	0.0049	0.0077	0.0004	0.0004	0.0000	0.09	0.75	0.90	0.24	0.67	0.78	0.22
g_Lactobacillus	0.0308	0.0283	0.0128	0.0128	0.0000	0.42	0.80	0.86	0.13	0.61	0.63	0.13
g_Bifidobacterium	0.0140	0.0220	0.0296	0.0296	0.0000	2.12	0.90	0.93	0.16	0.92	0.94	0.17
g_Faecalibacterium	0.0388	0.0430	0.0867	0.0867	0.0000	2.23	0.91	0.89	0.79	0.88	0.88	0.70
g_Parabacteroides	0.0051	0.0078	0.0117	0.0117	0.0000	2.29	0.87	0.88	0.29	0.76	0.81	0.22
g_Paraprevotella	0.0005	0.0015	0.0012	0.0012	0.0000	2.61	0.88	0.91	0.05	0.83	0.88	0.04
g_Akkermansia	0.0167	0.0184	0.0481	0.0481	0.0000	2.88	0.93	0.90	0.71	0.95	0.94	0.70
g_Ruminococcus	0.0078	0.0159	0.0235	0.0235	0.0000	3.02	0.97	0.93	1.00	0.94	0.94	1.00
g_Adlercreutzia	0.0008	0.0018	0.0023	0.0023	0.0000	3.03	0.88	0.91	0.16	0.96	0.97	0.00
g_Butyricimonas	0.0005	0.0015	0.0017	0.0017	0.0000	3.54	0.91	0.93	0.16	0.92	0.94	0.13
g_Odoribacter	0.0007	0.0022	0.0030	0.0030	0.0000	4.32	0.96	0.96	0.92	0.90	0.84	0.96
g_Coproccoccus	0.0103	0.0128	0.0453	0.0453	0.0000	4.42	0.97	0.94	0.97	1.00	1.00	1.00
g_Anaerostipes	0.0002	0.0008	0.0013	0.0013	0.0000	5.29	0.93	0.93	0.74	0.96	0.94	0.57
g_Blautia	0.0059	0.0071	0.0499	0.0499	0.0000	8.50	1.00	1.00	1.00	1.00	1.00	1.00
g_Bacteroides	0.0300	0.0240	0.2583	0.2583	0.0000	8.60	1.00	1.00	1.00	1.00	1.00	1.00
g_Epulopiscium	0.0001	0.0003	0.0012	0.0012	0.0000	10.70	0.99	0.98	0.95	0.96	0.91	0.96
g_Escherichia	0.0001	0.0002	0.0814	0.0814	0.0000	887.16	1.00	1.00	1.00	1.00	1.00	1.00

[0086] 실시예 5. 소변에서 분리한 세균유래 소포 메타게놈 분석 기반 당뇨병 진단모형

[0087] 상기 실시예 3의 방법으로, 당뇨병환자 60명과 나이와 성별을 매칭한 정상인 134명의 소변에서 소포를 분리한 후 메타게놈 시퀀싱을 수행하였다. 진단모형 개발은 먼저 t-test에서 두 군 사이의 p값이 0.05 이하이고, 두 군 사이에 2배 이상 차이가 나는 균주를 선정하고 난 후, logistic regression analysis 방법으로 진단적 성능 지표인 AUC(area under curve), 민감도, 및 특이도를 산출하였다.

[0088] 소변 내 세균유래 소포를 문(phylum) 수준에서 분석한 결과, Tenericutes 문 세균 바이오마커로 진단모형을 개발하였을 때, 당뇨병에 대한 진단적 성능이 유의하게 나타났다 (표 7 및 도 7 참조).

표 7

[0089]

Taxon	대조군		당뇨병		t-test		Training Set			Test Set		
	Mean	SD	Mean	SD	p-value	Ratio	AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity
p_Tenericutes	0.0051	0.0101	0.0013	0.0013	0.0001	0.26	0.71	0.97	0.21	0.50	0.89	0.05

[0091]

소변 내 세균유래 소포를 강(class) 수준에서 분석한 결과, Mollicutes, Coriobacteriia, Deltaproteobacteria, 및 Epsilonproteobacteria 강 세균에서 선택되는 하나 이상의 바이오마커로 진단모형을 개발하였을 때, 당뇨병에 대한 진단적 성능이 유의하게 나타났다 (표 8 및 도 8 참조).

표 8

[0092]

Taxon	대조군		당뇨병		t-test		Training Set			Test Set		
	Mean	SD	Mean	SD	p-value	Ratio	AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity
c_Mollicutes	0.0051	0.0101	0.0013	0.0013	0.0001	0.3	0.71	0.97	0.21	0.50	0.89	0.05
c_Coriobacteriia	0.0130	0.0112	0.0370	0.0370	0.0027	2.9	0.80	0.93	0.42	0.76	0.92	0.41
c_Deltaproteobacteria	0.0015	0.0025	0.0045	0.0045	0.0012	3.0	0.72	0.94	0.26	0.48	0.97	0.14
c_Epsilonproteobacteria	0.0003	0.0013	0.0049	0.0049	0.0003	15.0	0.77	0.97	0.37	0.58	1.00	0.23

[0094]

소변 내 세균유래 소포를 목(order) 수준에서 분석한 결과, Stramenopiles, Pseudomonadales, Coriobacteriales, Desulfovibrionales, 및 Campylobacterales 목 세균에서 선택되는 하나 이상의 바이오마커로 진단모형을 개발하였을 때, 당뇨병에 대한 진단적 성능이 유의하게 나타났다 (표 9 및 도 9 참조).

