



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0093511  
(43) 공개일자 2024년06월24일

- |   |   |
|---|---|
| <p>(51) 국제특허분류(Int. Cl.)<br/>G01N 33/50 (2017.01) C12Q 1/6883 (2018.01)</p> <p>(52) CPC특허분류<br/>G01N 33/5091 (2013.01)<br/>C12Q 1/6883 (2022.01)</p> <p>(21) 출원번호 10-2024-7013946</p> <p>(22) 출원일자(국제) 2022년10월26일<br/>심사청구일자 없음</p> <p>(85) 번역문제출일자 2024년04월25일</p> <p>(86) 국제출원번호 PCT/AU2022/051283</p> <p>(87) 국제공개번호 WO 2023/070153<br/>국제공개일자 2023년05월04일</p> <p>(30) 우선권주장<br/>2021903424 2021년10월26일 오스트레일리아(AU)</p> | <p>(71) 출원인<br/>레스피라다임 피티와이 엘티디<br/>오스트레일리아, 웨스턴 오스트레일리아6163, 스피어우드 1/235 로킹업 로드</p> <p>(72) 발명자<br/>보스코 안토니<br/>오스트레일리아, 웨스턴 오스트레일리아 6163, 스피어우드, 1/235 로킹업 로드, 씨/- 레스피라다임 피티와이 엘티디</p> <p>리드 제임스<br/>오스트레일리아, 웨스턴 오스트레일리아 6163, 스피어우드, 1/235 로킹업 로드, 씨/- 레스피라다임 피티와이 엘티디</p> <p>(74) 대리인<br/>특허법인한얼</p> |
|---|---|

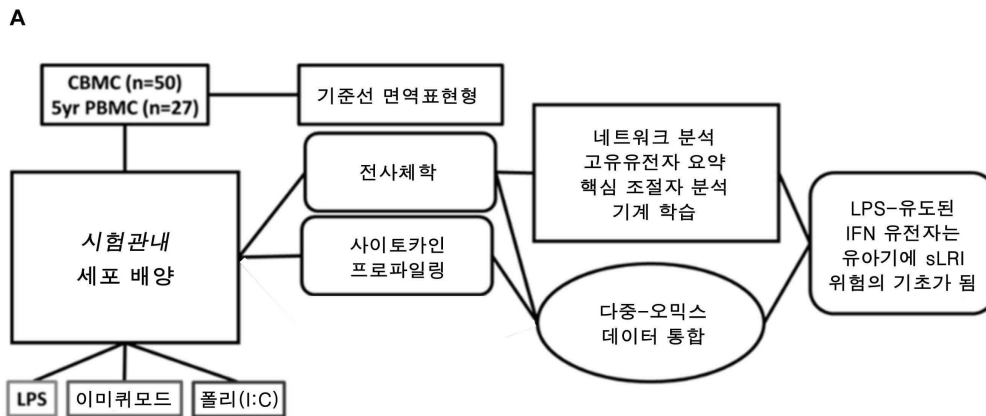
전체 청구항 수 : 총 44 항

(54) 발명의 명칭 **호흡기 감염 위험을 결정하기 위한 방법**

(57) 요약

본 발명은 호흡기 감염에 대한 감수성을 결정하기 위한 방법에 관한 것이다. 다른 측면에서, 본 발명은 개체가 호흡기 감염에 대한 증가된 감수성을 갖는지 여부를 결정하기 위한 방법을 제공하며, 상기 방법은 개체로부터의 체대혈 단핵세포(CBMC)를 TLR4 작용제와 접촉시키는 단계; CBMC에서 바이오마커의 발현 수준을 측정하는 단계를 포함하며; 여기서 TLR4 작용제와 접촉되지 않은 CBMC와 비교하여 TLR4 작용제와 접촉된 CBMC에서 바이오마커의 차등적 발현은 개체가 호흡기 감염에 대한 증가된 감수성을 가짐을 나타낸다.

대표도



(52) CPC특허분류

*G01N 33/5005* (2013.01)  
*G01N 33/5041* (2013.01)  
*G01N 33/5094* (2013.01)  
*C12Q 2600/118* (2013.01)  
*C12Q 2600/158* (2013.01)  
*G01N 2333/70596* (2013.01)  
*G01N 2800/12* (2013.01)  
*G01N 2800/52* (2021.08)  
*G01N 2800/60* (2013.01)

---

## 명세서

### 청구범위

#### 청구항 1

개체가 호흡기 감염에 대한 증가된 감수성을 갖는지 여부를 결정하기 위한 방법으로서, 상기 방법은

- 개체로부터의 제대혈 단핵세포(CBMC)를 포함하는 샘플을 TLR4 작용제와 접촉시키는 단계;
- CBMC에서 바이오마커 KLHC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1의 발현 수준을 측정하는 단계를 포함하며;

여기서 TLR4 작용제와 접촉되지 않은 CBMC와 비교하여 TLR4 작용제와 접촉된 CBMC에서 바이오마커 KLHC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1의 차등적 발현은 개체가 호흡기 감염에 대한 감수성이 증가됨을 나타내는 방법.

#### 청구항 2

개체가 호흡기 감염에 대한 증가된 감수성을 갖는지 여부를 결정하기 위한 방법으로서, 상기 방법은

- 개체로부터의 제대혈 단핵세포(CBMC)를 포함하는 샘플을 TLR4 작용제와 접촉시키는 단계;
- CBMC에서 인터페론 모듈 바이오마커의 발현 수준을 측정하는 단계를 포함하며;

여기서 TLR4 작용제와 접촉되지 않은 CBMC와 비교하여 TLR4 작용제와 접촉된 CBMC에서 인터페론 모듈 바이오마커의 차등적 발현은 개체가 호흡기 감염에 대한 감수성이 증가됨을 나타내는 방법.

#### 청구항 3

개체가 호흡기 감염에 대한 증가된 감수성을 갖는지 여부를 결정하기 위한 방법으로서, 상기 방법은

- 개체로부터의 제대혈 단핵세포(CBMC)를 포함하는 샘플을 TLR4 작용제와 접촉시키는 단계;
- CBMC에서 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커의 발현 수준을 측정하는 단계를 포함하며;

여기서 TLR4 작용제와 접촉되지 않은 CBMC와 비교하여 TLR4 작용제와 접촉된 CBMC에서 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커의 차등적 발현은 개체가 호흡기 감염에 대한 감수성이 증가됨을 나타내는 방법.

#### 청구항 4

개체가 호흡기 감염에 대한 증가된 감수성을 갖는지 여부를 결정하기 위한 방법으로서, 상기 방법은

- 개체로부터의 제대혈 단핵세포(CBMC)를 포함하는 샘플을 TLR4 작용제와 접촉시키는 단계;
- CBMC에서 바이오마커 KLHC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1의 발현 수준을 측정하는 단계;

- 개체로부터의 바이오마커의 발현 수준을 참조 데이터 세트와 비교하는 단계로서, 상기 참조 데이터 세트는 호흡기 감염에 대한 증가된 감수성을 갖거나 갖지 않는 개체의 개체들 중 하나로부터의 TLR4 작용제와 접촉된 CBMC에서 동일한 바이오마커의 발현 수준에 대한 정보를 포함하는 것인, 단계를 포함하며;

여기서 참조 데이터 세트와 비교하여 개체에서 바이오마커 KLHC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1의 발현 수준은 개체가 호흡기 감염에 대한 감수성이 증가됨을 나타내는 방법.

#### 청구항 5

개체가 호흡기 감염에 대한 증가된 감수성을 갖는지 여부를 결정하기 위한 방법으로서, 상기 방법은

- 개체로부터의 제대혈 단핵세포(CBMC)를 포함하는 샘플을 TLR4 작용제와 접촉시키는 단계;

- CBMC에서 인터페론 모듈 바이오마커의 발현 수준을 측정하는 단계;
  - 개체로부터의 바이오마커의 발현 수준을 참조 데이터 세트와 비교하는 단계로서, 상기 참조 데이터 세트는 호흡기 감염에 대한 증가된 감수성을 갖거나 갖지 않는 개체의 개체들 중 하나로부터의 TLR4 작용제와 접촉된 CBMC에서 동일한 바이오마커의 발현 수준에 대한 정보를 포함하는 것인, 단계를 포함하며;
- 여기서 참조 데이터 세트와 비교하여 개체에서 인터페론 모듈 바이오마커의 발현 수준은 개체가 호흡기 감염에 대한 감수성이 증가됨을 나타내는 방법.

**청구항 6**

- 개체가 호흡기 감염에 대한 증가된 감수성을 갖는지 여부를 결정하기 위한 방법으로서, 상기 방법은
- 개체로부터의 제대혈 단핵세포(CBMC)를 포함하는 샘플을 TLR4 작용제와 접촉시키는 단계;
  - CBMC에서 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커의 발현 수준을 측정하는 단계;
  - 개체로부터의 바이오마커의 발현 수준을 참조 데이터 세트와 비교하는 단계로서, 상기 참조 데이터 세트는 호흡기 감염에 대한 증가된 감수성을 갖거나 갖지 않는 개체의 개체들 중 하나로부터의 TLR4 작용제와 접촉된 CBMC에서 동일한 바이오마커의 발현 수준에 대한 정보를 포함하는 것인, 단계를 포함하며;
- 여기서 참조 데이터 세트와 비교하여 개체에서 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커의 발현 수준은 개체가 호흡기 감염에 대한 감수성이 증가됨을 나타내는 방법.

**청구항 7**

제2항 또는 제5항 중 어느 한 항에 있어서, 상기 인터페론 모듈 바이오마커는 표 3에 나열된 것인, 방법.

**청구항 8**

제3항 또는 제6항에 있어서, 상기 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커는 표 4에 나열된 것인, 방법.

**청구항 9**

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 방법은 기계 학습 알고리즘을 바이오마커의 차등적 또는 절대적 발현에 적용하고, 이에 따라 개체가 호흡기 감염에 대한 감수성이 증가됨을 나타내는 단계를 추가로 포함하는 것인, 방법.

**청구항 10**

제9항에 있어서, 상기 기계 학습 알고리즘은 랜덤 포레스트 분석인, 방법.

**청구항 11**

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 발현 수준은 차등적 발현이고, 차등적 발현은 1.5배 초과 증가 또는 감소인, 방법.

**청구항 12**

제4항 내지 제6항 중 어느 한 항에 있어서, 상기 증가된 감수성은 참조 데이터 세트와 비교하여 적어도 1.10, 적어도 1.11, 적어도 1.12, 적어도 1.13, 적어도 1.14, 적어도 1.15, 적어도 1.16, 적어도 1.17, 적어도 1.18, 적어도 1.19, 적어도 1.20, 적어도 1.21, 적어도 1.22, 적어도 1.23, 적어도 1.24, 적어도 1.25, 적어도 1.30, 적어도 1.35, 적어도 1.40, 적어도 1.45, 적어도 1.50, 적어도 1.55, 적어도 1.60, 적어도 1.65, 적어도 1.70, 적어도 1.75, 또는 적어도 1.80의 상대적 위험도 또는 오즈비(odds ratio)인, 방법.

**청구항 13**

제1항 내지 제12항 중 어느 한 항에 있어서, 상기 개체는 1세 이하인, 방법.

**청구항 14**

제1항 내지 제13항 중 어느 한 항에 있어서, 상기 개체는 1일, 2일, 3일, 4일, 5일, 6일, 7일, 2주, 1개월, 3개월, 6개월 또는 1세인, 방법.

**청구항 15**

제1항 내지 제14항 중 어느 한 항에 있어서, 상기 방법은 개체가 약 2세, 약 3세, 약 4세 또는 약 5세; 또는 2세, 3세, 4세, 5세일 때 개체가 호흡기 감염에 대한 감수성이 증가됨을 나타내는 것인, 방법.

**청구항 16**

제1항 내지 제15항 중 어느 한 항에 있어서, 상기 방법은 개체가 적어도 5세일 때 개체가 호흡기 감염에 대한 감수성이 증가됨을 나타내는 것인, 방법.

**청구항 17**

제1항 내지 제16항 중 어느 한 항에 있어서, 상기 호흡기 감염은 하부 호흡 기도 감염인, 방법.

**청구항 18**

제1항 내지 제17항 중 어느 한 항에 있어서, 상기 호흡기 감염은 박테리아 또는 바이러스성 호흡기 감염인, 방법.

**청구항 19**

제18항에 있어서, 상기 바이러스성 호흡기 감염은 인플루엔자 속, 파라인플루엔자, 코로나바이러스, 아데노바이러스, 메타뉴모바이러스, 라이노바이러스 및 호흡기 세포융합 바이러스 호흡기 감염으로 이루어진 군으로부터 선택되고, 바람직하게는 바이러스성 호흡기 감염은 라이노바이러스 또는 호흡기 세포융합 바이러스 호흡기 감염인, 방법.

**청구항 20**

제18항에 있어서, 상기 박테리아성 호흡기 감염은 헤모필루스 인플루엔자(*Haemophilus Influenzae*), 스타필로코커스 아우레우스(*Staphylococcus aureus*) 또는 모락셀라 종(*Moraxella spp.*) 호흡기 감염인, 방법.

**청구항 21**

제1항 내지 제20항 중 어느 한 항에 있어서, 상기 호흡기 감염은 중증 하부 호흡 기도 감염인, 방법.

**청구항 22**

제1항 내지 제21항 중 어느 한 항에 있어서, 상기 차등적 증가를 갖는 바이오마커는 IFI6, IFI27, RHEBL1, CCDC194, BATF2, CARD16, IFIT1, ISG20, IFITM3, VAMP5, TNFSF13B, SAMD9, RNF213-AS1, IFIT2, XRN1, CD38, LRRN2 및 CCDC194 이외에 표 3 및 4에 나열된 것인, 방법.

**청구항 23**

제1항 내지 제22항 중 어느 한 항에 있어서, 상기 CBMC는 적어도 4, 6, 12, 18 또는 24시간 동안 TLR4 작용제와 접촉되었거나 접촉된 적이 있는 것인, 방법.

**청구항 24**

제1항 내지 제23항 중 어느 한 항에 있어서, 상기 바이오마커는 핵산 또는 증폭 산물인, 방법.

**청구항 25**

제1항 내지 제24항 중 어느 한 항에 있어서, 상기 방법은 호흡기 감염에 대한 감수성을 낮추는 치료제를 개체에 투여하는 단계를 추가로 포함하는 것인, 방법.

**청구항 26**

제25항에 있어서, 상기 치료제는 팔리비주맙, 프레드니솔론, 오말리주맙 또는 폴리박테리아 제형, 또는 이들의

임의의 조합인, 방법.

**청구항 27**

제1항 내지 제26항 중 어느 한 항에 있어서, 상기 샘플은 제대혈인, 방법.

**청구항 28**

제1항 내지 제26항 중 어느 한 항에 있어서, 상기 CBMC는 제대혈로부터 정제되는 것인, 방법.

**청구항 29**

제1항 내지 제28항 중 어느 한 항에 있어서, 상기 CBMC는 B 세포 및 T 세포, 바람직하게는 CD4+ 및/또는 CD8+ T 세포를 포함하는 것인, 방법.

**청구항 30**

제29항에 있어서, 상기 T 세포는 CD4+ 중심 기억 T 세포 및/또는 CD8+ 중심 기억 T 세포인, 방법.

**청구항 31**

제29항 또는 제30항에 있어서, 상기 B 및 T 세포 단독에서 바이오마커의 발현 수준은 CBMC에서 측정되는 것인, 방법.

**청구항 32**

제29항 내지 제31항 중 어느 한 항에 있어서, 상기 CBMC는 CD14+ 단핵구 및 통상형 수지상세포(cDC) 및 선택적으로 형질세포양 DC(pDC)를 추가로 포함하는 것인, 방법.

**청구항 33**

제1항 내지 제32항 중 어느 한 항에 있어서, 상기 TLR4 작용제는 지질다당류(LPS), 모노포스포릴 지질 A(MPLA), 열 충격 단백질, S100A8, S100A9, RSV F 단백질, 피브리노겐, 헤파린 설페이트 또는 이의 단편, 히알루론산 또는 이의 단편, 니켈, 오피오이드, α1-산성 당단백질(AAG), 아미노알킬 글루코사미나이드 4-포스페이트(AGP), RC-529, 뮤린 β-디펜신 2 및 완전 프로인트 보조제(CFA)로 이루어진 군으로부터 선택되고, 바람직하게는 TLR4 작용제는 LPS인, 방법.

**청구항 34**

제33항에 있어서, 상기 TLR4 작용제는 박테리아로부터 유래되고, 선택적으로 TLR4 작용제는 정제되거나 박테리아 제제에 함유되는 것인, 방법.

**청구항 35**

제33항 또는 제34항에 있어서, 상기 TLR4 작용제는 LPS이고, 이는 TLR4 활성화를 자극하기에 효과적인 농도로 제공되며, 바람직하게는 LPS의 농도는 0.025 ng/ml 내지 100 ng/ml이고, 보다 바람직하게는 LPS의 농도는 1 ng/ml인, 방법.

**청구항 36**

제1항 내지 제35항 중 어느 한 항에 있어서, 상기 CBMC는 TLR4 작용제와 접촉되기 전에  $1 \times 10^6$  세포/mL로 현탁되는 것인, 방법.

**청구항 37**

각각의 하기 바이오마커와 결합하거나 개별적으로 복합체화되는 진단 시약을 포함하는 키트, 패널 또는 마이크로어레이:

- KLHC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLR5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1;

- 표 3에 나열된 인터페론 모듈 바이오마커; 또는
- 표 4에 나열된 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커.

**청구항 38**

각각의 하기 바이오마커와 결합하거나 개별적으로 복합체화되는 진단 시약을 포함하는, 제1항 내지 제28항 중 어느 한 항의 방법에 따라 사용하기 위함이거나 사용될 때의 키트, 패널 또는 마이크로어레이:

- KLHC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1;

- 표 3에 나열된 인터페론 모듈 바이오마커; 또는
- 표 4에 나열된 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커.

**청구항 39**

하기를 포함하는 검정:

- 개체로부터의 제대혈에서의 B 및 T 세포를 TLR4 작용제와 접촉시키는 단계; 및
- B 및 T 세포에서 바이오마커 KLHC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1의 발현 수준을 측정하는 단계.

**청구항 40**

하기를 포함하는 검정:

- 개체로부터의 제대혈에서의 B 및 T 세포를 TLR4 작용제와 접촉시키는 단계; 및
- B 및 T 세포에서 인터페론 모듈 바이오마커의 발현 수준을 측정하는 단계.

**청구항 41**

하기를 포함하는 검정:

- 개체로부터의 제대혈에서의 B 및 T 세포를 TLR4 작용제와 접촉시키는 단계; 및
- B 및 T 세포에서 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커의 발현 수준을 측정하는 단계.

**청구항 42**

하기를 포함하는 검정:

- 개체로부터의 제대혈 단핵세포(CBMC)를 TLR4 작용제와 접촉시키는 단계; 및
- CBMC에서 바이오마커 KLHC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1의 발현 수준을 측정하는 단계.

**청구항 43**

하기를 포함하는 검정:

- 개체로부터의 제대혈 단핵세포(CBMC)를 TLR4 작용제와 접촉시키는 단계; 및
- CBMC에서 인터페론 모듈 바이오마커의 발현 수준을 측정하는 단계.

**청구항 44**

하기를 포함하는 검정:

- 개체로부터의 제대혈 단핵세포(CBMC)를 TLR4 작용제와 접촉시키는 단계; 및
- CBMC에서 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커의 발현 수준을 측정하는 단계.

## 발명의 설명

### 기술 분야

[0001] 본 발명은 생의 초기 단계에서 호흡기 감염, 예를 들어 중증 하부 호흡 기도 감염에 대해 민감하거나, 이에 대해 취약하거나, 이에 대한 소인을 갖거나, 이에 대한 증가된 위험을 갖는 개체의 식별을 위한 방법에 관한 것이다. 이는 선제적 요법의 형태로 개입에 대한 기회를 제공한다.

[0002] **관련 출원**

[0003] 본 출원은 2021년 10월 26일자로 출원된 호주 가출원 제2021903424호의 우선권을 주장하며, 이의 전체 내용은 그 전체가 본원에 참조로 포함된다.

### 배경 기술

[0004] 중증 하부 호흡 기도 감염(sLRI)은 유아 및 아동의 응급실 방문의 주요 원인이며, 천식의 발병 및 천명음에 대한 주요 위험 요소이다. 일련의 전향적 출생 코호트로부터의 연구는 sLRI 및 천식 간의 연관성이 라이노바이러스(RV) 천명음 및 초기 공기 알레르겐(aeroallergen) 감염과 함께 아동에서 가장 강력함이 밝혀졌다. 그러나, RV는 유의미한 증상의 부재 시 천식 아동에서 정례적으로 검출될 수 있으며, 이는 RV가 필요하지만 sLRI의 발병 기전을 유도하기에 충분하지 않을 수 있음을 시사한다. 이와 관련하여, 이는 바이러스 검출과 동시에 및/또는 이전에 모락셀라(*Moraxella*), 스트렙토코커스(*Streptococcus*) 및 헤모필루스(*Haemophilis*) 종을 포함한 박테리아성 병원체의 존재가 뒤따른 기도 증상을 현저하게 증폭시키고, 천식의 후속 발병에 대한 위험을 증가시킬 수 있음을 입증하였다. 반대로, 초기 아동기 동안 미생물과 이들의 산물에 대한 노출은 또한 천식에 대해 예방하는 것으로 나타났으며, 이는 아마도 "농가 효과(farm effect)"를 통해 가장 명쾌하게 기술될 수 있다.

[0005] 일부 개체는 생애 초기, 및 후속 천식에서 sLRI에 대해 보다 민감한 이유를 결정하는 근본적인 면역학적 메커니즘이 잘 이해되지 않는다.

[0006] 면역 시스템이 아직 완전히 발달되지 않은 시기에, sLRI에 대한 감수성이 증가된 개체, 특히 생의 초기-단계의 개체를 식별할 필요성이 남아 있다.

[0007] 본 명세서에서 임의의 선행 기술에 대한 언급은 이 선행 기술이 임의의 관할권에서 공통적인 일반 지식의 일부를 형성하거나, 이 선행 기술이 당업계의 숙련된 자에 의해 이해되고, 관련된 것으로 간주되고/되거나 선행 기술의 다른 부분과 조합될 것으로 합리적으로 예상할 수 있음을 인정하거나 제안하는 것이 아니다.

### 발명의 내용

[0008] **발명의 요약**

[0009] 본 발명은 생물학적 조절(전사체 및 프로테옴)의 다층에 걸친 자극(LPS, 폴리(I:C), 이미퀴모드)의 패널에 대한 제대혈의 선천적 면역 반응을 특성화하고, 생의 초기 단계에서 sLRI에 대한 위험과 연관된 선천적 면역 반응 패턴을 식별하는 시스템 생물학 접근법을 제공한다.

[0010] 다른 측면에서, 본 발명은 개체가 호흡기 감염에 대한 증가된 감수성을 갖는지 여부를 결정하기 위한 방법을 제공하며, 상기 방법은

[0011] - 개체로부터의 제대혈에서의 B 및 T 세포를 TLR4 작용제(agonist)와 접촉시키는 단계;

[0012] - CBMC에서 바이오마커 KLHC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1의 발현 수준을 측정하는 단계를 포함하며;

[0013] 여기서 TLR4 작용제와 접촉되지 않은 CBMC와 비교하여 TLR4 작용제와 접촉된 CBMC에서 바이오마커 KLHC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1의 차등적 발현은 개체가 호흡기 감염에 대한 증가된 감수성을 가짐을 나타낸다.

[0014] 임의의 측면에서, T 세포는 CD4+ 및/또는 CD8+ T 세포일 수 있다. 일 구현예에서, CD4+ T 세포는 중심 기억 세포이다. 일 구현예에서, CD8+ T 세포는 중심 기억 세포이다.

[0015] 본 발명은 개체가 호흡기 감염에 대한 증가된 감수성을 갖는지 여부를 결정하기 위한 방법을 제공하며, 상기 방

법은

- [0016] - 개체로부터의 제대혈 단핵세포(CBMC)를 TLR4 작용제와 접촉시키는 단계;
- [0017] - CBMC에서 바이오마커 KLHDC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1의 발현 수준을 측정하는 단계를 포함하며;
- [0018] 여기서 TLR4 작용제와 접촉되지 않은 CBMC와 비교하여 TLR4 작용제와 접촉된 CBMC에서 바이오마커 KLHDC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1의 차등적 발현은 개체가 호흡기 감염에 대한 증가된 감수성을 가짐을 나타낸다.
- [0019] 다른 측면에서, 본 발명은 개체가 호흡기 감염에 대한 증가된 감수성을 갖는지 여부를 결정하기 위한 방법을 제공하며, 상기 방법은
- [0020] - 개체로부터의 제대혈에서의 B 및 T 세포를 TLR4 작용제와 접촉시키는 단계;
- [0021] - CBMC에서 바이오마커 KLHDC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1의 발현 수준을 측정하는 단계;
- [0022] - 개체로부터의 바이오마커의 발현 수준을 참조 데이터 세트와 비교하는 단계로서, 상기 참조 데이터 세트는 호흡기 감염에 대한 증가된 감수성을 갖거나 갖지 않는 개체의 개체들 중 하나로부터의 TLR4 작용제와 접촉된 CBMC에서 동일한 바이오마커의 발현 수준에 대한 정보를 포함하는 것인, 단계를 포함하며;
- [0023] 여기서 참조 데이터 세트와 비교하여 개체에서 바이오마커 KLHDC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1의 발현 수준은 개체가 호흡기 감염에 대한 감수성이 증가됨을 나타낸다.
- [0024] 다른 측면에서, 본 발명은 개체가 호흡기 감염에 대한 증가된 감수성을 갖는지 여부를 결정하기 위한 방법을 제공하며, 상기 방법은
- [0025] - 개체로부터의 제대혈 단핵세포(CBMC)를 TLR4 작용제와 접촉시키는 단계;
- [0026] - CBMC에서 바이오마커 KLHDC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1의 발현 수준을 측정하는 단계;
- [0027] - 개체로부터의 바이오마커의 발현 수준을 참조 데이터 세트와 비교하는 단계로서, 상기 참조 데이터 세트는 호흡기 감염에 대한 증가된 감수성을 갖거나 갖지 않는 개체의 개체들 중 하나로부터의 TLR4 작용제와 접촉된 CBMC에서 동일한 바이오마커의 발현 수준에 대한 정보를 포함하는 것인, 단계를 포함하며;
- [0028] 여기서 참조 데이터 세트와 비교하여 개체에서 바이오마커 KLHDC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1의 발현 수준은 개체가 호흡기 감염에 대한 감수성이 증가됨을 나타낸다.
- [0029] 다른 측면에서, 본 발명은 개체가 호흡기 감염에 대한 증가된 감수성을 갖는지 여부를 결정하기 위한 방법을 제공하며, 상기 방법은
- [0030] - 개체로부터의 제대혈에서의 B 및 T 세포를 TLR4 작용제와 접촉시키는 단계;
- [0031] - 인터페론 모듈 바이오마커의 발현 수준을 측정하는 단계를 포함하며;
- [0032] 여기서 TLR4 작용제와 접촉되지 않은 B 및 T 세포와 비교하여 TLR4 작용제와 접촉된 B 및 T 세포에서 인터페론 모듈 바이오마커의 차등적 발현은 개체가 호흡기 감염에 대한 감수성이 증가됨을 나타낸다.
- [0033] 다른 측면에서, 본 발명은 개체가 호흡기 감염에 대한 증가된 감수성을 갖는지 여부를 결정하기 위한 방법을 제공하며, 상기 방법은
- [0034] - 개체로부터의 제대혈 단핵세포(CBMC)를 TLR4 작용제와 접촉시키는 단계;
- [0035] - 인터페론 모듈 바이오마커의 발현 수준을 측정하는 단계를 포함하며;
- [0036] 여기서 TLR4 작용제와 접촉되지 않은 CBMC와 비교하여 TLR4 작용제와 접촉된 CBMC에서 인터페론 모듈 바이오마커의 차등적 발현은 개체가 호흡기 감염에 대한 감수성이 증가됨을 나타낸다.
- [0037] 다른 측면에서, 본 발명은 개체가 호흡기 감염에 대한 증가된 감수성을 갖는지 여부를 결정하기 위한 방법을 제공하며,

공하며, 상기 방법은

- [0038] - 개체로부터의 제대혈에서의 B 및 T 세포를 TLR4 작용제와 접촉시키는 단계;
- [0039] - B 및 T 세포에서 인터페론 모듈 바이오마커의 발현 수준을 측정하는 단계;
- [0040] - 개체로부터의 바이오마커의 발현 수준을 참조 데이터 세트와 비교하는 단계로서, 상기 참조 데이터 세트는 호흡기 감염에 대한 증가된 감수성을 갖거나 갖지 않는 개개의 개체들 중 하나로부터의 TLR4 작용제와 접촉된 B 및 T 세포에서 동일한 바이오마커의 발현 수준에 대한 정보를 포함하는 것인, 단계를 포함하며;
- [0041] 여기서 참조 데이터 세트와 비교하여 개체에서 인터페론 모듈 바이오마커의 발현 수준은 개체가 호흡기 감염에 대한 감수성이 증가됨을 나타낸다.
- [0042] 다른 측면에서, 본 발명은 개체가 호흡기 감염에 대한 증가된 감수성을 갖는지 여부를 결정하기 위한 방법을 제공하며, 상기 방법은
- [0043] - 개체로부터의 제대혈 단핵세포(CBMC)를 TLR4 작용제와 접촉시키는 단계;
- [0044] - CBMC에서 인터페론 모듈 바이오마커의 발현 수준을 측정하는 단계;
- [0045] - 개체로부터의 바이오마커의 발현 수준을 참조 데이터 세트와 비교하는 단계로서, 상기 참조 데이터 세트는 호흡기 감염에 대한 증가된 감수성을 갖거나 갖지 않는 개개의 개체들 중 하나로부터의 TLR4 작용제와 접촉된 CBMC에서 동일한 바이오마커의 발현 수준에 대한 정보를 포함하는 것인, 단계를 포함하며;
- [0046] 여기서 참조 데이터 세트와 비교하여 개체에서 인터페론 모듈 바이오마커의 발현 수준은 개체가 호흡기 감염에 대한 감수성이 증가됨을 나타낸다.
- [0047] 임의의 측면에서, 인터페론 모듈 바이오마커는 표 3에 나열된 것들이다.
- [0048] 다른 측면에서, 본 발명은 개체가 호흡기 감염에 대한 증가된 감수성을 갖는지 여부를 결정하기 위한 방법을 제공하며, 상기 방법은
- [0049] - 개체로부터의 제대혈에서의 B 및 T 세포를 TLR4 작용제와 접촉시키는 단계;
- [0050] - IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커의 발현 수준을 측정하는 단계를 포함하며;
- [0051] 여기서 TLR4 작용제와 접촉되지 않은 B 및 T 세포와 비교하여 TLR4 작용제와 접촉된 B 및 T 세포에서 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커의 차등적 발현은 개체가 호흡기 감염에 대한 감수성이 증가됨을 나타낸다.
- [0052] 다른 측면에서, 본 발명은 개체가 호흡기 감염에 대한 증가된 감수성을 갖는지 여부를 결정하기 위한 방법을 제공하며, 상기 방법은
- [0053] - 개체로부터의 제대혈 단핵세포(CBMC)를 TLR4 작용제와 접촉시키는 단계;
- [0054] - IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커의 발현 수준을 측정하는 단계를 포함하며;
- [0055] 여기서 TLR4 작용제와 접촉되지 않은 CBMC와 비교하여 TLR4 작용제와 접촉된 CBMC에서 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커의 차등적 발현은 개체가 호흡기 감염에 대한 감수성이 증가됨을 나타낸다.
- [0056] 다른 측면에서, 본 발명은 개체가 호흡기 감염에 대한 증가된 감수성을 갖는지 여부를 결정하기 위한 방법을 제공하며, 상기 방법은
- [0057] - 개체로부터의 제대혈에서의 B 및 T 세포를 TLR4 작용제와 접촉시키는 단계;
- [0058] - B 및 T 세포에서 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커의 발현 수준을 측정하는 단계;
- [0059] - 개체로부터의 바이오마커의 발현 수준을 참조 데이터 세트와 비교하는 단계로서, 상기 참조 데이터 세트는 호흡기 감염에 대한 증가된 감수성을 갖거나 갖지 않는 개개의 개체들 중 하나로부터의 TLR4 작용제와 접촉된 B 및 T 세포에서 동일한 바이오마커의 발현 수준에 대한 정보를 포함하는 것인, 단계를 포함하며;
- [0060] 여기서 참조 데이터 세트와 비교하여 개체에서 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커의 발현 수준은 개체가 호흡기 감염에 대한 감수성이 증가됨을 나타낸다.
- [0061] 다른 측면에서, 본 발명은 개체가 호흡기 감염에 대한 증가된 감수성을 갖는지 여부를 결정하기 위한 방법을 제

공하며, 상기 방법은

- [0062] - 개체로부터의 제대혈 단핵세포(CBMC)를 TLR4 작용제와 접촉시키는 단계;
- [0063] - CBMC에서 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커의 발현 수준을 측정하는 단계;
- [0064] - 개체로부터의 바이오마커의 발현 수준을 참조 데이터 세트와 비교하는 단계로서, 상기 참조 데이터 세트는 호흡기 감염에 대한 증가된 감수성을 갖거나 갖지 않는 개체의 개체들 중 하나로부터의 TLR4 작용제와 접촉된 CBMC에서 동일한 바이오마커의 발현 수준에 대한 정보를 포함하는 것인, 단계를 포함하며;
- [0065] 여기서 참조 데이터 세트와 비교하여 개체에서 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커의 발현 수준은 개체가 호흡기 감염에 대한 감수성이 증가됨을 나타낸다.
- [0066] 임의의 측면에서, 바이오마커의 발현 수준은 절대적 수준 또는 차등적 발현 수준일 수 있다.
- [0067] 임의의 측면에서, IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커는 표 4에 나열된 것들이다.
- [0068] 임의의 측면에서, 상기 방법은 기계 학습 알고리즘, 바람직하게는 랜덤 포레스트 분석을 바이오마커의 차등적 발현 또는 절대적 발현의 수준에 적용하고, 이에 따라 개체가 호흡기 감염에 대한 고위험의 감수성을 나타내는 단계를 추가로 포함한다.
- [0069] 임의의 측면에서, 개체는 1세 미만 또는 2세 이하이다. 임의의 구현예에서, 개체는 1일, 2일, 3일, 4일, 5일, 6일, 7일, 2주, 1개월, 2개월, 3개월, 6개월, 1세 또는 2세이다. 임의의 구현예에서, 개체는 적어도 1일, 적어도 2일, 적어도 3일, 적어도 4일, 적어도 5일, 적어도 6일, 적어도 7일, 적어도 2주, 적어도 1개월, 적어도 2개월, 적어도 3개월, 적어도 6개월, 적어도 1세 이상, 2세 이하이다. 임의의 구현예에서, 개체는 약 1일 내지 약 7일, 약 1일 내지 약 2주, 약 1주 내지 약 6개월, 약 1개월 내지 약 6개월, 약 1개월 내지 약 3개월, 약 6개월 내지 약 1세, 약 1개월 내지 2세, 약 6개월 내지 2세, 약 1세 내지 2세이다.
- [0070] 임의의 측면에서, 상기 방법은 개체가 약 2세, 약 3세, 약 4세 또는 약 5세; 또는 2세, 3세, 4세, 5세일 때 호흡기 감염에 대한 증가된 감수성을 결정한다. 임의의 구현예에서, 상기 방법은 개체가 약 2세 내지 약 5세, 약 2세 내지 약 4세, 약 2세 내지 약 3세, 약 3세 내지 약 5세 또는 약 4세 내지 약 5세일 때 호흡기 감염에 대한 증가된 감수성을 결정한다.
- [0071] 임의의 측면에서, 상기 방법은 개체가 적어도 5세일 때 호흡기 감염에 대한 증가된 감수성을 결정한다.
- [0072] 임의의 측면에서, 상기 호흡기 감염은 하부 호흡 기도 감염이다. 임의의 측면에서, 상기 호흡기 감염은 박테리아 또는 바이러스성 호흡기 감염이다. 임의의 측면에서, 상기 호흡기 감염은 중증 하부 호흡 기도 감염이다. 임의의 구현예에서, 상기 바이러스성 호흡기 감염은 라이노바이러스 및 호흡기 세포융합 바이러스(RSV)를 포함하고, 상기 박테리아성 호흡기 감염은 헤모필루스 인플루엔자(*Haemophilus Influenzae*), 스타필로코쿠스 아우레우스(*Staphylococcus aureus*) 및 모락셀라 종(*Moraxella spp.*)를 포함한다.
- [0073] 임의의 측면에서, 차등적 발현은 1.5배 초과와 증가 또는 감소이다. 임의의 구현예에서, 차등적 감소를 갖는 바이오마커는 IFI6, IFI27, RHEBL1, CCDC194, BATF2, CARD16, IFIT1, ISG20, IFITM3, VAMP5, TNFSF13B, SAMD9, RNF213-AS1, IFIT2, XRN1, CD38, LRRN2 및 CCDC194를 포함한다. 임의의 구현예에서, 차등적 증가를 갖는 바이오마커는 IFI6, IFI27, RHEBL1, CCDC194, BATF2, CARD16, IFIT1, ISG20, IFITM3, VAMP5, TNFSF13B, SAMD9, RNF213-AS1, IFIT2, XRN1, CD38, LRRN2 및 CCDC194 이외에 표 3 및 4에 나열된 것을 비롯하여, 본원에 기술된 바이오마커를 포함한다.
- [0074] 다른 측면에서, 본 발명은 TLR4 작용제로 자극된 CBMC에서 바이오마커 KLHDC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLR5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1의 발현 수준을 측정하는 방법을 제공한다.
- [0075] 다른 측면에서, 본 발명은 TLR4 작용제로 자극된 B 및 T 세포, 또는 CBMC에서 인터페론 모듈 바이오마커의 발현 수준을 측정하는 방법을 제공한다.
- [0076] 다른 측면에서, 본 발명은 TLR4 작용제로 자극된 B 및 T 세포, 또는 CBMC에서 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커의 발현 수준을 측정하는 방법을 제공한다.
- [0077] 다른 측면에서, 본 발명은 하기를 포함하는 검정(assay)을 제공한다:

- [0078] - 개체로부터의 제대혈에서의 B 및 T 세포를 TLR4 작용제와 접촉시키는 단계; 및
- [0079] - B 및 T 세포에서 바이오마커 KLHDC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1의 발현 수준을 측정하는 단계.
- [0080] 다른 측면에서, 본 발명은 하기를 포함하는 검정을 제공한다:
- [0081] - 개체로부터의 제대혈에서의 B 및 T 세포를 TLR4 작용제와 접촉시키는 단계; 및
- [0082] - B 및 T 세포에서 인터페론 모듈 바이오마커의 발현 수준을 측정하는 단계.
- [0083] 다른 측면에서, 본 발명은 하기를 포함하는 검정을 제공한다:
- [0084] - 개체로부터의 제대혈에서의 B 및 T 세포를 TLR4 작용제와 접촉시키는 단계; 및
- [0085] - B 및 T 세포에서 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커의 발현 수준을 측정하는 단계.
- [0086] 다른 측면에서, 본 발명은 하기를 포함하는 검정을 제공한다:
- [0087] - 개체로부터의 제대혈 단핵세포(CBMC)를 TLR4 작용제와 접촉시키는 단계; 및
- [0088] - CBMC에서 바이오마커 KLHDC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1의 발현 수준을 측정하는 단계.
- [0089] 다른 측면에서, 본 발명은 하기를 포함하는 검정을 제공한다:
- [0090] - 개체로부터의 제대혈 단핵세포(CBMC)를 TLR4 작용제와 접촉시키는 단계; 및
- [0091] - CBMC에서 인터페론 모듈 바이오마커의 발현 수준을 측정하는 단계.
- [0092] 다른 측면에서, 본 발명은 하기를 포함하는 검정을 제공한다:
- [0093] - 개체로부터의 제대혈 단핵세포(CBMC)를 TLR4 작용제와 접촉시키는 단계; 및
- [0094] - CBMC에서 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커의 발현 수준을 측정하는 단계.
- [0095] 다른 측면에서, 본 발명은 TLR4 작용제와 접촉된 CBMC에서 바이오마커 KLHDC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1의 발현 수준을 측정하는 단계를 포함하는 방법을 제공한다.
- [0096] 다른 측면에서, 본 발명은 TLR4 작용제와 접촉된 B 및 T 세포, 또는 CBMC에서 인터페론 모듈 바이오마커의 발현 수준을 측정하는 단계를 포함하는 방법을 제공한다.
- [0097] 다른 측면에서, 본 발명은 TLR4 작용제와 접촉된 B 및 T 세포, 또는 CBMC에서 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커의 발현 수준을 측정하는 단계를 포함하는 방법을 제공한다.
- [0098] 임의의 측면에서, B 및 T 세포, 또는 CBMC는 적어도 4, 6, 12, 18 또는 24시간 동안 TLR4 작용제와 접촉되었거나 접촉된 적이 있다.
- [0099] 임의의 측면에서, CBMC는 TLR4 작용제와 접촉되었거나 접촉된 다음, B 및 T 세포 단독에서 관련 바이오마커의 발현 수준이 측정되거나 결정된다.
- [0100] 임의의 측면 또는 구현예에서, 바이오마커는 단백질, 핵산, 예를 들어 RNA 또는 증폭 산물이다. 상기 바이오마커가 핵산 또는 증폭 산물인 경우, 상기 방법은 유전자 또는 RNA의 발현 수준 또는 양을 결정하는 단계를 포함한다. 바람직하게는, 바이오마커는 유전자 또는 유전자의 RNA 전사체로부터의 뉴클레오티드 서열을 포함하는 하나 이상의 핵산이다.
- [0101] 본원에 기술된 임의의 측면 또는 구현예에서, RNA 또는 이의 증폭 산물에 대한 언급이 있는 경우, 본 발명은 또한 경우에 따라 (RNA로부터 번역된) 상응하는 단백질의 존재, 수준 또는 양을 결정하거나 측정하는 단계를 포함한다.
- [0102] 본 발명의 임의의 측면에서, 하나 이상의 바이오마커의 수준 또는 양은 RNA의 수준 또는 양일 수 있다. 바람직하게는, 상기 RNA는 전(pre)-mRNA 또는 성숙 mRNA 중 어느 하나이고, 상기 RNA의 수준 또는 양에 대한 변화는 RNA 시퀀싱을 비롯하여, 본원에 기술된 임의의 방법을 사용하여 결정될 수 있다.

