

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2022년 6월 2일 (02.06.2022)



(10) 국제공개번호
WO 2022/114887 A1

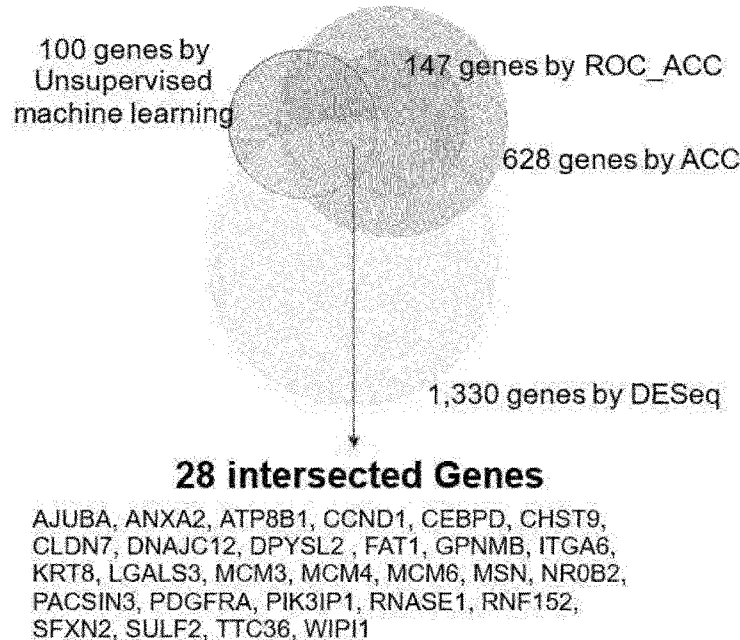
- (51) 국제특허분류:
C12Q 1/6883 (2018.01) *G01N 33/68* (2006.01)
- (21) 국제출원번호: PCT/KR2021/017751
- (22) 국제출원일: 2021년 11월 29일 (29.11.2021)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보:
10-2020-0162459 2020년 11월 27일 (27.11.2020) KR
10-2021-0167144 2021년 11월 29일 (29.11.2021) KR
- (71) 출원인: 국립암센터 (NATIONAL CANCER CENTER) [KR/KR]; 10408 경기도 고양시 일산동구 일산로 323 (마두동), Gyeonggi-do (KR). 동아대학교 산학협력단 (DONG-A UNIVERSITY RESEARCH FOUNDATION FOR INDUSTRY-ACADEMY COOPER-

ATION) [KR/KR]; 49315 부산시 사하구 낙동대로550번길 37, Busan (KR). 서울대학교산학협력단 (SEOUL NATIONAL UNIVERSITY R&DB FOUNDATION) [KR/KR]; 08826 서울시 관악구 관악로 1, Seoul (KR). 숙명여자대학교산학협력단 (SOOKMYUNG WOMEN'S UNIVERSITY INDUSTRY-ACADEMIC COOPERATION FOUNDATION) [KR/KR]; 04310 서울시 용산구 청파로47길 100, Seoul (KR). 의료법인 온그림의료재단 (ON HOSPITAL) [KR/KR]; 47261 부산시 부산진구 가야대로 721, Busan (KR).

- (72) 발명자: 이연수 (LEE, Yeon Su); 10085 경기도 김포시 김포한강3로 290-16, 802동 401호, Gyeonggi-do (KR). 이용선 (LEE, Yong Sun); 10416 경기도 고양시 일산동구 일산로 241, 103동 402호, Gyeonggi-do (KR). 백양현 (BAEK, Yang Hyun); 49201 부산시 서구 대신공원로 26, 동아대학교병원 간센터, Busan (KR). 성노현 (SEONG, Rho Hyun); 13589 경기도 성남시 분당구 중

(54) Title: COMPLEX MARKER FOR DIAGNOSING CHRONIC LIVER DISEASE BASED ON INTEGRATED TRANSCRIPTOME ANALYSIS AND USE THEREOF

(54) 발명의 명칭: 통합전사체분석 기반 만성 간질환 진단용 복합마커 및 이의 용도



(57) Abstract: The present invention relates to a complex marker for diagnosing chronic liver disease on the basis of integrated transcriptome analysis and use thereof and, more specifically, to: 28 complex markers selected through integrated transcriptome analysis and machine learning analysis on patients with steatosis and non-alcoholic steatohepatitis; and a method for providing information for diagnosing chronic liver disease by using same. In the present invention, 28 genes co-selected from among: genes selected through differentially expressed gene (DEG) analysis by performing integrated transcriptome analysis on patients with steatosis and NASH; and genes selected using a feature selection process, were finally selected, and it was confirmed that, when 28 complex markers or an ATPase phospholipid transporting 8B1 (ATP8B1) marker was used, the ability to diagnose chronic liver disease was significantly



WO 2022/114887 A1

양공원로 20, 426동 202호, Gyeonggi-do (KR). 유경현 (RYU, Kyung Hyun); 04354 서울시 용산구 새창로 70, 113동 605호, Seoul (KR). 박종훈 (PARK, Jong Hoon); 06289 서울시 강남구 선릉로 120, 10동 601호, Seoul (KR). 오수민 (OH, Su Min); 17589 경기도 안성시 중앙 3길 54, Gyeonggi-do (KR). 한상영 (HAN, Sang Young); 52543 경상남도 사천시 서포면 서포로 103, Gyeongsang-nam-do (KR).

(74) 대리인: 파도특허법인 유한회사 (PADO IP LAW PLLC); 06194 서울시 강남구 테헤란로78길 16, 9층 (대치동, 노벨빌딩), Seoul (KR).