표 9

[0095]

Taxon	대조군		당뇨병		t-test		Training Set			Test Set		
	Mean	SD	Mean	SD	p-value	Ratio	AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity
o_Stramenopiles	0.0022	0.0051	0.0000	0.0000	0.0010	0.00	0.76	0.92	0.16	0.49	0.86	0.18
o_Pseudomonadales	0.1477	0.1423	0.0664	0.0664	0.0000	0.45	0.75	0.88	0.30	0.63	0.67	0.59
o_Coriobacteriales	0.0130	0.0112	0.0370	0.0370	0.0027	2.86	0.81	0.91	0.42	0.78	0.93	0.29
o_Desulfovibrionales	0.0007	0.0018	0.0043	0.0043	0.0002	5.85	0.74	0.95	0.37	0.65	0.88	0.18
o_Campylobacterales	0.0003	0.0013	0.0049	0.0049	0.0003	14.97	0.76	0.98	0.33	0.57	1.00	0.29

[0097] 소변 내 세균유래 소포를 과(family) 수준에서 분석한 결과, Bradyrhizobiaceae, Cellulomonadaceae, Pseudomonadaceae, Moraxellaceae, Comamonadaceae, Enterococcaceae, Clostridiaceae, Coriobacteriaceae, Rikenellaceae, Desulfovibrionaceae, 및 Helicobacteraceae 과 세균에서 선택되는 하나 이상의 바이오마커로 진단모형을 개발하였을 때, 당뇨병에 대한 진단적 성능이 유의하게 나타났다 (표 10 및 도 10 참조).

표 10

[0098]

Taxon	대조군		당뇨병		t-test		Training Set			Test Set		
	Mean	SD	Mean	SD	p-value	Ratio	AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity
f_Bradyrhizobiaceae	0.0017	0.0043	0.0004	0.0004	0.0009	0.21	0.74	0.97	0.24	0.53	0.80	0.21
f_Cellulomonadaceae	0.0010	0.0025	0.0002	0.0002	0.0011	0.24	0.72	0.93	0.07	0.44	0.98	0.05
f_Pseudomonadaceae	0.0789	0.0754	0.0347	0.0347	0.0000	0.44	0.75	0.89	0.39	0.70	0.83	0.32
f_Moraxellaceae	0.0687	0.0941	0.0317	0.0317	0.0001	0.46	0.70	0.91	0.20	0.51	0.93	0.11
f_Comamonadaceae	0.0057	0.0091	0.0124	0.0124	0.0001	2.18	0.70	0.95	0.22	0.61	0.95	0.21
f_Enterococcaceae	0.0106	0.0108	0.0233	0.0233	0.0050	2.20	0.69	0.97	0.15	0.61	0.95	0.16
f_Clostridiaceae	0.0137	0.0122	0.0331	0.0331	0.0000	2.42	0.80	0.95	0.51	0.76	0.83	0.47
f_Coriobacteriaceae	0.0130	0.0112	0.0370	0.0370	0.0027	2.86	0.78	0.89	0.37	0.84	1.00	0.37
f_Rikenellaceae	0.0017	0.0028	0.0074	0.0074	0.0000	4.49	0.79	0.96	0.41	0.59	0.85	0.32
f_Desulfovibrionaceae	0.0007	0.0018	0.0043	0.0043	0.0002	5.85	0.76	0.97	0.37	0.57	0.93	0.16
f_Helicobacteraceae	0.0002	0.0012	0.0047	0.0047	0.0003	27.36	0.76	0.98	0.32	0.57	1.00	0.26

[0100] 소변 내 세균유래 소포를 속(genus) 수준에서 분석한 결과, Rhizobium, Cupriavidus, Acinetobacter, Pseudomonas, Lactobacillus, Citrobacter, Enterococcus, Paracoccus, Klebsiella, SMB53, Allobaculum, Desulfovibrio, AF12, 및 Flexispira 속 세균에서 선택되는 하나 이상의 바이오마커로 진단모형을 개발하였을 때, 당뇨병에 대한 진단적 성능이 유의하게 나타났다 (표 11 및 도 11 참조).

표 11

[0101]

Taxon	대조군		당뇨병		t-test		Training Set			Test Set		
	Mean	SD	Mean	SD	p-value	Ratio	AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity
g_Rhizobium	0.0030	0.0043	0.0000	0.0000	0.0000	0.00	0.91	0.86	0.74	0.86	0.83	0.67
g_Cupriavidus	0.0137	0.0551	0.0007	0.0007	0.0078	0.05	0.74	0.92	0.19	0.54	0.76	0.11

g_Acinetobacter	0.0564	0.0928	0.0157	0.0157	0.0000	0.28	0.70	0.96	0.17	0.64	1.00	0.17
g_Pseudomonas	0.0751	0.0734	0.0317	0.0317	0.0000	0.42	0.77	0.88	0.45	0.74	0.83	0.44
g_Lactobacillus	0.0367	0.0433	0.0736	0.0736	0.0008	2.00	0.64	0.98	0.26	0.75	0.93	0.28
g_Citrobacter	0.0011	0.0018	0.0028	0.0028	0.0045	2.47	0.65	0.95	0.26	0.48	0.98	0.06
g_Enterococcus	0.0090	0.0097	0.0227	0.0227	0.0023	2.51	0.68	0.95	0.29	0.78	0.88	0.17
g_Paracoccus	0.0026	0.0043	0.0067	0.0067	0.0015	2.62	0.70	0.94	0.14	0.72	0.93	0.22
g_Klebsiella	0.0019	0.0034	0.0051	0.0051	0.0003	2.72	0.61	0.98	0.14	0.66	0.98	0.33
g_SMB53	0.0031	0.0048	0.0182	0.0182	0.0000	5.92	0.86	0.97	0.64	0.88	0.93	0.89
g_Allobaculum	0.0002	0.0007	0.0019	0.0019	0.0045	8.88	0.62	0.97	0.21	0.70	0.95	0.33
g_Desulfovibrio	0.0005	0.0016	0.0042	0.0042	0.0001	8.90	0.71	0.98	0.29	0.69	1.00	0.28
g_AF12	0.0002	0.0012	0.0041	0.0041	0.0000	22.63	0.63	0.98	0.26	0.84	1.00	0.50
g_Flexispira	0.0000	0.0000	0.0040	0.0040	0.0000		0.70	1.00	0.38	0.63	1.00	0.39