- [0103] 임의의 측면에서, 상기 방법은 호흡기 감염에 대한 감수성을 낮추는 치료제를 개체에게 투여하는 단계를 추가로 포함한다. 바람직하게는 치료제는 팔리비주맙, 프레드니솔론, 오말리주맙 또는 폴리박테리아 제형, 또는 이들의 임의의 조합이다.
- [0104] 임의의 측면에서, B 및 T 세포, 또는 CBMC는 제대혈로부터 정제될 수 있다. 대안적으로, 임의의 측면에서, B 및 T 세포, 또는 CBMC는 개체로부터의 제대혈 샘플에 존재할 수 있다. 임의의 구현예에서, B 및 T 세포, 또는 CBMC는 제대혈에 존재하므로, B 및 T 세포 또는 CBMC를 TLR4 작용제와 접촉시키는 단계에 대한 본원의 임의의 참조는 B 및 T 세포 또는 CBMC를 함유하는 제대혈 (예를 들어, 미처리, 전체 또는 미정제된 제대혈)을 TLR4 작용제와 접촉시키는 단계를 포함한다. 일 구현예에서, 제대혈은 적혈구가 고갈될 수 있다. 일 구현예에서, 적혈구는 B 및 T 세포, 또는 CBMC가 TLR4 작용제와 접촉되는 경우, 존재하지 않거나 유의미한 수준으로 존재하지 않는다. 또 다른 구현예에서, 적혈구는 정상 수준으로 존재하며, 즉, 이들은 제대혈로부터 고갈되지 않았다.
- [0105] 임의의 측면에서, CBMC는 림프구(T 및 B 세포)를 포함하거나 이로 구성된다. 바람직하게는, CBMC는 CD14+ 단핵구 및 통상형 수지상세포(cDC) 및 선택적으로 형질세포양 DC(pDC)를 추가로 포함한다. 임의의 구현예에서, CBMC는 CD4+ T 세포, CD8+ T 세포 및 B 세포를 포함하거나 이로 구성된다.
- [0106] 임의의 측면에서, TLR4 작용제는 본원에 기술된 임의의 하나이다. 바람직하게는, TLR4 작용제는 박테리아로부터 유래된다. 보다 바람직하게는, TLR4 작용제는 LPS이다. 임의의 구현예에서, LPS는 정제될 수 있다. 대안적으로, LPS는 박테리아 제제에 함유될 수 있다.
- [0107] 다른 측면에서, 본 발명은 본원에 기술된 적어도 2개의 진단 시약을 포함하는 키트, 패널 또는 마이크로어레이를 제공하며, 각 시약은 본원에 기술된 상이한 바이오마커를 식별한다. 일 구현예에서, 키트는 2개, 3개, 4개, 5개, 6개, 7개, 8개, 9개, 10개, 11개, 12개, 13개, 14개 또는 그 이상의 바이오마커와 결합하거나 개별적으로 복합체화되는 진단 시약을 포함한다. 일 구현예에서, 키트는 각각의 하기 바이오마커와 결합하거나 개별적으로 복합체화되는 진단 시약을 포함할 수 있다:
- [0108] - KLHC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1;
- [0109] - 표 3에 나열된 인터페론 모듈 바이오마커; 또는
- [0110] - 표 4에 나열된 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커.
- [0111] 다른 측면에서, 본원에 기술된 적어도 2개의 진단 시약을 포함하는, 본 발명의 방법에 따라 사용되거나 사용될 때의 키트가 제공되며, 각 시약은 본원에 기술된 상이한 바이오마커를 식별한다. 일 구현예에서, 키트는 2개, 3개, 4개, 5개, 6개, 7개, 8개, 9개, 10개, 11개, 12개, 13개, 14개 또는 그 이상의 바이오마커와 결합하거나 개별적으로 복합체화되는 진단 시약을 포함한다. 일 구현예에서, 키트는 각각의 하기 바이오마커와 결합하거나 개별적으로 복합체화되는 진단 시약을 포함할 수 있다:
- [0112] - KLHC7B, IFNG, CASZ1, PSMB9, PARP3, ACOT7, NUB1, USP18, NLRC5, CCDC194, GCH1, PARP11, CXCL11 및 PMAIP1;
- [0113] - 표 3에 나열된 인터페론 모듈 바이오마커; 또는
- [0114] - 표 4에 나열된 IRF1 레귤론에 의해 조절되는 IFN 모듈 바이오마커.
- [0115] 본원에 사용된 바와 같이, 문맥상 달리 요구되는 경우를 제외하고, 용어 "포함한다" 및 "포함하는", "포함하다" 및 "포함되는"과 같은 용어의 변형은 추가의 첨가제, 성분, 정수 또는 단계를 배제하려는 의도가 아니다.
- [0116] 본 발명의 추가 측면 및 이전 단락에 기술된 측면의 추가 구현예는 예로서 그리고 첨부 도면을 참조하여 주어진, 다음의 설명으로부터 명백해질 것이다.

### 도면의 간단한 설명

- [0117] **도 1.** 다중-오믹스 데이터세트의 차원성 감소. (a) 실험 및 분석 설계의 개략적인 표현. (b) 기준선 CBMC 샘플의 면역표현형. Y-축은 CBMC로부터 식별된 총(total) 세포 유형 중 세포 유형 비율을 나타낸다. 산점도는 95% CI를 갖는 중앙값을 나타낸다. (c) 유전자 발현(PCA), 사이토카인(PCA) 데이터세트에 대한 다단계 차원성 감소. 축은 제1 (x) 및 제2 (y) 주성분 또는 정준 변수에 설명된 총 변이(%)의 비율을 나타낸다. (d-e) 제1 (맨 위 행) 및 제2 (맨 아래 행) 주성분에 대한 상위 기여 특징 또는 상응하는 (상기) 차원성 감소 플롯에 대한 교차

검증된 정준 변수를 나타내는 세로 막대 플롯. x-축은 절대 기여도(%) / 로딩을 나타내며; 적색은 양수/높음을 나타내고 청색은 음수/낮음을 나타낸다.

**도 2.** 인터페론 및 전염증성 유전자 발현이 식별된 차등적 발현 및 네트워크 분석은 선천적 CBMC 반응을 특징화한다. (a-c; 좌측 패널) LPS, 이미퀴모드 및 폴리(I:C) 반응에 대해 각각 일치된 비자극 샘플과 비교하여 유의미하게 상향조절(우측; 적색) 및 하향조절된(좌측; 청색) 유전자를 나타내는 볼케이노 플롯(Volcano plot). 화살표는 상향조절 및 하향조절된 유전자의 수를 나타낸다. (a-c; 우측 패널) 일치된 비자극 대조군과 비교하여 각각 CBMC LPS, 이미퀴모드 및 폴리(I:C) 반응에 대해 유의미하게 상향조절된 유전자로부터 과도하게 표현되는 (overrepresented) 경로. (d-f) LPS, 이미퀴모드 및 폴리(I:C) 반응의 네트워크 분석(WGCNA)으로부터 각각 식별된 모듈. 그룹은 맨 위에서 맨 아래로 나타낸 바와 같이 주요 지표 순이다. 모듈은 조절된 t-통계량 (limma/voom)으로 플롯팅되고, 중앙값, 제25 및 제75 사분위수,  $\pm 1.5 \times IQR$  및 이상치(outlier)를 나타낸다. 적색 선(맨 윗선)(조절된 t-통계량 = 2) 위에 중앙값을 갖는 모듈은 유의미하게 상향조절된 것으로 간주되며, 청색 선(맨 아랫선)(-2) 아래의 것들은 유의미하게 하향조절된 것으로 간주된다. 모듈은 범례와 동일한 순서로 좌측에서 우측으로 표지된다. (g) 각각의 반응에 대한 인터페론 및 전염증성 모듈에서의 유전자 수의 막대 플롯. (h) CBMC 반응 모듈 유전자 간의 랭크된 발현 및 연결성(connectivity)에 대한 스피어만의 로우 (Spearman's Rho) 값을 나타내는 히트맵. 각 반응의 인터페론 및 전염증성 모듈로부터의 구성원 유전자의 발현은 다른 반응으로부터의 동일한 유전자의 발현에 대해 상관관계를 보였다. 모든 상관관계와 관련된 p 값은 < 0.01이었다.

**도 3.** IFN 모듈 유전자 연결성 및 CBMC 반응 및 출생 5세 연령 사이의 드라이버. (a) LPS, 이미퀴모드 및 폴리(I:C) CBMC 반응 IFN 모듈 연결성 각각에 대한 밀도 플롯. 점선은 평균(밝은 회색) 및 중앙값(어두운 회색)을 나타낸다. 릴리포스(Liliefors) p 값 > 0.05는 정규적으로 분포된 연결성을 나타낸다. (b-d; 좌측 패널) LPS, 이미퀴모드 및 폴리(I:C) CBMC IFN 모듈 각각에 대한 상위 20개의 가장 많이 연결된 유전자의 네트워크 배선 다이어그램. 노드(node) 크기는 총 네트워크 중에 연결 수(정도)를 나타내며, 에지(edge)는 연결 강도를 나타낸다 (적색 에지, 즉 더 어두운 선은 상관관계 > 0.8을 나타냄). (b-d; 우측 패널) LPS, 이미퀴모드 및 폴리(I:C) CBMC IFN 모듈에 대한 VIPER 분석에 의해 식별된 상위 10개의 핵심 조절자(master regulator). 막대 플롯은 유의미하게 활성화(NES>2, 적색 선) 또는 비활성/억제된(NES<-2, 청색 선) 전사 인자에 대한 정규화된 농축 점수 (NES, normalized enrichment score)를 나타낸다. 회색 음영은 조정된 P 값 < 0.05를 나타낸다. (e) 일치된 CBMC(i) 및 5세 PBMC(ii) 샘플로부터 상위 20개의 가장 많이 연결된 제대혈 LPS-유도 IFN 모듈 유전자의 네트워크 배선 다이어그램. 네트워크 특성은 상기와 동일하다(도 3b-d). (f) LPS 자극에 대해 일치된 CBMC 및 5세 PBMC 반응의 제대혈 인터페론 모듈 유전자 연결성에 대한 네트워크 연결성 밀도 플롯. (g) 일치된 CBMC (i) 및 5세 PBMC (ii, n=9 유의미한 드라이버) 샘플에 대해 식별된 제대혈 인터페론 모듈 유전자의 상위 유의미한 드라이버. 막대 플롯 특성은 상기와 동일하다(도 3c). (h&i) 이미퀴모드 및 폴리(I:C) 자극에 대해 각각 일치된 CBMC 및 5세 PBMC 반응의 제대혈 인터페론 모듈 유전자 연결성에 대한 네트워크 연결성 밀도 플롯.

**도 4.** LPS-유도 IFN 유전자는 출생 시 sLRI 감수성을 예측한다. (a) 랜덤 포레스트 분류기는 랜덤으로 선택된 25명(50%) 연구 대상체로부터의 LPS-, 이미퀴모드 및 폴리(I:C)-유도 IFN 모듈 유전자에 대해 훈련되었고, 나머지 25명(50%) 대상체에서 검증되었다. 각 RF 모델은 각 분할에서 사용된 유전자 수 및 성장된 트리(tree) 수에 대해 최적화되었다. 플롯은 거짓(x-축, 1-특이성) 및 참(y-축, 민감성) 양성 비율로 정의된 수신자 조작 특성 (ROC) 곡선하 면적을 나타낸다. (b) RF 모델 예측은 훈련/유효성 검증 세트(50/50 랜덤 할당)를 2,000회 재-샘플링함으로써 반복하였다. 플롯은 중앙값(실선) 및 95% CI(점선)과 함께, 각 재-샘플에 대한 ROC 곡선하 면적을 나타낸다. (c) LPS-(i), 이미퀴모드-(ii) 및 폴리(I:C)-(iii) 유도 IFN 모듈 유전자 네트워크에 대한 네트워크 연결성 밀도 플롯은 생후 첫 해에 sLRI를 기록하고(주황색; 즉 밝은 선) 기록하지 않은(회색) 개체에 의해 계층화된다. (d-e(i)) 유아기에 sLRI에 대해 저항성(D(i)) 및 감수성(E(i))인 개체로부터의 CBMC 샘플에서 LPS-유도 IFN 모듈 유전자 네트워크 중 상위 20개의 가장 많이 연결된 유전자의 네트워크 배선 다이어그램. 노드 및 에지 특성은 도 3b와 동일하다. (d-e(ii)) 유아기에 sLRI에 대해 저항성(D(ii)) 및 감수성(E(ii))인 개체를 위한 CBMC LPS-유도 IFN 반응 모듈에 대한 VIPER 분석에 의해 식별된 상위 10개의 핵심 조절자. 막대 플롯 특성은 도 3c와 동일하다. (f) 생후 첫 해에 sLRI를 기록하고(주황색; 즉 밝은 선) 기록하지 않은(회색) 개체로 그룹화된, 제대혈 LPS-유도 IFN 모듈 고유 유전자의 박스(box) 및 위스커(whisker) 플롯. 박스는 중앙값, 제25 및 제75 사분위수,  $\pm 1.5 \times IQR$  및 이상치(outlier)를 나타내고; P 값은 맨-휘트니(Mann-Whitney) U 테스트에 의해 결정된다. (g) 좌측에서 우측으로 IFN 모듈 고유 유전자의 플롯: LPS(녹색; 그룹 1, 2, 7 및 8), 이미퀴모드(청색; 그룹 3, 4, 9 및 10) 및 폴리(I:C)(적색; 그룹 5, 6, 11 및 12) 유아기에 LRI 및 sLRI에 대해 저항성(-) 및 감수성(+)인 개체에 의해 그룹화된 CBMC 반응. 맨-휘트니 U 테스트에 의해 결정된 P 값 및 유의미한 결과는 도 4g

를 반영한다. 플롯은 중앙값(심볼) 및 95% CI(막대)를 나타낸다.

**도 5.** 외부 유전자 발현 데이터세트에서의 시험관내 CBMC 배양 IFN 모듈 유전자의 검증, LPS-유도 생물학적 특성의 다중-오믹 통합 및 IRF1 유전자 발현 상관관계. (a-c) 랜덤 포레스트 분류기는 비자극 및 LPS 또는 이미퀴모드/폴리(I:C) CBMC 유전자 발현 데이터(n=50)에 대해 훈련되었으며 예측에 사용되었다; (a) 혈액-유래 유전자 발현 프로파일로부터의 건강한 대조군(n=52)으로부터 각각 박테리아(n=52) 및 바이러스 감염(n=92)으로 입원한 아동(<17세), (b) PBMC 샘플로부터의 회복기와 비교하여 급성 바이러스 호흡기 감염으로 병원에 방문한 유아(<18개월; n=15) 및 아동(18개월-5세; n=16); 및 (c) 비강-유래 유전자 발현 프로파일로부터 검출 가능한 기도 바이러스 감염이 있고(n=193) 없었(n=105), 감기-유사 증상을 가진 천식 아동(6-17세). 플롯은 ROC 곡선하 면적을 나타낸다. (d) 다중-오믹 데이터 통합으로부터 결정된 유아기에 sLRI 감수성에 대한 다중-층 위험 프로파일. 층 간 공동-발현이 최대화되었으며, 각각 ±0.8보다 강한 양성(적색) 및 음성(청색) 상관관계가 나타났다. 주변선은 생후 첫 해에 sLRI에 대해 저항성(회색) 또는 감수성(주황색; 즉 밝은 선)인 개체로부터의 특징의 상대적 발현을 나타낸다. 입력 데이터는 일치된 비자극 샘플에 대해 조정되었다(기준선 면역표현형 데이터 제외). "R<sub>-</sub>"을 유전자 명칭(청색)과 구별하기 위해 전사 인자 ID(녹색)에 추가하였다. (e-g) LPS-유도 IRF1 유전자 발현과 IFN 및 전염증성 유전자(g), 바이러스-관련 수용체(h) 및 케모/사이토카인(i) 간의 연관성에 대한 플롯. 데이터는 일치된 비자극 샘플에 대해 조정되었고, 플롯은 스피어만의 로우 값(심볼) 및 95% CI(막대, 1000 부트스트랩)를 나타내며; 적색 및 청색 데이터 점(point)/표지는 BH-조정된 FDR < 0.05인 양성 및 음성 상관관계(즉, 비-중첩된 95% 신뢰 구간은 각각 0 초과 또는 미만임)를 나타낸다. 본 명세서에 개시되고 정의된 본 발명은 텍스트 또는 도면에서 언급되거나 명백한 개별 특징 중 2개 이상의 모든 대안적인 조합으로 확장된 것으로 이해될 것이다. 이러한 모든 상이한 조합은 본 발명의 다양한 대안적인 측면을 구성한다.

**도 6.** (a) LPS에 대한 CBMC 반응의 IFN 모듈 내에 포함된 유전자에 대해 선택된 10개의 유의미하게 과도하게 표현되는 경로(InnateDB). (b) LPS CBMC 반응 인터페론 모듈의 원래 VIPER 플롯(모티프 결합 부위로 트리밍 전). 플롯은 (L-R) p 값, 양성(우측; 적색) 및 음성(좌측; 청색) 상호작용, 전사 인자 유전자 명칭, 활성화(상단 패널; 적색) 또는 비활성화/억제된(하단 패널, 청색) 상태(NES), 및 TF 유전자의 상대적 발현을 나타낸다.

**도 7.** 입력 유전자가 LPS-유도 IFN 모듈에 공통적인 유전자로만 제한되는 경우 제대혈 CBMC 반응의 밀도 플롯 및 추정 드라이버 분석(이전에 기술된 바와 같음).

**도 8.** 일치된 CBMC(좌측) 및 5세 PBMC(우측) 샘플에 대한 CBMC 이미퀴모드(a) 및 폴리(I:C)(b) 유도 IFN 모듈용 드라이버.

**도 9.** 각각 CBMC LPS-(a), 이미퀴모드-(b) 및 폴리(I:C)-유도(c) IFN 모듈 랜덤 포레스트 분류기에 대해 가장 중요한 유전자인 상위 30개 유전자. x-축은 각 변수를 제외하여 각 모델에 대한 정확도 손실을 나타낸다.

**도 10.** (a, b) RF 모델 예측은 일관성을 유지하는 각 모델에 대해 원래 (최적화된) RF 파라미터를 사용하여 훈련/유효성 검증 세트(50/50 랜덤 할당, 2000회 재-샘플)를 재-샘플링함으로써 반복하였다. 밀도로서 각 재-샘플에 대한 AUC-ROC(a). 실선은 AUC-ROC 중앙값을 나타내고, 점선은 95% 신뢰 구간의 높고 낮음을 나타낸다. 이들 값은 다음과 같다: LPS: 중앙값=0.619,  $CI^{lo}$ =0.615,  $CI^{hi}$ =0.623; 이미퀴모드: 중앙값=0.338,  $CI^{lo}$ =0.335,  $CI^{hi}$ =0.341; 폴리(I:C): 중앙값=0.393,  $CI^{lo}$ =0.389,  $CI^{hi}$ =0.397. 회색 점선은 0.5의 AUC-ROC(랜덤 확률)를 나타낸다. (b) 재-샘플링된 AUC-ROC 값은 박스-및-위스커 플롯으로 표시된다. 일치된 재-샘플 순열은 회색 선으로 연결된다. 일치된 재-샘플 AUC-ROC 값은 모든 그룹 간에 유의미하게 상이하였다(Wilcoxon SRT p 값 < 0.00001).

**도 11.** (a) 유아기에 sLRI에 대해 저항성 및 감수성인 개체들 간의 별도의 연결성 측정(스피어만의 로우)을 사용한 평가에 의해 확인된 CBMC LPS-유도 IFN 모듈 차등적 연결성. (b,c) 입력 유전자를 LPS-유도 IFN 모듈에 공통적인 것들로만 제한한 후 이미퀴모드-유도(b) 및 폴리(I:C)-유도(c) IFN 모듈 유전자 네트워크에 대한 네트워크 연결성 밀도 플롯은 생후 첫 해에 sLRI를 기록하고(주황색; 즉 밝은 선) 기록하지 않은(회색) 개체에 의해 계층화된다

**도 12.** (a) 5세 연령에 천식/비-천식인 개체에 의해 그룹화된, 제대혈 LPS-유도 IFN 모듈 고유 유전자의 박스 및 위스커 플롯. (b) 생후 5년째에 천명음이 있거나 없는 개체에 의해 그룹화된, 제대혈 LPS-유도 IFN 모듈 고유 유전자의 박스 및 위스커 플롯.

**도 13.** (a) 열성(febrile) 박테리아 및 바이러스 감염으로 입원한 아동의 혈액-유래 유전자 발현 프로파일의 주

성분 분석(GSE72809). 유전자 발현 데이터 세트는 각각 이용 가능한 CBMC LPS-유도, 이미퀴모드-유도 및 폴리(I:C)-유도 IFN 모듈 유전자로 제한되었다. (b) 각각 CBMC LPS-, 이미퀴모드- 및 폴리(I:C)-유도 IFN 모듈 랜덤 포레스트 분류기에 대해 가장 중요한 유전자인 상위 30개 유전자.

**도 14.** (a-c) 생후 첫 해에 sLRI 감수성에 대한 IFN 및 전염증성 매개체 및 바이러스-관련 수용체 유전자의 분석. 데이터는 일치된 비자극 샘플에 대해 조정되었고, 플롯은 유아기에 sLRI에 대한 저항성과 비교하여 감수성인 개체의 CBMC 데이터에 대한 맨-휘트니 U 테스트 추정치 및 95% CI를 나타낸다. 적색 데이터 점/표지(더 어두운 표지)는  $p$  값  $< 0.05$ 인 발현 증가를 나타낸다. (d) LPS로 자극된 CBMC( $n=50$ ) 중 IFIH1 및 IRF1/STAT1 간의 스피어만의 상관관계 및 관련  $p$  값. 청색 점선은 데이터의 손실 적합성을 나타낸다.

**도 15.** LPS 처리 유무에 관계없이 제대혈로부터 수집된 골수성(상단 패널) 및 림프성(하단 패널) 세포를 비교하는 단일 세포 RNA 시퀀싱 분석으로부터의 차등적 유전자 발현. 적색(더 어두운 점)으로 색칠된 유전자는 상향조절된 것으로 간주되며, 청색으로 색칠된 유전자는 하향조절된 것으로 간주된다.

### 발명을 실시하기 위한 구체적인 내용

- [0118] **구현예에 대한 상세한 설명**
- [0119] 본 명세서에 개시되고 정의된 본 발명은 텍스트 또는 도면에서 언급되거나 명백한 개별 특징 중 2개 이상의 모든 대안적인 조합으로 확장된 것으로 이해될 것이다. 이러한 모든 상이한 조합은 본 발명의 다양한 대안적인 측면을 구성한다.
- [0120] 본 발명의 추가 측면 및 이전 단락에 기술된 측면의 추가 구현예는 예로서 그리고 첨부 도면을 참조하여 주어진, 다음의 설명으로부터 명백해질 것이다.
- [0121] 본 발명의 특정 구현예들에 대한 참조는 이제 상세하게 이루어질 것이다. 본 발명은 구현예와 함께 기술될 것이지만, 본 발명을 이러한 구현예로 제한하려는 의도가 아님을 이해할 것이다. 반대로, 본 발명은 청구범위에 의해 정의된 바와 같이 본 발명의 범주 내에 포함될 수 있는 모든 대안, 변형 및 등가물을 포괄하도록 의도된다.
- [0122] 본 발명은 천식의 후속 발병과 관련된 바이러스 감염을 포함하여 어떤 아동이 생애 초기에 호흡기 감염을 경험할 것인지 출생 시 예측할 수 있는 방법의 개발을 기반으로 한다. 본 발명자들의 결과는 생후 첫 해에 중증 호흡기 바이러스 감염(예를 들어, sLRI)에 대한 감수성이 주로 항-박테리아성 대 항-바이러스성 선천적 면역 경로에 의해 결정됨을 시사하고, 초기 개입을 위해 위험에 처한 유아의 식별을 위한 근거를 제공한다. 이와 관련하여, 본원에 제시된 데이터는 병원성 박테리아에 대한 반응이 바이러스 자극에 대한 반응보다 sLRI 감수성의 더 중요한 결정 요인임을 시사한다.
- [0123] 출생 시 강화된 박테리아 매개 TLR4-유도 IFN 반응/유전자 네트워크 연결성 패턴이 유아기에 바이러스 sLRI의 위험을 부여한다는 본원에 기술된 결과는 IFN 반응이 급성 바이러스 감염 동안 거의 보편적으로 보호된다는 점을 고려하면 놀랍다.
- [0124] 매년, 수천 명의 아동이 중증 호흡기 바이러스 감염으로 응급실을 방문한다. 중증 바이러스 하부 호흡 기도 감염(sLRI)은 유아 및 아동의 입원의 주요 원인이며, 후속 천식 발병에 대한 주요 위험 요소를 구성한다.
- [0125] 더욱이, 이들 아동 중 하위 집합은 전 세계적으로 3억 명의 사람들에게 영향을 미치는 기도의 만성 염증성 질환인 천식이 발병할 것이다. 특히, 자궁 내 존재 및 생후 처음 몇 년 동안 천식에 대한 체계적 면역 시스템이 기능적으로 미성숙하고 감염에 매우 민감한 가소성이 강화되는 중요한 기간을 나타낸다. 초기 유아기에 면역 시스템의 가소성은 면역 시스템을 재프로그래밍하고 질환 위험을 최소화하기 위한 면역조절성 약물의 투여에 이상적인 "기회의 창(window of opportunity)"을 제공한다. 본 발명은 선천적 면역을 조절하고 중증 호흡기 바이러스 감염 및 후속 천식의 발병을 감소시키거나 예방하기 위한 적절한 개입으로 치료될 수 있는 고-위험 유아의 매우 초기 식별을 가능하게 한다.
- [0126] 따라서, 고-위험 유아의 초기 식별은 이러한 아동이 면역조절 약물로 치료될 수 있게 하며, 이는 결국 중증 호흡기 바이러스 감염 및 후속 천식의 발병에 대한 위험을 최소화할 것이다. 중증 호흡기 바이러스 감염으로 인한 응급실 방문의 예방은 건강 관리 시스템에 수십억 달러를 절약하고, 수백만 명의 아동 및 이의 가족에 대한 삶의 질을 향상시킬 것이다.
- [0127] **일반**
- [0128] 본 명세서 전반에 걸쳐, 달리 구체적으로 명시되지 않거나 문맥상 달리 요구되지 않는 한, 단일 단계, 물질의

구성, 단계의 그룹 또는 물질의 구성 그룹에 대한 언급은 하나 및 복수 (즉, 하나 이상)의 단계, 물질의 구성, 단계의 그룹 또는 물질의 구성 그룹을 포함하는 것으로 간주되어야 한다. 따라서, 본원에 사용된 바와 같이, 단수형 "한(a)", "하나(an)" 및 "그(the)"는 문맥상 명백히 달리 명시되지 않는 한 복수 측면을 포함하며, 그 반대도 마찬가지이다. 예를 들어, "한(a)"에 대한 언급은 단일뿐만 아니라 둘 이상을 포함한다; "하나(an)"에 대한 언급은 단일뿐만 아니라 둘 이상을 포함한다; "그(the)"에 대한 언급은 단일뿐만 아니라 둘 이상 등등을 포함한다.

- [0129] 당업계의 숙련된 자는 본 발명이 구체적으로 기술된 것 이외의 변화 및 변형에 민감함을 인식할 것이다. 본 발명은 이러한 모든 변화 및 변형을 포함하는 것으로 이해되어야 한다. 본 발명은 또한 개별적으로 또는 집합적으로 본 명세서에 언급되거나 나타난 모든 단계, 특징, 조성물 및 화합물, 및 상기 단계 또는 특징 중 임의의 모든 조합 또는 임의의 둘 이상을 포함한다.
- [0130] 당업계의 숙련된 자는 본 발명의 실시예에 사용될 수 있는, 본원에 기술된 것과 유사 또는 동등한 많은 방법 및 재료를 인식할 것이다. 본 발명은 기술된 방법 및 재료에 결코 제한되지 않는다.
- [0131] 본원에 언급된 모든 특허 및 간행물은 그 전체가 참조로 포함된다.
- [0132] 본 발명은 단지 예시의 목적으로 의도되는, 본원에 기술된 특정 실시예에 의해 범주가 제한되지 않는다. 기능적으로-동등한 산물, 조성물 및 방법은 명백히 본 발명의 범주 내에 있다.
- [0133] 본원에서 본 발명의 임의의 실시예 또는 구현에는 달리 구체적으로 언급되지 않는 한 본 발명의 임의의 다른 실시예 또는 구현에 준용하여 적용되는 것으로 간주되어야 한다.
- [0134] 달리 구체적으로 정의되지 않는 한, 본원에 사용된 모든 기술적 및 과학적 용어는 (예를 들어, 세포 배양, 분자 유전학, 면역학, 면역조직화학, 단백질 화학, 생화학에서) 당업계의 통상의 기술자에 의해 일반적으로 이해하는 것과 동일한 의미를 갖는 것으로 간주되어야 한다.
- [0135] 달리 지시되지 않는 한, 본 개시에서 활용되는 재조합 단백질, 세포 배양 및 면역학적 기술은 당업계의 숙련된 자에게 널리 알려져 있는 표준 절차이다. 이러한 기술은 J. Perbal, *A Practical Guide to Molecular Cloning*, John Wiley and Sons (1984), J. Sambrook 등, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbour Laboratory Press (1989), T.A. Brown (편집자), *Essential Molecular Biology: A Practical Approach*, Volumes 1 and 2, IRL Press (1991), D.M. Glover 및 B.D. Hames (편집자), *DNA Cloning: A Practical Approach*, Volumes 1-4, IRL Press (1995 및 1996), 및 F.M. Ausubel 등 (편집자), *Current Protocols in Molecular Biology*, Greene Pub. Associates and Wiley-Interscience (1988, 현재까지의 모든 업데이트 포함), Ed Harlow 및 David Lane (편집자) *Antibodies: A Laboratory Manual*, Cold Spring Harbour Laboratory, (1988), 및 J.E. Coligan 등 (편집자) *Current Protocols in Immunology*, John Wiley & Sons (현재까지의 모든 업데이트 포함)과 같은 출처의 문헌 전반에 걸쳐 기술되고 설명된다.
- [0136] 용어 "및/또는", 예를 들어 "X 및/또는 Y"는 "X 및 Y" 또는 "X 또는 Y"를 의미하는 것으로 이해되어야 하며, 둘 모두의 의미 또는 의미 중 어느 하나에 대한 분명한 지지를 제공하는 것으로 간주되어야 한다.
- [0137] 본원에 사용된 바와 같이, 용어 "로부터 유래된"은 지정된 정수가 비록 해당 공급원으로부터 반드시 직접적이지 않을지라도 특정 공급원으로부터 수득되는 것임을 나타내는 것으로 간주되어야 한다.
- [0138] 본원에 사용된 바와 같이, TLR4 작용제는 지질다당류(LPS), 모노포스포릴 지질 A(MPLA), 열 충격 단백질, S100A8, S100A9, RSV F 단백질, 피브리노겐, 헤파린 설페이트 또는 이의 단편, 히알루론산 또는 이의 단편, 니켈, 오피오이드, α1-산성 당단백질(AAG), 아미노알킬 글루코사미나이드 4-포스페이트(AGP), RC-529, 뮤린 β-디펜신 2 및 완전 프로인트 보조제(complete Freund's adjuvant, CFA)로 이루어진 군으로부터 선택될 수 있다.
- [0139] 본원에 기술된 바와 같이, 용어 "감수성"은 특정 상태(예를 들어, 특정 특성, 표현형 또는 질환)의 발병으로 향하거나, 평균 개체보다 특정 상태에 저항할 수 없는 것으로 향하는 개체의 성향을 지칭한다. 이 용어는 증가된 감수성 및 감소된 감수성을 모두 포함한다. 따라서, 본원에 기술된 것을 포함하는 특정 바이오마커는 예를 들어 특정 바이오마커에 대해 1 초과의 상대적 위험도 (relative risk, RR) 또는 오즈비(odds ratio, OR)를 특징으로 하는, 호흡기 감염(예를 들어, sLRI)의 증가된 감수성(즉, 증가된 위험)의 특징일 수 있다. 대안적으로, 바이오마커는 1 미만의 상대적 위험도를 특징으로 하는 호흡기 감염(예를 들어, sLRI)의 감소된 감수성(즉, 감소된 위험)을 특징으로 한다.
- [0140] 감수성 또는 위험의 측정은 본원에 더 상세히 기술된 바와 같이, 상대적 위험도(RR), 오즈비(OR), 및 절대적 위

험도(AR)와 같은 측정을 포함한다.