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

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

공개:

- 국제조사보고서와 함께 (조약 제21조(3))
- 청구범위 보정 기한 만료 전의 공개이며, 보정서를 접수하는 경우 그에 관하여 별도 공개함 (규칙 48.2(h))
- 명세서의 서열목록 부분과 함께 (규칙 5.2(a))

improved. Thus, the complex marker of the present invention has been confirmed to be applicable as a biomarker for diagnosing chronic liver disease.

(57) 요약서: 본 발명은 통합전사체분석 기반 만성 간질환 진단용 복합마커 및 이의 용도에 관한 것으로, 보다 상세하게는 지방증, 비알코올성 지방간염 환자를 대상으로 통합전사체 분석 및 머신러닝 분석을 통해 선별한 28종 복합 마커 및 이를 이용한 만성 간질환 진단에 대한 정보를 제공하는 방법에 관한 것이다. 본 발명에서는 지방증, NASH 환자를 대상으로 통합전사체 분석을 수행하여 차별 발현 유전자(Differentially Expressed Gene:DEG) 분석을 통해 선별된 유전자 및 특징 선택(Feature selection) 프로세스를 이용하여 선별된 유전자 중에서 공동으로 선별된 28종의 유전자를 최종 선별하였으며, 28종의 복합마커 또는 ATP8B1(ATPase phospholipid transporting 8B1) 마커를 이용하는 경우 만성 간질환 진단능이 현저하게 향상된 것을 확인하였으므로, 본 발명의 복합마커는 만성 간질환 진단을 위한 바이오마커로 적용할 수 있는 것을 확인하였다.

명세서

발명의 명칭: 통합전사체분석 기반 만성 간질환 진단용 복합마커 및 이의 용도

기술분야

- [1] 본 발명은 통합전사체분석 기반 만성 간질환 진단용 복합마커 및 이의 용도에 관한 것으로, 보다 상세하게는 지방증, 비알코올성 지방간염 환자를 대상으로 통합전사체 분석 및 머신러닝 분석을 통해 선별한 28종 복합 마커 및 이를 이용한 만성 간질환 진단에 대한 정보를 제공하는 방법에 관한 것이다.

[2]

배경기술

- [3] 만성 간질환으로 인한 국내의 사회경제적 부담은 2010년 기준 약 3조 7000억 원으로 가장 심각한 질환 중 하나이다. 한국에서 발병빈도가 매우 높고 특히 40-50대의 경우 간암 및 간질환에 의한 사망률이 가장 높은 것으로 알려져 있다.
- [4] 만성 간질환 중 하나인 비-알코올성 지방간질환(non-alcoholic fatty liver disease; NAFLD)은 단순한 지방증으로부터 비-알코올성 지방간염((non-alcoholic steatohepatitis; NASH)에 이르는 범위의 진행성 간질환이다. 특히, 비-알코올성 지방간염(NASH)은 조직학적으로 알코올성 간염과 비슷한 지방산 축적, 간세포 손상 및 염증에 의해 특징화되는 간의 진행성 질환으로, 간 지방증으로부터 간경변증 및 간부전으로 확장되는 공정의 주요 단계이며, 최근 NASH의 발병이 최근 여러 해 동안 증가하고 있으며, NASH로 진행되는 환자는 간-관련된 유병률 및 사망률이 증가하고 있는 추세이다.
- [5] 이런 NASH를 포함하는 만성간질환의 활성, 단계 또는 중증도를 진단하기 위한 표준 방법으로 간생검(liver biopsy) 표본의 조직학적 검사를 사용하고 있으나, 간생검은 침습적 방법이라는 단점을 갖고 있다. 또한, 계속해서 증가하는 간 질환 환자들을 대상으로 모두 조직검사를 수행하는 것은 여러 가지 한계가 있고, 간생검은 통증, 출혈 및 아주 드문 경우 사망에 이를 수도 있는 부작용이 있다 (Rana L Smalling *et al.*, *Am J Physiol Gastrointest Liver Physiol.*, 305(5):G364-74, 2013; 대한민국공개특허 제 10-2020-0051676호).
- [6] 따라서, 침습적인 간생검을 대체할 수 있는 신뢰할만한 진단방법의 개발이 요구되고 있으며, 이를 위해 다양한 간질환 진단용 바이오마커가 발굴 및 연구되고 있다
- [7] 이에, 본 발명에서는 보다 효과적인 만성 간질환 진단용 바이오마커를 개발하기 위해 노력한 결과, 지방증과 NASH 환자 그룹을 대상으로 통합전사체 분석을 수행하여 선별된 유전자 및 다양한 분석방법을 통한 특징 선택(Feature selection) 프로세스를 이용하여 선별된 유전자 중에서 공동으로 선별된 28종의 유전자를 최종 선별하였으며, 28종의 복합마커 또는 ATP8B1 마커를 이용하는

경우 만성 간질환 진단능이 현저하게 향상된 것을 확인하고, 본 발명을 완성하였다.

[8]

발명의 상세한 설명

기술적 과제

[9] 본 발명의 목적은 28종의 마커를 포함하는 만성 간질환 진단용 복합마커 조성물 및 복합마커의 mRNA 또는 단백질 수준을 측정하는 제제를 포함하는 만성간질환 진단용 조성물을 제공하는 데 있다.

[10] 본 발명의 다른 목적은 상기 만성 간질환 진단용 조성물을 포함하는 만성 간질환 진단용 키트를 제공하는 데 있다.