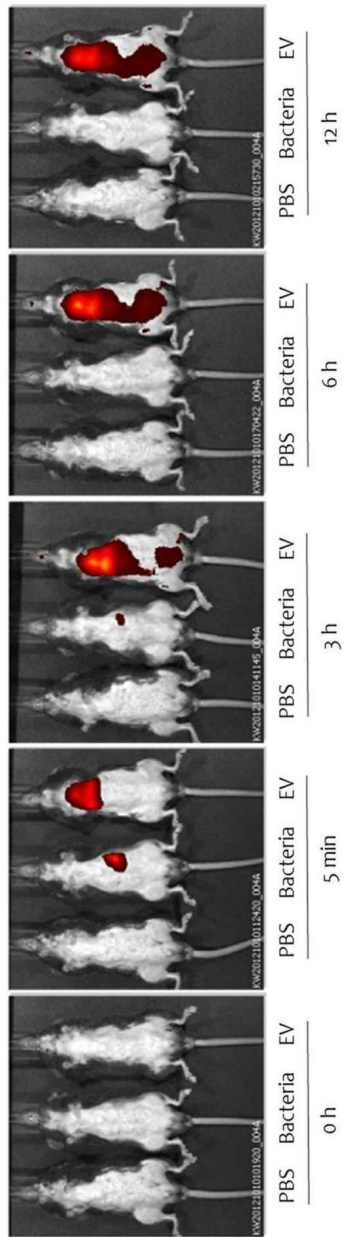
[0103]

상기 진술한 본 발명의 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다.

도면

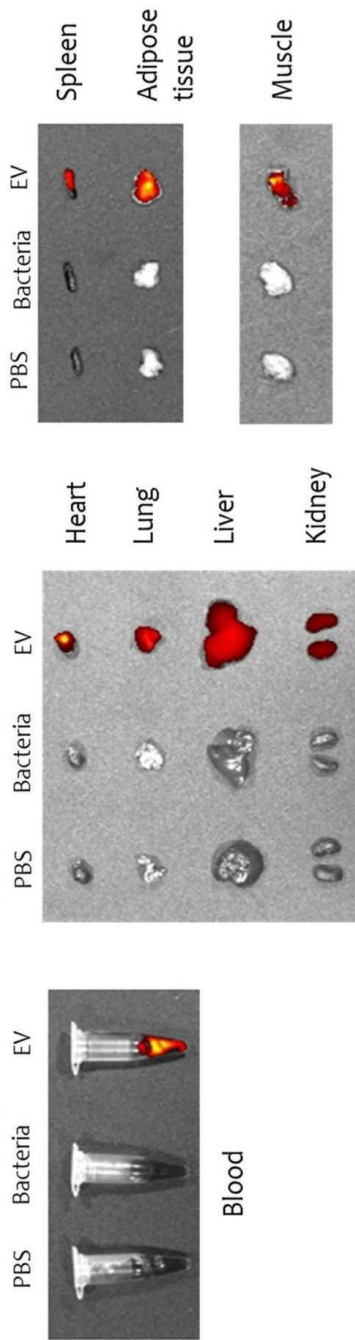
도면1a

In vivo IVIS imaging after oral application of gut microbe and microbial EVs

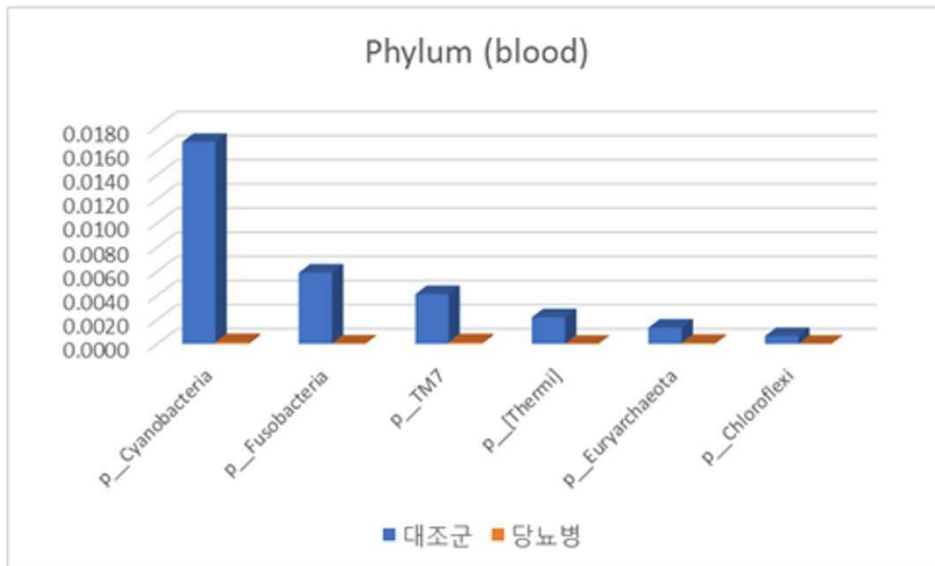
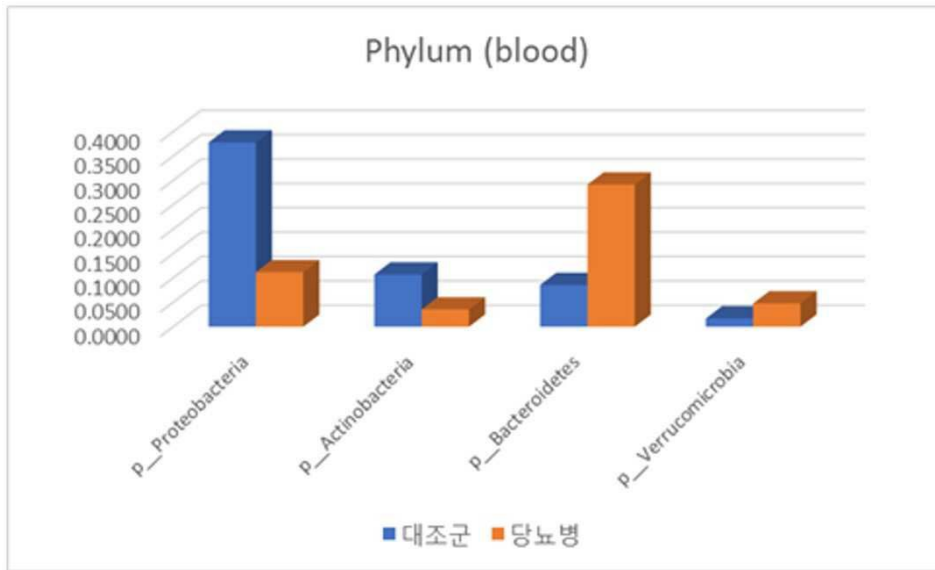