- [0141] 특정 구현예에서, 증가된 감수성은 적어도 1.10, 적어도 1.11, 적어도 1.12, 적어도 1.13, 적어도 1.14, 적어도 1.15, 적어도 1.16, 적어도 1.17, 적어도 1.18, 적어도 1.19, 적어도 1.20, 적어도 1.21, 적어도 1.22, 적어도 1.23, 적어도 1.24, 적어도 1.25, 적어도 1.30, 적어도 1.35, 적어도 1.40, 적어도 1.45, 적어도 1.50, 적어도 1.55, 적어도 1.60, 적어도 1.65, 적어도 1.70, 적어도 1.75, 및/또는 적어도 1.80의 RR 또는 OR 값을 갖는 위험도를 지칭한다. 1보다 큰 다른 비-정수 수치 값은 또한 위험을 특성화하는 데 가능하며, 이러한 수치 값은 또한 본 발명의 범위 내에 있다.
- [0142] 증가된 감수성은 또한 참조 데이터 세트와의 비교를 수반할 수 있다. 참조 데이터 세트는 (a) 호흡기 감염(예를 들어, sLRI)에 대해 증가되거나, 상승되거나, 높거나 더 높은 위험 또는 감수성인 것으로 결정되거나 (증가되거나 높은 위험 참조 데이터 세트로도 지칭됨) (b) 호흡기 감염에 대해 증가되지 않은 위험 또는 감수성인 것으로 결정된 (정상 또는 증가되지 않은 위험 참조 데이터 세트로도 지칭됨) 하나 이상의 개체로부터의 것일 수 있다. 따라서, 위험이 결정될 개체로부터의 샘플에서 바이오마커의 발현 수준이 증가되거나 높은 위험 참조 데이터 세트와 동일하거나 유의미하게 상이하지 않은 경우, 이후 개체가 호흡기 감염에 대해 증가되거나 높은 위험을 갖는다는 결정이 이루어질 수 있다. 대안적으로, 위험이 결정될 개체로부터의 샘플에서 바이오마커의 발현 수준이 증가되거나 높은 위험 참조 데이터 세트와 유의미하게 상이한 경우, 이후 개체가 호흡기 감염에 대해 증가되거나 높은 위험을 갖지 않는다는 결정이 이루어질 수 있다. 대안적으로, 위험이 결정될 개체로부터의 샘플에서 바이오마커의 발현 수준이 정상 또는 증가되지 않은 위험 참조 데이터 세트와 유의미하게 상이한 경우, 이후 개체가 호흡기 감염에 대해 증가되거나 높은 위험을 갖는다는 결정이 이루어질 수 있다.
- [0143] 본 발명의 임의의 방법 또는 용도에 있어서, 결정은 호흡기 감염(예를 들어, sLRI)의 증가되거나, 상승되거나, 높거나 더 높은 위험 또는 감수성일 수 있거나, 결정은 호흡기 감염의 증가되지 않거나, 상승되지 않거나, 더 높지 않거나 정상적인 위험 또는 감수성일 수 있다. 본원에 사용된 바와 같이, 호흡기 감염의 증가되거나, 상승되거나, 높거나 더 높은 위험 또는 감수성의 결정에 대한 참조는 개체가 선제적 요법의 형태로 개입을 필요로 한다는 결정에 대한 참조로 간주될 수 있다. 따라서, 본 발명의 임의의 방법 또는 용도에 있어서, 개체가 선제적 요법의 형태로 개입을 필요로 한다고 결정이 이루어진 경우, 상기 방법 또는 용도는 선제적 요법 (예를 들어, 본원에 기술된 임의의 선제적 요법)의 형태로 개입을 시행하는 단계를 추가로 포함한다.
- [0144] 용어 "단백질"은 단일 폴리펩티드 사슬, 즉 펩티드 결합에 의해 연결된 일련의 연속 아미노산, 또는 서로 공유적으로 또는 비-공유적으로 연결된 일련의 폴리펩티드 사슬(즉, 폴리펩티드 복합체)을 포함하는 것으로 간주되어야 한다. 예를 들어, 일련의 폴리펩티드 사슬은 적합한 화학적 또는 이황화 결합을 사용하여 공유적으로 연결될 수 있다. 비-공유 결합의 예는 수소 결합, 이온 결합, 반 데르 발스 힘 및 소수성 상호작용을 포함한다.
- [0145] 용어 "폴리펩티드" 또는 "폴리펩티드 사슬"은 전술한 단락으로부터 펩티드 결합에 의해 연결된 일련의 연속적인 아미노산을 의미하는 것으로 이해될 것이다.
- [0146] 용어 "마이크로어레이"는 기관 상에 결합된/복합체화된 어레이 요소 또는 리간드, 예를 들어, 항체의 정렬된 배열을 지칭한다.
- [0147] 단수형 또는 복수형으로 사용되는 경우, 용어 "폴리뉴클레오티드"는 일반적으로 비변형된 RNA 또는 DNA 또는 변형된 RNA 또는 DNA일 수 있는 임의의 폴리리보뉴클레오티드 또는 폴리데옥시리보뉴클레오티드를 지칭한다. 따라서, 예를 들어, 본원에 정의된 폴리뉴클레오티드는, 제한 없이, 단일- 및 이중-가닥 DNA, 단일- 및 이중-가닥 영역을 포함하는 DNA, 단일- 및 이중-가닥 RNA, 및 단일- 및 이중-가닥 영역을 포함하는 RNA, 단일-가닥 또는, 보다 전형적으로, 이중-가닥일 수 있거나 단일- 및 이중-가닥 영역을 포함할 수 있는 DNA 및 RNA를 포함하는 하이브리드 분자를 포함한다. 게다가, 본원에 사용된 용어 "폴리뉴클레오티드"는 RNA 또는 DNA, 또는 RNA 및 DNA 둘 다를 포함하는 삼중-가닥 영역을 지칭한다. 용어 "폴리뉴클레오티드"는 구체적으로 cDNA를 포함한다. 상기 용어는 하나 이상의 변형된 염기를 함유하는 DNA(cDNA를 포함함) 및 RNA를 포함한다. 일반적으로, 용어 "폴리뉴클레오티드"는 비변형된 폴리뉴클레오티드의 모든 화학적, 효소적 및/또는 대사적으로 변형된 형태뿐만 아니라 바이러스 및 단순 및 복합 세포를 포함하는 세포의 특징적인 DNA 및 RNA의 화학적 형태를 포괄한다.
- [0148] 용어 "올리고뉴클레오티드"는 제한 없이, 단일-가닥 데옥시리보뉴클레오티드, 단일- 또는 이중-가닥 리보뉴클레오티드, RNA:DNA 하이브리드 및 이중-가닥 DNA를 포함하는 20개 염기 미만의 비교적 짧은 폴리뉴클레오티드를 지칭한다. 단일-가닥 DNA 프로브 올리고뉴클레오티드와 같은 올리고뉴클레오티드는 종종 화학적 방법, 예를 들어 상업적으로 이용가능한 자동화된 올리고뉴클레오티드 합성기를 사용하여 합성된다. 그러나, 올리고뉴클레오

티드는 시험관내 제조합 DNA-매개 기술을 포함한 다양한 다른 방법과 세포 및 유기체에서의 DNA 발현에 의해 제조될 수 있다.

[0149] 본원에 사용된 바와 같이, 용어 "대상체"는 인간을 포함하는 임의의 동물, 예를 들어 포유동물을 의미하는 것으로 간주되어야 한다. 예시적인 대상체는 인간 및 비-인간 영장류를 포함하나 이에 제한되지는 않는다. 예를 들어, 대상체는 인간이다.

[0150] **샘플**

[0151] 임의의 측면에서, B 및 T 세포, 또는 CBMC는 제대혈로부터 정제될 수 있다. 대안적으로, 임의의 측면에서, B 및 T 세포, 또는 CBMC는 개체로부터 제대혈의 샘플에 존재할 수 있다. 임의의 구현예에서, B 및 T 세포, 또는 CBMC는 제대혈에 존재하여서, CBMC를 TLR4 작용제와 접촉시키는 것에 대한 본원의 임의의 참조는 B 및 T 세포 또는 CBMC를 함유하는 제대혈(예를 들어 미처리, 전체 또는 비-정제된 제대혈)을 TLR4 작용제와 접촉시키는 것이 포함된다. 일 구현예에서, 제대혈은 적혈구가 고갈될 수 있다. 일 구현예에서, 적혈구는 B 및 T 세포, 또는 CBMC가 TLR4 작용제와 접촉되는 경우, 존재하지 않거나 유의미한 수준으로 존재하지 않는다. 또 다른 구현예에서, 적혈구는 정상 수준으로 존재하며, 즉, 이들은 제대혈로부터 고갈되지 않았다.

[0152] 임의의 측면에서, 개체는 1세 미만 또는 2세 이하이다. 임의의 구현예에서, 개체는 1일, 2일, 3일, 4일, 5일, 6일, 7일, 2주, 1개월, 6개월, 1세 또는 2세이다. 임의의 구현예에서, 적어도 1일, 적어도 2일, 적어도 3일, 적어도 4일, 적어도 5일, 적어도 6일, 적어도 7일, 적어도 2주, 적어도 1개월, 적어도 2개월, 적어도 3개월, 적어도 6개월, 적어도 1세 이상, 2세 이하이다. 임의의 구현예에서, 개체는 약 1일 내지 약 7일, 약 1일 내지 약 2주, 약 1주 내지 약 6개월, 약 1개월 내지 약 6개월, 약 1개월 내지 약 3개월, 또는 약 6개월 내지 약 1세, 약 1개월 내지 2세, 약 6개월 내지 2세, 약 1세 내지 2세이다.

[0153] 임의의 구현예에서, 제대혈은 출생 24시간 이내에 개체로부터 유래된다. 임의의 구현예에서, 제대혈은 48시간 이내에 개체로부터 유래된다. 제대혈은 수득된 후 본 발명의 방법에 사용하기 전에 최대 2년 동안 동결시킬 수 있다. 예를 들어, 제대혈은 본원에 기술된 바와 같이 본 발명의 상기 방법의 대상이 되는 개체의 현재 연령 이하의 시간 동안 수득되고 동결될 수 있다. 따라서, CBMC는 개체의 출생 후 24시간 또는 48시간 이내의 신선한 제대혈 또는 최대 2년까지 동결된 제대혈로부터 수득될 수 있다.

[0154] 임의의 구현예에서, 제대혈 적혈구는 염화암모늄 용해, 밀도 구배 기술, 저장성(hypotonic) 용해, 면역자기적 세포 분리 또는 침강, 유세포 분석 분류(flow cytometric sorting), 또는 당업계의 숙련된 자에 의해 인식되는 동등한 방법에 의해 고갈된다.

[0155] 임의의 구현예에서, CBMC는 생리학적 조건 또는 그 근처의 영양 배지에서 배양된다. 예를 들어, CBMC는 37°C, 5% CO<sub>2</sub>에서 RPMI + 5% AB 비-가열 불활성화된 혈청에서 배양될 수 있다. 임의의 구현예에서, CBMC는 비-가열 불활성화된 혈청을 함유하는 영양 배지에서 배양된다.

[0156] 임의의 구현예에서, B 및 T 세포는 생리학적 조건 또는 그 근처의 영양 배지에서 배양된다. B 및 T 세포가 배양된 배지는 당업계에 알려져 있다.

[0157] **TLR4 작용제로 자극**

[0158] 본 발명의 임의의 측면에서, B 및 T 세포, 또는 CBMC 또는 제대혈은 TLR4 작용제로 자극될 수 있다.

[0159] B 및 T 세포, 또는 CBMC 또는 제대혈은 TLR4 작용제로 적어도 6시간, 적어도 12시간, 적어도 18시간 또는 적어도 24시간 동안 자극될 수 있다. 임의의 구현예에서, CBMC는  $1 \times 10^6$  개 세포/mL로 현탁되고 이후 TLR4 작용제로 자극될 수 있다.

[0160] 임의의 구현예에서, TLR4 작용제는 TLR4 활성화를 자극하는 데 유효 농도로 제공된다. 유효 농축물의 예는 0.025 ng/ml 내지 100 ng/ml, 바람직하게는 1 ng/ml의 LPS이다. 다른 TLR4 작용제에 대해, 유효 농도는 LPS 1 ng/ml와 동일한 TLR4 활성화를 제공하는 임의의 양이다. 본원에 기술된 바와 같이, 당업계의 숙련된 자는 TLR4 활성화를 자극하기 위한 임의의 TLR4 작용제의 유효 농도를 결정하는데 사용될 수 있는 방법을 또한 인식할 것이다. 예를 들어, 숙련된 자는 임의의 TLR4 작용제의 유효 농도에 대한 검정을 수행할 수 있다. 예시적인 검정은 B 및 T 세포, CBMC 또는 제대혈을 TLR4 작용제로 시험관내에서 밤새 자극하는 것과 비자극 B 및 T 세포, CBMC 또는 제대혈과 비교하여 NK-κB 매개 전사 활성을 측정하는 것을 각각 포함하며, 상기 NK-κB 매개 전사 활성의 증가는 TLR4 작용제의 유효 농도를 나타낸다.

- [0161] **바이오마커 검출 및 측정**
- [0162] 샘플 내 바이오마커는 당업계에 공지된 임의의 적합한 방법에 의해 측정될 수 있는 것으로 이해된다. 바이오마커의 발현 수준의 측정은 직접적이거나 간접적일 수 있다. 예를 들어, RNA 또는 단백질의 존재비 수준은 직접적으로 정량화될 수 있다. 대안적으로, 바이오마커의 양은 cDNA, 증폭된 RNA 또는 DNA의 존재비 수준을 측정함으로써, 또는 바이오마커의 발현 수준을 나타내는 RNA, 단백질 또는 기타 분자의 양 또는 활성을 측정함으로써 간접적으로 결정될 수 있다.
- [0163] 일 구현예에서, 바이오마커의 발현 수준은 바이오마커의 폴리뉴클레오티드 수준을 측정함으로써 결정된다. 특정 바이오마커 유전자의 전사 수준은 샘플에 존재하는 mRNA 또는 이로부터 유래된 폴리뉴클레오티드의 양으로부터 결정될 수 있다. 폴리뉴클레오티드는 마이크로어레이 분석, 중합효소 연쇄반응(PCR), 역전사효소 중합효소 연쇄반응(RT-PCR), 노던 블롯, 유전자 발현의 연속 분석(SAGE), 총(total) RNA-시퀀싱, mRNA 시퀀싱, Cap 분석 유전자 발현(CAGE) 시퀀싱, 단일 세포 RNA-시퀀싱 또는 NanoString nCounter를 포함하나, 이에 제한되지 않는 다양한 방법에 의해 검출되고 정량화될 수 있다. 예를 들어, Draghici Data Analysis Tools for DNA Microarrays, Chapman and Hall/CRC, 2003; Simon 등, Design and Analysis of DNA Microarray Investigations, Springer, 2004; Real-Time PCR: Current Technology and Applications, Logan, Edwards, and Saunders eds., Caister Academic Press, 2009; Bustin A-Z of Quantitative PCR (IUL Biotechnology, No. 5), International University Line, 2004; Velculescu 등 (1995) Science 270: 484-487; Matsumura 등 (2005) Cell. Microbiol. 7: 11-18; Serial Analysis of Gene Expression (SAGE): Methods and Protocols (Methods in Molecular Biology), Humana Press, 2008을 참조하며; 본원에 그 전체가 참조로 포함된다.
- [0164] 일 구현예에서, 마이크로어레이는 바이오마커의 수준을 측정하기 위해 사용된다. 마이크로어레이 분석의 장점은 각 바이오마커의 발현이 동시에 측정될 수 있다는 것이고, 마이크로어레이는 특정 질환 또는 병태, 레귤론 또는 네트워크에 대한 발현 프로파일을 제공하도록 구체적으로 설계될 수 있다.
- [0165] 마이크로어레이는 폴리뉴클레오티드 서열을 포함하는 프로브를 선택하고, 이후 이러한 프로브를 고체 지지체 또는 표면에 고정화함으로써 제조된다. 예를 들어, 프로브는 DNA 서열, RNA 서열, 또는 DNA 및 RNA의 공중합체 서열을 포함할 수 있다. 프로브의 폴리뉴클레오티드 서열은 또한 DNA 및/또는 RNA 유사체, 또는 이들의 조합을 포함할 수 있다. 예를 들어, 프로브의 폴리뉴클레오티드 서열은 게놈 DNA의 전체 또는 부분 단편일 수 있다. 프로브의 폴리뉴클레오티드 서열은 또한 합성 올리고뉴클레오티드 서열과 같은 합성된 뉴클레오티드 서열일 수 있다. 프로브 서열은 생체내에서 효소적으로, 시험관내에서 효소적으로(예를 들어, PCR에 의해), 또는 시험관내에서 비-효소적으로 합성될 수 있다.
- [0166] 본 발명의 방법에 사용되는 프로브는 다공성 또는 비-다공성일 수 있는 고체 지지체에 고정화되는 것이 바람직하다. 예를 들어, 상기 프로브는 폴리뉴클레오티드의 3' 또는 5' 말단에서 니트로셀룰로스 또는 나일론 막 또는 필터에 공유결합으로 부착된 폴리뉴클레오티드 서열일 수 있다. 이러한 혼성화 프로브는 당업계에 잘 알려져 있다 (예를 들어, Sambrook 등, Molecular Cloning: A Laboratory Manual (제3판, 2001 참조). 대안적으로, 고체 지지체 또는 표면은 유리 또는 플라스틱 표면일 수 있다. 일 구현예에서, 혼성화 수준은 폴리뉴클레오티드의 집단, 예컨대 DNA 또는 DNA 모방체의 집단, 또는 대안적으로 RNA 또는 RNA 모방체의 집단이 고정화된 표면에 고체 상으로 구성된 프로브의 마이크로어레이에 대해 측정된다. 상기 고체상은 비다공성, 또는 선택적으로 겔과 같은 다공성 물질일 수 있다.
- [0167] 일 구현예에서, 마이크로어레이는 각각 본원에 기술된 바이오마커 중 하나를 나타내는 결합(예를 들어, 혼성화) 부위 또는 "프로브"의 정렬된 어레이를 갖는 지지체 또는 표면을 포함한다. 바람직하게는 마이크로어레이는 주소 지정이 가능한 어레이이고, 보다 바람직하게는 위치적으로 주소 지정이 가능한 어레이이다. 보다 구체적으로, 어레이의 각 프로브는 바람직하게는 각 프로브의 동일성 (즉, 서열)이 어레이 내의 (즉, 지지체 또는 표면 상의) 이의 위치로부터 결정될 수 있도록 고체 지지체 상의 알려진, 미리 결정된 위치에 위치한다. 각 프로브는 바람직하게는 단일 부위에서 고체 지지체에 공유결합으로 부착된다.
- [0168] 마이크로어레이는 다수의 방식으로 제조될 수 있으며, 그 중 몇몇은 하기에 기술된다. 이들이 어떻게 생산되었든, 마이크로어레이는 특정 특성을 공유한다. 상기 어레이는 재현가능하여, 주어진 어레이의 다중 카피가 생성되도록 하고 서로 용이하게 비교될 수 있게 한다. 바람직하게는, 마이크로어레이는 결합 (예를 들어, 핵산 혼성화) 조건 하에 안정적인 물질로부터 제조된다. 마이크로어레이는 일반적으로 소형, 예를 들어, 1 cm<sup>2</sup> 내지 25 cm<sup>2</sup> 이다; 그러나, 더 큰 어레이가, 예를 들어, 스크리닝 어레이가 또한 사용될 수 있다. 바람직하게는, 마이크로어레이에서 주어진 결합 부위 또는 결합 부위들의 독특한 세트가 세포에서 단일 유전자의 산물에 (예를 들어, 특

정 mRNA에, 또는 이로부터 유래된 특정 cDNA에) 특이적으로 결합 (예를 들어, 혼성화)할 것이다.

- [0169] 그러나, 일반적으로 다른 관련되거나 유사한 서열은 주어진 결합 부위에 교차 혼성화될 것이다.
- [0170] 전술한 바와 같이, 특정 폴리뉴클레오티드 분자가 특이적으로 혼성화하는 "프로브"는 상보적인 폴리뉴클레오티드 서열을 포함한다. 마이크로어레이의 프로브는 전형적으로 1,000개 이하의 뉴클레오티드 서열로 구성된다. 일부 구현예에서, 어레이의 프로브는 10 내지 1,000개 뉴클레오티드의 뉴클레오티드 서열로 구성된다. 일 구현예에서, 프로브의 뉴클레오티드 서열은 10-200개 뉴클레오티드 범위의 길이이고, 유기체 중 한 종의 게놈 서열이어서, 이러한 유기체 종의 게놈에 대해 상보적이며 따라서 이에 혼성화될 수 있는 서열을 갖는 복수개의 상이한 프로브가 존재하고, 상기 게놈의 전체 또는 이의 일부분에 걸쳐 순차적으로 타일링된다. 다른 구현예에서, 프로브는 10-30개 뉴클레오티드 범위의 길이, 10-40개 뉴클레오티드 범위의 길이, 20-50개 뉴클레오티드 범위의 길이, 40-80개 뉴클레오티드 범위의 길이, 50-150개 뉴클레오티드 범위의 길이, 80-120개 뉴클레오티드 범위의 길이이거나, 또는 60개 뉴클레오티드 길이이다.
- [0171] 상기 프로브는 유기체 게놈의 일부에 상응하는 DNA 또는 DNA "모방체" (예를 들어, 유도체 및 유사체)를 포함할 수 있다. 다른 구현예에서, 마이크로어레이의 프로브는 상보적인 RNA 또는 RNA 모방체이다. DNA 모방체는 DNA와의 특이적인 왓슨-크릭-유사 혼성화 또는 RNA와의 특이적인 혼성화가 가능한 서브유닛으로 구성된 중합체이다. 핵산은 염기 모이어티, 당 모이어티 또는 인산염 골격(예를 들어, 포스포로티오에이트)에서 변형될 수 있다.
- [0172] DNA는 예를 들어, 게놈 DNA 또는 클로닝된 서열의 중합효소 연쇄반응(PCR) 증폭에 의해 수득될 수 있다. PCR 프라이머는 바람직하게는 게놈 DNA의 특정 단편의 증폭을 초래할 게놈의 공지된 서열을 기반으로 선택된다. 당업계에 잘 알려진 컴퓨터 프로그램은 올리고(Oligo) 버전 5.0(National Biosciences)과 같이 요구되는 특이성 및 최적의 증폭 특성을 갖춘 프라이머 설계에 유용하다. 전형적으로 마이크로어레이 상의 각 프로브는 10개 염기 내지 50,000개 염기, 일반적으로 300개 염기 내지 1,000개 염기 길이일 것이다. PCR 방법은 당업계에 잘 알려져 있으며, 예를 들어, Innis 등, eds., PCR Protocols: A Guide To Methods And Application, Academic Press Inc., San Diego, CA (1990)에 기술되어 있고; 그 전체가 본원에 참조로 포함된다. 제어된 로봇 시스템이 핵산을 분리하고 증폭하는데 유용하다는 것은 당업계의 숙련된 자에게 명백할 것이다.
- [0173] 폴리뉴클레오티드 프로브를 생성하기 위한 대안적인, 바람직한 수단은 예를 들어, N-포스포네이트 또는 포스포아미다이트 화학물질을 사용한, 합성 폴리뉴클레오티드 또는 올리고뉴클레오티드의 합성에 의한 것이다 (Froehler 등, Nucleic Acid Res. 14:5399-5407 (1986); McBride 등, Tetrahedron Lett. 24:246-248 (1983)). 합성 서열은 전형적으로 약 10개 내지 약 500개 염기 길이, 보다 전형적으로 약 20개 내지 약 100개 염기, 및 가장 바람직하게는 약 40개 내지 약 70개 염기 길이이다. 일부 구현예에서, 합성 핵산은 비-천연 염기, 예컨대 이노신을 포함하나 결코 이에 제한되지는 않는다. 전술한 바와 같이, 핵산 유사체는 혼성화를 위한 결합 부위로 사용될 수 있다. 적합한 핵산 유사체의 예는 펩티드 핵산이다(예를 들어, Egholm 등, Nature 363:566-568 (1993); 미국특허 제5,539,083호 참조).
- [0174] 프로브는 바람직하게는 결합 에너지, 염기 조성, 서열 복잡성, 교차-혼성화 결합 에너지 및 2차 구조를 고려하는 알고리즘을 사용하여 선택된다. 2001년 1월 25일자로 공개된, Friend 등, 국제특허 공개공보 W001/05935호; Hughes 등, Nat. Biotech. 19:342-7 (2001)을 참조한다.
- [0175] 숙련된 기술자는 또한 양성 대조군 프로브, 예를 들어, 표적 폴리뉴클레오티드 분자의 서열에 상보적이고 혼성화 가능한 것으로 알려진 프로브, 및 음성 대조군 프로브, 예를 들어, 표적 폴리뉴클레오티드 분자의 서열에 상보적이지 않고 혼성화 가능하지 않은 것으로 알려진 프로브가 어레이 상에 포함되어야 함을 인식할 것이다. 일 구현예에서, 양성 대조군은 어레이의 둘레를 따라 합성된다. 다른 구현예에서, 양성 대조군은 어레이 전체에 걸쳐 대각선 줄무늬로 합성된다. 또 다른 구현예에서, 각 프로브에 대한 역 상보체(reverse complement)는 프로브의 위치 옆에서 합성되어 음성 대조군으로서 역할을 한다. 또 다른 구현예에서, 다른 유기체 종으로부터의 서열은 음성 대조군 또는 "스파이크-인(spike-in)" 대조군으로서 사용된다.
- [0176] 상기 프로브는 예를 들어, 유리, 플라스틱(예를 들어, 폴리프로필렌, 나일론), 폴리아크릴아미드, 니트로셀룰로오스, 젤 또는 기타 다공성 또는 비다공성 물질로 만들어질 수 있는 고체 지지체 또는 표면에 부착된다. 표면에 핵산을 부착하기 위한 하나의 방법은 일반적으로 Schena 등, Science 270:467-470(1995)에 기술된 바와 같이 유리 플레이트 상에 프린팅하는 것이다. 이 방법은 cDNA의 마이크로어레이를 제조하는 데 특히 유용하다(또한, DeRisi 등, Nature Genetics 14:457-460(1996); Shalon 등, Genome Res. 6:639-645(1996); 및 Schena 등, Proc. Natl. Acad. Sci. U.S.A. 93:10539-11286(1995) 참조; 그 전체가 본원에 참조로 포함됨).

- [0177] 마이크로어레이를 제조하기 위한 두 번째 방법은 고-밀도 올리고뉴클레오티드 어레이를 생성한다. 원위치(in situ) 합성을 위한 포토리소그래피 기술(Fodor 등, 1991, Science 251:767-773; Pease 등, 1994, Proc. Natl. Acad. Sci. U.S.A. 91:5022-5026; Lockhart 등, 1996, Nature Biotechnology 14: 1675, 미국특허 제5,578,832호, 제5,556,752호, 및 제5,510,270호 참조; 그 전체가 본원에 참조로 포함됨) 또는 정의된 올리고뉴클레오티드의 신속한 합성 및 증착을 위한 다른 방법(Blanchard 등, Biosensors & Bioelectronics 11:687-690; 그 전체가 본원에 참조로 포함됨)을 사용하여 표면 상의 정의된 위치에서 정의된 서열에 상보적인 수천 개의 올리고뉴클레오티드를 포함하는 어레이를 생성하기 위한 기술이 알려져 있다. 이러한 방법을 사용할 때, 알려진 서열의 올리고뉴클레오티드(예를 들어, 60-mer)는 유도체화된 유리 슬라이드와 같은 표면 상에 직접적으로 합성된다. 일반적으로, 생성된 어레이는 중복성으로서, RNA당 여러 개의 올리고뉴클레오티드 분자를 갖는다.
- [0178] 예를 들어, 마스크(masking)에 의한 마이크로어레이를 제조하기 위한 다른 방법(Maskos and Southern, 1992, Nuc. Acids. Res. 20: 1679-1684; 그 전체가 본원에 참조로 포함됨)을 또한 사용할 수 있다. 원칙적으로, 임의의 유형의 어레이, 예를 들어, 나일론 혼성화 막 상의 도트 블롯(dot blot) (Sambrook 등, Molecular Cloning: A Laboratory Manual, 제3판, 2001 참조)을 사용할 수 있다. 그러나, 당업계의 숙련된 자에 의해 인식될 것인바, 혼성화 부피는 더 작을 것이므로 매우 작은 어레이가 종종 바람직할 것이다.
- [0179] 마이크로어레이는 또한 올리고뉴클레오티드 합성을 위한 잉크젯 인쇄 장치에 의해, 예를 들어, Blanchard의 미국특허 제6,028,189호; Blanchard 등, 1996, Biosensors and Bioelectronics 11:687-690; Blanchard, 1998, Synthetic DNA Arrays in Genetic Engineering, Vol. 20, J. K. Setlow, Ed., Plenum Press, New York, 페이지 111-123에 의해 기술된 방법 및 시스템을 사용하여 제조될 수 있으며; 본원에 그 전체가 참조로 포함된다. 구체적으로, 이러한 마이크로어레이에서 올리고뉴클레오티드 프로브는 프로필렌 카르보네이트와 같은 높은 표면 장력의 "마이크로액적(microdroplet)"에 개별 뉴클레오티드 염기를 연속적으로 증착시킴으로써 어레이 내, 예를 들어, 유리 슬라이드 상에서 합성된다. 마이크로액적은 부피가 작고 (예를 들어, 100 pL 이하, 보다 바람직하게는 50 pL 이하), 마이크로어레이 상에서 서로 분리되어 (예를 들어, 소수성 도메인에 의해), 어레이 요소 (즉, 상이한 프로브)의 위치를 정의하는 원형 표면 장력 웰(well)을 형성한다. 이러한 잉크-젯 방법으로 제작된 마이크로어레이는 전형적으로 밀도가 높고, 바람직하게는 1 cm 당 적어도 약 2,500개의 상이한 프로브의 밀도를 갖는다. 폴리뉴클레오티드 프로브는 폴리뉴클레오티드의 3' 또는 5' 말단에서 공유결합으로 지지체에 부착된다. 마이크로어레이 분석에 의해 측정될 수 있는 바이오마커 폴리뉴클레오티드는 발현된 RNA 또는 자연적으로 발생하는 핵산 분자뿐만 아니라 합성 핵산 분자를 비롯하여, 이로부터 유래된 핵산(예를 들어, cDNA 또는 RNA 중합효소 프로모터를 포함하는 cDNA로부터 유래된 증폭된 RNA)일 수 있다. 일 구현예에서, 표적 폴리뉴클레오티드 분자는 총 세포 RNA, 폴리(A)<sup>+</sup> 메신저 RNA(mRNA) 또는 이의 분획, 세포질 mRNA, 또는 cDNA로부터 전사된 RNA를 포함하나, 결코 이에 제한되지 않는 RNA(즉, cRNA; 예를 들어, Linsley & Schelter, 1999년 10월 4일자로 출원된 미국특허 출원번호 제09/411,074호, 또는 미국특허 제5,545,522호, 제5,891,636호 또는 제5,716,785호 참조)를 포함한다. 총(total) 및 폴리(A)<sup>+</sup> RNA를 제조하기 위한 방법은 당업계에 잘 알려져 있으며, 이는 일반적으로, 예를 들어, Sambrook, 등, Molecular Cloning: A Laboratory Manual (제3판, 2001)에 기술되어 있다. 구아니디늄 티오시아네이트 용해 후 CsCl 원심분리(Chirgwin 등, 1979, Biochemistry 18:5294-5299), 실리카 겔-기반 컬럼 (예를 들어, RNeasy (Qiagen, Valencia, Calif.) 또는 StrataPrep (Stratagene, La Jolla, Calif.))을 사용하거나, Ausubel 등, eds., 1989, Current Protocols In Molecular Biology, Vol. III, Green Publishing Associates, Inc., John Wiley & Sons, Inc., New York, pp. 13.12.1-13.12.5에 기술된 바와 같이 페놀 및 클로로포름을 사용하여 관심 세포로부터 RNA를 추출할 수 있다. 폴리(A)<sup>+</sup> RNA는 예를 들어, 올리고-dT 셀룰로오스를 사용한 선택에 의해 또는 대안적으로 전체 세포 RNA의 올리고-dT 프라이밍된 역전사에 의해 선택될 수 있다. RNA는 당업계에 알려진 방법에 의해, 예를 들어, ZnCl<sub>2</sub>와의 인큐베이션에 의해 단편화되어 RNA 단편을 생성할 수 있다.
- [0180] 일 구현예에서, 총 RNA, mRNA, 또는 이로부터 유래된 핵산은 자극된 샘플로부터 분리된다. 특정 세포에서 불량하게 발현되는 바이오마커 폴리뉴클레오티드는 정규화 기술을 사용하여 농축될 수 있다(Bonaldo 등, 1996, Genome Res. 6:791-806).
- [0181] 전술된 바와 같이, 바이오마커 폴리뉴클레오티드는 하나 이상의 뉴클레오티드에서 검출가능하게 표지될 수 있다. 당업계에 알려진 임의의 방법을 사용하여 표적 폴리뉴클레오티드를 표지할 수 있다. 바람직하게는, 이 표지화는 RNA의 길이를 따라 균일하게 표지를 통합시키고, 보다 바람직하게는 표지화는 고도의 효율로 수행된다. 예를 들어, 폴리뉴클레오티드는 올리고-dT 프라이밍된 역전사에 의해 표지될 수 있다. 랜덤 프라이머(예를

들어, 9-mer)는 폴리뉴클레오티드의 전장(full length)에 걸쳐 표지된 뉴클레오티드를 균일하게 통합하기 위해 역전사에 사용될 수 있다. 대안적으로, 랜덤 프라이머는 폴리뉴클레오티드를 증폭시키기 위해 PCR 방법 또는 T7 프로모터-기반 시험관내 전사 방법과 함께 사용될 수 있다.

[0182] 상기 검출가능한 표지는 발광성 표지일 수 있다. 예를 들어, 형광 표지, 생물발광 표지, 화학발광 표지 및 비색 표지가 본 발명의 실시예에 사용될 수 있다. 사용될 수 있는 형광 표지는 플루오레세인, 인광체, 로다민 또는 폴리메틴 염료 유도체가 포함되나, 이에 제한되지는 않는다.

[0183] 추가적으로, FluorePrime(Amersham Pharmacia, Piscataway, N.J.), FluoreDite(Millipore, Bedford, Mass.), FAM(ABI, Foster City, Calif.) 및 Cy3 또는 Cy5(Amersham Pharmacia, Piscataway, N.J.)와 같은 형광 포스포아미다이트를 포함하나, 이에 제한되지 않는 상업적으로 이용가능한 형광 표지가 사용될 수 있다. 대안적으로, 검출가능한 표지는 방사성표지된 뉴클레오티드일 수 있다.

[0184] 일 구현예에서, 샘플로부터의 바이오마커 폴리뉴클레오티드 분자는 참조 샘플의 상응하는 폴리뉴클레오티드 분자와 차등적으로 표지된다. 상기 참조는 정상적인 생물학적 샘플 (즉, 대조군 샘플, 예를 들어, sLRI에 민감하지 않은 개체로부터의 자극된 CMBC) 또는 참조 생물학적 샘플 (예를 들어, sLRI에 감수성이 있는 개체로부터의 자극된 CMBC)로부터의 폴리뉴클레오티드 분자를 포함할 수 있다.

[0185] 핵산 혼성화 및 세척 조건은 표적 폴리뉴클레오티드 분자가 어레이의 상보적 폴리뉴클레오티드 서열, 바람직하게는 상보적 DNA가 위치하는 특정 어레이 부위에 특이적으로 결합하거나 특이적으로 혼성화하도록 선택된다. 그 위에 위치하는 이중-가닥 프로브 DNA를 함유하는 어레이는 바람직하게는 표적 폴리뉴클레오티드 분자와 접촉하기 전에 DNA를 단일-가닥으로 만들기 위해 변성 조건에 적용된다. 단일-가닥 프로브 DNA(예를 들어, 합성 올리고데옥시리보핵산)를 함유하는 어레이는 예를 들어, 자기-상보적 서열로 인해 형성되는 헤어핀 또는 이합체를 제거하기 위해, 표적 폴리뉴클레오티드 분자와 접촉하기 전에 변성될 필요가 있을 수 있다.

[0186] 최적의 혼성화 조건은 프로브 및 표적 핵산의 길이(예를 들어, 올리고머 대 염기 200개를 초과하는 폴리뉴클레오티드) 및 유형(예를 들어, RNA 또는 DNA)에 따라 달라질 것이다. 당업계의 기술자는 올리고뉴클레오티드가 더 짧아짐에 따라, 만족스러운 혼성화 결과를 위해 비교적 균일한 용융 온도를 달성하도록 이의 길이를 조정하는 것이 필요할 수 있음을 이해할 것이다. 핵산에 대한 특정 (즉, 엄격한) 혼성화 조건에 대한 일반적인 파라미터는 Sambrook 등, *Molecular Cloning: A Laboratory Manual* (제3판, 2001) 및 Ausubel 등, *Current Protocols In Molecular Biology*, vol. 2, Current Protocols Publishing, New York (1994)에 기술되어 있다. Schena 등의 cDNA 마이크로어레이에 대한 전형적인 혼성화 조건은 65°C에서 4시간 동안 5× SSC + 0.2% SDS에서 혼성화하고, 이어서 25°C에서 낮은 엄격도의 세척 완충액 (1× SSC + 0.2% SDS)에서 세척하고, 이어서 25°C에서 10분 동안 더 높은 엄격도의 세척 완충액 (0.1× SSC + 0.2% SDS)에서 세척하는 것이다 (Schena 등, *Proc. Natl. Acad. Sci. U.S.A.* 93:10614 (1993)). 유용한 혼성화 조건은 예를 들어, *Hybridization With Nucleic Acid Probes*, Elsevier Science Publishers B.V.; 및 Kricka, 1992, *Nonisotopic Dna Probe Techniques*, Academic Press, San Diego, Calif에서 또한 제공된다. 특히 바람직한 혼성화 조건은 1 M NaCl, 50 mM MES 완충액 (pH 6.5), 0.5% 소듐 사르코신 및 30% 포름아미드 내에서의 프로브의 평균 용융 온도 또는 이에 가까운 온도 (예를 들어, 51°C 이내, 보다 바람직하게는 21°C 이내)에서의 혼성화를 포함한다.

[0187] 형광 표지된 유전자 산물이 사용되는 경우, 마이크로어레이의 각 부위에서의 형광 방출은 바람직하게는 스캐닝 공초점 레이저 현미경(scanning confocal laser microscopy)에 의해 검출될 수 있다. 일 구현예에서, 적절한 여기(excitation) 라인을 사용하는 별도의 스캔이 사용된 2개의 형광단 각각에 대해 수행된다. 대안적으로, 2개의 형광단에 대한 특정 파장에서 동시에 표본 조명을 허용하는 레이저를 사용할 수 있고, 2개의 형광단으로부터의 방출을 동시에 분석할 수 있다 (Shalon 등, 1996, "A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization," *Genome Research* 6:639-645를 참조하고, 이는 모든 목적을 위해 그 전문이 참조로 포함된다). 어레이는 컴퓨터로 제어되는 X-Y 스테이지 및 현미경 대물렌즈가 있는 레이저 형광 스캐너로 스캔될 수 있다. 2개의 형광단의 순차적인 여기는 다중-라인의 혼합 가스 레이저로 달성되고, 방출된 광은 파장에 의해 분할되고, 2개의 광전자증배관으로 검출된다. 형광 레이저 스캐닝 장치는 Schena 등, *Genome Res.* 6:639-645 (1996), 및 본원에 인용된 기타 참조문헌에 기술되어 있다. 대안적으로, Ferguson 등, *Nature Biotech.* 14:1681-1684 (1996)에 기술된 광섬유 다발은 다수의 부위에서 mRNA 존재비 수준을 동시에 모니터링하는 데 사용될 수 있다.