[11] 본 발명의 또 다른 목적은 상기 만성 간질환 진단용 조성물을 이용한 만성 간질환 진단을 위한 정보 제공방법을 제공하는 데 있다.

[12]

과제 해결 수단

[13] 상기 목적을 달성하기 위해,

[14] 본 발명은 AJUBA(ajuba LIM protein), ANXA2(annexin A2), ATP8B1(ATPase phospholipid transporting 8B1), CCND1(cyclin D1), CEBPD(CCAAT enhancer binding protein), CHST9(carbohydrate sulfotransferase 9), CLDN7(claudin 7), DNAJC12(DnaJ heat shock protein family (Hsp40) member C12), DPYSL2(dihydropyrimidinase like 2), FAT1(FAT atypical cadherin 1), GPNMB(glycoprotein nmb), ITGA6(integrin subunit alpha 6), KRT8(keratin 8), LGALS3(galectin 3), MCM3(minichromosome maintenance complex component 3), MCM4(minichromosome maintenance complex component 4), MCM6(minichromosome maintenance complex component 6), MSN(moesin), NR0B2(nuclear receptor subfamily 0 group B member 2), PACSIN3(protein kinase C and casein kinase substrate in neurons 3), PDGFRA(platelet derived growth factor receptor alpha), PIK3IP1(phosphoinositide-3-kinase interacting protein 1), RNASE1(ribonuclease A family member 1, pancreatic), RNF152(ring finger protein 152), SFXN2(sideroflexin 2), SULF2(sulfatase 2), TTC36(tetratricopeptide repeat domain 36) 및 WIPI1(WD repeat domain, phosphoinositide interacting 1)을 포함하는 복합마커; 또는

[15] ATP8B1(ATPase phospholipid transporting 8B1) 마커를 포함하는 만성 간질환 진단 마커 조성물을 제공한다.

[16]

[17] 또한, 본 발명은 상기 만성 간질환 진단 마커 조성물의 mRNA 또는 단백질 수준을 측정하는 제제를 포함하는 만성 간질환 진단용 조성물을 제공한다.

[18] 본 발명의 바람직한 일실시에에 있어서, 상기 만성 간질환은 지방증,

- 비-알코올성 지방간질환(non-alcoholic fatty liver disease; NAFLD), 비-알코올성 지방간염((non-alcoholic steatohepatitis; NASH) 또는 간경화증일 수 있다.
- [19] 본 발명의 바람직한 다른 일실시예에 있어서, 상기 mRNA 수준을 측정하는 제제는 마커에 특이적으로 결합하는 프라이머쌍, 프로브 또는 안티센스 뉴클레오타이드일 수 있다.
- [20] 본 발명의 바람직한 또 다른 일실시예에 있어서, 상기 단백질 수준을 측정하는 제제는 마커 또는 마커의 펩타이드 단편에 특이적으로 결합하는 항체, 상호작용 단백질, 리간드, 나노입자(nanoparticles) 또는 압타머(aptamer)를 포함할 수 있다.
- [21]
- [22] 또한, 본 발명은 상기 만성 간질환 진단용 조성물을 포함하는 만성 간질환 진단용 키트를 제공한다.
- [23] 본 발명의 바람직한 또 다른 일실시예에 있어서, 상기 키트는 RT-PCR(Reverse transcription polymerase chain reaction) 키트, DNA 칩 키트, ELISA(enzyme linked immunosorbent assay) 키트, 단백질 칩 키트, 래피드(rapid) 키트 또는 MRM(Multiple reaction monitoring) 키트일 수 있다.
- [24]
- [25] 또한, 본 발명은 (a) 환자의 생물학적 시료로부터 ATP8B1 마커 또는 상기 복합마커의 mRNA 또는 단백질 발현 수준을 측정하는 단계; 및
- [26] (b) 상기 mRNA 또는 단백질 발현 수준을 대조군 시료와 비교하는 단계를 포함하는 만성 간질환 진단을 위한 정보 제공방법을 제공한다.
- [27] 본 발명의 바람직한 일실시예에 있어서, 상기 만성 간질환 진단을 위한 정보 제공방법은 CEBPD(CCAAT enhancer binding protein), DNAJC12(DnaJ heat shock protein family (Hsp40) member C12), NR0B2(nuclear receptor subfamily 0 group B member 2), PACSIN3(protein kinase C and casein kinase substrate in neurons 3), RNF152(ring finger protein 152) 및 SFXN2(sideroflexin 2), TTC36(tetratricopeptide repeat domain 36) 마커의 mRNA 또는 이의 단백질 발현 수준이 대조군에 비해 감소하고,
- [28] AJUBA(ajuba LIM protein), ANXA2(annexin A2), ATP8B1(ATPase phospholipid transporting 8B1), CCND1(cyclin D1), CHST9(carbohydrate sulfotransferase 9), CLDN7(claudin 7), DPYSL2(dihydropyrimidinase like 2), FAT1(FAT atypical cadherin 1), GPNMB(glycoprotein nmb), ITGA6(integrin subunit alpha 6), KRT8(keratin 8), LGALS3(galectin 3), MCM3(minichromosome maintenance complex component 3), MCM4(minichromosome maintenance complex component 4), MCM6(minichromosome maintenance complex component 6), MSN(moesin), PDGFRA(platelet derived growth factor receptor alpha), PIK3IP1(phosphoinositide-3-kinase interacting protein 1), RNASE1(ribonuclease A family member 1, pancreatic), SULF2(sulfatase 2), 및 WIPI1(WD repeat domain, phosphoinositide interacting 1) 마커의 mRNA 또는 이의 단백질 발현 수준이

대조군에 비해 증가하면, 만성 간질환인 것으로 정보를 제공하는 단계를 추가로 포함할 수 있다.