도면1b

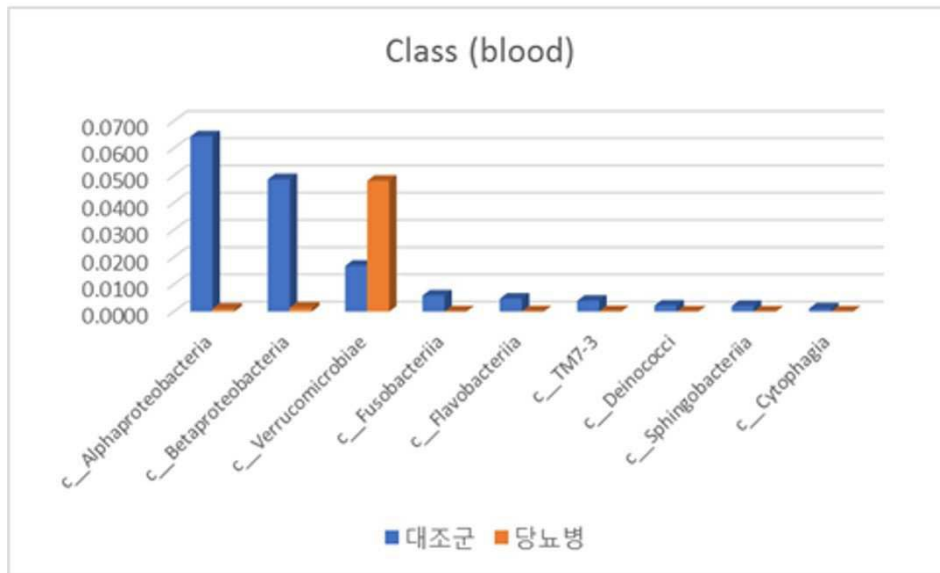
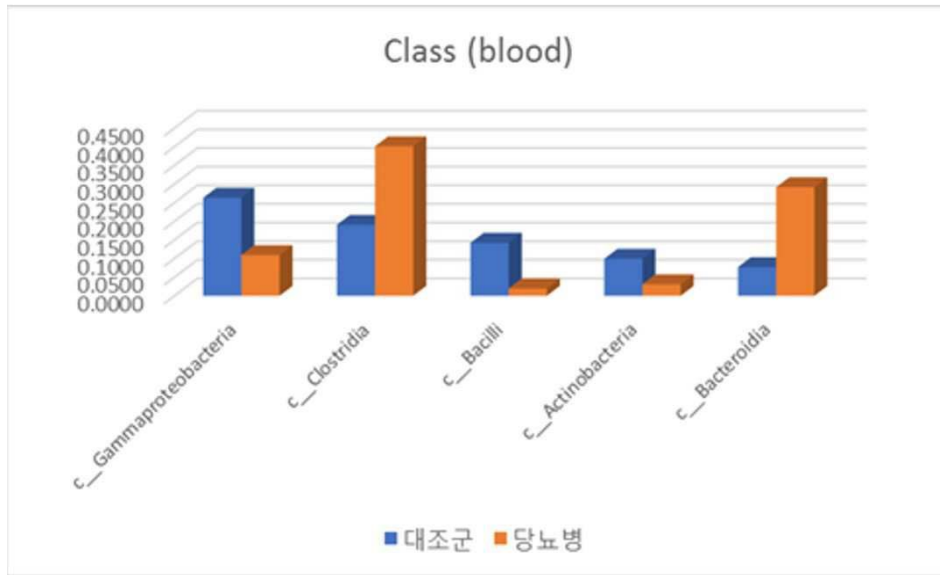
IVIS Imaging 12 h after oral application of gut microbe and microbial EVs



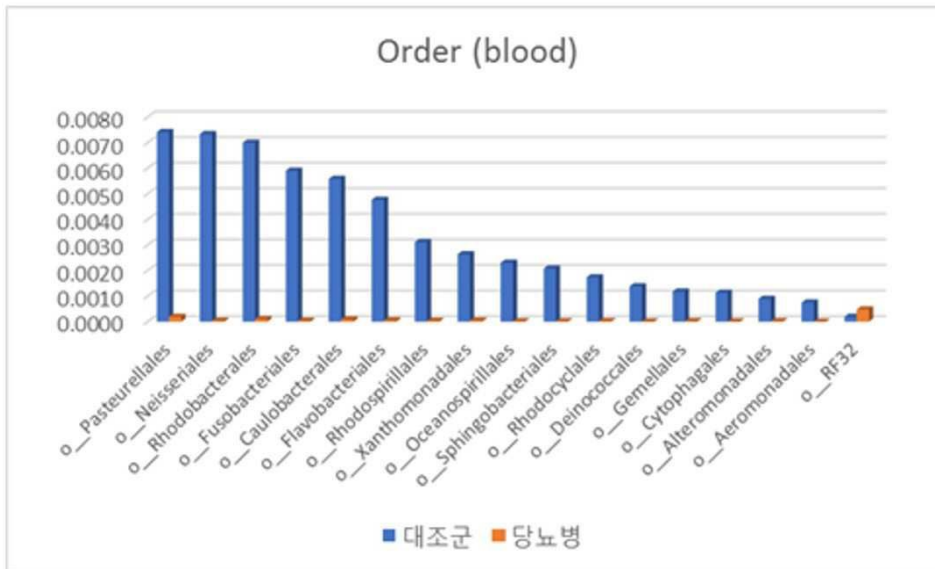
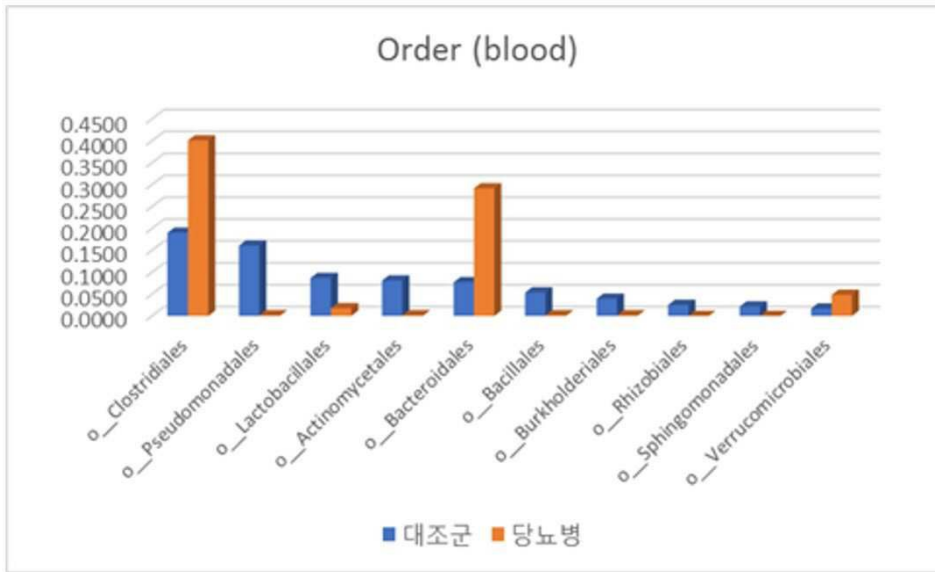
도면2