[0188] 폴리뉴클레오티드는 또한 노던 블롯팅, 뉴클레아제 보호 검정, RNA 지문분석(fingerprinting), 중합효소 연쇄반응, 리가아제 연쇄반응, Q베타 레플리카아제, 등은 증폭 방법, 가닥 변위 증폭, 전사 기반 증폭 시스템, 뉴클레

아제 보호 (S 1 뉴클레아제 또는 RNase 보호 검정), SAGE 뿐만 아니라 국제특허 공개공보 WO 88/10315호 및 WO 89/06700호, 및 국제출원 PCT/US 87/00880호 및 PCT/US89/01025호에 개시된 방법을 포함하나, 이에 제한되지 않는 다른 방법에 의해 분석될 수 있으며; 그 전체가 본원에 참조로 포함된다.

[0189] 표준 노던 블롯 검정은 당업계의 통상의 기술자에게 알려진 종래의 노던 혼성화 기술에 따라, RNA 전사체 크기를 확인하고, 대안적으로 스플라이싱된 RNA 전사체 및 샘플 내 mRNA의 상대적인 양을 식별하는 데 사용될 수 있다. 노던 블롯에 있어서, RNA 샘플은 변성 조건 하에 아가로스 겔에서 전기영동에 의해 크기별로 먼저 분리된다. RNA는 이후 막으로 트랜스퍼되고, 교차-결합되며, 표지된 프로브와 혼성화된다. 랜덤-프라이밍, 역-번역 또는 PCR-생성 DNA 프로브, 시험관내 전사 RNA 프로브 및 올리고뉴클레오티드를 포함하여 비동위원소 또는 높은 특정 활성 방사성 표지된 프로브가 사용될 수 있다. 추가적으로, 부분적인 상동성만을 갖는 서열(예를 들어, 상이한 종으로부터의 cDNA 또는 엑손을 함유할 수 있는 게놈 DNA 단편)이 프로브로서 사용될 수 있다. 표지된 프로브, 예를 들어, 전장, 단일 가닥 DNA 또는 해당 DNA 서열의 단편을 함유하는 방사성 표지된 cDNA는 적어도 20개, 적어도 30개, 적어도 50개, 또는 적어도 100개 길이의 연속 뉴클레오티드일 수 있다. 프로브는 당업계의 숙련된 자에게 알려진 임의의 많은 상이한 방법에 의해 표지될 수 있다. 이러한 연구에 가장 일반적으로 사용되는 표지는 방사성 원소, 효소, 자외선에 노출되었을 때 형광을 발하는 화학물질 등이다. 다수의 형광 물질이 알려져 있으며 표지로서 활용될 수 있다. 이들은 이에 제한되지는 않으나, 플루오레세인, 로다민, 아우라민, 텍사스 레드, AMCA 블루 및 루시퍼 엘로우를 포함한다. 특정 검출 물질은 염소에서 제조되고 이소티오시아네이트를 통해 플루오레세인과 접합된 항-토끼 항체이다. 단백질은 또한 방사성 원소 또는 효소로 표지될 수 있다. 방사성 표지는 임의의 현재 이용 가능한 계수 절차에 의해 검출될 수 있다. 사용될 수 있는 동위원소는 이에 제한되지는 않으나, <sup>3</sup>H, <sup>14</sup>C, <sup>32</sup>P, <sup>35</sup>S, <sup>36</sup>Cl, <sup>35</sup>Cr, <sup>57</sup>Co, <sup>58</sup>Co, <sup>59</sup>Fe, <sup>90</sup>Y, <sup>125</sup>I, <sup>131</sup>I, 및 <sup>186</sup>Re를 포함한다. 효소 표지도 마찬가지로 유용하며, 현재 활용되는 임의의 비색법, 분광 광도법, 형광 분광 광도법, 전류 측정법 또는 가스 측정 기술에 의해 검출될 수 있다. 효소는 카르보다이미드, 디이소시아네이트, 글루타르알데히드 등과 같은 가교 분자와 반응함으로써 선택된 입자에 접합된다. 당업계의 기술자에게 알려진 임의의 효소가 활용될 수 있다. 이러한 효소의 예는 이에 제한되지는 않으나, 퍼옥시다제, 베타-D-갈락토시다제, 우레아제, 글루코스 옥시다제 + 퍼옥시다제 및 알칼리성 포스파타제를 포함한다. 미국특허 제3,654,090호, 제3,850,752호 및 제4,016,043호는 대안적인 표지 물질 및 방법의 개시를 위한 예로서 참조된다.

[0190] 뉴클레아제 보호 검정(리보뉴클레아제 보호 검정 및 S1 뉴클레아제 검정을 모두 포함함)은 특정 mRNA를 검출하고 정량화하기 위해 사용될 수 있다. 뉴클레아제 보호 검정에서, 안티센스 프로브(예를 들어, 방사성 표지 또는 비동위원소로 표지됨)는 용액에서 RNA 샘플에 혼성화된다. 혼성화 후, 단일-가닥의 비혼성화된 프로브 및 RNA는 뉴클레아제에 의해 분해된다. 아크릴아미드 겔은 나머지 보호된 단편을 분리하는 데 사용된다. 전형적으로, 용액 혼성화는 막-기반 혼성화보다 더 효율적이며, 이는 최대 20-30 μg의 블롯 혼성화와 비교하여 최대 100 μg의 샘플 RNA를 수용할 수 있다. 뉴클레아제 보호 검정의 가장 일반적인 유형인 리보뉴클레아제 보호 검정은 RNA 프로브의 사용을 필요로 한다. 올리고뉴클레오티드 및 기타 단일-가닥 DNA 프로브는 S1 뉴클레아제를 함유하는 검정에서만 사용될 수 있다. 단일-가닥의 안티센스 프로브는 전형적으로 뉴클레아제에 의한 프로브:표적 하이브리드의 절단을 방지하기 위해 표적 RNA와 완전히 상동성이어야 한다.

[0191] 연속 분석 유전자 발현(SAGE)은 또한 세포 샘플에서 RNA 존재비를 결정하는 데 사용될 수 있다. 예를 들어, Velculescu 등, 1995, Science 270:484-7; Carulli 등, 1998, Journal of Cellular Biochemistry Supplements 30/31:286-96을 참조하며, 본원에 그 전체가 참조로 포함된다. SAGE 분석은 검출을 위해 특별한 장치가 필요하지 않으며, 이는 다수의 전사 산물의 발현을 동시에 검출하는데 바람직한 분석 방법 중 하나이다. 먼저, 세포로부터 폴리 A<sup>+</sup> RNA를 추출한다. 다음으로, RNA는 비오티닐화된 올리고(dT) 프라이머를 사용하여 cDNA로 변환되고, 4-염기 인식 제한 효소(앵커링 효소: AE)로 처리되어 이들의 3' 말단에 비오틴 그룹을 함유하는 AE-처리된 단편이 생성된다. 다음으로, AE-처리된 단편은 결합을 위해 스트렙타비딘과 인큐베이션된다. 결합된 cDNA는 2개의 분획으로 나누어지고, 각 분획은 상이한 이중-가닥 올리고뉴클레오티드 어댑터(링커) A 또는 B에 연결된다. 이러한 링커는 다음으로 구성된다: (1) 앵커링 효소의 작용에 의해 형성된 돌출된 부분의 서열과 상보적인 서열을 갖는 돌출된 단일 가닥 부분, (2) 태깅 효소(TE) 역할을 하는 IIS-유형 제한 효소의 5' 뉴클레오티드 인식 서열(인식 부위로부터 20 bp 이하 떨어진 미리 결정된 위치에서 절단됨), 및 (3) PCR-특이적 프라이머를 구성하기 위한 충분한 길이의 추가 서열. 링커-연결된 cDNA는 태깅 효소를 사용하여 절단되고, 링커-연결된 cDNA 서열 부분만 남겨지며, 이는 짧은-가닥의 서열 태그 형태로 존재한다. 다음으로, 두 가지 상이한 유형의 링커로부터의 짧은-가닥 서열 태그의 풀(pool)이 서로 연결되고, 이어서 링커 A 및 B에 특이적인 프라이머를 사용한 PCR 증폭을 따른다. 결과적으로, 증폭 산물은 링커 A 및 B에 결합된 2개의 인접한 서열 태그(디태크

(ditag)의 무수한 서열을 포함하는 혼합물로서 획득된다. 증폭 산물은 앵커링 효소로 처리되고, 유리(free) 디태크 부분은 표준 연결 반응에서 가닥에 연결된다. 이후 증폭 산물이 복제된다. 클론의 뉴클레오티드 서열의 결정은 일정한 길이의 연속적인 디태그의 판독을 획득하는 데 사용될 수 있다. 각 태그에 상응하는 mRNA의 존재는 이후 클론의 뉴클레오티드 서열 및 서열 태그에 대한 정보로부터 식별될 수 있다.

[0192] 정량적 역전사 효소 PCR(qRT-PCR)은 또한 바이오마커의 발현 프로파일을 결정하는 데 사용될 수 있다(예를 들어, 미국특허 공개공보 제2005/0048542A1호 참조; 본원에 그 전체가 참조로 포함됨). RT-PCR에 의한 유전자 발현 프로파일링의 제1 단계는 RNA 주형을 cDNA로 역전사시키고, 이후 PCR 반응에서 기하급수적인 증폭을 따른다. 가장 일반적으로 사용되는 두 가지 역전사 효소는 조류 골수아세포증 바이러스 역전사 효소(AMV-RT) 및 몰로니 뮤린 백혈병 바이러스 역전사 효소(MLV-RT)이다. 역전사 단계는 전형적으로 상황 및 발현 프로파일링의 목표에 따라, 특정 프라이머, 랜덤 헥사머 또는 올리고-dT 프라이머를 사용하여 프라임된다. 예를 들어, 추출된 RNA는 제조사의 지시에 따라 GeneAmp RNA PCR 키트(Perkin Elmer, CA, USA)를 사용하여 역전사될 수 있다. 이후, 파생된 cDNA는 후속 PCR 반응에서 주형으로 사용될 수 있다.

[0193] 상기 PCR 단계는 다양한 내열성 DNA-의존성 DNA 중합효소를 사용할 수 있지만, 이는 전형적으로 5'-3' 뉴클레아제 활성을 갖지만 3'-5' 교정 엔도뉴클레아제 활성은 결여된 Taq DNA 중합효소를 사용한다. 따라서, TAQMAN PCR은 전형적으로 Taq 또는 Tth 중합효소의 5'-뉴클레아제 활성을 활용하여 이의 표적 애플리콘에 결합된 혼성화 프로브를 가수분해하지만, 동등한 5' 뉴클레아제 활성을 갖는 임의의 효소가 사용될 수 있다. 2개의 올리고뉴클레오티드 프라이머는 PCR 반응의 전형적인 애플리콘을 생성하는 데 사용된다. 제3 올리고뉴클레오티드 또는 프로브는 두 PCR 프라이머 사이에 위치한 뉴클레오티드 서열을 검출하도록 설계되었다. 프로브는 Taq DNA 중합효소에 의해 연장 가능하지 않으며, 이는 리포터 형광 염료 및 퀸처(quencher) 형광 염료로 표지된다. 리포터 염료로부터 임의의 레이저-유도 방출은 두 염료가 프로브 상에 있음으로서 서로 가까이 위치할 때 퀸칭 염료에 의해 퀸칭된다. 증폭 반응 동안, Taq DNA 중합효소는 주형-의존적 방식으로 프로브를 절단한다. 생성된 프로브 단편은 용액에서 분리되고, 방출된 리포터 염료로부터의 신호는 상기 제2 형광단의 퀸칭 효과가 없다. 한 분자의 리포터 염료가 합성된 각각의 신규 분자에 대해 유리되고, 미퀸칭된(unquenched) 리포터 염료의 검출은 상기 데이터의 정량적 해석을 위한 기초를 제공한다.

[0194] TAQMAN RT-PCR은 예를 들어, ABI PRISM 7700 서열 검출 시스템 (Perkin-Elmer- Applied Biosystems, Foster City, Calif., USA), 또는 Lightcycler (Roche Molecular Biochemicals, Mannheim, Germany)와 같은 상업적으로 이용가능한 장비를 사용하여 수행될 수 있다. 바람직한 구현예에서, 5' 뉴클레아제 절차는 ABI PRISM 7700 서열 검출 시스템과 같은 실시간 정량적 PCR 장치 상에서 실행된다. 시스템은 열순환기(thermocycler), 레이저, 전하-결합 장치(CCD), 카메라 및 컴퓨터로 구성된다. 시스템은 기기를 실행하고 데이터를 분석하기 위한 소프트웨어를 포함한다. 5'-뉴클레아제 검정 데이터는 초기에 Ct로 또는 역치 주기로 표현된다. 형광 값은 매 주기 동안 기록되며, 증폭 반응의 해당 지점까지 증폭된 산물의 양을 나타낸다. 형광 신호가 통계적으로 유의미한 것으로 처음 기록되는 시점이 역치 주기(Ct)이다.

[0195] 오류 및 샘플-대-샘플 변화의 영향을 최소화하기 위해, RT-PCR은 일반적으로 내부 표준을 사용하여 수행된다. 이상적인 내부 표준은 상이한 조직 중에 일정한 수준으로 발현되며, 실험적 치료에 의해 영향을 받지 않는다. 유전자 발현 패턴을 정규화하는 데 가장 자주 사용되는 RNA는 하우스키핑 유전자인 글리세르알데히드-3-인산염-탈수소효소(GAPDH) 및 베타-액틴에 대한 mRNA이다.

[0196] RT-PCR 기술의 보다 최근의 변화는 실시간 정량적 PCR이며, 이는 이중-표지된 형광성 프로브(즉, TAQMAN 프로브)를 통해 PCR 산물 축적을 측정한다. 실시간 PCR은 각 표적 서열에 대한 내부 경쟁자를 정규화에 사용하는 정량적 경쟁 PCR, 및 샘플 내에 함유된 정규화 유전자 또는 RT-PCR을 위한 하우스키핑 유전자를 사용하는 정량적 비교 PCR과 모두 호환가능하다. 추가의 세부사항의 경우, 예를 들어, Held 등, Genome Research 6:986-994 (1996)을 참조한다.

[0197] RNA-시퀀싱(RNA-seq)과 같은 차세대 시퀀싱 방법은 또한 샘플 세포에서 RNA 존재비를 평가하기 위해 사용될 수 있다. RNA-seq을 수행하는 예시적인 프로토콜은 RNA ACCESS® 프로토콜 또는 TRUSEQ® RIBO-ZERO® 프로토콜 (ILLUMINA®)을 포함한다. 당업계는 숙련된 자는 또한 총(total) RNA 시퀀싱, mRNA 시퀀싱, 3' mRNA 시퀀싱, 5' mRNA 시퀀싱, CAGE-Seq를 포함하나, 이에 제한되지 않는 RNA-seq를 수행하는 다양한 방법을 인식할 것이다.

[0198] 바이오마커 데이터는 환자가 sLRI에 감수성이 있는지 여부를 평가하기 위해 바이오마커를 식별하고 테스트 및 참조 발현 프로파일 간의 관찰된 바이오마커 수준 차이의 통계적 유의성을 결정하는 다양한 방법에 의해 분석될 수 있다. 특정 구현예에서, 환자 데이터는 다변량 선형 관별 분석(LDA), 수신자 조작 특성(ROC) 분석, 주성분

분석(PCA), 앙상블 데이터 마이닝 방법, 베이지안 일반화 선형 모델, 가우시안 프로세스, 나이브 베이즈, 엘라스틱 넷(elastic net),  $k$ -최근접 이웃( $k$ -nearest neighbors), 라쏘(lasso), 페널티 로지스틱 회귀(penalized logistic regression), 부분 최소 제곱법(partial least squares), 마이크로어레이 예측 분석(PAM), 포아송 선형 판별 분석(poisson linear discriminant analysis), 음이항 선형 판별 분석(negative-Binomial linear discriminant analysis), 신경망, 서포트 벡터 머신, 마이크로어레이의 유의성 분석(SAM), 마이크로어레이의 세포 특이적 유의성 분석(csSAM), 밀도-정규화 이벤트의 스페닝-트리 진행 분석(SPADE), 및 다차원 단백질 식별 기술(MUDPIT) 분석을 포함하나, 이에 제한되지 않는 하나 이상의 방법에 의해 분석된다. (예를 들어, Hilbe (2009) Logistic Regression Models, Chapman & Hall/CRC Press; McLachlan (2004) Discriminant Analysis and Statistical Pattern Recognition. Wiley Interscience; Zweig 등 (1993) Clin. Chem. 39:561-577; Pepe (2003) The statistical evaluation of medical tests for classification and prediction, New York, NY: Oxford; Sing 등 (2005) Bioinformatics 21:3940-3941; Tusher 등 (2001) Proc. Natl. Acad. Sci. U.S.A. 98:5116-5121; Oza (2006) Ensemble data mining, NASA Ames Research Center, Moffett Field, CA, USA; English 등 (2009) J. Biomed. Inform. 42(2):287-295; Zhang (2007) Bioinformatics 8: 230; Shen-Orr 등 (2010) Journal of Immunology 184: 144-130; Qiu 등 (2011) Nat. Biotechnol. 29(10):886-891; Ru 등 (2006) J. Chromatogr. A. 1111(2): 166- 174, Jolliffe Principal Component Analysis (Springer Series in Statistics, 제2판, Springer, NY, 2002), Koren 등 (2004) IEEE Trans Vis Comput Graph 10:459-470 참조; 그 전체가 본원에 참조로 포함됨).

[0199] 상기 바이오마커 데이터, 즉 유전자 발현 데이터를 분석하는 바람직한 방법은 랜덤 포레스트 분류기에 의한 것이다.

[0200] **랜덤 포레스트 분류기**

[0201] 랜덤 포레스트 분류기는 많은 결정 트리로 구성되는 앙상블 분류기이며, 개별 트리별로 출력되는 클래스의 모드인 클래스를 출력한다. 랜덤 포레스트는 교차-유효성 검증 대신 부트스트래핑을 활용한다. 각 반복에 대해, 랜덤 샘플(대체 포함)이 추출되고 가능한 가장 큰 트리가 성장한다. 각 트리는 최종 클래스 예측에서 투표를 받는다. 랜덤 포레스트에 맞추기 위해, 트리 수(예를 들어, 부트스트랩 반복)가 지정된다. 랜덤 포레스트 알고리즘은 훈련 정확도의 평균 감소에 의해 바이오마커 중요도를 측정한다. 랜덤 포레스트 방법은 다수의 상이한 결정 트리를 사용한다. 바이오마커는 유의미한 랜덤 포레스트 분석으로부터의 결정 트리의 결정 분기 역할을 하는 경우 식별력 있는 유의성을 갖는 것으로 간주된다.

[0202] 랜덤 포레스트 (또는 랜덤 포레스트들)는 많은 결정 트리로 구성되는 앙상블 분류기이며, 개별 트리별로 출력되는 클래스의 모드인 클래스를 출력한다(Breiman, Leo (2001). "Random Forests". Machine Learning 45 (1): 5-32). 랜덤 포레스트는 이용가능한 가장 정확한 학습 알고리즘 중 하나이며, 즉, 데이터 세트에 대해 매우 정확한 분류기를 생성한다. (Caruana, Rich; Karampatziakis, Nikos; Yessenalina, Ainur (2008). "An empirical evaluation of supervised learning in high dimensions." Proceedings of the 25th International Conference on Machine Learning (ICML)). 이 방법은 제어된 변화로 결정 트리의 집합을 구성하기 위해 "배깅 (bagging)" 및 특징의 랜덤 선택을 조합한다. 특징에 대한 랜덤 하위 집합의 선택은 확률적인 판별을 구현하는 방법인 랜덤 부분공간(subspace) 방법의 예이다. 부트스트랩 분포는 원본 데이터에 기초하여 통계량의 변화를 추정하는 방법으로 사용된다. 부트스트랩 샘플, 예를 들어, 150 또는 500에서 성장한 각 트리에 대해, 부트스트랩 샘플에서 제외된 관측치의 오류율이 모니터링된다. 이를 "아웃-오브-백(Out-of-bag)" 오류율로 칭한다.

[0203] 각 트리는 다음 알고리즘을 사용하여 구성된다: (1) 훈련 사례의 수를  $N$ 으로 하고, 분류기의 변수의 수가  $M$ 이 되도록 한다; (2) 트리의 노드에서 결정을 결정하기 위해 입력 변수의 수  $m$ 이 사용되도록 하며;  $m$ 은  $M$ 보다 훨씬 작아야 한다; (3) 모든  $N$ 개의 이용가능한 훈련 사례로부터 대체물로  $n$ 회를 선택함으로써 이 트리에 대한 훈련 세트를 선택하고(즉, 부트스트랩 샘플을 취함), 나머지 사례를 사용하여 이들의 클래스를 예측함으로써 트리의 오류를 추정한다; (4) 트리의 각 노드에 대해, 해당 노드에서 결정의 기반이 되는  $m$ 개의 변수를 랜덤으로 선택한다. 훈련 세트에서 이들  $m$ 개의 변수를 기반으로 최상의 분할을 계산한다; 그리고 (5) (정상적인 트리 분류기를 구성하는 데 있어서 수행될 수 있는 바와 같이) 각 트리는 완전히 성장하고 전지(pruned) 하지 않는다.

[0204] 예측을 위해 새로운 샘플이 트리로 푸시다운(push down) 된다. 이는 종료되는 단말 노드에 훈련 샘플의 표지가 할당된다. 이 절차는 앙상블의 모든 트리에 걸쳐 반복되며, 모든 트리의 모드 투표는 랜덤 포레스트 예측으로서 기록된다.

[0205] 일 구현예에서, 랜덤 입력을 사용한 트리 포레스트를 기반으로 분류 및 회귀를 수반하는 랜덤 포레스트 분석은

"randomForest: Breiman and Cutler's random forests for classification and regression" (중속성: R (>=2.5.0), 통계) (버전: 4.6-6) (2012-01-06) (Leo Breiman 및 Adele Cutler에 의한 포트란(Fortran) 원본, Andy Liaw 및 Matthew Wiener에 의한 R 포트(port))를 사용하여 수행된다. A. Liaw 및 M. Wiener (2002). Classification and Regression by randomForest. R News 2(3), 18-22를 참조한다.

[0206] 랜덤 포레스트는 Liaw 및 Wiener, R News Vol. 2/3, 2002년 12월, pgs. 18-22; Dfaz-Uriarte 및 Alvarez, BMC Bioinformatics. 2006년 1월 6일; 7:3); Statnikov 등, BMC Bioinformatics. 2008년 7월 22일; 9:319; Shi 등, Mod Pathol. 2005년 4월; 18(4):547-57, Breiman, 1999, "Random Forests—Random Features," Technical Report 567, Statistics Department, U.C. Berkeley, 1999년 9월에 추가로 기술되어 있고, 이는 그 전체가 참조로 본원에 포함되어 있으며, 이들 각각은 그 전체가 본원에 참조로 포함된다.

[0207] **호흡기 감염**

[0208] 본 발명은 호흡기 감염에 대한 감수성을 결정하기 위한 방법을 제공한다. 바람직하게는, 호흡기 감염은 하부 호흡 기도 감염이다. 임의의 구현예에서, 감염은 박테리아 또는 바이러스 감염일 수 있다. 박테리아 감염은 본원에 기술된 바와 같은 임의의 것일 수 있다. 바이러스 감염은 본원에 기술된 바와 같은 임의의 것일 수 있다.

[0209] 본원에 사용된 바와 같이, 용어 호흡기 감염은 호흡 기도 어느 곳이든 바이러스 또는 박테리아에 의한 감염을 의미한다. 호흡기 감염의 예는 감기, 부비동염, 인후염, 편도염, 후두염, 기관지염, 폐렴 또는 세기관지염을 포함하나 이에 제한되지는 않는다. 바람직하게는, 본 발명의 임의의 구현예에서 호흡기 감염은 감기이다.

[0210] 개체는 바이러스 테스트에 의해 호흡 기도 감염을 갖는 것으로 식별될 수 있으며, 가려운 유루안(itchy watery eye), 콧물, 코막힘, 재채기, 인후염, 기침, 두통, 발열, 불쾌감, 피로 및 허약의 증상을 나타낼 수 있다. 일 측면에서, 호흡기 감염을 갖는 대상체는 임의의 다른 호흡기 병태를 갖지 않을 수 있다. 바이러스의 존재 또는 양의 검출은 임상 샘플(비강 세척물, 가래, BAL) 또는 혈청으로부터 분리된 RNA의 PCR/시퀀싱에 의한 것일 수 있다.

[0211] 인플루엔자(일반적으로 "독감"으로 지칭됨)는 조류 및 포유동물에 영향을 미치는 오르토믹소바이러스과(Orthomyxoviridae)의 RNA 바이러스(인플루엔자 바이러스)에 의해 야기되는 감염성 질환이다. 질환의 가장 흔한 증상은 오한, 발열, 인후통, 근육통, 중증 두통, 기침, 허약/피로 및 전반적인 불편함이다.

[0212] 인플루엔자 바이러스는 오르토믹소바이러스과(Orthomyxoviridae)의 5개 속 중 3개를 구성한다. 인플루엔자 A형 및 B형 바이러스는 계절적인 유행 동안 공-순환하며, 중증 인플루엔자 감염을 야기할 수 있다. 인플루엔자 C형 바이러스 감염은 덜 흔하지만 중증일 수 있으며 지역적 유행을 야기할 수 있다.

[0213] 인플루엔자 A형 바이러스는 이들 바이러스에 대한 항체 반응을 기반으로 상이한 혈청형 또는 아형으로 세분화될 수 있다. 인플루엔자 A 바이러스는 바이러스 표면 상의 두 단백질을 기반으로 아형으로 나뉜다: 헤마글루티닌(H) 및 뉴라미니다아제(N). 18개의 상이한 헤마글루티닌 아형 및 11개의 상이한 뉴라미니다아제 아형(각각 H1 내지 H18 및 N1 내지 N11)이 있다. 인간에서 확인된 아형은 H1N1, H1N2, H2N2, H3N2, H5N1, H7N2, H7N3, H7N7, H9N2 및 H10N7이다.

[0214] 인플루엔자는 이환율 및 심지어 사망률을 포함하는 파괴적인 건강 문제 외에도 심각한 경제적 영향으로 공중 보건에 막대한 영향을 미친다. 따라서, 개체의 감염을 예방하거나 감염의 중증도를 감소시킬 수 있는 치료제가 필요하다.

[0215] 임의의 구현예에서, 예방이 필요한 인플루엔자 감염은 인플루엔자 A형, B형 또는 C형으로 이루어진 군으로부터 선택된 바이러스에 의한 감염이다. 인플루엔자 A형 바이러스는 이들 바이러스에 대한 항체 반응을 기반으로 상이한 혈청형 또는 아형으로 세분화될 수 있다. 인플루엔자 A 바이러스는 바이러스 표면 상의 두 단백질을 기반으로 아형으로 나뉜다: 헤마글루티닌(H) 및 뉴라미니다아제(N). 18개의 상이한 헤마글루티닌 아형 및 11개의 상이한 뉴라미니다아제 아형(각각 H1 내지 H18 및 N1 내지 N11)이 있다. 인간에서 확인된 아형은 H1N1, H1N2, H2N2, H3N2, H5N1, H7N2, H7N3, H7N7, H9N2 및 H10N7이다.

[0216] 본 발명의 임의의 측면에서, 병태는 라이노바이러스 또는 호흡기 세포융합 바이러스(RSV)에 의해 유발될 수 있다. 추가로, 본 발명의 임의의 측면에서, 바이러스 매개 악화는 라이노바이러스 또는 RSV 매개이다. 라이노바이러스 또는 RSV는 본원에 기술된 임의의 혈청형일 수 있다. 전형적으로, 라이노바이러스는 RV-A, RV-B 또는 RV-C 라이노바이러스 종의 구성원이다.

[0217] 본 발명의 다른 측면에서, 병태는 인플루엔자 바이러스 과/속, 파라인플루엔자, 코로나바이러스, 아데노바이러

스 및 메타뉴모바이러스에 의해 유발될 수 있다.

[0218] **치료, 투여, 투여량 및 제형**

[0219] 본 발명은 생애 초기에 호흡기 감염, 예를 들어 중증 하부 호흡 기도 감염에 감수성이 있는 개체의 식별을 허용한다. 이는 선제적 요법의 형태로 개입에 대한 기회를 제공한다.

[0220] 예시적인 선제적 치료제는 팔리비주맙, 프레드니솔론, 오말리주맙 또는 폴리박테리아 제형 및 이들의 임의의 조합을 포함한다.

[0221] 용어 '호흡기'는 코, 목구멍, 후두, 기관, 기관지 및 폐를 포함하는 신체 시스템을 통해 산소가 체내로 유입되고 이산화탄소가 배출되는 과정을 지칭한다.

[0222] 본원에 사용된 바와 같이, 상부 호흡 기도는 다음 영역을 포함할 수 있다: 코 및 비강, 부비동, 인두 및 성대 주름 (성대) 위의 후두 부분. 전형적으로, 하부 호흡 기도는 다음 영역을 포함한다: 성대 아래의 후두 부분, 기관, 기관지 및 세기관지. 폐는 하부 호흡 기도에 포함될 수 있으며, 호흡 세기관지, 폐포관, 폐포낭 및 폐포를 포함한다.

[0223] 용어 '호흡기 질환' 또는 '호흡기 병태'는 염증을 수반하고, (비강, 인두 및 후두 포함하는) 상부 및 (기관, 기관지 및 폐를 포함하는) 하부 호흡 기도를 포함하는 호흡기계(respiratory system)의 구성 요소에 영향을 미치는 여러 질병 중 임의의 하나를 지칭한다.

[0224] 호흡기 질환의 증상에는 기침, 과도한 가래 생성, 숨가쁨의 느낌 또는 천명음이 들리는 흉부 답답함이 포함될 수 있다. 운동 능력은 상당히 제한될 수 있다. 천식에 있어서 체중, 키 및 연령을 기준으로 노모그래프적으로 (nomographically) 예측된 것의 백분율로서 FEV<sub>1.0</sub>(1초 내의 강제호기량)은 강제호기에서 최고조 호기 유량일 수 있는 바와 같이 감소될 수 있다. COPD에 있어서 FVC의 비율로서 FEV<sub>1.0</sub>은 전형적으로 0.7 미만으로 감소된다. 이러한 각 조건의 영향은 직장/학교 결석 일수, 수면 방해, 기관지확장제 약물에 대한 요건, 경구용 글루코코르티코이드를 포함한 글루코코르티코이드에 대한 요건에 의해 또한 측정될 수 있다.

[0225] 호흡기 질환의 존재, 개선, 치료 또는 예방은 대상체의 임의의 임상적 또는 생화학적 관련 방법 또는 이로부터의 생검에 의해 결정될 수 있다. 예를 들어, 측정된 파라미터는 폐 기능의 존재 또는 정도, 폐색의 징후 및 증상; 운동 관용; 야간 각성; 학교 또는 직장에 대한 결석일; 기관지확장제 사용; 흡입된 코르티코스테로이드 (ICS) 용량; 경구용 (글루코코르티코이드) GC 사용; 다른 의약품에 대한 필요성; 의학 치료에 대한 필요성; 병원 입원일 수 있다.

[0226] 용어, 대상체의 "치료" 또는 "치료하는"은 질환 또는 병태, 질환 또는 병태의 증상, 또는 질환 또는 병태의 위험 (또는 이에 대한 감수성)을 지연시키거나, 둔화시키거나, 안정화시키거나, 치유하거나, 치료하거나, 경감시키거나, 완화시키거나, 변경하거나, 구체하거나, 덜 악화시키거나, 개량하거나, 개선시키거나, 영향을 미치기 위한 목적으로 본원에 기술된 치료 화합물의 적용 또는 투여를 포함한다. 용어, "치료하는"은 경감과 같은 임의의 객관적 또는 주관적 파라미터를 비롯하여, 부상, 병리 또는 병태의 치료 또는 완화에 있어서의 임의의 성공 지표; 관해; 악화 속도의 완화; 질환의 중증도 완화; 안정화, 증상의 감소 또는 대상체에 대한 부상, 병리 또는 병태를 더 관용가능하게 만드는 것; 퇴화 또는 쇠퇴 속도의 둔화; 퇴화의 최종점을 덜 쇠퇴하게 만드는 것; 또는 대상체의 신체적 또는 정신적 웰빙(well-being)을 개선하는 것을 지칭한다.

[0227] 요법에 대한 양성 반응은 또한 호흡기 바이러스 감염 후 호흡기 증상의 악화, 예를 들어, 천식 증상 (악화)의 예방 또는 약화일 수 있다. 이는 주니퍼 천식 제어 설문지(Juniper Asthma Control Questionnaire) (ACQ-6)를 기반으로 기준선에서 연구 기간의 종료까지의 질환 점수의 평균 변화에 대한 비교에 의해 평가될 수 있으며, 감염/감기 증상의 발병 후 매일 하부 호흡기 증상 점수(LRSS - 흉부 답답함, 천명음, 호흡 곤란 및 기침)를 평가할 수도 있다. 기준선 폐 함수(최고조 호기 유량 PEF)로부터 변화를 또한 평가할 수 있으며, 요법에 대한 양성 반응은 감소된 PEF의 유의미한 약화일 수 있다. 예를 들어, 위약 치료군은 악화의 최고조에서 15%의 오전 PEF의 유의미한 감소를 나타낸 반면, 치료군은 기준선으로부터 15% 미만의 PEF 변화로 유의미하지 않은 감소를 나타냈다.

[0228] 본 발명의 방법에 따라 사용하기 위한 요법은 유효량으로 투여되는 것이다. 문구 '치료학적 유효량' 또는 '유효량'은 (i) 특정 질환, 병태 또는 장애를 치료하거나, (ii) 특정 질환, 병태 또는 장애의 하나 이상의 증상을 약화시키거나, 개선하거나, 제거하거나, (iii) 본원에 기술된 특정 질환, 병태 또는 장애의 하나 이상의 증상의 발병을 지연시키는 본원에 기술된 바와 같은 요법을 지칭한다. 바람직하지 않은 효과, 예를 들어, 부작용은 때

때로 원하는 치료 효과와 함께 나타나며; 따라서, 의사는 적절한 "유효량"이 무엇인지 결정하는 데 있어서 잠재적 위험에 대한 잠재적 이익의 균형을 유지한다.

[0229] 필요한 정확한 양은 대상체의 종, 연령 및 일반적인 병태, 투여 방식 등에 따라 대상체마다 다를 것이다. 따라서, 정확한 "유효량"을 지정하는 것이 가능하지 않을 수 있다. 그러나, 임의의 개별 경우에 적절한 "유효량"은 정례적인 실험만을 사용하여 당업계의 통상의 기술자에 의해 결정될 수 있다.

[0230] 본원에 기술된 바와 같은 치료제는 건조 분말, 스프레이, 미스트 또는 에어로졸을 포함하는, 비강내 투여용으로 제형화될 수 있다. 이는 호흡기 감염의 치료에 특히 바람직할 수 있다.

[0231] 예를 들어, 비강 스프레이 또는 점비액과 같은 투여를 위한 담체가 액체인 적합한 제형은 활성 성분에 대해 수성 또는 유성 용액을 포함한다. 대안적으로, 치료제는 건조 분말로서 제공될 수 있으며, 단지 본원에 정의된 바와 같이 상부 호흡 기도에 투여될 수 있다.

[0232] 적절한 담체의 선택은 고려되는 특정 투여 유형에 따라 달라진다. 상부 호흡 기도, 예를 들어 비점막 표면을 통한 투여를 위해, 본원에 기술된 치료의 활성 화합물은 용액, 예를 들어, 물 또는 등장성 식염수, 완충액 또는 비완충액 또는 현탁액, 비강내 투여를 위해 점적제(drops) 또는 스프레이로 제형화될 수 있다. 바람직하게는, 이러한 용액 또는 현탁액은 비강 분비물에 대해 등장성이고, 예를 들어 약 pH 4.0 내지 약 pH 7.4 또는 pH 6.0 내지 pH 7.0 범위와 같이 거의 동일한 pH이다. 완충액은 생리학적으로 호환가능하며, 단순히 예를 들어 인산염 완충액을 포함하여야 한다. 예를 들어, 대표적인 코막힘 완화제는 약 6.2의 pH로 완충되는 것으로 기술되어 있다(Remington's, Id., 페이지 1445). 물론, 통상의 기술자는 비강 및/또는 상부 호흡기 투여를 위해 무해한 수성 담체에 적합한 식염수 함량 및 pH를 용이하게 결정할 수 있다.

[0233] 제형에 추가적인 점도, 수분 유지 및 기분 좋은 질감 및 향을 제공하기 위해, 당업계에 알려진 방부제, 착색제, 윤활성 또는 점성 미네랄 또는 식물성 오일, 향료, 아로마 오일과 같은 천연 또는 합성 식물 추출물, 및 예를 들어, 글리세롤과 같은 보습제 및 점도 강화제와 같은 기타 성분도 포함될 수 있다. 본원에 기술된 치료제의 비강 투여를 위해, 점적제, 드롭렛(droplets) 및 스프레이의 생성을 위한 다양한 장치가 당업계에서 이용가능하다. 예를 들어, 본원에 기술된 치료제는 수동으로 구동되는 펌프, 예를 들어, 한쪽 말단에 부착된 유연한 고무 벌브(bulb)에 의해 제공되는 공기압에 의해 내용물이 한 방울씩 배출되는 유리, 플라스틱 또는 금속 분배 튜브를 포함하는 간단한 점적기 (또는 파이켓)에 의해 비강에 투여될 수 있다.