[29] 본 발명의 바람직한 또 다른 일실시예에 있어서, 상기 만성 간질환 진단을 위한 정보 제공방법은 ATP8B1(ATPase phospholipid transporting 8B1) 마커의 mRNA 또는 이의 단백질 발현 수준이 대조군에 비해 증가하면, 만성 간질환인 것으로 정보를 제공하는 단계를 추가로 포함할 수 있다.

[30] 본 발명의 바람직한 또 다른 일실시예에 있어서, 상기 mRNA 발현 정도 측정은 역전사효소 중합효소반응, 경쟁적 역전사효소 중합효소반응, 실시간 역전사효소 중합효소반응, RNase 보호 분석법, 노던 블랏팅 또는 DNA 칩을 이용하여 수행할 수 있다.

[31] 본 발명의 바람직한 또 다른 일실시예에 있어서, 상기 단백질 발현 정도 측정은 단백질 칩 분석, 면역측정법, 리간드 바인딩 어세이, MALDI-TOF(Matrix Desorption/Ionization Time of Flight Mass Spectrometry)분석, SELDI-TOF(Surface Enhanced Laser Desorption/Ionization Time of Flight Mass Spectrometry)분석, 방사선 면역분석, 방사 면역 확산법, 오우크테로니 면역 확산법, 로케트 면역전기영동, 조직면역 염색, 보체 고정 분석법, 2차원 전기영동 분석, 액상 크로마토그래피-질량분석(liquid chromatography-Mass Spectrometry, LC-MS), LC-MS/MS(liquid chromatography-Mass Spectrometry/ Mass Spectrometry), 웨스턴 블랏 및 ELISA(enzyme linked immunosorbent assay)를 이용하여 수행할 수 있다.

[32]

발명의 효과

[33] 본 발명에서는 지방증, NASH 환자를 대상으로 통합전사체 분석을 수행하여 차별 발현 유전자(Differentially Expressed Gene:DEG) 분석을 통해 선별된 유전자 및 특징 선택(Feature selection) 프로세스를 이용하여 선별된 유전자 중에서 공동으로 선별된 28종의 유전자를 최종 선별하였으며, 28종의 복합마커 또는 ATP8B1(ATPase phospholipid transporting 8B1) 마커를 이용하는 경우 만성 간질환 진단능이 현저하게 향상된 것을 확인하였으므로, 본 발명의 복합마커는 만성 간질환 진단을 위한 바이오마커로 적용할 수 있는 것을 확인하였다.

[34]

도면의 간단한 설명

[35] 도 1은 만성 간질환 진단을 위한 복합마커 선별 과정을 나타낸 모식도이다.

[36] 도 2는 통합전사체 분석 후, 차별 발현 유전자(Differentially Expressed Gene:DEG) 분석을 수행한 것으로, 지방증과 NASH 환자 그룹의 유전자 발현패턴을 분석한 히트맵 및 유전자 선별과정을 나타낸 도면이다.

[37] 도 3은 마커 유전자 선별을 위한 분석 과정을 나타낸 모식도이다.

[38] 도 4는 모델링에 따라 선정된 특징 사례 및 특징 사례 선별 과정을 나타낸 도면이다.

- [39] 도 5는 차별 발현 유전자(DEG) 분석을 통해 선별된 1330개 유전자 및 특징사례 분석을 통해 선별된 유전자들 중에서 공통으로 발현된 유전자를 선별 과정을 나타낸 모식도로, 총 28종의 유전자를 선별하였다.
- [40] 도 6은 복합마커로 선별된 28종 유전자의 간질환 단계별 발현 패턴을 나타낸 도면이다.
- [41] 도 7은 본 발명의 28종 복합마커를 이용하여 만성 간질환을 진단하였을 때, 진단의 정확성을 ROC 곡선을 통해 나타낸 도면이다.

[42]

발명의 실시를 위한 최선의 형태

- [43] 이하, 본 발명을 상세하게 설명한다.

[44]

- [45] 본 발명은 일관점에서, AJUBA(ajuba LIM protein; NCBI Gene ID: 84962), ANXA2(annexin A2; NCBI Gene ID: 302), ATP8B1(ATPase phospholipid transporting 8B1; NCBI Gene ID: 5205), CCND1(cyclin D1; NCBI Gene ID: 595), CEBPD(CCAAT enhancer binding protein; NCBI Gene ID: 1052), CHST9(carbohydrate sulfotransferase 9; NCBI Gene ID: 83539), CLDN7(claudin 7; NCBI Gene ID: 1366), DNAJC12(DnaJ heat shock protein family (Hsp40) member C12; NCBI Gene ID: 56521), DPYSL2(dihydropyrimidinase like 2; NCBI Gene ID: 1808), FAT1(FAT atypical cadherin 1; NCBI Gene ID: 2195), GPNMB(glycoprotein nmb; NCBI Gene ID: 10457), ITGA6(integrin subunit alpha 6; NCBI Gene ID: 3655), KRT8(keratin 8; NCBI Gene ID: 3856), LGALS3(galectin 3; NCBI Gene ID: 3958), MCM3(minichromosome maintenance complex component 3; NCBI Gene ID: 4172), MCM4(minichromosome maintenance complex component 4; NCBI Gene ID: 4173), MCM6(minichromosome maintenance complex component 6; NCBI Gene ID: 4175), MSN(moesin; NCBI Gene ID: 4478), NR0B2(nuclear receptor subfamily 0 group B member 2; NCBI Gene ID: 8431), PACSIN3(protein kinase C and casein kinase substrate in neurons 3; NCBI Gene ID: 29763), PDGFRA(platelet derived growth factor receptor alpha; NCBI Gene ID: 5156), PIK3IP1(phosphoinositide-3-kinase interacting protein 1; NCBI Gene ID: 113791), RNASE1(ribonuclease A family member 1, pancreatic; NCBI Gene ID: 6035), RNF152(ring finger protein 152; NCBI Gene ID: 220441), SFXN2(sideroflexin 2; NCBI Gene ID: 118980), SULF2(sulfatase 2; NCBI Gene ID: 55959), TTC36(tetratricopeptide repeat domain 36; NCBI Gene ID: 143941) 및 WIPI1(WD repeat domain, phosphoinositide interacting 1; NCBI Gene ID: 55062)을 포함하는 복합마커; 또는