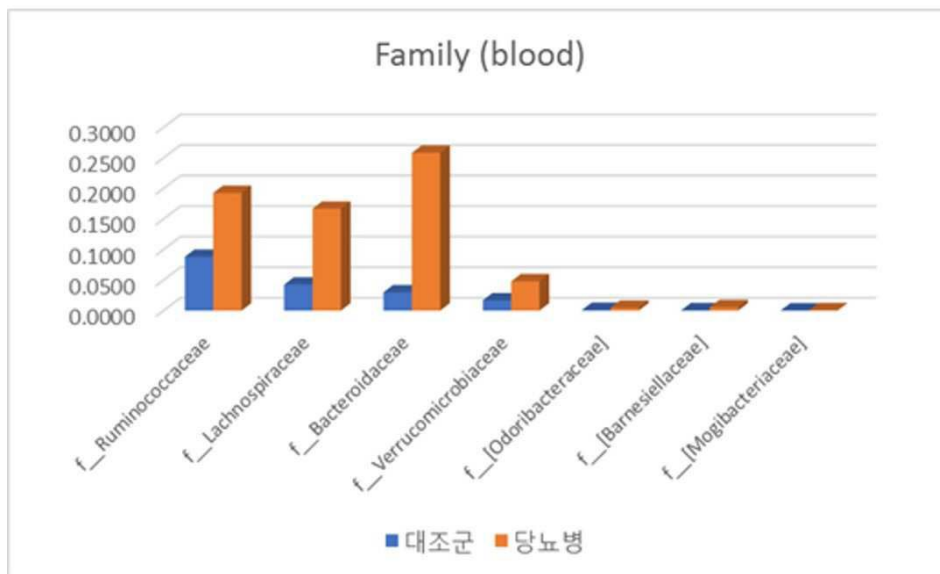
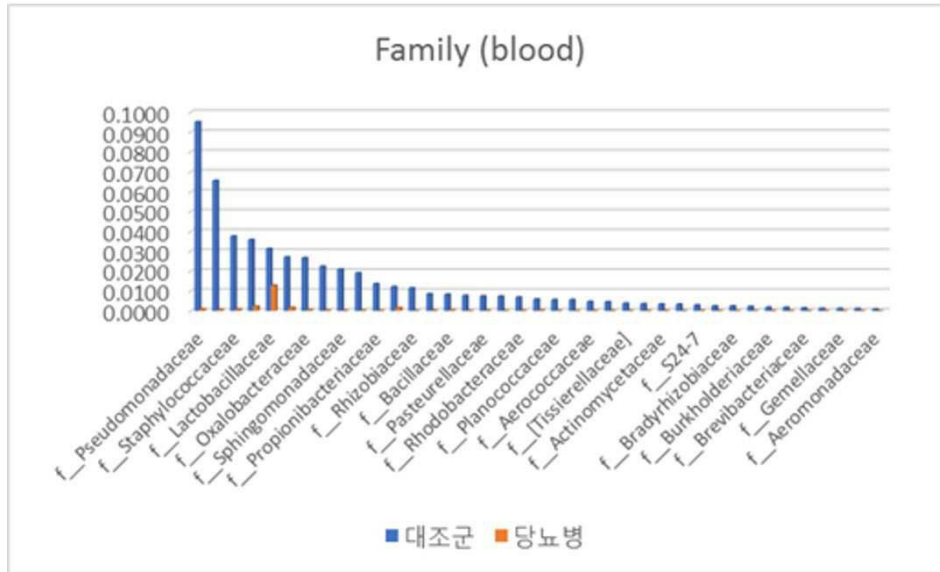
도면3