[0234] **실시예**

[0235] **실시예 1 - 재료 및 방법**

[0236] **연구 모집단**

[0237] 대상체는 이전에 기술된 바와 같이 천식 발병의 높은 위험에 대해 출생 전 등록된 10년 전향적(prospective) 출생 코호트인, 아동기 천식 연구로부터의 50 개체의 하위 집합이었다(Kusel 등, *J Allergy Clin Immunol*, 2007, 119:1105-1110; Holt 등, *J Allergy Clin Immunol*, 2019, 143:1176-1182 e1175; Kusel 등, *Pediatr Infect Dis J*, 2006, 25:680-686; Kusel 등, *Eur Respir J*, 2012, 39:876-882; Holt 등, *J Allergy Clin Immunol*, 2010, 125:653-659; Kusel 등, *J Allergy Clin Immunol*, 2005, 116:1067-1072). 흉부 가랑거림 외에 천명음 및 /또는 발열이 존재하는 경우 급성 호흡기 감염을 sLRI로 간주하였다. 가랑거림(흉부 가랑거림)은 아동의 흉부로부터 들리는 습한 시끄러운 숨소리로 정의되는 반면, 천명음은 가청, 호기, 고음의 휘파람 소리로 정의된다. 발열은 호흡기 감염 증상의 시작 48시간 동안 1시간 이상 간격으로 두 번 측정하여 온도 >38°C(디지털 온도계)를 기록하여 정의하였다. 호흡기 바이러스 감염 병력은 증상 발생 48시간 이내에 가정 방문 동안 수집된 상세한 평가 및 비인두 흡인물(RT-PCR)로부터 결정되었다(Kusel 등, *J Allergy Clin Immunol*, 2007, 119:1105-1110; Kusel 등, *Pediatr Infect Dis J*, 2006, 25:680-686). 5세때 현재 진행되는 천명음(Current wheeze at 5 years, Crwz5)은 5년 추적 관찰 전 12개월 내 기록된 임의의 천명음 이벤트(부모 평가)로 정의되었다. 5세때 천식은 천식에 대해 의사 진단, 천식 의약품에 대한 처방, 5세때 현재 진행되는 천명음을 갖는 것으로 정의되었다. 5세때 비-천식 결정은 이러한 기준이 전혀 없었다. 제대혈(Umbilical cord blood)은 출생 시 수집되었고, 말초혈액은 (가능한 출생일에 가까운) 0.5세, 1세, 2세, 3세, 4세, 5세 및 10세에 수집되었다.

[0238] **면역표현형**

[0239] 동결보존된 CBMC를 해동하고 10% 비-가열-불활성화된 FBS(Serana Australia)를 함유한 RPMI 1640(Gibco)으로 세척하였다. 10 μl의 세포 혼합물을 트리판 블루로 염색하고 혈구계수기로 계수하였다. 대략 1×10<sup>6</sup> 개 세포는

면역표현형을 위해 분취되었고,  $0.25 \times 10^6$  개 세포는 각 샘플에 대한 비염색된 대조군을 위해 분취되었다. 세포를 4°C에서 5분 동안 1500rpm(~500g)으로 원심분리하여 펠릿화하고, 진공 흡인에 의해 과잉 배지를 제거하였다. 각 샘플을 차가운 FACS 완충액(PBS + 1% BSA) 중의 단클론 항체 마스터 믹스(CD19-FITC *RRID: AB\_395812*, CD3-AF700 *RRID: AB\_396952*, CD4-V500 *RRID: AB\_1937323*, CD14-APC-Cy7 *RRID: AB\_1645464*, HLA-DR-PerCP-Cy5.5 cat #347364, CD25-BV421 *RRID: AB\_11154578*, CD127-BV605 *RRID: AB\_2738138*, CD123-CF594 *RRID: AB\_11153664*, CD11c-PE-Cy7 *RRID: AB\_10611859* [BD Bioscience] 및 FcεRIα-APC *RRID: AB\_10671394* [eBioscience]) 50 μl와 함께 30분 동안 4°C로 암실에서 인큐베이션하였다. 세포를 세척하고, (Cytotfix/Cytoperm 완충액(BD Biosciences))으로 1시간 동안 고정하고, 투과시키고, FoxP3-PE (세포내, BD Biosciences)와 함께 30분 동안 인큐베이션하였다. 동일한 항체 배치를 제조사가 권장하는 희석으로 모든 샘플에 대해 사용하였다. 개별 세포는 각 세포 계측법 실행 전 품질 관리 평가(Rainbow 보정 및 CS&T 비드(BD Biosciences)) 이후 FACSDiva 소프트웨어(BD Biosciences)가 포함된 LSR-Fortessa 플랫폼을 사용하여 획득되었으며, 비염색된 대조군은 각 샘플에 대해 포함되었다. 처음에, 샘플을 FlowJo 10.3 소프트웨어를 사용하여 보정하고 게이팅하였다. 보정된 FCS 파일을 R(3.6.2) 통계적 환경으로 불러오고 *flowWorkspace* 및 *flowCore* 패키지로 사전-처리하였다. 로직 변환(*flowCore*) 및 배치 보정(*sva*)을 모든 샘플에 적용하였다. 비모수적 대응표본 테스트(윌콕슨 부호 랭크 테스트) 또는 독립표본 테스트(맨-휘트니 U 테스트)를 그룹간 차이를 결정하기 위해 사용하였다.

[0240] 시험관내 세포 배양

[0241] 샘플은 랜덤화된 블록을 할당하고 일관된 시약/자극 스톡(stock)을 사용하여 동일한 직원에 의해 순차적으로 배양되었다. 체대혈 적혈구는 면역자기적으로 고갈시켰으며(EasySep 키트, StemCell), 각 샘플은 일치된 비자극 대조군과 함께 LPS (Enzo Biochem, 1 ng/ml), 이미퀴모드 (Invivogen, 5 μl/ml) 및 폴리(I:C) (Invivogen, 50 μl/ml)를 포함한 RPMI + 5% AB 혈청 (비-가열 불활성화됨, Sigma-Aldrich)으로 18시간 동안 배양(37°C, 5% CO<sub>2</sub>) 하였다. 배양 상층액의 분취액은 사이토카인 정량화를 위해 -20°C에서 보관하였다. 세포 펠릿은 RNA 추출을 위해 -20°C에서 트리졸(Trizol) (Invitrogen)에 보관하였다.

[0242] 데이터 생성

[0243] *RNA-Seq*: RNA를 RNeasy MinElute 키트(Qiagen)를 사용하여 배치별로 추출하고 추출 배치 정보를 기록하였다. RNA 농도를 측정하였고(Bioanalyzer, Agilent, Santa Clara, USA) 양호한 품질인 것으로 확인되었다(RIN 점수; 평균 = 8.514, 95%CI = 8.46-8.567). NEBNext Ultra II 키트 (New England BioLabs, Massachusetts, USA)로 제조된 시퀀싱 라이브러리에 저수율 프로토콜이 사용되었으며, 시퀀싱을 위해 호주 게놈 연구소(Australian Genome Research Facility) (AGRF, Melbourne, Australia)에서 NovaSeq 6000 (Illumina, San Diego, USA) 플랫폼을 사용하여 시퀀싱하였다 (100 bp의 페어드-엔드(paired-end)).

[0244] *사이토카인*: 48개의 사이토카인(Bio-plex Pro, BioRad)의 농도는 Luminex 200 시스템(Luminex)으로 동시에 정량화하였다. 분석물 정량화(pg/ml)는 표준 곡선에 대한 정렬에 의해 결정되었다. 사이토카인 패널에는 *CTACK*, FGF 염기성, 에오타신, G-CSF, GM-CSF, GRO-α, *HGF*, IFN-α2, IFN-γ, IL-1β, IL-1ra, IL-1α, IL-2, IL-2Rα, IL-3, IL-4, IL-5, IL-6, IL-7, *IL-8*, IL-9, IL-10, IL-12(p40), IL-12(p70), IL-13, IL-15, IL-16, IL-17A, *IL-18*, IP-10, LIF, MCP-3, MCP-1, M-CSF, MIF, MIG, MIP-1α, MIP-1β, β-NGF, *PDGF-BB*, SCF, *SCGF-β*, SDF-1α, RANTES, TNF, TRAIL, 및 VEGF가 포함되었다. 9개의 사이토카인은 샘플의 >20%에서 검출의 한계 밖에 있고 제거되었다.

[0245] 데이터 사전-처리

[0246] *RNAseq*: 이진 염기 콜(binary base call, BCL) 서열 파일은 bcl2fastq 파이프라인(Illumina)을 사용하여 fastq 파일로 변환되었다. 서열 데이터는 메디컬 서열 분석 파이프라인(MEDical Sequence Analysis Pipeline, MESA P)으로 처리되었고, HISAT2(Pertea 등, *Nat Protoc*, 2016, 11:1650-1667)로 hg38 게놈에 정렬되었으며, 수(count)는 GenomicAlignments R 패키지로부터의 summariseOverlaps 함수로 정량화되었다. 사전- 및 사후-정렬 QC는 각각 FastQC 및 SAMStat로 평가되었다.

[0247] *사이토카인*: 사이토카인은 샘플(비자극 샘플 제외)의 30% 이상이 범위 외(OOR) 값으로 기록된 경우 제외되었다. 이는 추가 분석으로부터 *CTACK*, IL-3, IL-7, IL-8, IL-13, IL-18, *PDGF-BB*, *SCGF-β*, SDF-1α의 제거를 초래하였다. 나머지 OOR 값은 부분 정보가 대치(imputation)에 대한 검출의 한계 이하/이상의 값을 정의하는 데 사용되도록, truncnorm R 패키지로부터의 rtruncnorm 함수로 절단된 정규 분포를 기반으로 이하/이상 최저/최대

대치되었다. 가장 적절한 변환 및 정규화의 방법을 테스트하였으며(데이터는 나타내지 않음), ArcSinh 변환 및 Loess 정규화를 적용하였다. 배치 효과는 선형 모델링으로 제거되었다(limma R 패키지로부터의 BatchEffect 함수 제거).

[0248] 차원성 감소: 상기 실험 설계는 동일한 배치 내의 각 개체에 대해 4가지 조건(비자극, LPS-, 이미퀴모드- 및 폴리(I:C)-자극)으로부터 일치된 데이터가 생성되었음을 보장하였다. 이는 차원성 감소 분석(주성분 분석)을 위한 다단계 설계를 허용하며, 이에 따라 대상체 내 분산이 대상체 간 분산으로 분해되며, 이는 후속 다변량 분석의 성능 및 해석 가능성을 상당히 개선시킨다. 이 목적을 위해, withinsideVariation 함수가 R의 mixOmics 패키지로부터 채택되었다(Rohart 등, *PLoS Comput Biol*, 2017, 13: e1005752).

[0249] 전사체: 사전-처리 후, 50 개체에 대한 일치된 비자극, LPS-, 이미퀴모드- 및 폴리(I:C)-자극 샘플을 위한 CBMC 유전자 발현 데이터는 유의미하게 가변적인 유전자에 대해서만 노이즈를 감소시키기 위해 R의 varianceBasedfilter 함수로 필터링하였다. 이 분석을 위해, 유전자는 0.05/유전자의 수(n=17356)로 결정된  $2.88 \times 10^{-6}$ 의 엄격한 역치보다 낮은 p 값으로 유의미한 것으로 간주되었다. 이로 인해 차원성 감소를 위한 5,885 개의 유전자가 생성되었다. 대상체 내 변화는 withinVariation 함수로 계산되었고, 유전자는 단위 분산으로 조정되었으며, FactoMineR 패키지로부터의 PCA 함수는 주성분 분석에 사용되었고(Le 등, *J of Statistical Software*, 2008, 25: 1-18), 주성분 점수 및 변수 기여도는 플롯에 사용되었다.

[0250] 사이토카인: 사전-처리 후, 대상체 내 변화는 50 개체에 대한 일치된 비자극, LPS-, 이미퀴모드- 및 폴리(I:C)-자극 샘플을 위한 CBMC 사이토카인 농도 데이터(n=39)로부터 (상기와 같이) 계산되었다. 주성분 분석은 상기 기술된 바와 같이 적용되었다.

[0251] 전사체 분석

[0252] EdgeR: 총 50,019개의 원시 전사체는 사전-처리 후 분석을 위해 이용가능하였다. 원시 전사체는 임의의 샘플에서 수(count)가 없거나, 주석이 부족하거나, ≤25개 샘플에서 백만 개당 <0.5 수를 갖는 경우 제거되었다. 이 전략은 분석을 위해 17,363개의 전사체를 생성하였다. 데이터는 M-값의 절사 평균(trimmed mean)으로 정규화되었다(TMM 정규화(Robinson 등, *Genome Biol*, 2010, 11: R25)). 실험 설계는 (연령 및 자극에 대한) 블록 랜덤화를 배치(batch)로 수반하므로, 이러한 대응표본(paired) 비교에 대한 세포 배양 배치 번호와 관련된 배치 효과는 없었다. 독립표본(unpaired) 비교를 위한 식별 가능한 배치 효과가 관찰되지 않았지만, 배양 배치는 관계 없이 독립표본 분석(즉, 유아기에 sLRI 감수성)을 위한 공변량으로서 포함되었다. 일련의 실증적 대조군 유전자(임의의 관심대상 비교 간에 유의미하게 상이하지 않음)를 모델링하는, RUVSeq R 패키지로부터의 RUVg 함수로 원치않는 변화를 식별하고 제거하여(Risso 등, *Nat Biotechnol*, 2014, 32: 896-902), 추정되는 기술적 효과(레인, 시퀀싱 등)를 결정한다. 설계된 대응표본은 RUVg에 의해 공변량으로서 식별된 추세를 포함하는, 일치된 자극 샘플 및 비자극 샘플 간의 분석을 위해 사용되었다. EdgeR(Robinson 등, *Bioinformatics*, 2010, 26: 139-140) 파이프라인은 estimateDisp, glmQLFit, glmLRT 및 topTags를 포함하는 기본 파라미터로 실행되었으며, 각 유전자를 위한 수에 대한 음이항 일반화된 로그-선형 모델을 맞추고, genewise 우도비(likelihood ratio) 테스트를 수행한다. 연령 5세때 CBMC 및 일치된 샘플(n=27) 간의 분석을 위해, 일치된 비자극 샘플 및 상응하는 시뮬레이션된 샘플 간의 차이뿐만 아니라 RUVg 추세를 모델링한 대응표본 설계가 사용되었다. 비자극/자극 및 RUVg 추세를 모델링한 대응표본 설계를 사용하여 주요 결과 간의 차이를 결정하였다. 유전자는 FDR-조정된 Benjamini-Hochberg p 값이 < 0.01이고, Log2 배수 변화가 1 이상(상향조절됨) 또는 -1 이하(하향조절됨)를 기록한 경우 유의미하게 상이한 것으로 간주되었다. 보정된 유전자 수의 매트릭스는 다운스트림 네트워크 분석을 위해 생성되었다. 크기 인자는 estimateSizeFactors 함수가 포함된 중앙값 비율 방법을 사용하여 추정되었으며, 분산 안정화 변환(variance stabilizing transformation, VST)은 DESeq2 패키지로부터의 varianceStabilizingTransformation 함수로 수 데이터에 적용되었다(Love 등, *Genome Biol*, 2014, 15:550). 기술적 변화 및 배양 배치와 관련된 RUVg 추세는 limma 패키지로부터의 RemoveBatchEffect 함수가 포함된 선형 모델을 사용하여 공변량으로서 제거되었다.

[0253] Limma-voom: 전사체 필터, 정규화 및 모델 설계는 EdgeR 분석에 대해 기술된 바와 동일하게 수행되었다. 데이터는 백만 개당 Log2-수로 변환되었으며, 평균-분산 관계는 voom 함수를 사용하여 가중치를 생성하도록 추정되었다. limma(Ritchie 등, *Nucleic Acids Res*, 2015, 43: e47) 파이프라인은 lmFit, Contrasts.fit, eBayes 및 topTable 함수를 포함하는 기본 파라미터로 실행되었다. EdgeR 분석과 동일한 기준을 적용하여 유전자 유의성을 결정하였다. 이 분석으로부터, 각 유전자에 대해 계산된 조절된 t-통계량은 시뮬레이션된 샘플 및 일치된 비자극 샘플 간에 어떤 모듈이 차등적으로 조절되는지 표시하는 방법으로 각 네트워크에 대해 모듈별로 플롯팅되었

다. 조절된 t-통계량은 모든 유전자에 걸쳐 "조절된"(실증적 베이즈(Bayes)), 표준 오류에 대한 M-값(Log<sub>2</sub>-배수 변화)의 비율이다. 조절된 t-통계량 대신 모듈 고유 유전자를 적용하는 것은 모듈 상향-/하향-조절과 관련하여 동일한 전체 결과가 산출되었다(데이터는 나타내지 않음). 조절된 t-통계량 2 이상의 중앙값을 갖는 모듈은 유의미하게 상향조절된 것으로 간주되며, -2 이하인 모듈은 유의미하게 하향조절된 것으로 간주된다.

[0254] 가장 유전자 공-발현 네트워크 분석: (상기 기술된) 보정된 수 데이터는 분석을 위해 17,363개의 유전자를 포함하는 입력으로서 사용되었다. 이 분석을 위해, 3개의 섭동 네트워크가 생성되었고, 이들 각각은 비자극 샘플 및 상응하는 자극된 샘플(즉, LPS, 이미퀴모드 및 폴리(I:C) 네트워크)을 포함하였다(WGCNA(Zhang 등, *Stat Appl Genet Mol Biol*, 2005, 4: Article 17; Langfelder 등, *BMC Bioinformatics*, 2008, 9: 559)). varianceBasedfilter 함수는 각 조건에 대해 유의미하게 가변적인 유전자(p 값 < 0.01), 및 비자극 샘플 및 각각의 자극된 샘플 간의 조합 유전자를 필터링하는 데 사용되었다. 이 전략은 LPS, 이미퀴모드 및 폴리(I:C) 네트워크에 이용가능한 6561개, 6757개 및 6764개의 유전자를 각각 생성하였다. 소프트 파워(soft power)는 "부호화(signed)"로 설정된 networkType 파라미터가 포함된 pickSoftThreshold 함수로 계산하였다. 이로 인해 LPS, 이미퀴모드 및 폴리(I:C) 네트워크에 대해 각각 7, 8 및 7의 소프트 파워가 초래되었다. 인접 및 토폴로지 중첩 행렬(TOM)은 각각 지정된 "부호화(signed)" 네트워크를 갖는 인접 및 TOMsimilarity 함수로 생성되었다. TOM 비 유사성 행렬은 1에서 TOM 유사성 행렬(1-TOM)을 뺀 것으로 계산하였다. 모듈은 hclust 함수(방법 = "평균")로 계층적 클러스터링으로 식별되고, cutreeDynamic 함수(방법 = "하이브리드", deepSplit = 2, minClusterSize = 50)로 전지(pruned) 되었다. TOM 플롯은 TOMplot 함수로 생성되었다. 모듈 고유 유전자는 moduleEigengenes 함수로 계산되었다. 모듈은 유사한 경우 병합되었으며, 고유 유전자, 계층적 클러스터링 및 mergeCloseModules 함수로 0.1에서 절단된 계통도의 상관관계에 의해 결정되었다. 네트워크 통계량 및 모듈 내 연결성은 intramodularConnectivity 함수로 계산되었다. 모듈은 유전자 온톨로지 용어 농축(GOenrichmentAnalysis), ReactomePA(Yu 등, *Mol Biosyst*, 2016, 12: 477-479) 및 clusterProfiler(Yu 등, *Omic*s, 2012, 16:284-287) R 패키지, InnateDB(Breuer 등, *Nucleic Acids Res*, 2013, 41: D1228-1233), 및 상위 모듈 유전자의 식별(Log<sub>2</sub>-FC/유전자 연결성)로부터 유의미하게 농축된 경로를 평가함으로써 합의 접근법으로 주석이 달렸다. 네트워크 간의 모듈 보존은 200개의 순열, "부호화(signed)"로 설정된 networkType, 및 각 비교를 위한 평균 모듈 크기로 설정된 "골드" (랜덤) 모듈 크기가 포함된 modulePreservation 함수로 계산되었다. 랭크된 발현은 모든 샘플에 걸쳐 각 유전자의 (랭크) 평균 발현으로 계산되었고, 랭크된 연결성은 "부호화(signed)"로 설정된 유형 및 상응하는 네트워크의 소프트 파워로 설정된 파워가 포함된 (랭크) SoftConnectivity 함수로 계산되었다. 소프트 연결성은 다른 모든 유전자에 대한 네트워크에서 각 유전자의 인접(공-발현 측정)의 합으로 정의된다. 연결성 밀도는 밀도 함수로 결정되었으며, Sheather-Jones의 순화 대역폭(smoothing bandwidth) 방법이 사용되었다. 연결성 밀도는 릴리포스(Liliefors) 정규성 테스트(lillie.test 함수)로 정규 분포에 대해 평가되었다. 스피어만의 상관관계 행렬은 다른 모든 유전자에 대한 각 유전자의 상관관계 값의 합으로 정의되는, 모듈 내 연결성을 개별적으로 평가하기 위해 각 모듈에 대해서도 계산되었다. 상위 20개의 가장 많이 연결된 유전자의 네트워크 배선 다이어그램은 igraph R 패키지로부터의 graph\_from\_adjacency\_matrix 함수로 구성되었다. 노드(node) 크기는 총 네트워크 중에 연결 수(정도)를 나타내며, 에지(edge)는 연결 강도를 나타낸다(적색 에지는 상관관계 > 0.8을 나타냄).

[0255] 핵심 조절자 분석

[0256] 유전자 조절자 네트워크는 ARACNe로 역설계되었으며(Margolin 등, *BMC Bioinformatics*, 2006, 7 Suppl 1, S7), 전사 인자 활성화는 VIPER로 추론되었다(Alvarez 등, *Nat Genet*, 2016, 48:838-847). 유의미한(p<0.05) TF는 RcisTarget에 의해 결정된 레귤론 표적 유전자의 영역에서 알려진 결합 모티프가 있는 경우 반응의 드라이버로 간주되었다(Aibar 등, *Nat Methods*, 2017, 14:1083-1086). VIPER으로부터 출력된 정규화된 발현 점수(NES)는 다운스트림 분석을 위해 유지되었다.

[0257] 기계 학습

[0258] 유전자 발현 데이터는 훈련(50%) 및 유효성 검증(50%) 세트에 랜덤으로 할당되었으며, 각 분석에 대해 각각의 모듈 유전자에 대해서만 필터링되었다. 모든 모델에 대해 동일한 랜덤 할당이 적용되었다. 유효성 검증 모델의 경우, CAS 코호트 데이터가 각각의 모듈 유전자에 대해 필터링되고 훈련 세트로 사용되었으며, 외부 유전자 발현 데이터가 유효성 검증에 사용되었다(동일한 입력 유전자에 대해 필터링됨).

[0259] RandomForest 패키지는 모델 구축을 위해 R에서 사용되었고, 아웃-오브-백(Out-of-bag) 오류율에 따라 결정 트리 수(n<sub>tree</sub>) 및 후보 변수(m<sub>try</sub>)가 최적화되었다. 랜덤 포레스트(RF) 분석은 모듈 구성원 유전자가 RF 해석에

영향을 미치는 것으로 인식된 높은 다중공선성을 나타내도록, 공-발현에 따라 유전자를 설계적으로 클러스터링하는 WGCNA에 의해 정의된 모듈을 사용하여 수행되었다(Strobl 등, *BMC Bioinformatics*, 2008, 9: 307; Tolosi 등, *Bioinformatics*, 2011, 27: 1986-1994). 그러나, 공선성은 주로 가변 중요도의 해석에 영향을 미치며, 전체 모델 예측 정확도에는 영향을 미치지 않는다. 이러한 이유로, 본 연구에 사용된 RF 분류기는 결과를 예측하기 위한 IFN 모듈 유전자의 유용성을 테스트하기 위해 주로 사용되었으며, (도면에 기록되어 있지만) 가변 중요도 측정은 이들의 실제 값에 대한 과소평가로 간주되어야 한다.

[0260] CAS 코호트: 개체 간의 기준선/비자극 CBMC 유전자 발현의 잠재적인 차이를 설명하기 위해, 일치된 자극 유전자 발현 프로파일로부터  $\Delta$  값을 취하고(예를 들어, LPS-자극 유전자 발현 - 일치된 비자극 유전자 발현 = 조정된 LPS-자극 유전자 발현 매트릭스), 이를 입력으로 사용하였다. 유전자는 상응하는 반응의 IFN 모듈에 존재하는 것들에 대해서만 필터링되었다. 대상체(n=50)는 테스트 또는 유효성 검증 세트(50/50 분할)에 랜덤으로 할당되었으며, 동일한 랜덤 할당은 LPS, 이미퀴모드 및 폴리(I:C) 데이터세트에 적용되었다. RandomForest 함수(randomForest R 패키지)는 유아기에 sLRI를 경험한 개체 및 경험하지 않은 개체를 분류하기 위해, 각 분할에서 후보로서 랜덤으로 샘플링된 변수의 수("mtry") 및 성장할 결정 트리의 수("ntree")와 관련하여 각 RF 모델을 최적화하는 것이었다. mtry 파라미터의 경우, 10보다 작은 것부터 또는 입력 유전자 수의 제곱근부터 입력 유전자 수의 제곱근의 최대 5배까지 1씩 증가하는 시퀀스가 정의되었다. RF 분류기는 1000으로 설정된 ntree를 포함한, 시퀀스의 모든 수에 대해 구축되었으며, 가장 낮은 아웃-오브-백 오류율(OOBer)이 기록된 mtry 값을 최적으로 선택하였다. OOBer에서 동점일 경우 가장 낮은 수가 선택되었다. ntree 파라미터의 경우, 500에서 10,000까지 100씩 증가하는 시퀀스가 정의되었으며, 상기 정의된 바와 같이 최적으로 설정된 mtry가 포함된, 시퀀스의 모든 수에 대해 RF 분류기가 구축되었다. 가장 작은 OOBer(동점일 경우 가장 낮은 수)를 생성하는 최적의 ntree가 선택되었다. 이 접근법은 LPS, 이미퀴모드 및 폴리(I:C) RF 모델에 대해 각각 11, 23 및 121의 mtry 및 5400, 1000 및 1000의 ntree 값을 초래하였다. 훈련 세트(70/30 분할)를 내부적으로 부트스트랩하는 훈련된 모델은 LPS, 이미퀴모드 및 폴리(I:C) RF 모델에 대해 각각 36%, 72% 및 52%의 OOBer를 기록하였다. 최적화 후, 최종 RF 분류기를 사용하여 예측 함수(stats R 패키지)로 상응하는 유효성 검증 세트(분류 표지 없음)에서 주요 결과 상태를 예측하였다. ROCR R 패키지로부터의 예측 및 성능 함수는 예측을 실제 값과 비교하고, 참 양성 비율 및 거짓 양성 비율을 결정하고, 수신자 조작 특성(ROC) 곡선화 면적을 계산하는 데 사용되었다.

[0261] 훈련/유효성 검증 세트 재-샘플링: RF 분류기의 재현성을 테스트하기 위해, 조정된 LPS, 이미퀴모드 및 폴리(I:C) 데이터세트는 훈련/유효성 검증 세트 할당(50/50 분할)과 관련하여, 각각 랜덤으로 2000개 재-샘플링되었다. 동일한 2000개의 랜덤 할당이 LPS, 이미퀴모드 및 폴리(I:C) 분류기에 대해 사용되었다. RF 분류기는 훈련 세트에 대해 구축되었으며, (상기, 이전에 결정된 최적의 파라미터를 사용하여) 각 재-샘플에 대해 (상기와 같이) 유효성 검증 세트에 대해 테스트되었으며, ROC 곡선화 면적은 예측 정확도를 결정하기 위해 매번 기록되었다. 훈련 및 유효성 검증 세트의 상이한 비례 할당이 유사한 결과를 생성했는지 여부를 평가하기 위해, 상기 조정된 LPS, 이미퀴모드 및 폴리(I:C) 데이터세트의 RF 분류기는 60%/40% 및 70%/30%의 훈련/유효성 검증 할당에서 (훈련/유효성 검증 세트 할당과 관련하여) 각각 랜덤으로 1000회 재-샘플링되었다. 동일한 1000개의 랜덤 할당이 LPS, 이미퀴모드 및 폴리(I:C) 분류기에 대해 사용되었으며, 이들은 상기 (최적화된) 파라미터를 사용하여 구축되었다. ROC 곡선화 면적은 예측 정확도를 결정하기 위해 각각의 재-샘플에 대해 기록되었다.

[0262] 외부 코호트에서 유효성 검증: 본 발명자들은 본 발명자들의 CBMC 데이터에 대해 RF 분류기를 훈련시키고, 이를 사용하여 유전자 발현 옴니버스(Gene Expression Omnibus)에서 공개적으로 이용가능한 일련의 데이터 세트로부터 유래된 샘플을 분류하였다. 일반적으로, 본 발명자들은 CAS 코호트로부터의 비자극 샘플을 사용하여 "건강한" 개체(감염 없음)를 나타내고, 자극된 샘플을 사용하여 감염이 확인된 유아/아동의 항-박테리아(LPS) 및 항-바이러스(이미퀴모드/폴리(I:C)) 선천적 면역 반응을 나타내었다. GSE72809: RF 분류기는 조정된 LPS- 및 이미퀴모드-/폴리(I:C) 자극된 CBMC 유전자 발현 데이터세트(각각 n=50)에 대해 훈련되었으며, 혈액-유래 유전자 발현 프로파일로부터 건강한 대조군(n=52)에서 각각 박테리아(n=52) 및 바이러스(n=92) 감염으로 입원한 아동을 예측하는 데 사용되었다. 모델 최적화가 구현되었으며, 모델 예측 정확도는 (상기) 이전에 기술된 바와 같이 외부 유전자 발현 프로파일에 대해 테스트되었다. GSE113211: RF 분류기는 조정된 이미퀴모드- 및 폴리(I:C)-자극된 CBMC 유전자 발현 데이터세트(각각 n=50)에 대해 훈련되었으며, PBMC 샘플로부터 회복 후 (증상 없음, 감염 후 8.8 ± 2.5주) 수집된 일치된 샘플에서 급성 바이러스성 세기관지염으로 입원한 유아(≤ 18개월, n=15) 및 영아(18개월-5세, n=16)를 예측하는 데 사용되었다. 모든 대상체에 대해 함께 뿐만 아니라 유아 및 아동에 대해 별도로, 모델 최적화가 구현되었고, 모델 예측 정확도는 (상기) 이전에 기술된 바와 같이 외부 유전자 발현 프로파일에 대해 테스트되었다. GSE115770: RF 분류기는 조정된 이미퀴모드- 및 폴리(I:C)-자극 CBMC 유전자 발현 데이터세트(각각 n=50)에 대해 훈련되었으며, 바이러스성-관련(n= 193) "감기"-유사 질병을

갖거나, 비-바이러스성 "감기"-유사 질병 (n=105) (샘플은 발병 후 1-6일에 채취됨)을 가진 천식 아동(6-17세)의 연구 방문을 예측하는 데 사용되었으며, 이들 중 일부는 나중에 악화를 경험하였다(58명은 경험하고 25명은 경험하지 않음). 감기 증상이 있는 바이러스 감염 및 악화되는 바이러스 감염에 대해, 비강- 및 혈액-유래 유전자 발현 프로파일에 대해 별도로, 모델 최적화가 구현되었고, 모델 예측 정확도는 (상기) 이전에 기술된 바와 같이 테스트되었다.

[0263] 다중-오믹 데이터 통합

[0264] DIABLO(Singh 등, *Bioinformatics*, 2019, 35:3055-3062) 모델은 지도(supervised) 다중-오믹 데이터 통합을 위해 구축되었으며, 이는 일치된 데이터세트 간의 공-발현을 최대화하기 위해 부분 최소 제곱(Partial Least Squares) 분석을 일반화한다. 모든 데이터세트(변역표현형 제외)는 분석 전에 기준선 조정되었으며, 유전자 발현 데이터는 노이즈를 감소시키기 위해 유의미하게 가변적인 유전자(n=6353)로 필터링되었다. 구성요소의 수 및 특징 선택 파라미터는 5× 교차 유효성 검증으로 조정되었다.

[0265] 통계적 분석

[0266] 모든 통계적 분석은 R 환경(버전 3.6.2)에서 계산되었으며, 그래프는 R 또는 Prism 소프트웨어(버전 8, GraphPad Software, La Jolla California USA)로부터 생성되었다. 비-모수적 통계적 방법은 그룹 차이(맨-휘트니 U 테스트(독립표본 분석) 및 윌콕슨 부호 랭크 테스트(대응표본 분석); wilcox.test 함수 [stats R 패키지]) 및 상관관계(스피어만의 랭크 상관 계수; cor.test 함수 [stats R 패키지])를 테스트하기 위해 적용되었다. 연구 모집단 특성의 비교를 위해, 피셔의 정확도 테스트(Fisher's Exact test)(fisher.test 함수 [stats R 패키지])를 사용하여 오즈비, 95% 신뢰 구간 및 범주형 변수에 대해 수반되는 P 값을 계산하였고, 맨-휘트니 U 테스트를 사용하여 연속형 변수에 대한 p 값을 결정하였다.

[0267] 실시예 2 - 연구 모집단의 임상 특성

[0268] 연구 모집단은 아동기 천식 연구(CAS) 코호트 내의 50명의 아동의 하위 집합으로 구성되었다. 23명의 대상체(46%)는 첫 해(유아기)에 적어도 한 번의 천명음 및/또는 열성 sLRI를 경험하였으며, 이는 주요 관심 결과였다(표 1). 연령 5세 때 현재 진행되는 천명음(OR=2.48) 및 천식(OR=2.86) 유병률은 감수성이 있는 개체에서 더 더 높았지만 이는 통계적으로 유의미하지 않았다. 주요 결과에 대한 성별, 임신 주수, 출생 체중, 피부 단자 테스트 양성(skin prick test positivity), 및 유아기에 URI와 관련하여 차이가 관찰되지 않았다. 전체적으로, 이 하위 집합은 주요 임상 특성과 관련하여 CAS 코호트(n=263)를 대표하는 것으로 확인되었다(표 1). 라이노바이러스는 이 하위 집합에서 생후 첫 해로부터 식별된 가장 빈번한 바이러스 병원체였으며 (감염성 비인두 샘플의 56.9%에 존재함), RSV(13.125%)가 그 뒤를 이었다(데이터는 나타내지 않음).

[0269] 표 1. 연구 모집단의 특성 및 표현

	CAS 서브세트 (n=50)	sLRIy1 음성 (n=27)	sLRIy1 양성 (n=23)	OR (95% CI)	P 값	CAS 코호트 (n=263)	OR (95% CI)	P 값
성별 (여성)	24/50 (48%)	14/27 (51.85%)	10/23 (43.48%)	0.72 (0.2-2.51)	0.58	115/251 (45.82%)	0.92 (0.48-1.77)	0.88
임신 기간(주(weeks)): 평균 [범위]	39.14 [36-41]	39.11 [36-41]	39.17 [37-41]	NA	0.96	39.03 [34-41]	NA	0.89
출생 체중(그램: 평균 [범위])	3496.52 [2755-4415]	3468.56 [2755-4270]	3529.35 [2920-4415]	NA	0.65	3406.17 [2085-5110]	NA	0.27
0.5, 2, 또는 5세에 SPT+	24/50 (48%)	16/27 (59.26%)	14/23 (60.87%)	1.07 (0.3-3.88)	1	118/198 (59.6%)	1.59 (0.82-3.13)	0.15
1세에 URI	47/50 (94%)	26/27 (96.3%)	21/23 (86.96%)	0.41 (0.01-8.41)	0.59	215/235 (91.49%)	0.69 (0.13-2.46)	0.78
1세에 LRI	39/50 (78%)	16/27 (59.26%)	23/23 (100%)	Inf (2.97-Inf)	0.0004	160/235 (68.08%)	0.6 (0.26-1.28)	0.18
1세에 sLRIy1	23/50 (46%)	0/27 (0%)	23/23 (100%)	Inf (77.85-Inf)	>0.0001	101/235 (42.98%)	0.88 (0.46-1.72)	0.75
5세에 현재 천명음	14/43 (32.56%)	6/25 (24%)	8/18 (44.44%)	2.48 (0.57-11.49)	0.2	56/198 (28.28%)	0.82 (0.38-1.8)	0.58
5세에 천식	9/43 (20.93%)	3/19 (15.79%)	5/14(35.71%)	2.86 (0.44-22.97)	0.24	37/198 (18.69%)	0.87 (0.37-2.24)	0.83
천명음을 내지 않음 (3-5세)	15/44 (34.09%)	17/24 (70.83)	12/18 (66.67%)	0.83 (0.18-3.82)	1	70/198 (35.35%)	1.06 (0.51-2.72)	1

[0270] 약어: CAS=아동기 천식 연구, OR=오즈비, CI=신뢰구간, URI=상부 호흡기 감염(바이러스), LRI=하부 호흡기 감염(바이러스). sLRIy1은 1차 결과(생후 첫 해의 sLRI 발병률)를 나타낸다. 범주형 변수의 경우, 오즈비, 95% CI 및 수반되는 P 값은 피셔의 정확도 테스트에 의해 결정되었다. 연속형 변수의 경우, P 값은 맨-휘트니 U 테스트에 의해 결정되었다. 볼드체로 나타낸 P 값은 통계적으로 유의미한 것으로 간주된다.

[0272] 실시예 3 - 기준선 유세포 분석

[0273] 11-색 유세포 분석 패널을 기준선 제대혈 단핵 세포(CBMC) 샘플에 적용하여 세포 구성을 평가하였다. 림프구(T 및 B 세포)는 CBMC 중에서 식별된 세포 유형의 대부분을 구성하였다(도 1b). CD14+ 단핵구 및 통상형 수지상세포(cDC)가 골수성 구획 중에서 식별되었고, 더 작은 비율의 형질세포양 DC(pDC) 및 호염기구가 또한 식별되었다(도 1b). 생후 첫 해에 sLRI와 관련하여 기준선 세포 구성에는 차이가 없었다(데이터는 나타내지 않음).

[0274] **실시예 4 - CBMC에서 선천적 면역 반응의 다중-오믹 프로파일링**

[0275] 모든 50명의 대상체로부터의 CBMC를 LPS(TLR4), 또는 이미퀴모드(TLR7) 또는 폴리(I:C)(TLR3)와 18시간 동안 배양하여 비자극 대조군과 함께 선천적 면역 반응을 유도하였다(도 1a). 이 시점은 즉각적 및 2차 반응 프로그램의 다운스트림에서 신호전달 연쇄반응(signalling cascade)을 포착하기 위해 선택되었다(Shalek 등, *Nature*, 2014, 510:363-369; Jovanovic 등, *Science*, 2015, 347:1259038; Lawlor 등, *Front Immunol*, 2021, 12:636720). 유전자 발현은 세포 펠릿(RNAseq)으로부터 프로파일링되었으며, 상층액은 사이토카인을 프로파일링하는 데 사용되었다(멀티플렉스 검정). 연령 5세 때 수집된 일치하는 PBMC 샘플은 대상체의 하위 집합(n=27)에 이용가능하며, 이들은 동일한 조건 하에서 동시에 배양하였다. 데이터 전-처리 및 필터링 후, 17,363개의 전사체 및 39개의 사이토카인은 분석을 위해 이용가능하였다(방법의 경우 실시예 1 참조). 비지도 주성분 분석(PCA) 차원 감소는 탐색적 데이터 분석을 위해 적용하였다. 각각의 오믹 층(omic layer)으로부터의 샘플은 예상대로 자극에 의해 클러스터링되었다(도 1c). 전사체 및 사이토카인의 경우, 처음 두 가지 주성분은 인터페론(IFN) 및 전염증성 특징(예를 들어, CXCL10/IP-10, IL-1 $\beta$ , IL-6)이 포착되었다(도 1d).