- [46] ATP8B1(ATPase phospholipid transporting 8B1) 마커를 포함하는 만성 간질환 진단 마커 조성물을 제공한다.

[47]

- [48] 본 발명에서 사용된 용어 "진단"은 병리 상태의 존재 또는 특징을 확인하는 것을 의미한다. 본 발명의 목적상, 진단은 만성 간질환, 특히, 지방증(Steatosis), 비-알코올성 지방간질환(non-alcoholic fatty liver disease; NAFLD), 비-알코올성 지방간염((non-alcoholic steatohepatitis; NASH) 또는 간경화증 여부를 확인하는 것이다.
- [49] 본 발명에서 사용된 용어 "진단용 바이오마커"란 정상 대조군에 비해 만성 간질환이 진행된 개체에 비해, 특정 유전자 발현 수준 또는 단백질 발현 수준의 유의적인 증가 또는 감소 양상을 보이는 폴리펩티드 또는 핵산(예: mRNA 등), 지질, 당지질, 당단백질, 당(단당류, 이당류, 올리고당류 등) 등과 같은 유기 생체 분자 등을 포함하며, 바람직하게는 ATP8B1 마커 또는 상기 28종의 복합마커이다.
- [50]
- [51] 본 발명의 구체적인 일구현예에서, 도 1에 나타난 모식도와 같은 방법으로 만성 간질환 진단을 위한 마커를 선별하였다. 지방증(Steatosis), 비-알코올성 지방간염인 NASH 환자를 대상으로 통합전사제 분석을 수행하여 차별 발현 유전자(Differentially Expressed Gene:DEG) 분석을 통해 선별된 유전자(도 2)와 생명정보학 분석의 특징 선택(Feature selection) 프로세스를 이용하여 선별된 유전자(도 3 및 도 4) 중에서 공통으로 나온 28종의 유전자를 최종 선별하였다(도 5).
- [52] 또한, 선별된 28종의 유전자를 지방증과 NASH 단계별로 나누어 유전자 발현 패턴을 확인한 결과(도 6), 각 단계별로 서로 상이한 패턴을 보이는 것으로 확인되었으므로, 이들의 조합을 통해 만성 간질환 각 단계별 진단이 가능하다고 판단되었다.
- [53] 본 발명의 구체적인 다른 일구현예에서, 본 발명에서 선별한 28종의 복합마커를 이용하여 만성 간질환에 대한 진단능을 평가한 결과, 도 7에 나타난 바와 같이, AUC 값은 모두 1로 통계적으로 매우 유의한 만성 간질환 진단능을 보이는 것을 확인하였다.
- [54] ROC 곡선은 진단법의 정확성을 비교하는 방법으로 널리 사용되는 방법 중 하나이다 (Akobeong, 2007). X축에는 '민감도', Y축에는 '1-특이도'로 하여 그래프를 그리고, 이때 대각선을 기준으로 곡선 아래 면적이 0.5이며, 대각선보다 ROC 곡선이 위에 있을 경우 좋은 모델이라 예측하게 된다. AUC는 0.5에서 1 사이 값을 갖게 되며, 1에 가까워질수록 곡선도 넓어지고 예측한 모델이 좋다는 것을 의미한다.
- [55] 또한, 선별한 28종의 복합마커 중에서 일부에 대해, qPCR을 이용하여 검증(validation)을 수행한 결과, ATP8B1, CEBPD, GPNMB, KRT8 및 RNASE1 마커를 포함하는 일부 마커들이 통계적으로 유의한 결과를 보여주는 것을 확인하였다(표 2). 즉, 본 발명의 만성 간질환 진단을 위한 ATP8B1 마커 또는 복합마커는, 만성 간질환 진단에 가장 적합한 마커인 것을 확인하였다.

[56]

[57] 따라서, 본 발명은 다른 관점에서, 상기 ATP8B1 마커 또는 복합마커의 mRNA 또는 단백질 수준을 측정하는 제제를 포함하는 만성 간질환 진단용 조성물에 관한 것이다.

[58] 본 발명에 있어서, 상기 만성 간질환은 지방증, 비-알코올성 지방간질환(non-alcoholic fatty liver disease; NAFLD), 비-알코올성 지방간염(non-alcoholic steatohepatitis; NASH) 또는 간경화증일 수 있다.

[59] 본 발명에 사용된 용어 "mRNA 발현수준 측정"이란 생물학적 시료에서 만성 간질환 진단을 위한 바이오마커의 mRNA 존재 여부와 발현 정도를 확인하는 과정으로 mRNA의 양을 측정한다. 이를 위한 분석 방법으로는 역전사 중합효소반응(RT-PCR), 경쟁적 역전사 중합효소반응(Competitive RT-PCR), 실시간 역전사 중합효소반응(Real-time RT-PCR), RNase 보호 분석법(RNAse protection assay), 노던 블랏팅(Northern blotting), DNA 칩 등이 있으나 이로 제한되는 것은 아니다.