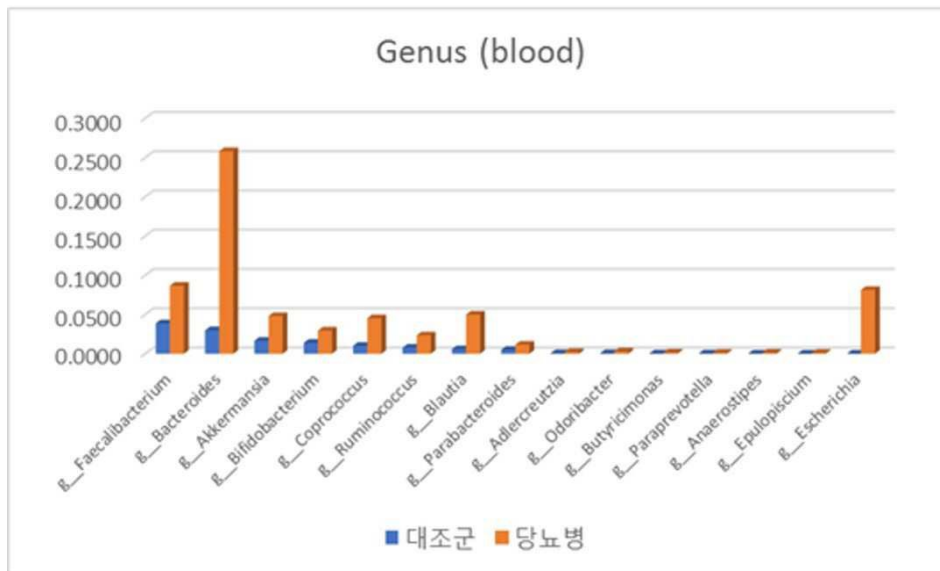
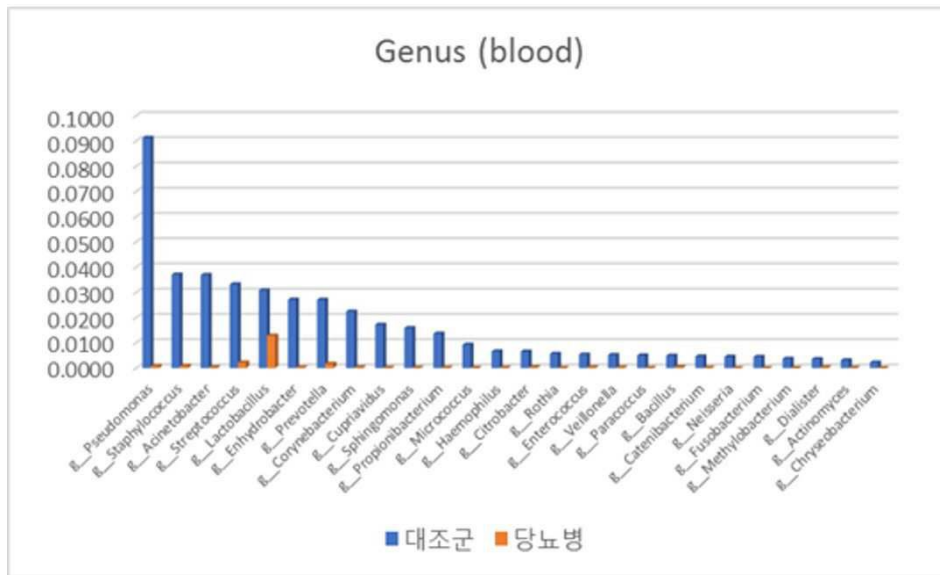
도면4