[0276] **실시예 5 - IFN 및 전염증성 유전자 발현 프로그램은 CBMC 반응에서 상향조절된다**

[0277] 해당 데이터가 게놈-전체 적용범위(genome-wide coverage)를 제공함에 따라 제1 분석은 전사체학 데이터에 중점을 두었다. 차등적 발현 분석을 사용하여, 641개의 상이하게 발현된 유전자(DEG)가 제대혈 LPS 반응에 대해 (Log<sub>2</sub>-배수 변화 > 1, 조정된-P 값 < 0.01), 그리고 1000개 초과 DEGs가 이미퀴모드 및 폴리(I:C) 반응에 대해 식별되었다(도 2a-c; 좌측 패널). 경로 분석(InnateDB; Breuer 등, *Nucleic Acids Res*, 2013, 41:D1228-1233)은 모든 반응에서 상향조절된 유전자로부터의 사이토카인 및 케모카인 신호전달 경로의 농축을 식별하였으며, IFN 신호전달 경로는 이미퀴모드 및 폴리(I:C) CBMC 반응에 대해 두드러졌다(도 2a-c; 우측 패널). 특히, 바이러스 자극은 429개의 상향조절된 유전자의 공통적인 세트를 유도하였으며, 이는 TLR3 및 TLR7 활성화 간에 공유된 핵심 항바이러스 반응을 구성하였다(데이터는 나타내지 않음). 또한, 462개 및 243개의 유전자는 각각 폴리(I:C) 또는 이미퀴모드에 반응하여 특이적으로 상향조절된 것으로 식별되었으며, 이는 TLR3 또는 TLR7 활성화의 다운스트림에 있는 특유의 신호전달 경로를 입증한다(데이터는 나타내지 않음). 다음 CIBERSORTx를 사용하여 RNA-Seq 데이터로부터 배양 후(post-culture)의 세포 구성을 추정하였다(Newman 등, *Nat Biotechnol*, 2019, 37:773-782). 두드러진 세포 유형에는 단핵구, B 세포 및 CD4+ T 세포가 포함되었다(데이터는 나타내지 않음). 적혈구 비율은 면역자기적 고갈(immunomagnetic depletion)의 결과로서 무시해도 될 정도였다(방법의 경우 실시예 1 참조). 세포 구성 변화는 자극 및 연령 간에 식별되었지만, 유아기에 서열 순서, 성별 또는 sLRI 접촉은 식별되지 않았다. 선천적 면역 유전자 발현의 변화에 대한 추가 조사는 출생 시 대 연령 5세 때 (연령/자극 당 n=27) 수집된 일치된 샘플에서도 수행되었다(데이터는 나타내지 않음). 흥미롭게도, 5세 때 LPS 반응은 출생과 비교하여 IRF1, STAT1 및 IFIT1-3을 포함한 IFN-관련 유전자의 상향조절에 의해 특징지어졌다(표 2). 대조적으로, IFN-관련 유전자의 차등적 발현은 이미퀴모드 또는 폴리(I:C) 자극 후 출생과 연령 5세 사이에 관찰되지 않았다(데이터는 나타내지 않음). 마지막으로, 본 분석으로부터의 임의의 조건에 대해 유아기에서 sLRI에 대한 저항성 및 감수성이 있는 개체 간에 유전자가 유의하게 상이하지 않았으며(데이터는 나타내지 않음), 이는 sLRI 위험이 개별 유전자 발현 크기만으로는 부여되지 않음을 시사한다.

[0278] 마지막으로, 이 분석으로부터의 임의의 조건에 대해 유아기의 sLRI에 대한 내성 및 감수성이 있는 개체 간에는 유전자가 유의하게 상이하지 않았으며(데이터는 나타내지 않음), 이는 sLRI 위험이 개체 유전자 발현 크기만으로는 부여되지 않음을 시사한다.

[0279] 표 2. 출생 시 및 LPS 자극 후 5세 때 IFN-관련 유전자의 차등적 발현.

심볼	logFC	FDR
IFNG	-1.841	3.45E-28
IFIT3	-1.779	3.29E-166
TLR3	-1.644	1.93E-16
IFIT1	-1.566	1.78E-75
OASL	-1.482	7.88E-58
IFIT2	-1.232	4.03E-82
MX1	-1.097	3.56E-107
OAS1	-1.097	8.10E-79
STAT1	-1.045	1.48E-89
IRF1	-1.044	3.05E-97

[0280]

[0281] 실시예 6 - 선천적 면역 기능을 기본으로 하는 공-발현 네트워크의 식별

[0282]

유전자는 별개로 기능하지 않으며, 이들은 네트워크에서 함께 작동한다. 가중 유전자 공-발현 네트워크 분석(WGCNA)은 유전자 발현의 전반적인 연결성 구조 및 기능적 구성을 밝히기 위해 사용되었다. 이 분석은 LPS, 이미퀴모드 및 폴리(I:C) 반응에 대해 각각 11, 11 및 8개의 공-발현 모듈을 식별하였다(도 2d-f). 모든 반응은 IFN 및 전염증성 모듈의 상향조절을 나타내었으며, 체대혈 선천적 반응의 필수적인 구성 요소가 이미 식별되었고, 따라서 이들은 다운스트림 분석을 위해 전진시켰다(도 2d-f, 도 6a). LPS 반응은 이미퀴모드(1114개 유전자) 및 폴리(I:C)(2201개 유전자)와 비교하여 가장 작은 IFN 모듈(180개 유전자; 표 3)을 가지며, 전염증성 모듈(LPS, 2297개 유전자; 이미퀴모드, 924개 유전자; 폴리(I:C), 646개 유전자)에 대해서는 그 반대였다(도 2g). 특히, 상이한 자극에 대한 IFN 및 전염증성 모듈 유전자 간에, 특히 폴리(I:C) IFN 및 LPS 전염증성 모듈(n=385 유전자) 간에 실질적인 중첩이 있었다(데이터는 나타내지 않음).

[0283] 표 3. LPS-유도 인터페론 모듈 유전자. 유전자 심볼 및 ENSEMBL 식별번호가 제공된다.

<i>IDO1</i> ENSG00000131203	<i>SOCS1</i> ENSG00000185338	<i>GBP4</i> ENSG00000162654	<i>CD80</i> ENSG00000121594	<i>GBP1</i> ENSG00000117228	<i>CXCL11</i> ENSG00000169248
<i>CXCL9</i> ENSG00000138755	<i>MT2A</i> ENSG00000125148	<i>TCAF2</i> ENSG00000170379	<i>STAT1</i> ENSG00000115415	<i>RTP4</i> ENSG00000136514	<i>GBP5</i> ENSG00000154451
<i>FBXO6</i> ENSG00000116663	<i>HAPLN3</i> ENSG00000140511	<i>CCL19</i> ENSG00000172724	<i>APOL1</i> ENSG00000100342	<i>PMAIP1</i> ENSG00000141682	<i>USP30-AS1</i> ENSG00000256262
<i>IL15RA</i> ENSG00000134470	<i>GBP2</i> ENSG00000162645	<i>PARP9</i> ENSG00000138496	<i>IRF1</i> ENSG00000125347	<i>TCAF2C</i> ENSG00000283528	<i>TNFSF10</i> ENSG00000121858
<i>APOL3</i> ENSG00000128284	<i>TCAF2P1</i> ENSG00000159860	<i>EPSTI1</i> ENSG00000133106	<i>UBE2L6</i> ENSG00000156587	<i>ETV7</i> ENSG00000010030	<i>TTC39A</i> ENSG00000085831
<i>CETP</i> ENSG00000087237	<i>XAF1</i> ENSG00000132530	<i>CMPK2</i> ENSG00000134326	<i>GBP1P1</i> ENSG00000225492	<i>HAS2</i> ENSG00000170961	<i>APOL2</i> ENSG00000128335
<i>IFI44</i> ENSG00000137965	<i>WARS1</i> ENSG00000140105	<i>RSAD2</i> ENSG00000134321	<i>OR2I1P</i> ENSG00000237988	<i>LGALS3BP</i> ENSG00000108679	<i>IFNG</i> ENSG00000111537
<i>NLRC5</i> ENSG00000140853	<i>OAS3</i> ENSG00000111331	<i>APOL6</i> ENSG00000221963	<i>PSME2</i> ENSG00000100911	<i>PTGES3P1</i> ENSG00000234518	<i>DLL4</i> ENSG00000128917
<i>TRIM69</i> ENSG00000185880	<i>ALPK1</i> ENSG00000073331	<i>PSMB9</i> ENSG00000240065	<i>NFE2L3</i> ENSG00000050344	<i>FAS</i> ENSG00000026103	<i>PLAAT4</i> ENSG00000133321
<i>HELZ2</i> ENSG00000130589	<i>BCL2L14</i> ENSG00000121380	<i>IFI35</i> ENSG00000068079	<i>TRIM21</i> ENSG00000132109	<i>UBD</i> ENSG00000213886	<i>IRF9</i> ENSG00000213928
<i>TNFSF4</i> ENSG00000117586	<i>RGS1</i> ENSG00000090104	<i>SSTR2</i> ENSG00000180616	<i>SMTNL1</i> ENSG00000214872	<i>TIFA</i> ENSG00000145365	<i>TMEM229B</i> ENSG00000198133
<i>PSME1</i> ENSG00000092010	<i>DDX60</i> ENSG00000137628	<i>CASZ1</i> ENSG00000130940	<i>HERC6</i> ENSG00000138642	<i>CASP17P</i> ENSG00000235505	<i>SAMD9L</i> ENSG00000177409
<i>USP18</i> ENSG00000184979	<i>NFIX</i> ENSG00000008441	<i>PSMB10</i> ENSG00000205220	<i>GBP6</i> ENSG00000183347	<i>DTX3L</i> ENSG00000163840	<i>PTPRK</i> ENSG00000152894
<i>IFIT3</i> ENSG00000119917	<i>KLHDC7B</i> ENSG00000130487	<i>TAPBP</i> ENSG00000231925	<i>PARP12</i> ENSG00000059378	<i>TAP2</i> ENSG00000204267	<i>RCN1</i> ENSG00000049449
<i>GPR174</i> ENSG00000147138	<i>TAP1</i> ENSG00000168394	<i>PARP10</i> ENSG00000178685	<i>GBP3</i> ENSG00000117226	<i>HLA-F</i> ENSG00000204642	<i>IFIH1</i> ENSG00000115267
<i>GPR155</i> ENSG00000163328	<i>PARP11</i> ENSG00000111224	<i>LINC02328</i> ENSG00000258733	<i>ERVK-28</i> ENSG00000267696	<i>MX1</i> ENSG00000157601	<i>TRIM22</i> ENSG00000132274

[0284]

<i>IL12RB1</i> ENSG0000009 6996	<i>TRAFD1</i> ENSG0000013 5148	<i>IRF1-AS1</i> ENSG0000019 7536	<i>OASL</i> ENSG0000013 5114	<i>ACOT7</i> ENSG0000009 7021	<i>GCH1</i> ENSG0000013 1979
<i>PARP3</i> ENSG0000004 1880	<i>OPTN</i> ENSG0000012 3240	<i>MOV10</i> ENSG0000015 5363	<i>ISG15</i> ENSG0000018 7608	<i>MDGA1</i> ENSG0000011 2139	<i>SAMD4A</i> ENSG0000002 0577
<i>SP110</i> ENSG0000013 5899	<i>IL15</i> ENSG0000016 4136	<i>CEACAM1</i> ENSG0000007 9385	<i>CASP4</i> ENSG0000019 6954	<i>PCGF5</i> ENSG0000018 0628	<i>NMI</i> ENSG0000012 3609
<i>SYNPO2</i> ENSG0000017 2403	<i>IFIT5</i> ENSG0000015 2778	<i>CXCL10</i> ENSG0000016 9245	<i>C4BPB</i> ENSG0000012 3843	<i>TRANK1</i> ENSG0000016 8016	<i>GSDMD</i> ENSG0000010 4518
<i>CD69</i> ENSG0000011 0848	<i>IFI6</i> ENSG0000012 6709	<i>DHX58</i> ENSG0000010 8771	<i>RAB1AP1</i> ENSG0000022 9023	<i>OAS2</i> ENSG0000011 1335	<i>UBA7</i> ENSG0000018 2179
<i>BISPR</i> ENSG0000028 2851	<i>PSMB8</i> ENSG0000020 4264	<i>APOBEC3F</i> ENSG0000012 8394	<i>PLAT</i> ENSG0000010 4368	<i>DDX58</i> ENSG0000010 7201	<i>HERC5</i> ENSG0000013 8646
<i>E2F3P1</i> ENSG0000026 7046	<i>DLG3</i> ENSG0000008 2458	<i>PARP14</i> ENSG0000017 3193	<i>CARD16</i> ENSG0000020 4397	<i>CALHM6</i> ENSG0000018 8820	<i>APOBEC3G</i> ENSG0000023 9713
<i>TOMM20P2</i> ENSG0000025 5987	<i>LAG3</i> ENSG0000008 9692	<i>HLA-A</i> ENSG0000020 6503	<i>TAFA2</i> ENSG0000019 8673	<i>IFI27</i> ENSG0000016 5949	<i>MX2</i> ENSG0000018 3486
<i>BTN3A1</i> ENSG0000002 6950	<i>IFIT1</i> ENSG0000018 5745	<i>USP41</i> ENSG0000016 1133	<i>RHEBL1</i> ENSG0000016 7550	<i>BTN3A3</i> ENSG0000011 1801	<i>TMEM140</i> ENSG0000014 6859
<i>RNF213</i> ENSG0000017 3821	<i>CCL8</i> ENSG0000010 8700	<i>ISG20</i> ENSG0000017 2183	<i>BST2</i> ENSG0000013 0303	<i>IFITM3</i> ENSG0000014 2089	<i>LINC01949</i> ENSG0000023 3828
<i>CLEC2B</i> ENSG0000011 0852	<i>LAP3</i> ENSG0000000 2549	<i>LY6E</i> ENSG0000016 0932	<i>CCDC194</i> ENSG0000026 9720	<i>BATF2</i> ENSG0000016 8062	<i>EIF2AK2</i> ENSG0000005 5332
<i>TNFSF13B</i> ENSG0000010 2524	<i>LINC02528</i> ENSG0000022 6004	<i>DGLUCY</i> ENSG0000013 3943	<i>VAMP5</i> ENSG0000016 8899	<i>NUB1</i> ENSG0000001 3374	<i>SAMD9</i> ENSG0000020 5413
<i>STAT2</i> ENSG0000017 0581	<i>RNF213-AS1</i> ENSG0000026 3069	<i>ZDHHC4P1</i> ENSG0000023 8025	<i>PGAP1</i> ENSG0000019 7121	<i>IFIT2</i> ENSG0000011 9922	<i>OAS1</i> ENSG0000008 9127
<i>CD38</i> ENSG0000000 4468	<i>XRN1</i> ENSG0000011 4127	<i>GVINP1</i> ENSG0000025 4838	<i>NT5C3A</i> ENSG0000012 2643	<i>LRRN2</i> ENSG0000017 0382	<i>FBXO39</i> ENSG0000017 7294

[0285]

[0286]

다음으로, 각각의 반응 간의 유전자 네트워크 패턴을 비교하였다. 첫째, 모듈 보존 통계량이 계산되었으며, 이와 더불어 결과는 LPS-유도 IFN 모듈이 이미퀴모드 및 폴리(I:C) 반응의 IFN 모듈 내에서 고도로 보존되었으나 그 반대는 아님을 나타낸다(데이터는 나타내지 않음). 이미퀴모드 및 폴리(I:C) 반응과 관련된 IFN 모듈은 서로 보존되었으며, 전염증성 모듈은 모든 반응 간에 보존되었다(데이터는 나타내지 않음). 둘째, 랭크된 유전자 발현 및 랭크된 연결성을 계산하여 모듈을 비교하였다. 체대혈 LPS-유도 IFN 모듈 유전자와 이미퀴모드 및 폴리(I:C) 자극 후 동일한 유전자 간의 발현 크기 ( $\rho = 0.88$  &  $0.82$ ) 및 모듈 내 연결성 ( $\rho = 0.57$  &  $0.59$ ) 간에 두드러진 차이가 각각 관찰되었다(도 2h). 모듈 내의 연결성을 조사하기 위해, 각 모듈의 모든 유전자에 걸친 연결성 밀도를 플롯팅하였으며, 이는 가장 많이 연결된 유전자를 식별하였다(도 3a, 도 3b-d; 좌측 패널). LPS-유도 IFN 모듈의 연결성은 정규 분포를 특징으로 하는 반면, 바이러스성 자극은 좌-편향된 분포를 생성하였다(도 3a). 주요 IFN 신호전달 유전자(예를 들어, IRF1, STAT1)는 LPS-유도 IFN 모듈 내에서 가장 많이 연결된 유전자 중에 존재하였지만, 가장 많이 연결된 유전자의 강도는 바이러스 자극의 IFN 모듈과 비교하여 감소하였다(도 3b-d; 좌측 패널). LPS-유도 전염증성 모듈은 이미퀴모드- 또는 폴리(I:C)-유도 염증성 모듈과 비교하여 더 큰 연결성을 나타내었다(데이터는 나타내지 않음). 선천적 면역/전염증성 사이토카인(예를 들어, IL1A/B, CXCL2/3/8)을 코딩하는 유전자는 출생 시 모든 반응의 전염증성 모듈에서 가장 많이 연결된 유전자 중 하나였다(데이터는 나타내지 않음). 요약하면, 바이러스 및 박테리아 자극이 전염증성 및 IFN 반응 유전자의 중첩 세트를 활성화하지만, 근본적인 네트워크 구조는 현저하게 상이하였다.

[0287]

**실시예 7 - 출생 시 및 연령 5세의 선천적 면역 기능의 핵심 조절자의 식별**

[0288]

VIPER(Alvarez 등, *Nat Genet*, 2016, 48:838-847) 분석을 사용하여 모듈 연결성 패턴을 유도할 것으로 예측되는 핵심 조절자를 식별하였다. 이 접근법은 LPS-유도 IFN 모듈이 출생 시 BATF3, STAT3 및 IRF1 전사 인자(TF)에 의해 추정적으로 구동되는 반면, 이미퀴모드- 및 폴리(I:C)-유도 IFN 모듈 상위 드라이버에는 여러 STAT(예

를 들어, STAT2) 및 IRF(예를 들어, IRF7) TF(도 3b-d; 우측 패널, 도 6b)가 포함되어 있음을 보여주었다. IRF1은 LPS-유도 IFN 모듈 유전자 180개 중 52개 유전자를 조절하는 것으로 확인되었다(표 4). 세 가지 반응 모두에 대한 전염증성 모듈은 CEBPB, AP-1(예를 들어, JUN, FOSL1) 및 NF-κB(예를 들어, NFKB2, RELB)에 대해 농축되었다(데이터는 나타내지 않음). 중요하게는, LPS 반응 IFN(169/180, 93.89%)(도 7) 및 전염증성(443/2297, 19.29%) 모듈(데이터는 나타내지 않음)로부터 보존된 것들로부터 제한된 입력 유전자를 사용한 반복 분석은 데이터를 변경하지 않았다. 마지막으로, 5세때 수집된 CMBC 및 일치된 PBMC 샘플(n=27) 간의 유전자 네트워크 패턴을 비교하였다. LPS-유도 IFN 모듈의 연결성은 출생과 비교하여 5세때 현저하게 높았으며, 이 모듈의 배선이 발달 조절의 대상임을 시사한다(도 3e 및 3f). 추가적으로, IRF1 농축은 체대혈에서만 식별되었다(도 3g). 대조적으로, 이미퀴모드 및 폴리(I:C) 자극에 의해 유발된 IFN 반응은 출생과 5세 사이에 비교적 유사한 연결성 패턴을 나타냈으며, 이를 뒷받침하는 것으로 추정되는 드라이버는 또한 출생과 5세 사이에 비교할만 하였다(예를 들어, STAT2, IRF7) (도 3h&i, 도 8). 이미퀴모드 및 폴리(I:C) 전염증성 모듈은 출생 시와 비교하여 5세때 수집된 혈액에서 모듈 내 연결성이 감소되는 것을 특징으로 한다(데이터는 나타내지 않음).

[0289] 표 4. IRF1 레귤론에 의해 조절되는 LPS-유도 IFN 모듈 유전자. 유전자 심볼(symbol) 및 ENSEMBL 식별번호가 제공된다.

<i>CLEC2B</i> ENSG00000110852	<i>CXCL9</i> ENSG00000138755	<i>ERVK-28</i> ENSG00000267696	<i>FAS</i> ENSG00000026103
<i>FBXO6</i> ENSG00000116663	<i>GBP1</i> ENSG00000117228	<i>GBP1P1</i> ENSG00000225492	<i>GBP2</i> ENSG00000162645
<i>GBP3</i> ENSG00000117226	<i>GBP4</i> ENSG00000162654	<i>GBP5</i> ENSG00000154451	<i>GBP6</i> ENSG00000183347
<i>GPR174</i> ENSG00000147138	<i>HAPLN3</i> ENSG00000140511	<i>IDO1</i> ENSG00000131203	<i>IFNG</i> ENSG00000111537
<i>IL12RB1</i> ENSG00000096996	<i>IL15</i> ENSG00000164136	<i>IL15RA</i> ENSG00000134470	<i>LINC02328</i> ENSG00000258733
<i>NFIX</i> ENSG00000008441	<i>NLRC5</i> ENSG00000140853	<i>OR211P</i> ENSG00000237988	<i>PARP3</i> ENSG00000041880
<i>PLAAT4</i> ENSG00000133321	<i>PSMB10</i> ENSG00000205220	<i>PSMB9</i> ENSG00000240065	<i>PSME1</i> ENSG00000092010
<i>PSME2</i> ENSG00000100911	<i>PTGES3P1</i> ENSG00000234518	<i>SAMD4A</i> ENSG00000020577	<i>SOCS1</i> ENSG00000185338
<i>STAT1</i> ENSG00000115415	<i>SYNPO2</i> ENSG00000172403	<i>TAP1</i> ENSG00000168394	<i>TAPBP</i> ENSG00000231925
<i>TCAF2</i> ENSG00000170379	<i>TCAF2C</i> ENSG00000283528	<i>TCAF2P1</i> ENSG00000159860	<i>TIFA</i> ENSG00000145365
<i>TRAFD1</i> ENSG00000135148	<i>UBD</i> ENSG00000213886	<i>UBE2L6</i> ENSG00000156587	<i>USP30-AS1</i> ENSG00000256262

[0290] 실시예 8 - 출생 시 선천적 면역 반응은 생후 첫 해에 sLRI를 예측한다

[0292] 출생 시 선천적 면역 반응이 생후 첫 해에 sLRI를 예측할 수 있는지 결정하기 위해, 데이터 세트는 훈련(50%, n=25) 및 유효성 검증 세트(50%, n=25)에 랜덤으로 할당되었으며, 랜덤 포레스트 분류기는 CBMC IFN 모듈에 대해 훈련되었다. 놀랍게도, LPS-유도 IFN 모듈 유전자에 대해 훈련된 분류기는 유효성 검증 데이터 세트에서 72%의 정확도로 생후 첫 해에 sLRI를 예측할 수 있었다(ROC 곡선하 면적 = 0.724)(도 4a, 도 9). 대조적으로, 이미퀴모드 또는 폴리(I:C) 데이터로부터 구축된 분류기는 생후 첫 해에 sLRI를 예측하지 못하였다(도 4a). 예측 랜덤 포레스트 모델은 또한 180개 유전자 중 일부가 다른 것보다 더 높은 정확도를 제공함을 나타내었으며, 이는 감소된 지니(Gini) 평균 또는 감소된 정확도 평균으로 정량화되었다(표 5 및 도 9a). 지니 중요도 스코어링의 경우, 통계적 역치 (중앙값보다 높은 2개의 중앙값 절대 편차)를 사용하여 모델로부터 어떤 유전자가 가장 예측적인지를 결정하였다. 14개의 유전자가 이 역치를 초과하였다(표 5).

[0293] 표 5. LPS-유도 IFN 모듈에서 지니 중요도(Gini Importance)에 의해 점수화된 상위 필터링된 유전자.

유전자 심볼	양상블 ID	평균 지니지수 감소량 (MeanDecreaseGini)
<i>KLHDC7B</i>	ENSG00000130487	0.398792717
<i>IFNG</i>	ENSG00000111537	0.248785492
<i>CASZ1</i>	ENSG00000130940	0.217869997
<i>PSMB9</i>	ENSG00000240065	0.206963824
<i>PARP3</i>	ENSG00000041880	0.188704302
<i>ACOT7</i>	ENSG00000097021	0.187326939
<i>NUB1</i>	ENSG00000013374	0.182121626
<i>USP18</i>	ENSG00000184979	0.179460628
<i>NLRC5</i>	ENSG00000140853	0.173439287
<i>CCDC194</i>	ENSG00000269720	0.168838173
<i>GCH1</i>	ENSG00000131979	0.164823734
<i>PARP11</i>	ENSG00000111224	0.162908602
<i>CXCL11</i>	ENSG00000169248	0.1562304
<i>PMAIP1</i>	ENSG00000141682	0.156106684

[0294]

[0295] 입력으로서 이용가능한 비교적 작은 샘플 수를 고려하여 이러한 결과가 재현 가능한지 여부를 테스트하기 위해, 상기 분석은 훈련/유효성 검증 세트에서 랜덤으로 대상체 구성원을 재-샘플링함으로써 반복되었으며, 다시 이를 LPS-유도 IFN 모듈 유전자만이 평균적으로 우연보다 유아기에 sLRI를 더 잘 예측할 수 있음을 확인하였다(도 10). 또한, 유아기에 sLRI 감수성과 관련하여 LPS-유도 IFN 모듈에 대해 현저하게 상이한 연결성 패턴이 관찰되었으며, 이는 이미퀴모드 또는 폴리(I:C) IFN 모듈로부터 분명하지 않았다(도 4b&c). 구체적으로, 감수성이 있는 개체는 비록 반응의 추정 드라이버가 비교할만 할지라도, LPS-유도 IFN 모듈에 대해 더 강한 유전자 네트워크 패턴을 보였다(IRF1, STAT3, BATF)(도 4b,d-e(i)). 또한, 이미퀴모드 및 폴리(I:C) 반응을 LPS-유도 IFN 모듈의 해당 유전자로만 제한하는 것은 유아기에 sLRI 감수성과 관련하여 연결성 패턴 또는 드라이버에서 눈에 띄는 차이가 나타나지 않았다(도 11). 감수성이 있는 개체에 대한 CBMC의 LPS-유도 IFN 모듈의 연결성 밀도 플롯(도 4b)은 5세때 PBMC 연결성의 전체 연결성 밀도(도 3f)와 유사하지만, 모듈 내 연결성은 관련이 없었으며(데이터는 나타내지 않음), 이는 상이한 프로세스로부터 유사성이 나타난다는 것을 시사한다. 모듈 고유 유전자는 또한 전체 모듈 발현을 요약하고 임상적 특성과 연관성을 보여주기 위해 계산되었다. 데이터는 제대혈 LPS-유도된 IFN 모듈 고유 유전자가 생후 첫 해에 sLRI에 감수성이 있는 개체(p=0.016)뿐만 아니라 연령 5세때 천식(p=0.015) 및 현재 진행되는 천명음(p=0.02)이 있는 개체가 계층화되었음을 나타내었다(도 4f, 도 12). 이 결과는 LPS 반응에 대해서만 유의미하였으며, 이는 IFN 모듈에 대해 특이적이었다(도 4g, 도 12).

[0296] 실시예 9 - 외부 코호트에서 출생 시 인터페론 반응의 유효성 검증

[0297] 배양물에서 CBMC에 의해 유도된 IFN 모듈 유전자 발현 프로파일과 관련된 상기 발견이 생체내 아동기 감염에 대한 자연적으로 발생하는 IFN 반응을 반영하는지 여부를 조사하기 위해, RF 분류기는 CBMC 데이터에 대해 훈련되었으며, 유전자 발현 옴니버스(Gene Expression Omnibus)에서 공개적으로 이용가능한 일련의 데이터 세트로부터 유래된 샘플을 분류하는 데 사용되었다. 제1 데이터 세트는 확인된 박테리아(n=52) 또는 바이러스(n=92) 감염으로 입원을 필요로 하는 열성 질환을 가진 아동(<17세) 대 건강한 대조군(n=52)으로부터의 전혈 유전자 발현 프로파일이 포함되었다(GSE72809; Herberg 등, *JAMA*, 2016, 316:835-845). LPS 및 이미퀴모드/폴리(I:C) 데이터에 대해 훈련된 RF 분류기는 각각 박테리아(AUC=0.889) 및 바이러스(AUC=0.874/0.838) 감염이 있는 아동을 정확하게 예측하는 것으로 확인되었다(도 5a-c, 도 13). 제2 데이터 세트는 급성 바이러스성 세기관지염으로 입원한

유아(<18개월, n=30) 및 영아(18개월-5세, n=32)로부터의 PBMC 샘플로 구성되었다(GSE113211; Jones 등, *Am J Respir Crit Care Med*, 2019, 199:1537-1549). 비자극 및 이미퀴모드(AUC=0.8) 또는 폴리(I:C)(AUC=0.877)로 처리된 CBMC에 대해 구축된 분류기는 연령에 관계없이, 회복 후 (증상 없음, 감염 후 8.8 ± 2.5주) 일치된 샘플과 비교하여 급성 질환 동안 수집된 샘플을 정확하게 계층화할 수 있다. 모델은 유아(AUC = 0.922, 폴리(I:C); AUC = 0.827, 이미퀴모드) 및 아동(AUC = 0.789, 폴리(I:C); AUC = 0.842, 이미퀴모드)에 대해 별도로 잘 수행되었다(데이터는 나타내지 않음). 제3 데이터 세트는 바이러스성-관련 또는 비-바이러스성 "감기"-유사 질환(발병 후 1-6일)이 있는 천식 아동(6-17세)의 연구 방문으로부터의 비강-유래 유전자 발현 프로파일로 구성되었으며, 이들 중 일부는 나중에 악화를 경험하였다(n=83, 58명은 바이러스-양성이었음)(GSE115770; Altman 등, *Nat Immunol*, 2019, 20:637-651). 호흡기 바이러스 감염을 가진 증상이 있는 아동은 이미퀴모드(AUC=0.8) 및 폴리(I:C)(AUC=0.832) 정의된 RF 분류기에서 증상이 있지만 바이러스-음성인 아동으로부터 정확하게 예측되었다(도 5e). 또한, 후속적으로 악화를 경험한 (증상 발병 후 10일 이내) 바이러스-양성 및 바이러스-음성 천식 아동을 분류하는 비교할만한 정확도가 있었다(데이터는 나타내지 않음). 동일한 연구에서, 바이러스 악화 대비 비-바이러스 악화의 말초 혈액-유래 유전자 발현 프로파일은 폴리(I:C)보다 이미퀴모드(AUC=0.671) 정의된 분류기(AUC=0.627)에서 더 정확하게 분류되었다(데이터는 나타내지 않음). 그러나, 이 코호트에서 혈액 발현 프로파일로부터의 악화와 관계없이 바이러스 양성 천식 대 바이러스 음성 천식을 예측할 때 성능이 좋지 않았다(데이터는 나타내지 않음).

[0298] 실시예 10 - 다중-오믹 통합

[0299] 마지막으로, 다중-오믹 데이터 통합(DIABLO; Singh 등, *Bioinformatics*, 2019, 35:3055-3062)을 사용하여 sLRI 위험을 부여하는 생물학적 층 전반에 걸쳐 상관된 분자적 특징을 식별하였다. 입력 데이터는 CBMC 기준선 면역 세포 유형 비율(n=8), 가변 mRNA 전사체(n=6353), VIPER-유래 레귤론 활성 점수(n=1224) 및 사이토카인/케모카인 단백질(n=39)로 구성되었다. LPS-유도 IFN-신호전달 전사(IRF9, STAT1, GBP2/4) 및 IRF1 활성은 T 세포, 단핵구 및 DC 세포 유형, 면역(HOXB4, NFIX) 조절제 및 전염증성 사이토카인/케모카인(IL-1β, MIP1α, MIF)과 조합하여, 생후 첫 해의 sLRI에 대한 위험의 주요 결정 요인인 것으로 데이터를 강화하였다(데이터는 나타내지 않음). 제대혈로부터 수집된 B 세포 및 T 세포는 또한 LPS에 의해 자극될 때 IRF1 및 STAT1의 상향조절에 의해 나타난 바와 같이 IFN 신호의 공급원인 것으로 나타났다(도 15). LPS-자극된 단핵구/수지상세포 및 NK 세포는 이들 유전자를 상향조절하지 않았다(데이터는 나타내지 않음).

[0300] LPS-유도 IRF1 활성은 네트워크, 핵심 조절자 및 통합 분석으로부터 별도로 식별되었으므로, 본 발명자들은 IRF1 유전자 발현 상관관계를 추가로 조사하였다. 출생 시 IRF1 유전자 발현은 선택적 STAT 및 IRF 계열 전사 인자(예를 들어, STAT1, IRF9), 전염증성 매개체(예를 들어, IL-1β, IL-6, CCL3/MIP-1α) 및 바이러스-관련 수용체(예를 들어, ICAM1, IFIH1)와 양의 상관관계가 있다(도 5d-f). 추가적으로, CBMC STAT1 및 IFIH1 유전자 발현은 유아기에 sLRI에 대해 감수성이 있는 개체 중에서 LPS에 더 높게 반응하였으며, IFIH1 발현은 IRF1 및 STAT1 발현과 상관관계가 있었다(도 14).

[0301] 요약하면, 중증 바이러스성 하부 호흡 기도 감염(sLRI)은 유아 및 아동의 입원의 주요 원인이며 후속 천식 발병의 증대한 위험 요소가 된다. 박테리아 및 바이러스 병원체가 상호작용하여 sLRI의 발병기전(pathogenesis)을 유도한다는 것이 점점 더 인식되지만, 근본적인 선천적 면역 메커니즘은 잘 이해되지 않는다. 이러한 지식 격차를 해결하기 위해, 본원의 실시예에 기술된 바와 같이 다중-오믹 접근법을 사용하여 출생 시 박테리아(LPS/TLR4) 및 바이러스(폴리(I:C)/TLR3; 이미퀴모드/TLR7) 자극에 대한 선천적 면역 반응을 체계적으로 프로파일링하였으며, 출생 후 첫 해에 sLRI에 대한 감수성과 관련된 반응 패턴을 조사하였다. 데이터는 자극의 패널에 대한 선천적 면역 반응이 전염증성 및 IFN-매개 유전자 발현 프로그램과 중첩되는 것을 포함하지만, 폴리(I:C)/이미퀴모드가 아닌 LPS 반응 프로파일이 sLRI를 예측하는 것임을 나타내었다. 더욱이, 감수성은 IFN 유전자 네트워크의 활성화에 의해 결정되었고, 제대혈 LPS 반응에서 이 네트워크의 연결성 패턴은 sLRI 위험이 있는 유아들 중에서 현저하게 과장되었다. 더욱이, 이들 유전자의 연결성 패턴은 또한 제대와 5세의 LPS 반응 사이에 매우 가변적이었으며, sLRI 위험을 결정하는 동일한 메커니즘이 발달 조절의 대상임을 시사한다. 이러한 결과는 LPS 유도된 IFN 반응에 대해 특이적이며, TLR3/7 자극에 의해 유발된 IFN 반응이나 테스트된 입양의 반응에 대한 전염증성 유전자에서는 관찰되지 않았으며, 이는 유아기에 sLRI에 대한 위험이 강화된 아동에서 LPS 반응의 배선이 특별히 변경되었음을 시사한다. LPS-유도 IFN 모듈의 발현은 경미한(비-천명음/비-열성) 하부 호흡 기도 감염과 관련 없음이 주목할 만하며, 이는 특히 감염 중증도와 이러한 결과의 관계를 강조한다. 핵심 조절자 분석은 출생 시 LPS-유도 IFN 반응의 주요 드라이버로서 IRF1을 식별하였다. 5세까지, IRF7을 포함한 IRF 전사 인자 계열의 다른 구성원에 의해 IRF1의 추정 활성이 대체되었다. 대조적으로, IRF7은 출생 시 및 5세때 TLR3/7

IFN 반응의 지배적인 드라이버였다. 마지막으로, DIABLO를 사용하여 본 발명자들은 sLRI 위험과 관련된 다중-오믹 시그니처를 식별하였으며, 이는 전염증성 사이토카인 및 면역 조절제와 관련된 IRF1 조절 및 IFN 유전자를 특징으로 한다. 요약하면, 본 발명자들의 결과는 생후 첫 해에 sLRI에 대한 감수성이 주로 항-박테리아 대 항-바이러스 선천적 면역 경로에 의해 결정되고, 초기 개입을 위해 위험에 처한 유아의 식별에 대한 근거를 제공하며, 약물 개발을 위한 표적을 식별함을 시사한다.

[0302] 유아기에 라이노바이러스-유도된 천명음의 재발성 에피소드는 CAS (Kusel 등, *J Allergy Clin Immunol*, 2007, 119:1105-1110) 및 COAST (Jackson 등, *Am J Respir Crit Care Med*, 2008, 178:667-672)를 포함한 전향적 코호트 연구로부터 입증된 바와 같이, 종종 초기-발병 천식의 첫 번째 징후이다. 그러나, 바이러스 감염은 종종 증상이 없는 천식 아동에서 정례적으로 검출되며, 생애 초기에 전반에 걸친 재발성 sLRI만으로는 천식 발병을 유도하기에 충분하지 않으며, 이는 추가적인 질환 보조-인자에 대한 관여를 시사한다. 이와 관련하여, 본원에 제시된 데이터는 병원성 박테리아에 대한 반응이 바이러스 자극에 대한 반응보다 sLRI 감수성에 대해 더 중요한 결정 요인임을 시사한다. 특히, 생후 첫 5년 동안 기도 미생물군집에서 이러한 박테리아 군집의 존재비를 증가시키는 방향으로의 전환은 종종 바이러스 검출 또는 호흡기 증상의 발병보다 앞서는 것으로 이전에 나타내었다. 이는 병원성 박테리아의 존재가 후속 바이러스 감염 동안 종종 증상의 발현에 대해 기도를 "프라이밍(prime)"할 수 있음을 시사한다.