[60] 상기 ATP8B1 마커 또는 복합마커의 mRNA 수준을 측정하는 제제는 상기 마커의 유전자에 특이적으로 결합하는 프라이머쌍, 프로브 또는 안티센스 뉴클레오타이드인 것을 특징으로 하며, 상기 유전자들의 핵산 정보가 GeneBank 등에 알려져 있으므로 당업자는 상기 서열을 바탕으로 이들 프라이머쌍, 프로브 또는 안티센스 뉴클레오타이드를 디자인할 수 있다.

[61] 본 발명에서 사용된 용어 "프라이머"는 표적 유전자 서열을 인지하는 단편으로서, 정방향 및 역방향의 프라이머 쌍을 포함하나, 바람직하게는, 특이성 및 민감성을 가지는 분석 결과를 제공하는 프라이머 쌍이다.

[62] 본 발명에서 사용된 용어 "프로브"란 시료 내의 검출하고자 하는 표적 물질과 특이적으로 결합할 수 있는 물질을 의미하며, 상기 결합을 통하여 특이적으로 시료 내의 표적 물질의 존재를 확인할 수 있는 물질을 의미한다. 프로브의 종류는 당업계에서 통상적으로 사용되는 물질로서 제한은 없으나, 바람직하게는 PNA(peptide nucleic acid), LNA(locked nucleic acid), 펩타이드, 폴리펩타이드, 단백질, RNA 또는 DNA 일 수 있으며, 가장 바람직하게는 PNA이다.

[63] 본 발명에서 사용된 용어 "안티센스"는 안티센스 올리고머가 왓슨-크릭 염기쌍 형성에 의해 RNA 내의 표적 서열과 혼성화되어, 표적서열 내에서 전형적으로 mRNA와 RNA:올리고머 헤테로이중체의 형성을 허용하는 뉴클레오타이드 염기의 서열 및 서브유닛간 백본을 갖는 올리고머를 의미한다. 올리고머는 표적 서열에 대한 정확한 서열 상보성 또는 근사 상보성을 가질 수 있다.

[64]

[65] 본 발명에 사용된 용어 "단백질 발현수준 측정"이란 생물학적 시료에서 만성 간질환 진단을 위한 바이오마커로부터 발현된 단백질의 존재 여부와 발현 정도를 확인하는 과정이다. 상기 유전자의 단백질 또는 펩타이드 단편에

특이적으로 결합하는 항체, 상호작용 단백질, 리간드, 나노입자(nanoparticles) 또는 aptamer(aptamer)를 이용하여 단백질의 양을 확인할 수 있다.

- [66] 상기 단백질 발현 수준 측정 또는 비교 분석 방법으로는 단백질 칩 분석, 면역측정법, 리간드 바인딩 어세이, MALDI-TOF(Matrix Desorption/Ionization Time of Flight Mass Spectrometry)분석, SELDI-TOF(Surface Enhanced Laser Desorption/Ionization Time of Flight Mass Spectrometry)분석, 방사선 면역분석, 방사 면역 확산법, 오우크테로니 면역 확산법, 로케트 면역전기영동, 조직면역 염색, 보체 고정 분석법, 2차원 전기영동 분석, 액상 크로마토그래피-질량분석(liquid chromatography-Mass Spectrometry, LC-MS), LC-MS/MS(liquid chromatography-Mass Spectrometry/ Mass Spectrometry), 웨스턴 블랏 및 ELISA(enzyme linked immunosorbent assay)등이 있으나 이로 제한되는 것은 아니다.
- [67] 본 발명에 사용된 용어 "항체"는 항원과 특이적으로 결합하여 항원-항체 반응을 일으키는 물질을 가리킨다. 본 발명의 목적상, 항체는 본 발명의 바이오마커에 대해 특이적으로 결합하는 항체를 의미한다. 본 발명의 항체는 다클론 항체, 단클론 항체 및 재조합 항체를 모두 포함한다.
- [68]
- [69] 본 발명은 또 다른 관점에서, 상기 만성 간질환 진단용 조성물을 포함하는 만성 간질환 진단용 키트에 관한 것이다.
- [70] 상기 키트는 당업계에 알려져 있는 통상의 제조방법에 의해 제조될 수 있다. 상기 키트는 예를 들면, 동결 건조 형태의 항체와 완충액, 안정화제, 불활성 단백질 등을 포함할 수 있다.
- [71] 상기 키트는 검출 가능한 표지를 더 포함할 수 있다. 용어 "검출 가능한 표지"는 표지가 없는 동일한 종류의 분자들 중에서 표지를 포함하는 분자를 특이적으로 검출하도록 하는 원자 또는 분자를 의미한다. 상기 검출 가능한 표지는 상기 단백질 또는 그의 단편에 특이적으로 결합하는 항체, 상호작용 단백질, 리간드, 나노입자, 또는 aptamer에 부착된 것일 수 있다. 상기 검출 가능한 표지는 방사종(radionuclide), 형광원(fluorophore), 효소(enzyme)를 포함할 수 있다.
- [72] 상기 키트는 당업계에 알려진 다양한 면역분석법 또는 면역염색법에 따라 이용될 수 있다. 상기 면역분석법 또는 면역염색법은 방사능면역분석, 방사능면역침전, 면역침전, ELISA, 캡처-ELISA, 억제 또는 경쟁 분석, 샌드위치 분석, 유세포 분석, 면역형광염색 및 면역친화성 정제를 포함할 수 있다. 바람직하게, 상기 키트는 RT-PCR(reverse transcription polymerase chain reaction) 키트, DNA 칩 키트, ELISA(enzyme linked immunosorbent assay) 키트, 단백질 칩 키트, 래피드(rapid) 키트 또는 MRM(multiple reaction monitoring)인 것일 수 있다.
- [73]
- [74] 본 발명은 또 다른 관점에서, (a) 환자의 생물학적 시료로부터 상기 ATP8B1 마커 또는 복합마커의 mRNA 또는 단백질 발현 수준을 측정하는 단계; 및