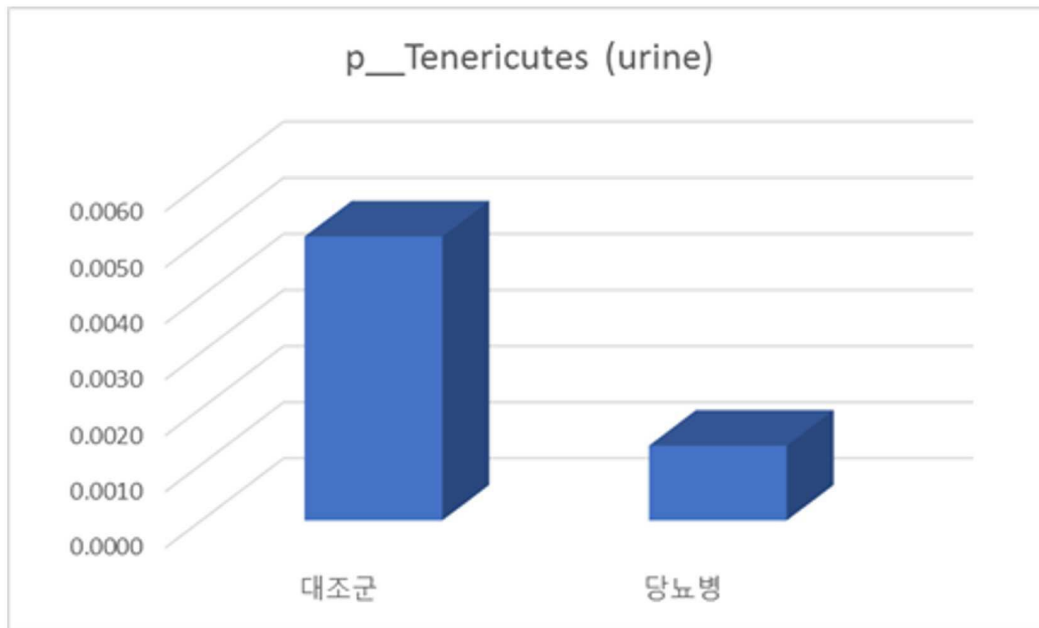
도면5



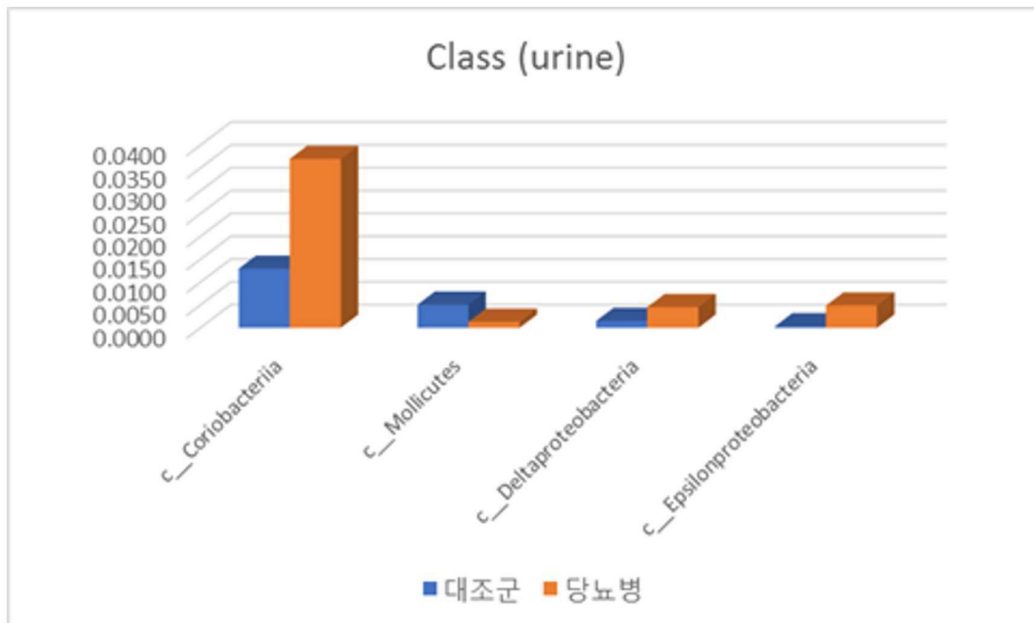
도면6



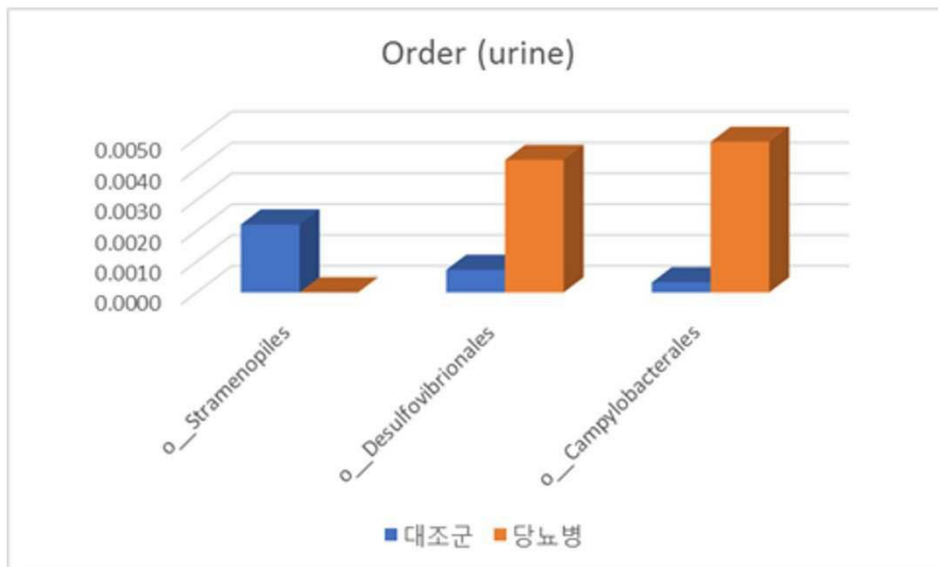
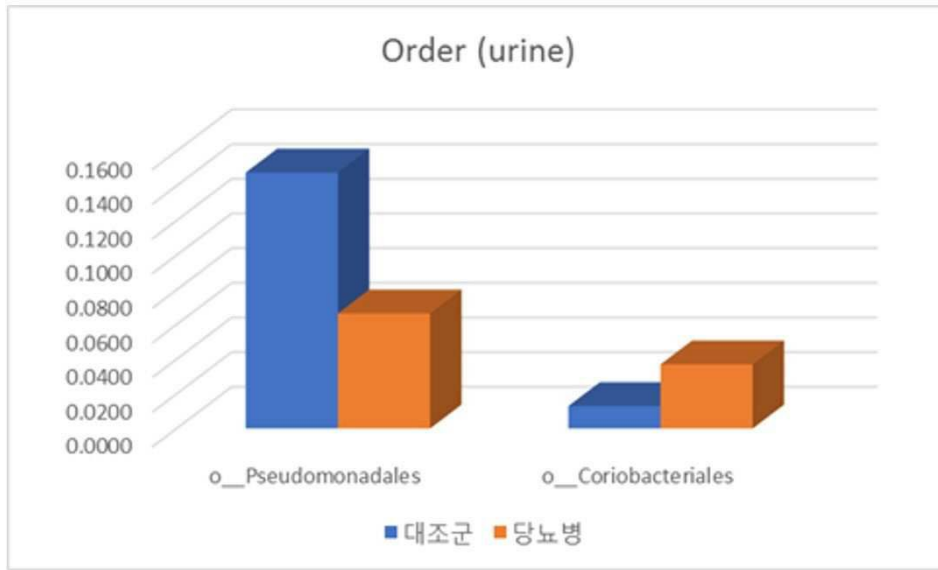
도면7



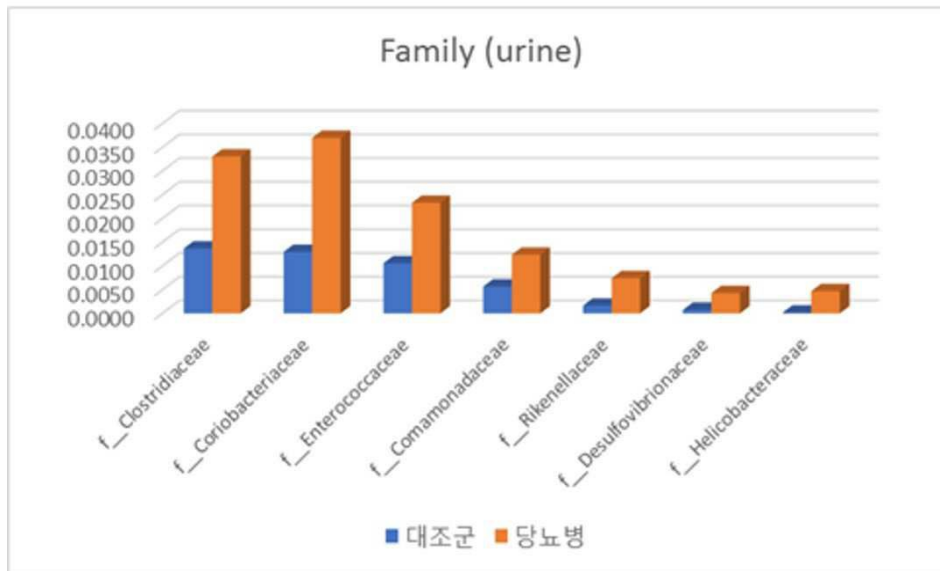
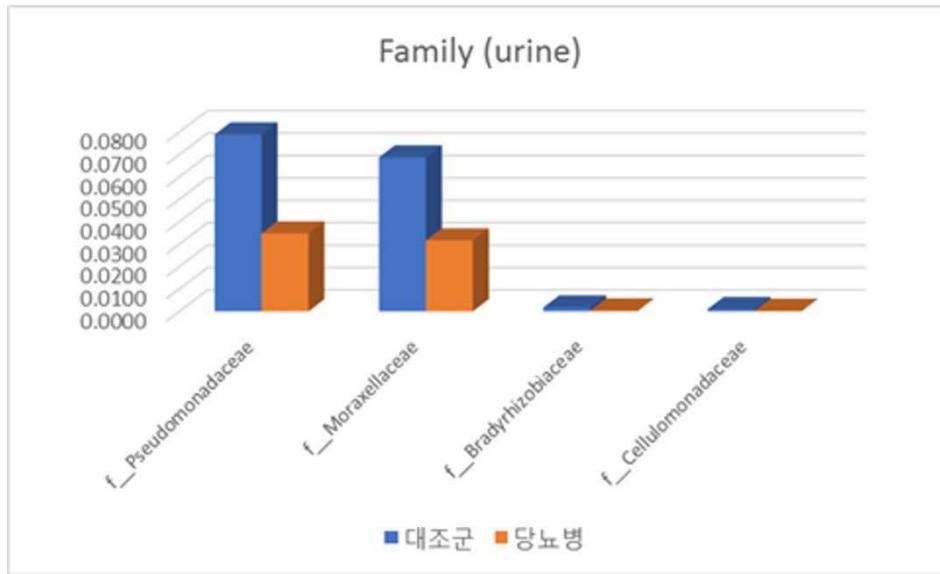
도면8