[0303] 본 연구는 출생 시 강화된 LPS-유도 IFN 반응/유전자 네트워크 연결성 패턴이 유아기에 바이러스 sLRI 위험을 부여한다는 것을 확인하였으며, 이는 IFN 반응이 급성 바이러스 감염 동안 거의 보편적으로 보호된다는 점을 고려하면 놀랍다. 본 연구에서, 시스템 생물학은 LPS-유도 IFN 네트워크의 핵심 조절자로서 IRF1을 식별하기 위해 사용되었다. IRF1은 기준선에서 인터페론-매개 항바이러스 프로그램의 구성적 발현 및 호흡기 바이러스 감염에 의해 유발되는 이러한 프로그램의 유도성 발현을 촉진한다. 그러나, IRF1은 인터페론 프로그램의 유도에 필수적인 것은 아니며, 항바이러스 면역에 필수적인 역할보다는 상호 보완적인 역할을 제공한다. 본원에 기술된 바와 같은 실시예는 출생 시 LPS-유도 IRF1 유전자 발현이 별개의 IFN-신호전달 매개체(예를 들어, STAT1이지만 JAK1/TYK2는 아님)뿐만 아니라 전염증성 유전자/사이토카인 발현(예를 들어, CXCL9/10/11, IL1B)과 연관되어 있음을 나타낸다(도 5e-g). 이는 제1형 IFN-활성화 IRF1이 특정 전-염증성 유전자의 유도를 촉진한다는 증거와 일치한다. 또한, 바이러스 감지 및 부착 관련 수용체(도 5f)와 IRF1 연관성은 이러한 수용체의 IRF1 제어를 위한 역할을 지원하며, 이는 증가된 IRF1 활성화가 어떻게 라이노바이러스 감염에 대한 과장된 반응에 대해 선천적 면역 시스템을 프라이밍(prime) 할 수 있지를 부분적으로 설명할 수 있다.

[0304] 선천적 면역 반응은 다층의 분자 조절에 걸쳐 다수의 세포 유형의 조직화된 활성화에 의해 지배된다. 이러한 이유로, 유아기에 sLRI 위험과 관련된 집합적 반응을 포착하기 위해 공-발현된 특징의 통합된 바이오마커 프로파일 이 생성되었다. 놀랍게도, IFN-신호전달 유전자는 체대혈의 LPS 자극 후 선택된 다중-오믹 위험 프로파일의 전사체 중에서 두드러졌으며, 이는 네트워크-기반 접근법(IRF1, STAT1)으로부터의 주요 결과와 일치한다. 중요한 바이러스 인식 수용체 MDA5를 코딩하는 IFIH1도 식별되었다. 추가로, IFIH1 발현은 민감성이 있는 개체에서 유의미하게 상승되었으며, IRF1 및 STAT1과 상관관계가 있으며, 이는 위험에 처한 개체가 TLR4 활성화 시 주요 바이러스 감지 수용체를 조절하지 못할 수 있음(dysregulate)을 시사한다. DC 교차-제시에 의한 CD8<sup>+</sup> T 세포 활성화에 중요한 MHC 클래스 I 조절(NLRC5, RFX5)을 포함하는 다른 중요한 IRF1-의존성 선천적 경로가 상기 데이터에서 명백하다. 선택된 사이토카인 중에서, IL-1β는 특히 IFN 관련 전사체와 강력하게 연결되어, TLR4 활성화의 다운스트림에 연결된 항바이러스 및 전염증성 반응 프로그램을 강조하였다. 마지막으로, 유아기에 sLRI 위험에 대한 통합된 프로파일에서 선택된 많은 특징은 천식과 직접적으로 관련되어 있다. 이들은 천식 위험 유전자(IRF1, P2RY14, ABO)의 전사체 및 리모델링(MMP7), 기도 염증(NLRC5), Th2-조절장애(LGALS3BP), 뿐만 아니라 세포 유형(T 세포, 단핵구/DC, pDC) 및 사이토카인(예를 들어, IL-1β, IL-16, MIF)과 관련된 특징을 포함한다. 본원에 사용된 바와 같이 지도(supervised) 데이터 통합 접근법은 생물학적 층들에 걸쳐 위험 프로파일을 확장한다. 이는 여러 생물학적 수준으로부터 의미있는 정보를 추출하는 통합 다중-오믹스 방법의 고유한 능력을 강조한다.

[0305] 본 연구는 단일 출생 코호트에서의 CBMC 선천적 면역 반응에 초점을 맞추었으며, 신생아의 면역 시스템이 생후 첫 주 및 수 개월 동안 급격한 발달 변화를 겪는 것을 고려하면, CBMC 반응이 아동기 동안 후기에 발생하는 감염에 대한 면역 반응을 어느 정도 반영하는지 의문을 제기하는 것이 타당하다. 이 문제를 해결하기 위해, 랜덤 포레스트 분류기는 제공된 시험관내 데이터에 대해 훈련되었으며, 외부 코호트로부터 유래된 감염-관련 숙주 반응 데이터에 적용되었다. LPS 또는 이미퀴모드/폴리(I:C) 자극 후 유도된 IFN 반응을 사용하여 전혈 샘플로부터 각각 현재 박테리아성 및 바이러스성 열성 감염으로 병원에 방문한 아동을 정확하게 계층화할 수 있는 것으로

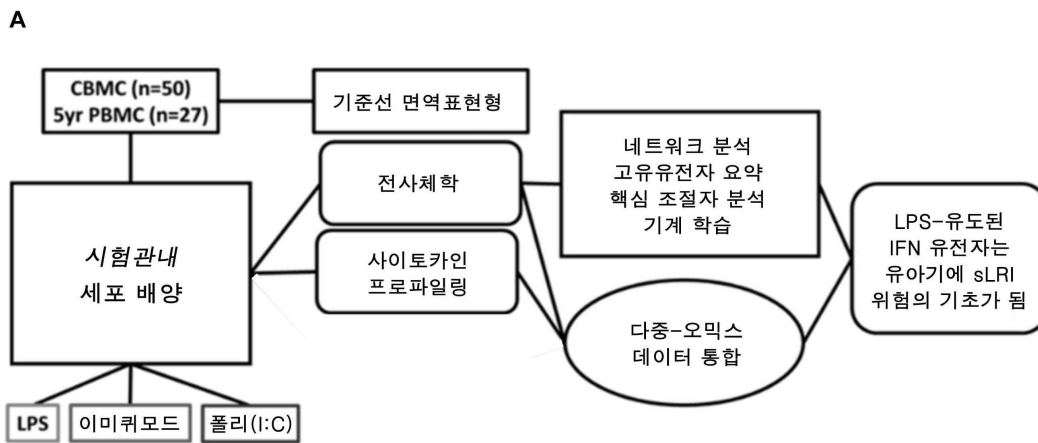
확인되었다. 이러한 데이터는 생체내 호흡기 감염에 대해 본원에 기술된 시험관내 모델의 관련성을 주장한다. 더욱이, 이미퀴모드 또는 폴리(I:C)에 의해 유도된 CBMC 반응은 또한 혈액 및 기도에서 바이러스성 세기관지염 및 천식 악화가 있는 유아 및 아동의 호흡기 바이러스 감염을 예측하였으며, 이는 시그니처가 순환 혈액 기도 조직 간 세포 구성의 변화에 대해 어느 정도 강력하다는 것을 시사한다. 또한, 유아(<18개월)를 예측할 때 더 어리거나(18개월~5세) (GSE113211; Jones 등, *AM J Respir Crit Care*, 2019, 199:1537-1549) 또는 더 나이가 많은 (6-17세) 아동 (GSE115770; Altman 등, *Nat Immunol*, 2019, 20:637-651)에 비해 랜덤 포레스트 모델의 정확도가 더 높은 것으로 확인되었다. 종합하면, 이는 재대혈에 대한 시험관내 조사로부터 식별된 IFN 유전자 네트워크가 실제 세계 상황에서 감염의 진짜(*bona fide*) 반응 매개체임을 뒷받침한다.

[0306]

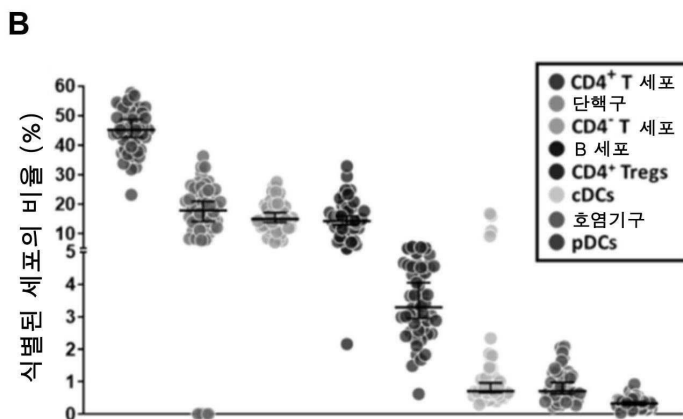
요약하면, 본원에 기술된 결과는 출생 시 LPS-유도 IFN 반응이 생후 첫 해에 sLRI의 위험을 예측하고, sLRI 감수성 및 아동 천식 발병에 대한 선천적 면역 궤적을 변형하는 데 잠재적인 유용성을 갖는 세포 및 분자 표적을 식별함을 증명한다.