- [75] (b) 상기 mRNA 또는 단백질 발현 수준을 대조군 시료와 비교하는 단계를 포함하는 만성 간질환 진단을 위한 정보 제공방법에 관한 것이다.
- [76] 상기 방법에서 "생물학적 시료(biological sample)"란 조직, 세포, 혈액, 혈청, 혈장, 타액, 뇌척수액 또는 뇨와 같은 시료 등을 의미하며, 바람직하게는 혈액, 혈장, 혈청을 의미한다.
- [77] 상기 만성 간질환 진단에 대한 정보 제공방법에서 ATP8B1 마커 또는 복합마커의 mRNA 또는 이의 단백질의 발현 수준 측정 방법은 상술한 바와 같다.
- [78] 본 발명의 복합마커를 이용하여 만성 간질환 진단을 위한 정보를 제공하는 경우, 상기 만성 간질환 진단을 위한 정보 제공방법은 CEBPD(CCAAT enhancer binding protein), DNAJC12(DnaJ heat shock protein family (Hsp40) member C12), NR0B2(nuclear receptor subfamily 0 group B member 2), PACSIN3(protein kinase C and casein kinase substrate in neurons 3), RNF152(ring finger protein 152) 및 SFXN2(sideroflexin 2), TTC36(tetratricopeptide repeat domain 36) 마커의 mRNA 또는 이의 단백질 발현 수준이 대조군에 비해 감소하고,
- [79] AJUBA(ajuba LIM protein), ANXA2(annexin A2), ATP8B1(ATPase phospholipid transporting 8B1), CCND1(cyclin D1), CHST9(carbohydrate sulfotransferase 9), CLDN7(claudin 7), DPYSL2(dihydropyrimidinase like 2), FAT1(FAT atypical cadherin 1), GPNMB(glycoprotein nmb), ITGA6(integrin subunit alpha 6), KRT8(keratin 8), LGALS3(galectin 3), MCM3(minichromosome maintenance complex component 3), MCM4(minichromosome maintenance complex component 4), MCM6(minichromosome maintenance complex component 6), MSN(moesin), PDGFRA(platelet derived growth factor receptor alpha), PIK3IP1(phosphoinositide-3-kinase interacting protein 1), RNASE1(ribonuclease A family member 1, pancreatic), SULF2(sulfatase 2), 및 WIPI1(WD repeat domain, phosphoinositide interacting 1) 마커의 mRNA 또는 이의 단백질 발현 수준이 대조군에 비해 증가하면, 만성 간질환인 것으로 정보를 제공하는 단계를 추가로 포함할 수 있다.
- [80] 본 발명의 ATP8B1 마커를 이용하여 만성 간질환 진단을 위한 정보를 제공하는 경우, 상기 만성 간질환 진단을 위한 정보 제공방법은 ATP8B1(ATPase phospholipid transporting 8B1) 마커의 mRNA 또는 이의 단백질 발현 수준이 대조군에 비해 증가하면, 만성 간질환인 것으로 정보를 제공하는 단계를 추가로 포함할 수 있다.
- [81]
- [82] 이하, 실시예를 통하여 본 발명을 더욱 상세히 설명하고자 한다. 이들 실시예는 오로지 본 발명을 예시하기 위한 것으로서, 본 발명의 범위가 이들 실시예에 의해 제한되는 것으로 해석되지 않는 것은 당업계에서 통상의 지식을 가진 자에게 있어서 자명할 것이다.
- [83]

[84] 실시예 1 : 환자 선정

[85] 본 발명에서는 만성 간질환 진단을 위한 마커를 선별하기 위해, 지방증(Steatosis) 환자 42명, 비-알코올성 지방간염인 NASH 환자 56명을 선별하여 Biopsy 또는 수술 후 시료로부터 환자의 조직을 채취하였다.

[86] [표1]

환자 임상 정보

	Steatosis	Nash
성별 (n)		
여성	15	25
남성	26	31
나이 (연)		
평균	46.6	48.1
Median	48.0	48.5
BMI		
평균	28.8	29.4
Median	29.3	28.9
AST (IU/L)	30.0	59.1
ALT (IU/L)	38.0	72.4

[87]

[88] 실시예 2 : 차별 발현 유전자(Differentially Expressed Gene:DEG) 분석

[89] 실시예 1에서 수득한 환자의 조직을 이용하여 트리아졸(Trizol)로 RNA를 분리/정제한 후 나노드랍(nanodrop)과 피코그린(picogreen)을 이용하여 정량분석하고 애질런트(Agilent)사의 Bioanalyzer를 이용하여 정성분석을 수행했다. 이후 실험에 적합한 양과 품질의 샘플로 일루미나(Illumina)사의 TruSeq Stranded Total RNA LT Sample Prep Kit (Gold)를 사용하여 라이브러리를 구축하고 Novaseq 6000을 사용하여 염기서열분석(sequencing)을 수행하였다. 이로부터 도출된 FASTQ 파일로 다음과 같이 생명정보분석을 수행하였다.