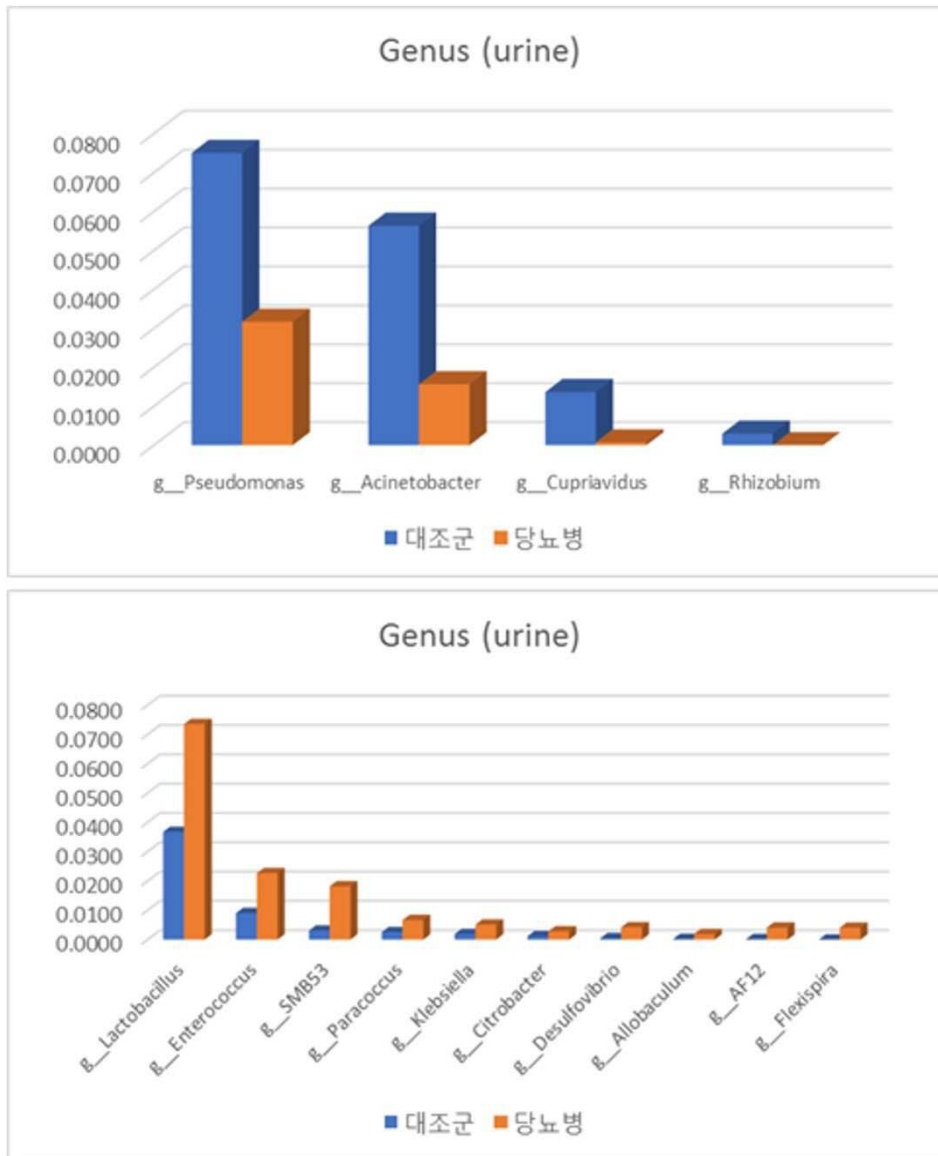
도면9



도면10



도면11



서열 목록

- <110> MD Healthcare Inc.
- <120> Method for diagnosis of diabetes using analysis of bacteria metagenome
- <130> MP17-352
- <150> KR 10-2017-0025085
- <151> 2017-02-24
- <160> 2
- <170> KoPatentIn 3.0
- <210> 1
- <211> 50

<212> DNA

<213> Artificial Sequence

<220><223> 16S\_V3\_F

<400> 1

tcgtcggcag cgtcagatgt gtataagaga cagcctacgg gnggcwgcag 50

<210> 2

<211> 55

<212> DNA

<213> Artificial Sequence

<220><223> 16S\_V4\_R

<400> 2

gtctcgtggg ctcggagatg tgtataagag acaggactac hvvggtatct aatcc 55