**도면**

**도면1a**

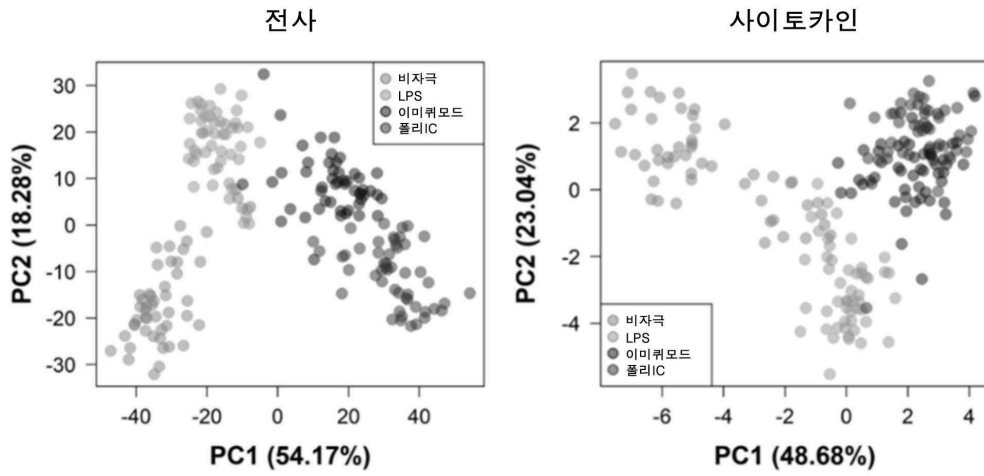


**도면1b**



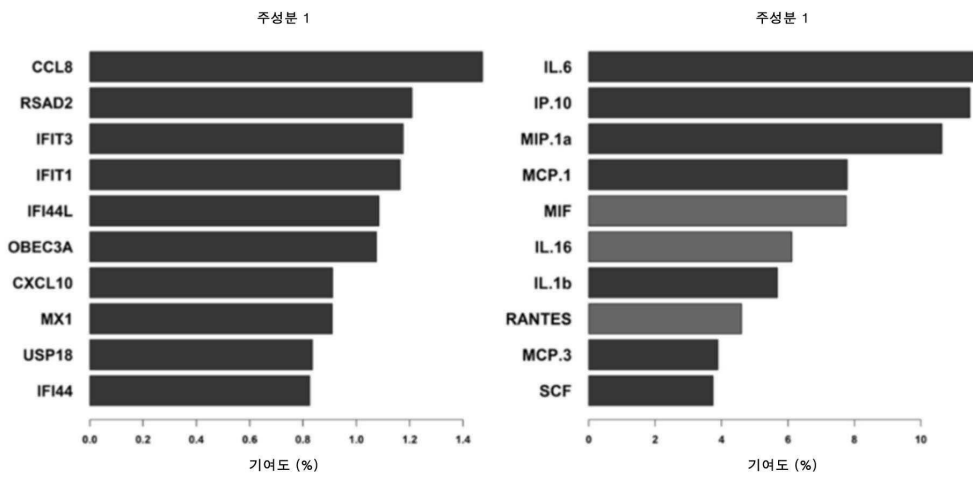
도면1c

C



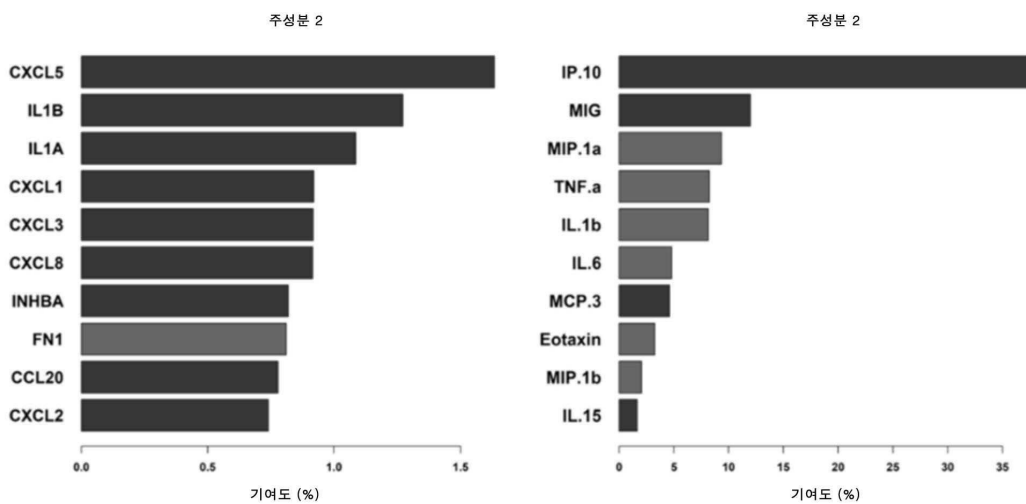
도면1d

D



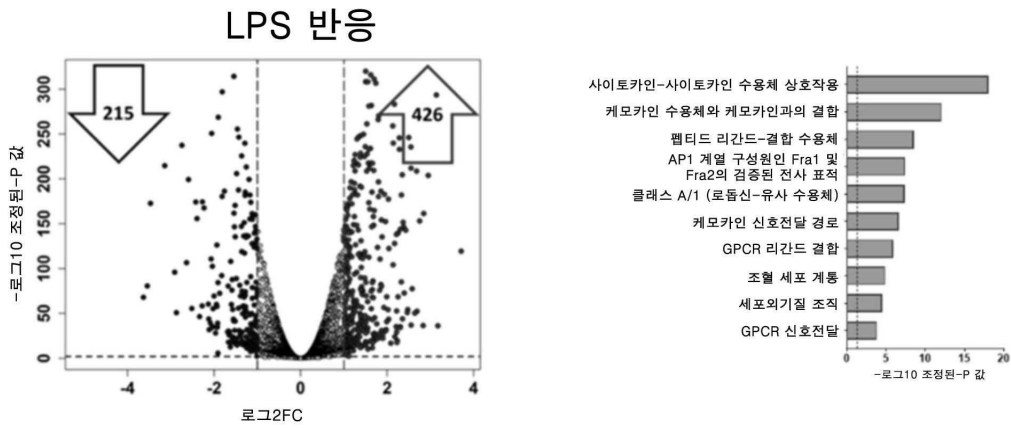
도면1e

E



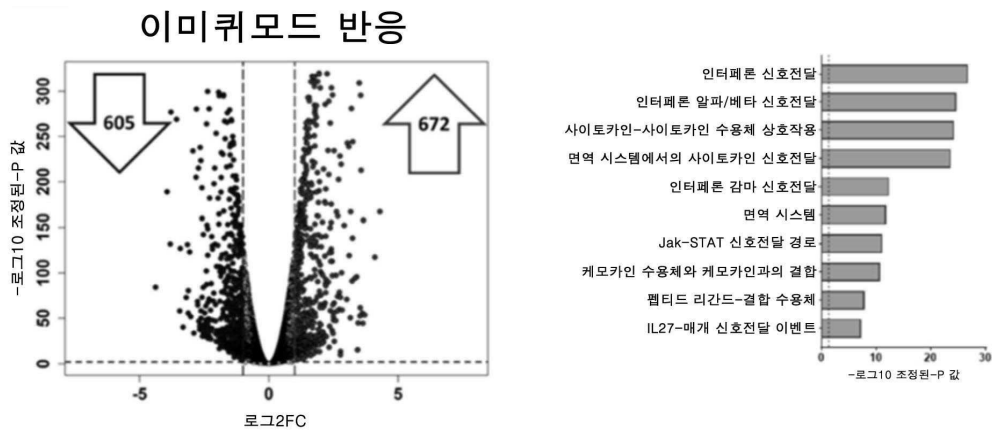
도면2a

A



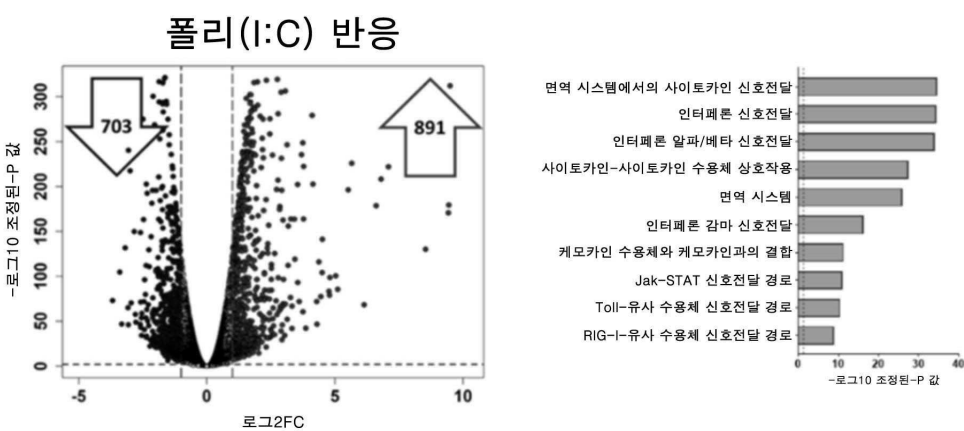
도면2b

B



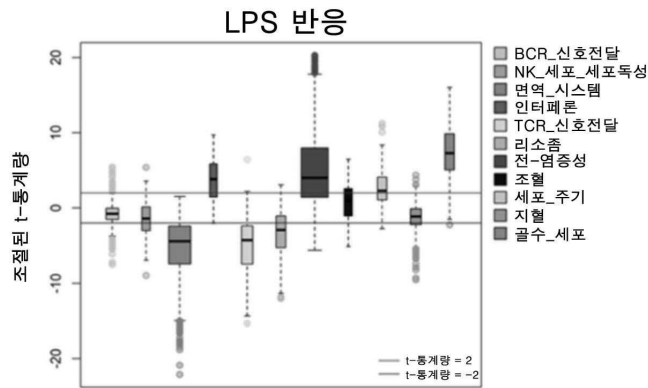
도면2c

C



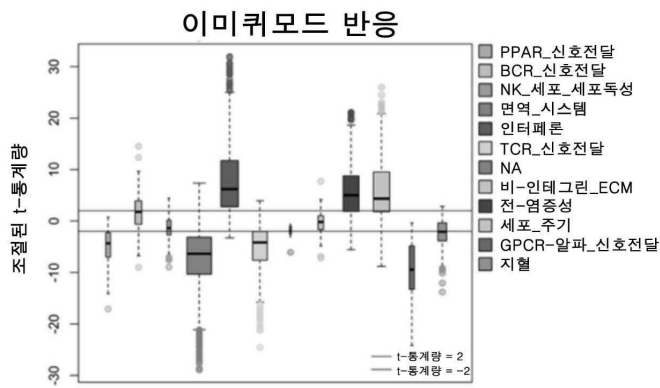
도면2d

D



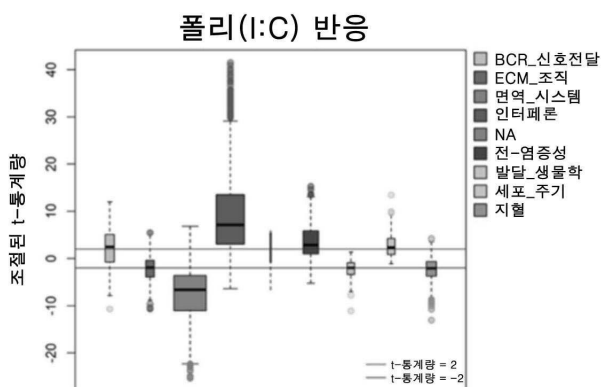
도면2e

E



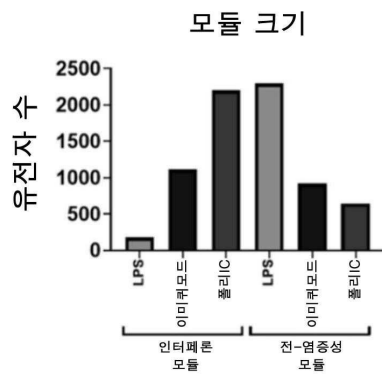
도면2f

F



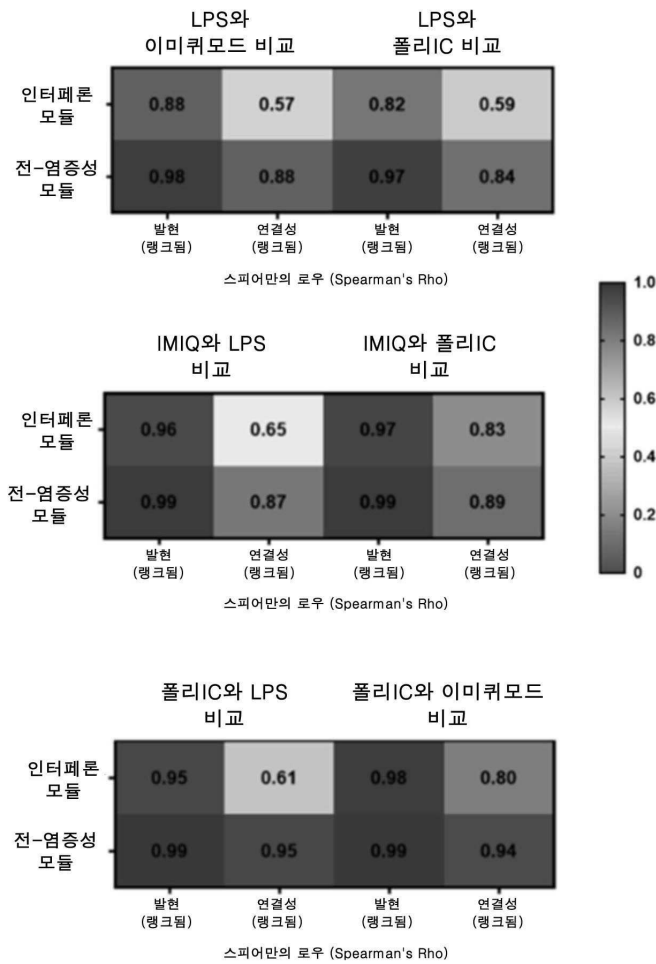
도면2g

G



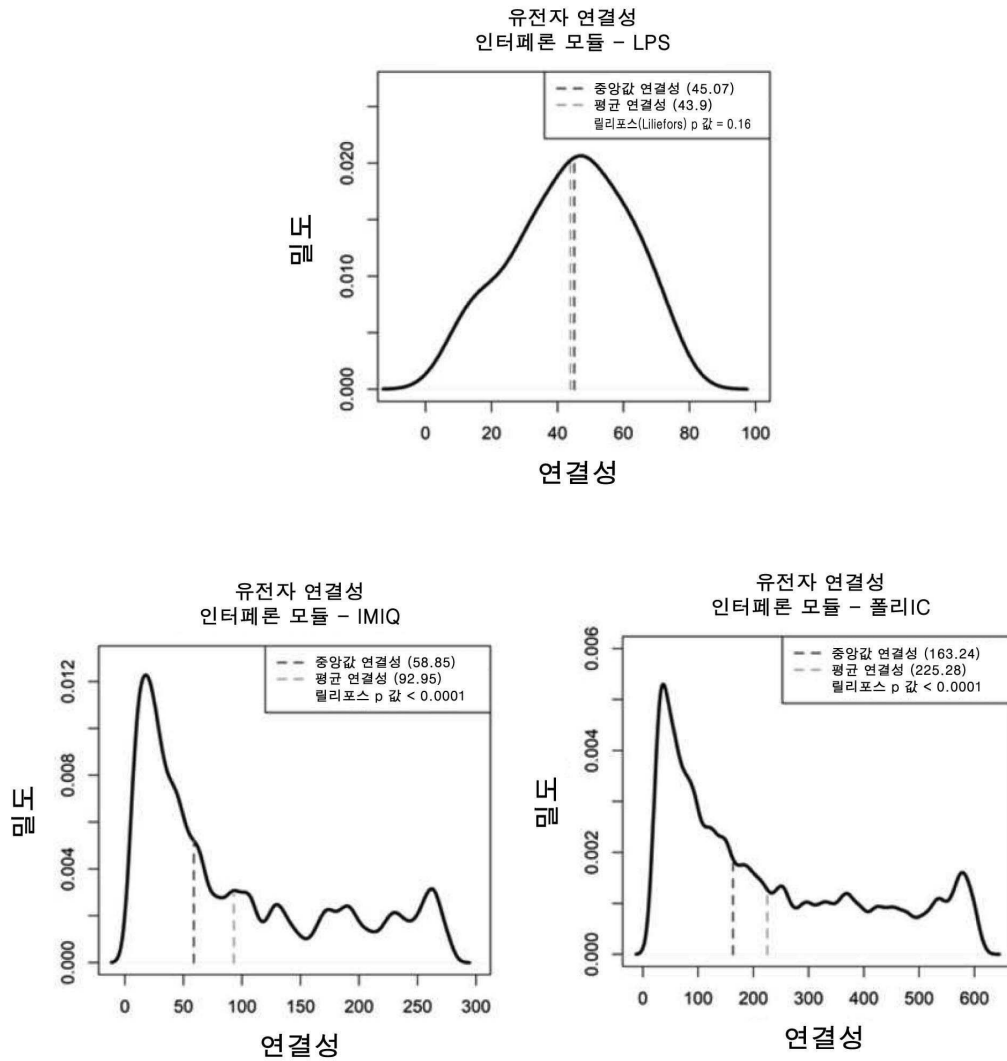
도면2h

H



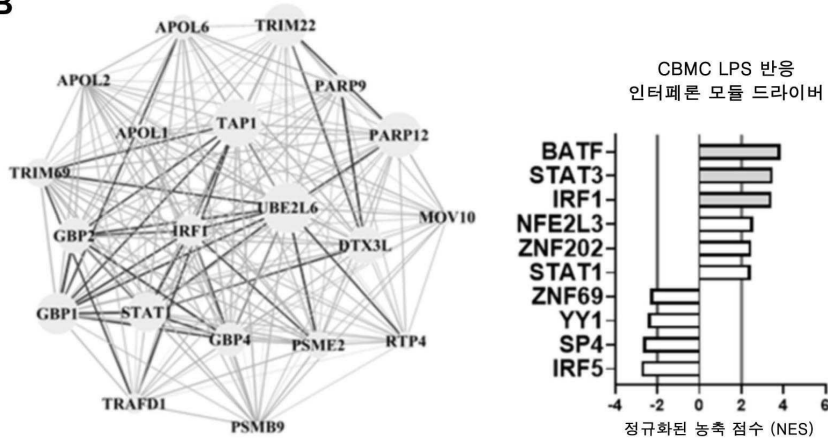
도면3a

A



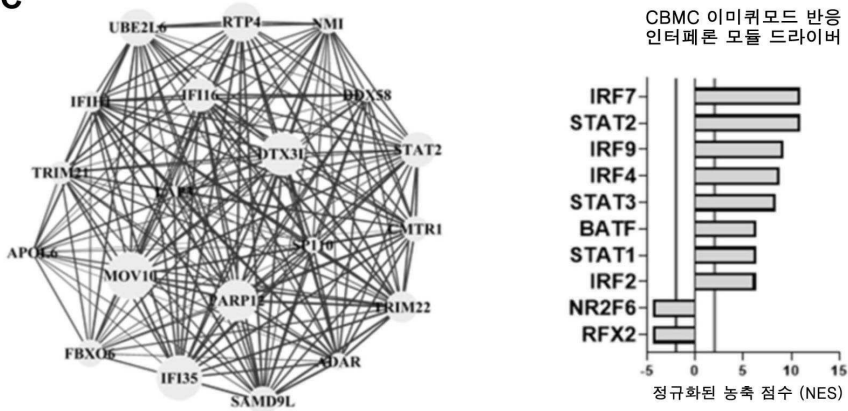
도면3b

B



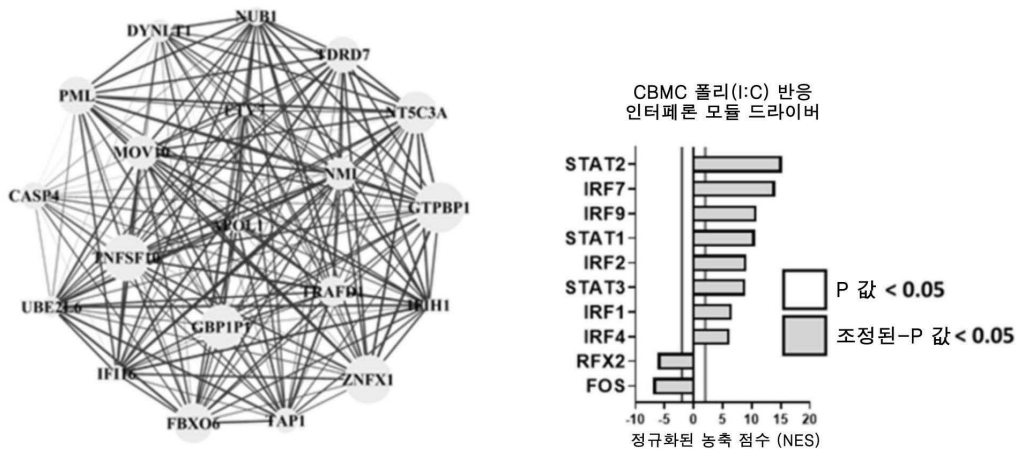
도면3c

C



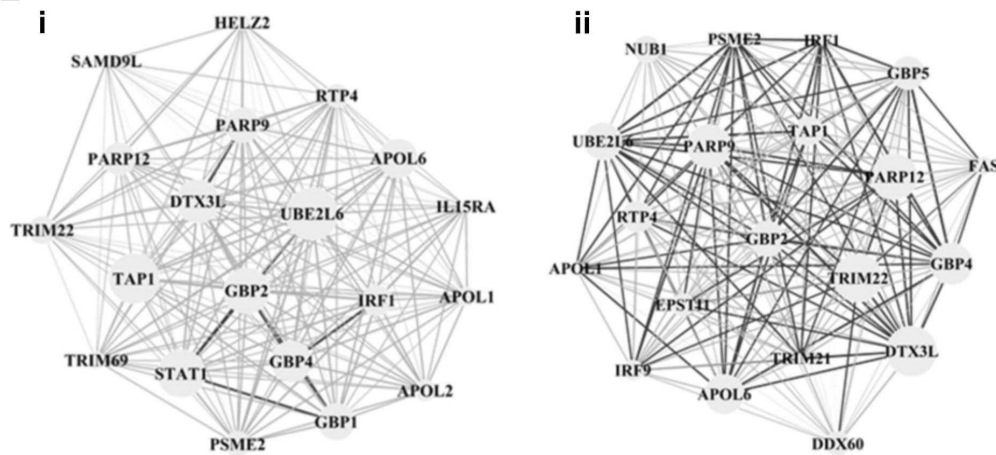
도면3d

D

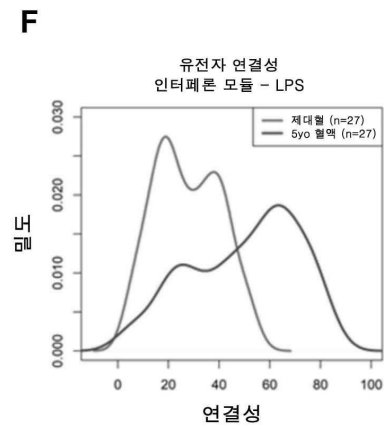


도면3e

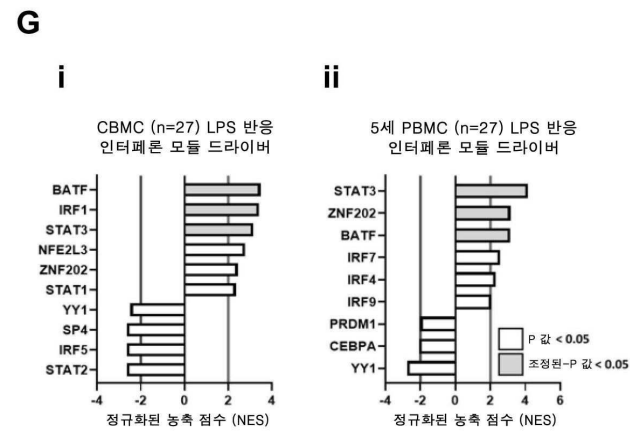
E



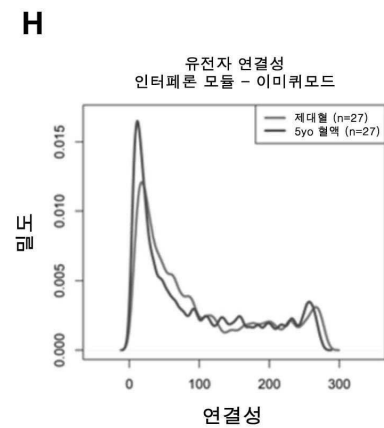
도면3f



도면3g

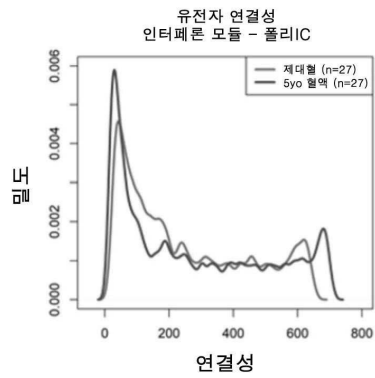


도면3h



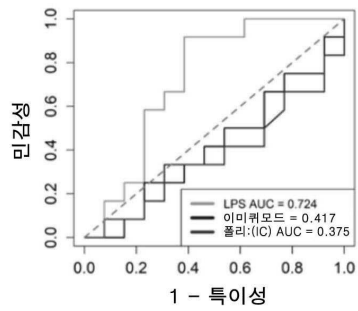
도면3i

I



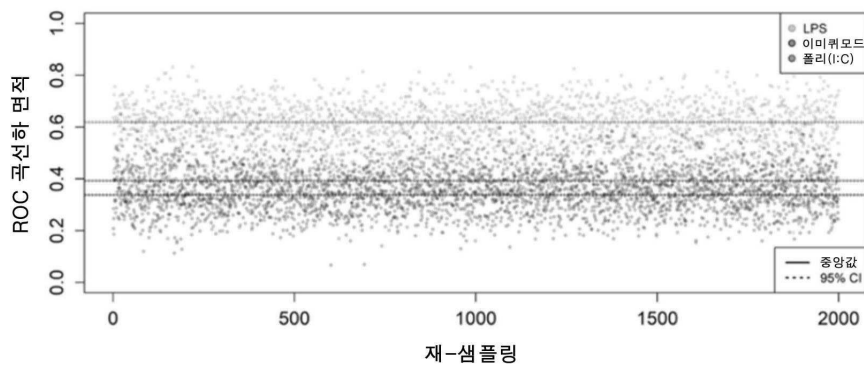
도면4a

A

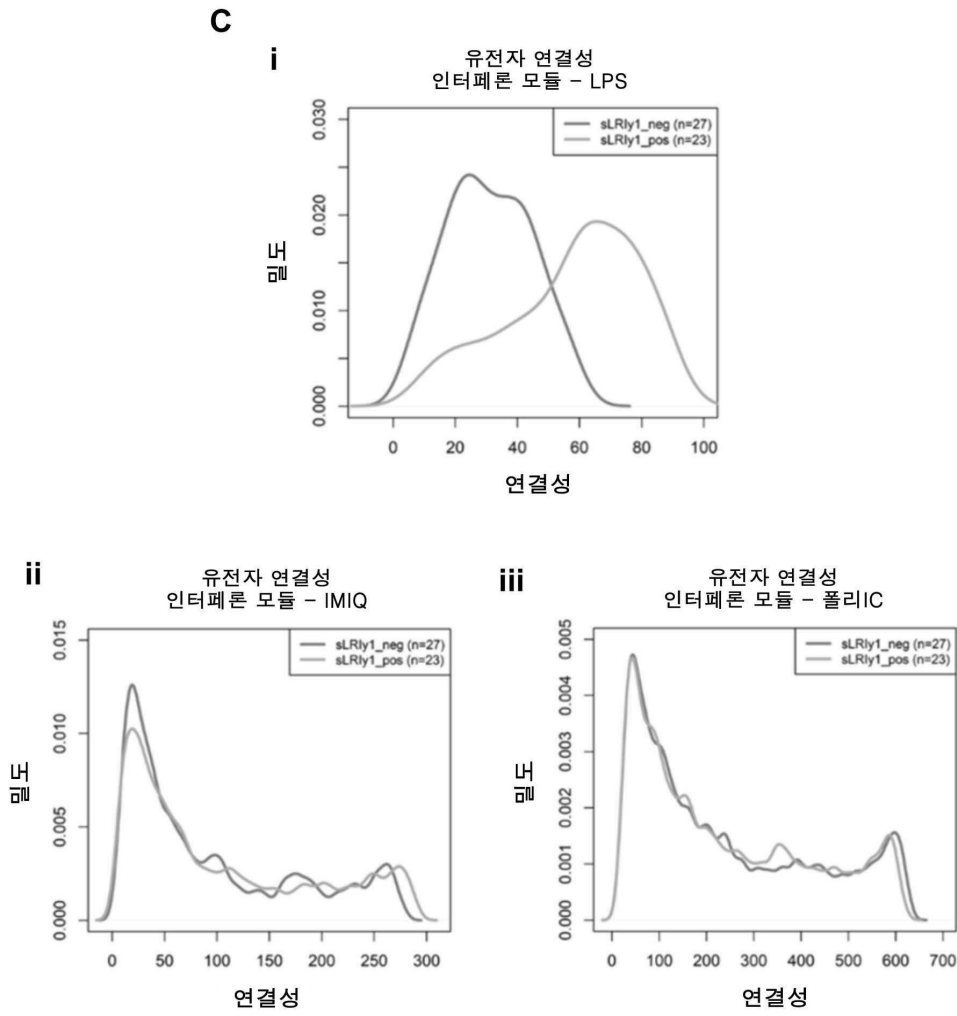


도면4b

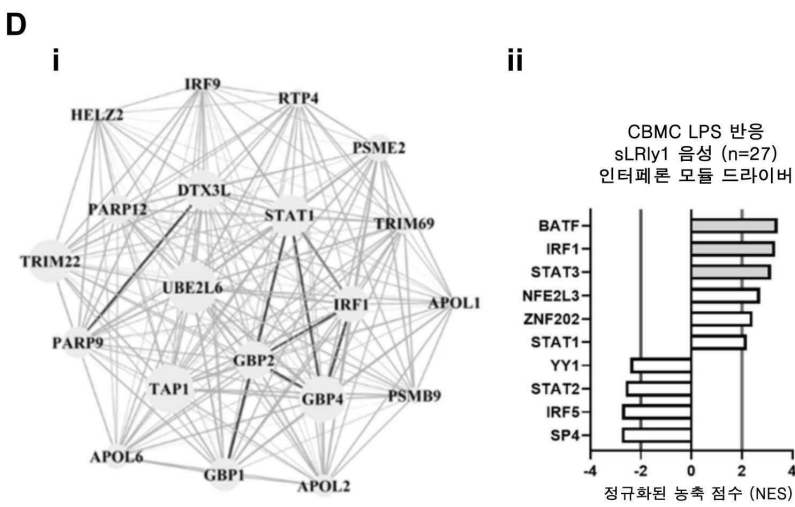
B



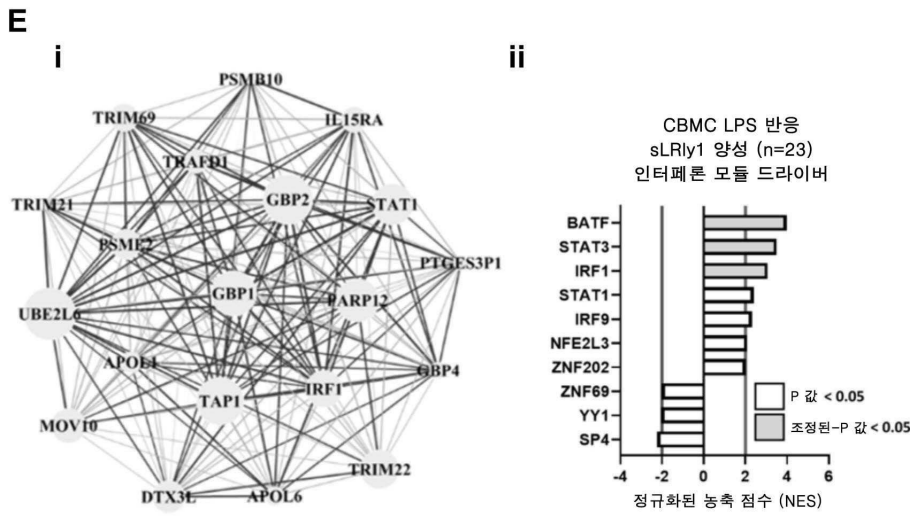
도면4c



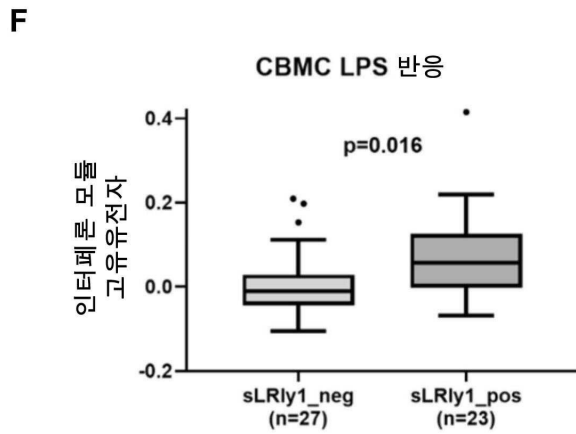
도면4d



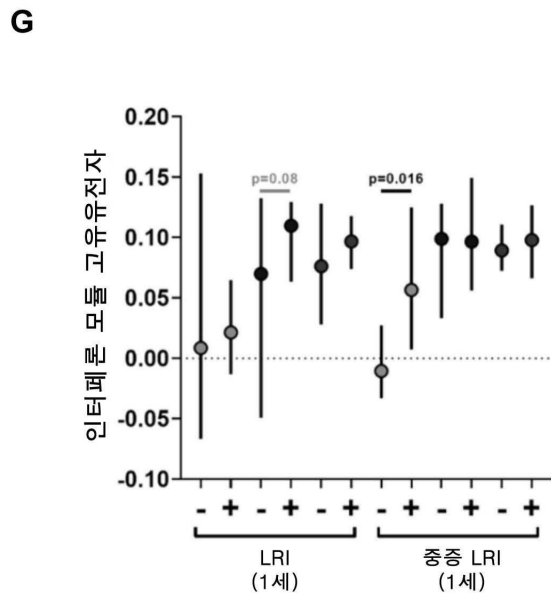
도면4e



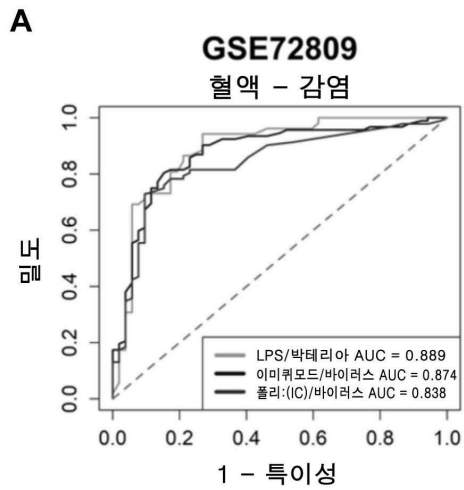
도면4f



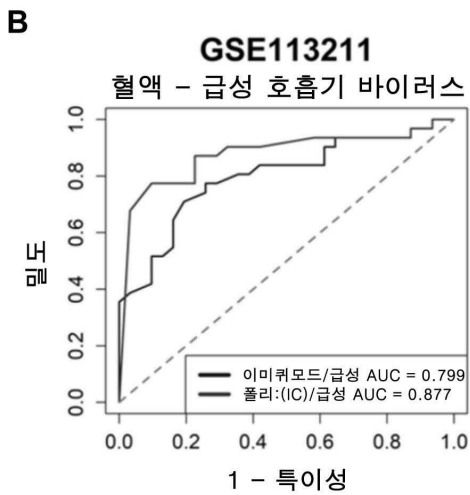
도면4g



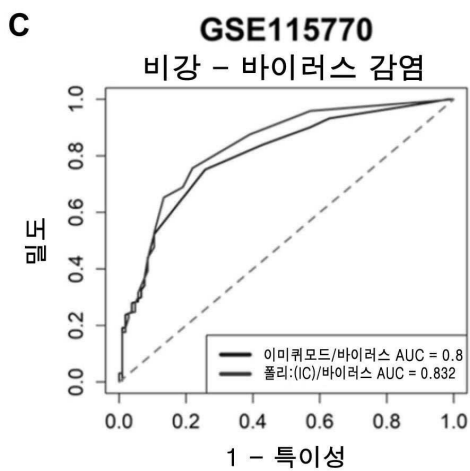
도면5a



도면5b

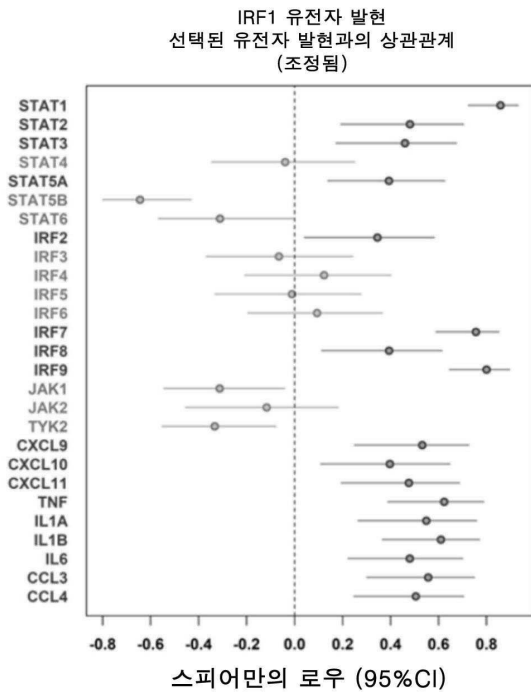


도면5c



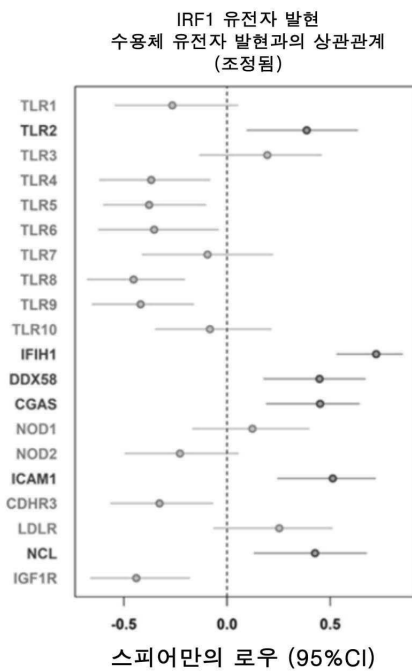
도면5d

D



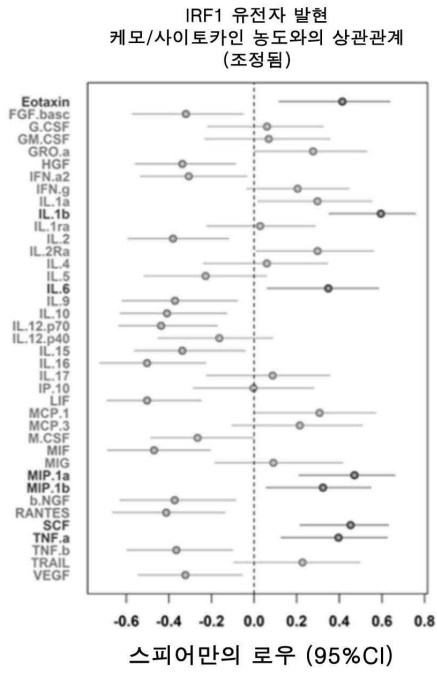
도면5e

E

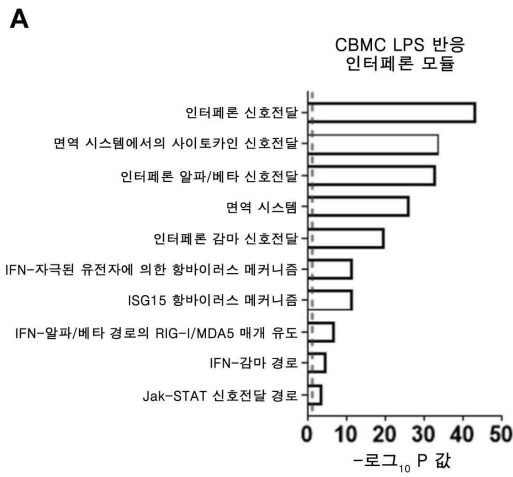


도면5f

F

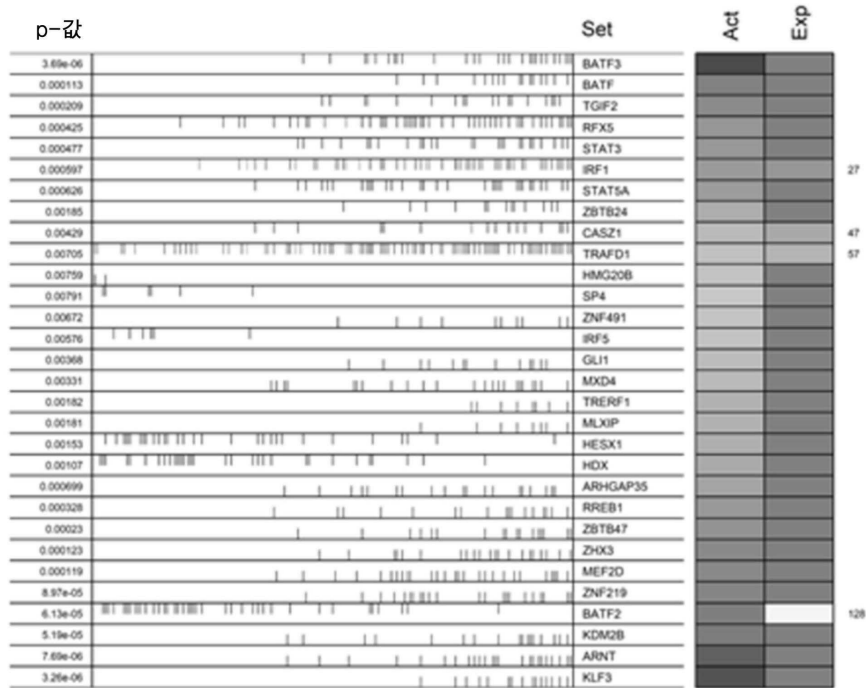


도면6



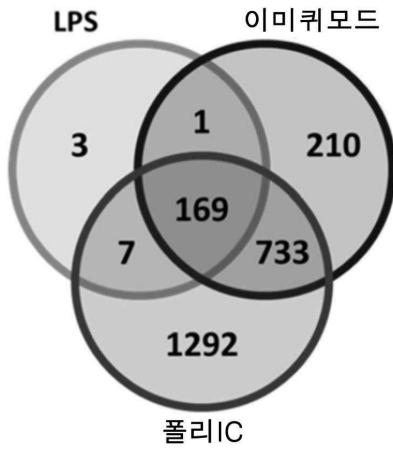
**B**

LPS – 인터페론 모듈

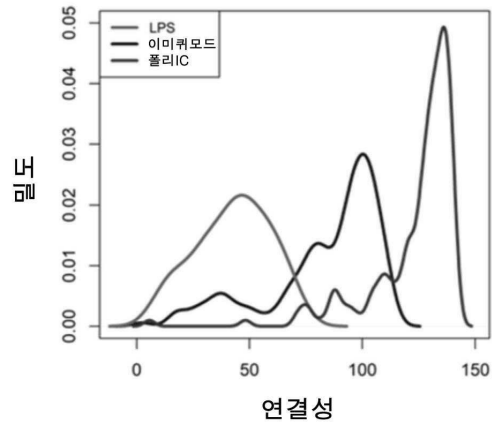


도면7

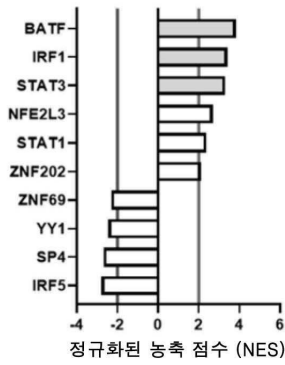
인터페론 모듈



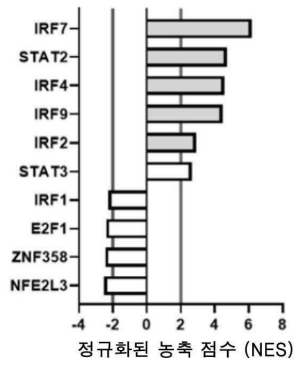
유보존된 유전자의 연결성  
인터페론 모듈



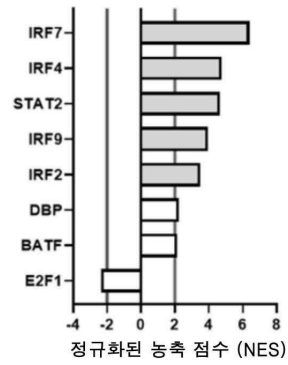
CBMC LPS 반응  
인터페론 모듈 드라이버  
(보존된 유전자, n=169)



CBMC 이미퀴모드 반응  
인터페론 모듈 드라이버  
(보존된 유전자, n=169)

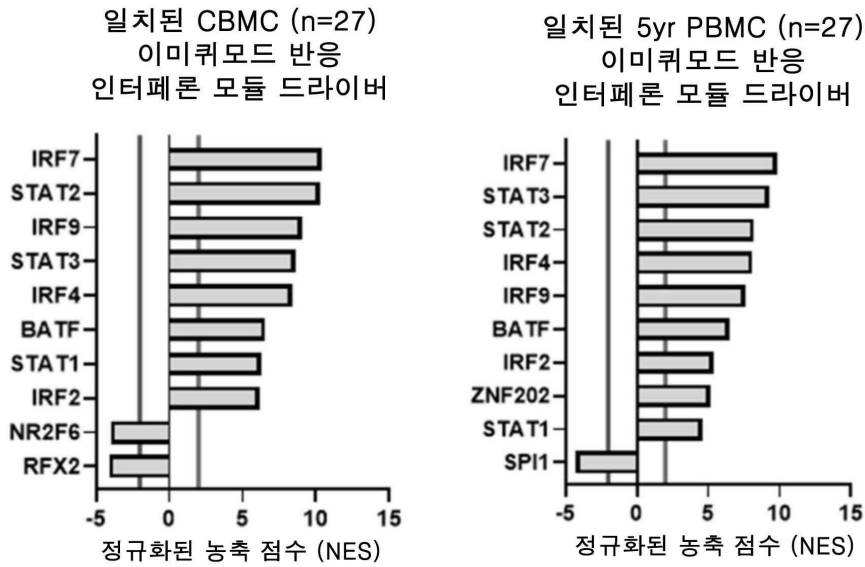


CBMC 폴리IC 반응  
인터페론 모듈 드라이버  
(보존된 유전자, n=169)

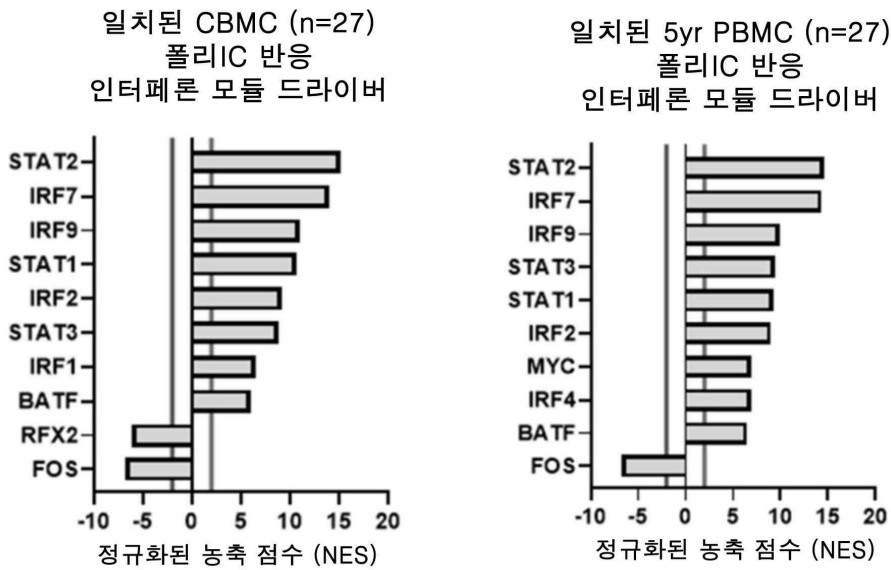


도면8

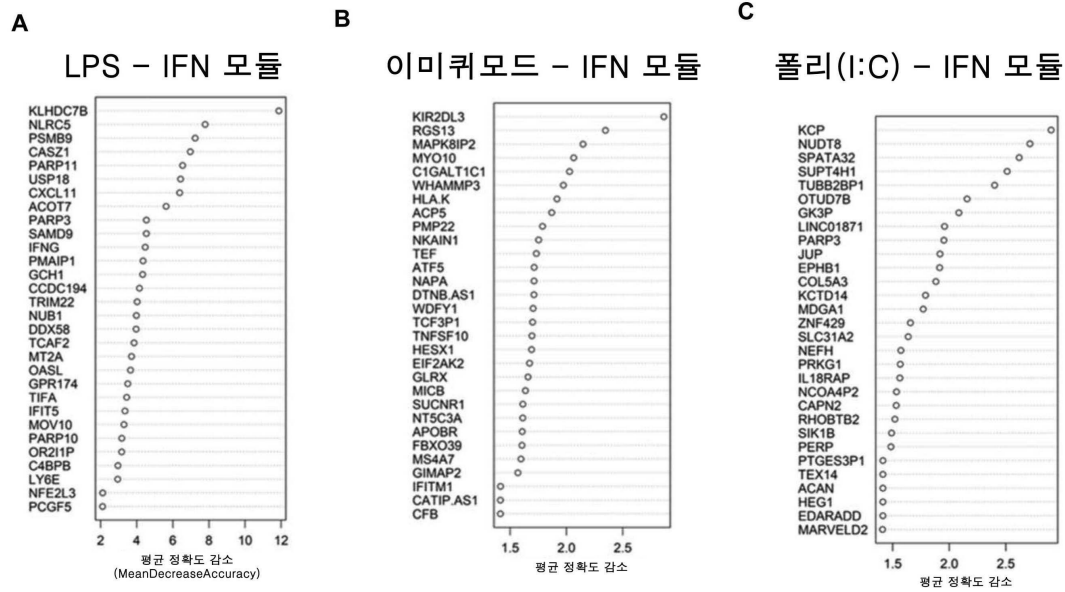
A



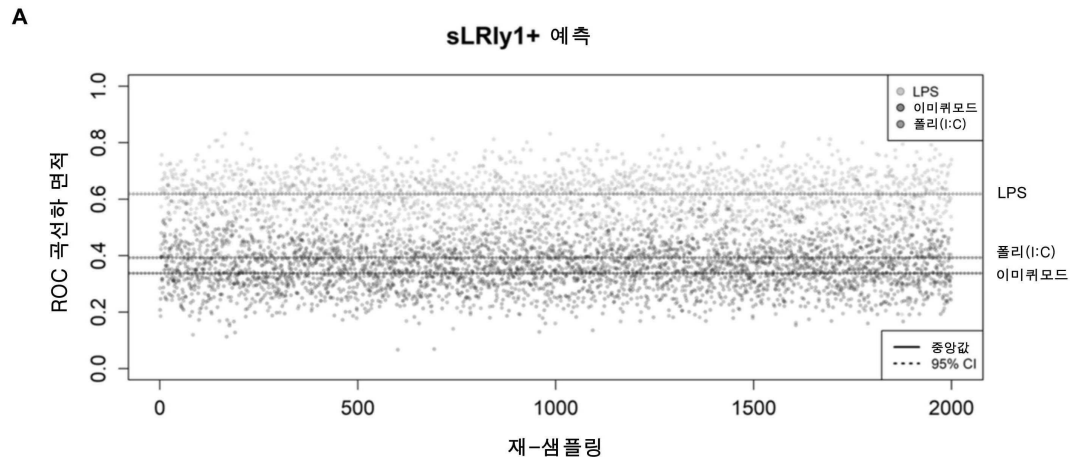
B



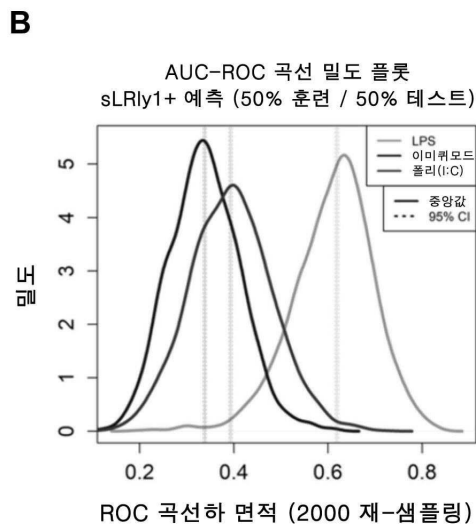
도면9



도면10a

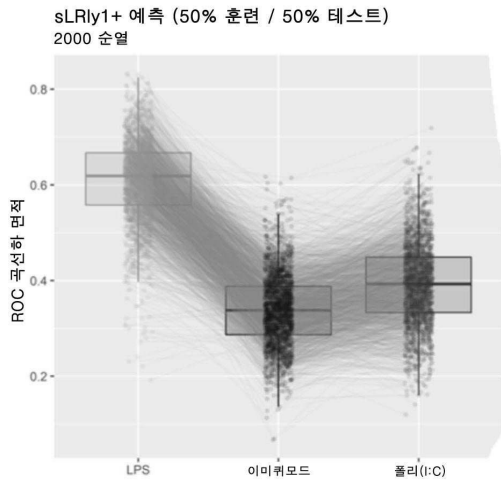


도면10b



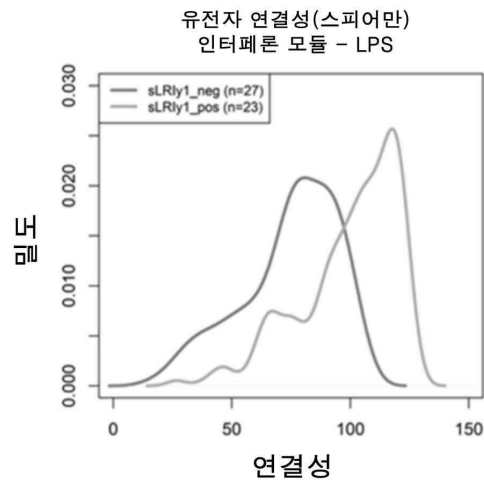
도면10c

C

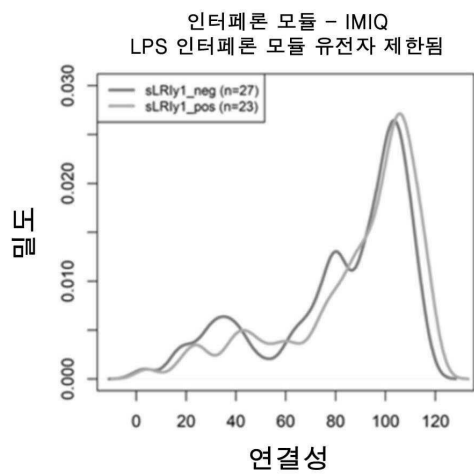


도면11

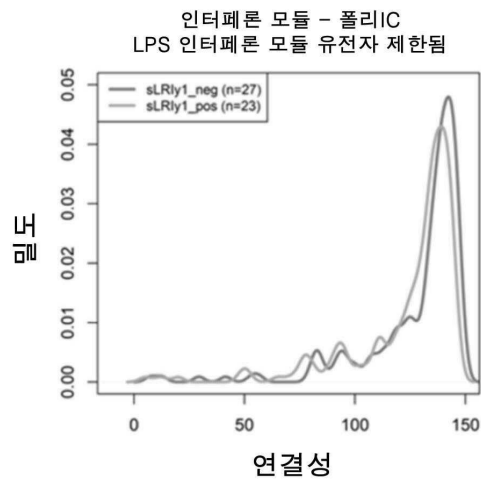
A



B

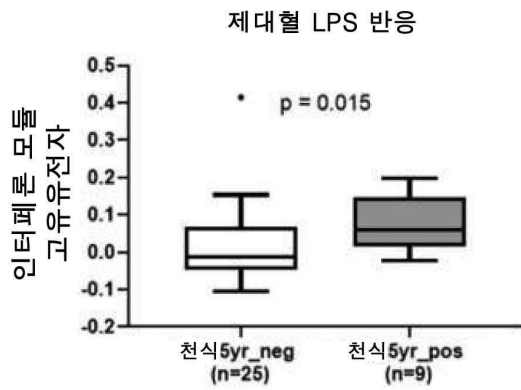


C

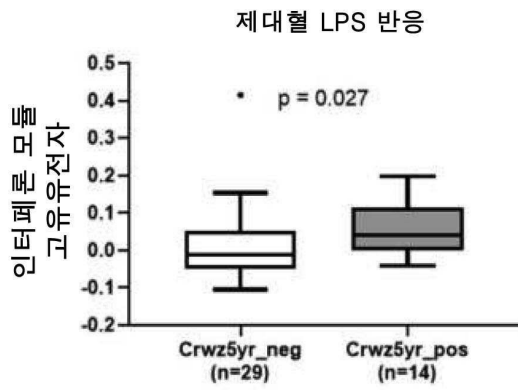


도면12a

A

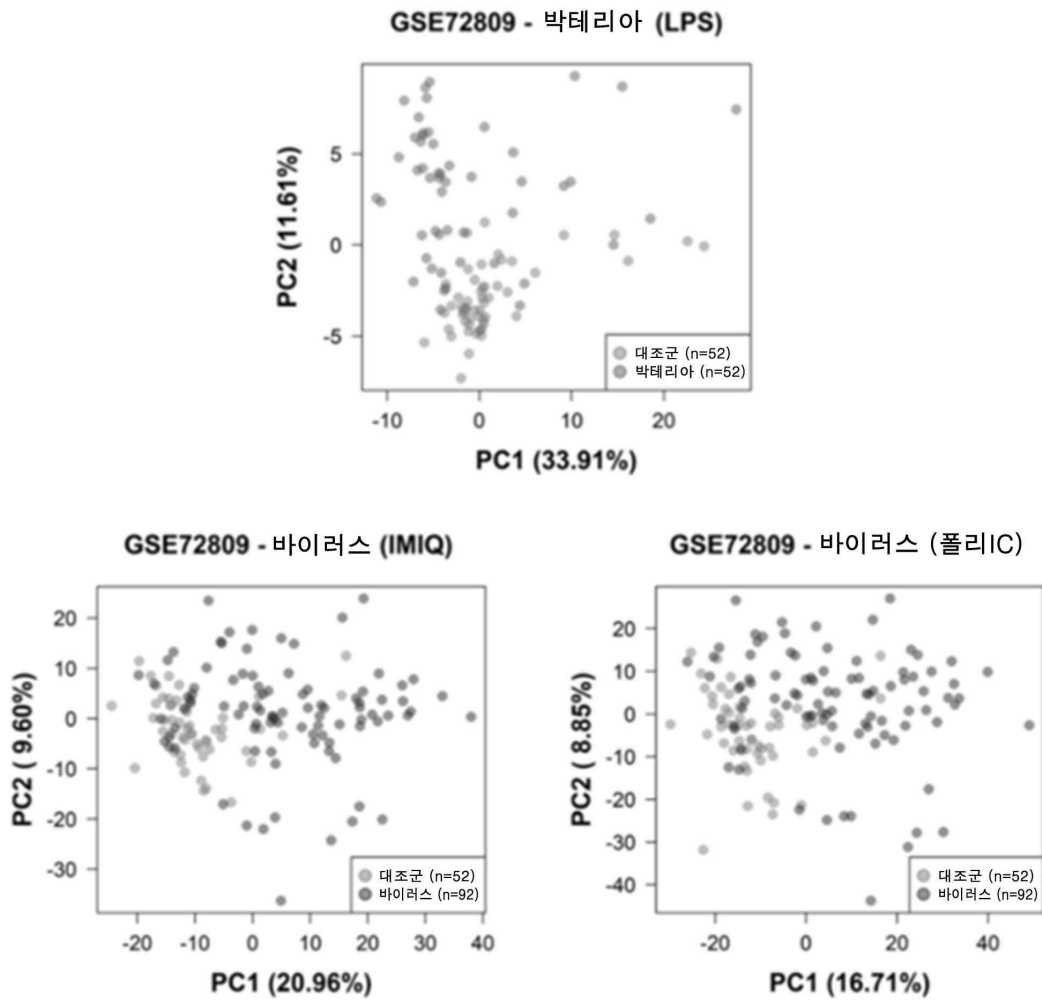


B



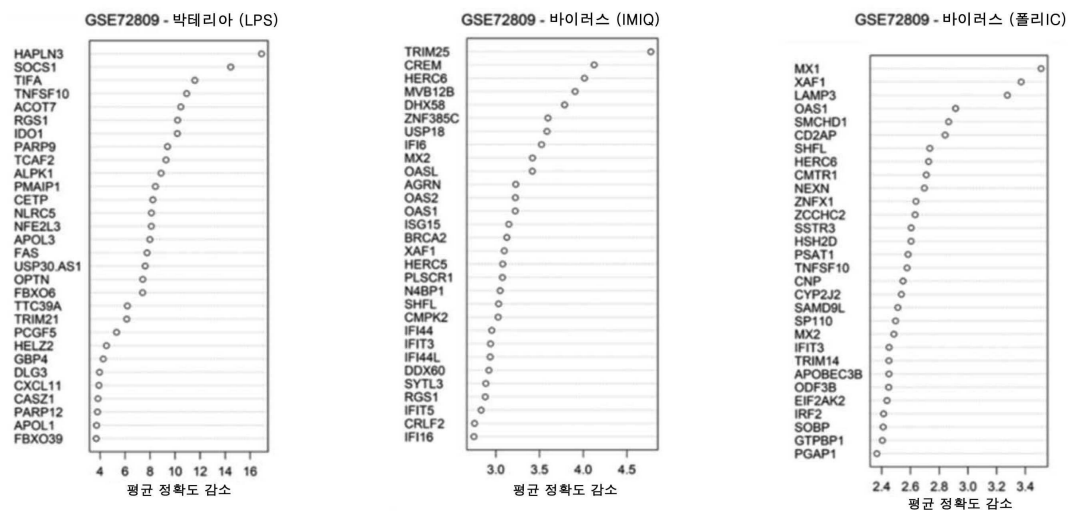
도면13a

A



도면13b

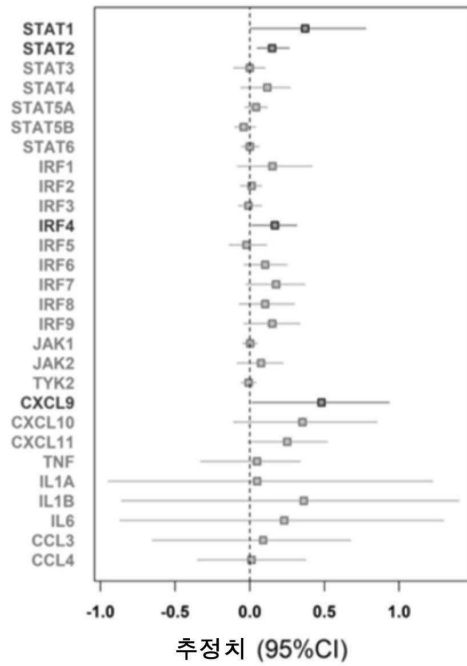
B



도면14a

**A**

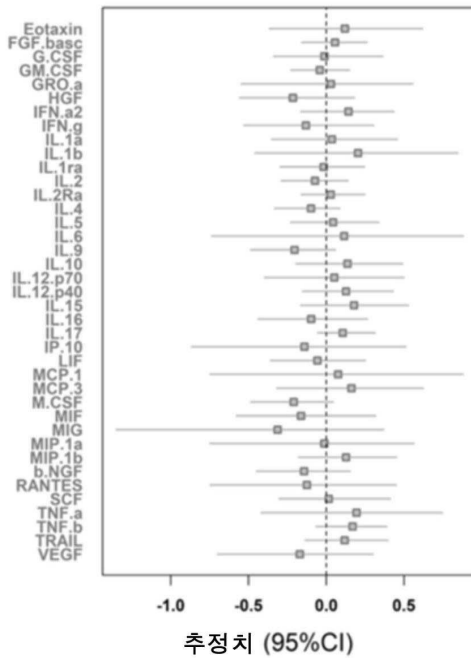
sLRly1+ 대 sLRly1-  
 맨-휘트니 U 테스트  
 (기준선 조정됨)



도면14b

**B**

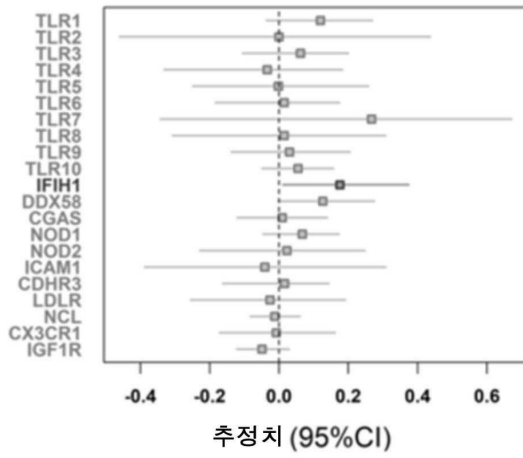
sLRly1+ 대 sLRly1-  
 맨-휘트니 U 테스트  
 (기준선 조정됨)



도면14c

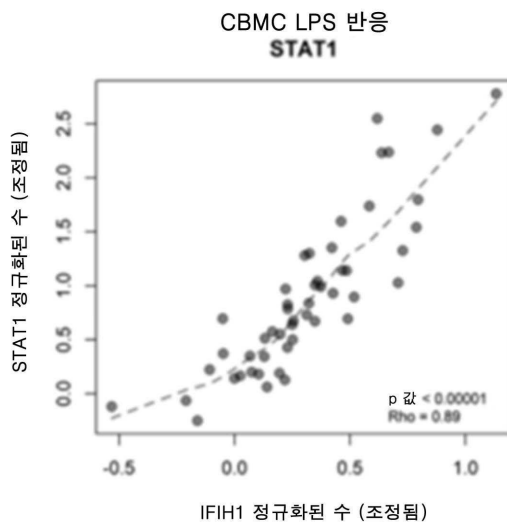
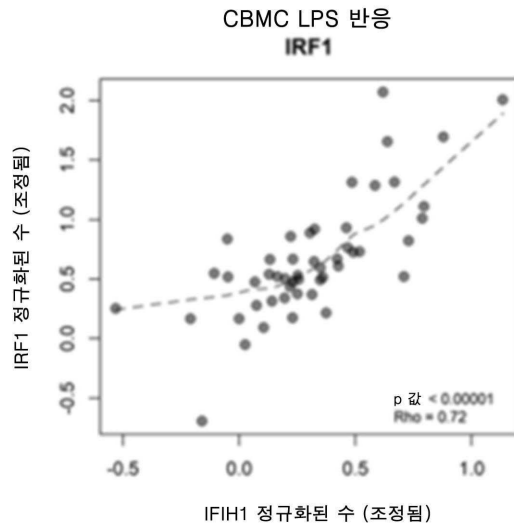
C

sLRly1+ 대 sLRly1-  
 맨-휘트니 U 테스트  
 (기준선 조정됨)



도면14d

D



도면15