[90] 총 RNA 서열분석 로우 데이터(total RNA-seq raw data(.fastq))의 품질을 확인하고 이를 향상시키기 위해 Trim Galore를 이용하여 결과를 획득하였다. 이 과정에서 내부적으로 FastQC로 리드(read)의 품질을 확인하고 컷어댑트(Cutadapt)로 낮은 품질의 리드(read)나 남은 어댑터 시퀀스(adapter

seq)를 제거하였다. 이후 리드(read)를 레퍼런스(reference)인 휴먼 지놈(genome)에 얼라인먼트(alignment) 하기 위해 STAR alignment tool을 이용하였다. 추후 전사체 발현량 분석을 통한 예후 예측 마커 발굴의 정확도를 높이기 위해 레퍼런스(reference)에 얼라인먼트(alignment)된 결과 파일(.bam)은 Picard mark duplication tool을 이용하여 중복(duplication)이 제거된 결과 파일(.bam)로 획득하였다. 이후, 질병의 진행 단계별 유전자 발현량을 계산하기 위해, 환자들을 지방증(Steatosis)과 NASH로 그룹핑하고 커프티프(cuffdiff)를 진행하여 유전자 수준에서 그룹별, 시료별 발현 수준을 비교 분석할 수 있도록 정규화된(normalized) 결과를 획득하였다.

- [91] 또한, 통합전사체 분석을 통해 수집된 데이터를 도 2에 나타난 바와 같이, DEG (Differentially expressed gene)의 발현 차이에 따라 히트맵으로 시각화하였으며, (1) 발현 레벨(Expression Level (FPKM)) 컷오프 > 1, (2) p-value < 0.05, (3) Log2 Fold Change > 11.31 및 (4) Standard Deviation < Mean 조건에 맞는 유전자 1330 개를 선별하였다.

[92]

[93] **실시예 3 : 특징 선택 (Feature selection)에 의한 유전자 선별**

- [94] 유전자 선별을 위해서는 도 4에 기재된 특징 선택(Feature selection) 프로세스와 특징중요도(Feature importance) 방식의 분석을 수행하였으며, 도 3에 기재된 워크플로우(Workflow)와 같이 진행하였다.

- [95] 특징 선택(Feature selection) 방식에서는 파이썬(python)으로 개발된 selectKBest 방법을 이용하여 반복적으로 유전자를 제거하면서 선형 지지 벡터 머신(linear Support Vector Machine) 모델에 의해 잘 분류되는 유전자 세트를 선택하였다 (참조 : <https://scikit-learn.org/>). 유의미한 유전자 그룹 평가에 있어서 AUC(area under the curve)와 ROC(Receiver operating characteristic) ACC(accuracy) 분석을 수행한 바, AUC의 경우에는 정확도(accuracy) 방식으로 계산한 후 ANOVA를 이용하여 통계적으로 두 그룹의 차이가 있는 지를 확인하고 이를 통해 628개의 유전자를 선별하였고, ROC ACC의 경우에는 ROC AUC 방식으로 계산한 후 ANOVA(Analysis of variance)를 이용하여 통계적으로 두 그룹의 차이가 있는 지 확인하고 147개의 적합한 유전자를 선별하였다.

- [96] 또한 특징중요도(Feature importance) 방식에서는, 파이썬으로 개발된 lightGBM 방법(unsupervised machine learning)을 이용하여 회귀 모델(regression model) 기법으로 계산하였을 때, 두 그룹을 분류하는데 있어 가장 잘 분류된다고 계산되는 순서대로 100개의 유전자를 선별하였다 (참조 : <https://lightgbm.readthedocs.io/en/latest/>).

[97]

[98] **실시예 4 : 복합마커 선별 및 만성 간질환 단계별 발현 패턴 분석**

- [99] 본 발명에서는 도 5에 나타난 모식도와 같이, 차별 발현 유전자(Differentially Expressed Gene:DEG) 분석을 통해 선별된 유전자와 특징 선택(Feature selection)

프로세스를 이용하여 선별된 유전자 중에서 공통으로 선별된 28종의 유전자를 최종 선별하였다.

[100] 또한, 선별된 28종의 유전자를 지방증(Steatosis) 단계와 NASH 단계로 나누어 유전자 발현 패턴을 확인한 결과, 도 6 및 표 3에 나타난 바와 같이, 각 단계별로 서로 상이한 패턴을 보이는 것으로 확인되었으므로, 이들의 조합을 통해 만성 간질환 각 단계별 진단이 가능한 것을 확인하였다.

[101]

[102] [표2]

복합마커로 선별된 28종 유전자의 간질환 단계별 발현 패턴

	steatosis_ave	NASH_ave	Fold change	P-value
AJUBA	11.737	19.982	1.702	5.5004E-09
ANXA2	30.265	61.127	2.020	2.5006E-07
ATP8B1	9.454	13.147	1.391	7.7905E-09
CCND1	34.658	51.279	1.480	2.4338E-06
CEBPD	37.628	24.979	0.664	3.6391E-06
CHST9	5.428	10.294	1.896	8.1691E-08
CLDN7	4.535	6.403	1.412	2.1963E-08
DNAJC12	25.228	18.762	0.744	5.1324E-06
DPYSL2	16.829	23.891	1.420	9.1687E-08
FAT1	13.149	26.209	1.993	2.693E-11
GPNMB	13.640	30.623	2.245	1.3566E-09
ITGA6	10.457	15.596	1.491	7.1599E-09
KRT8	121.044	181.429	1.499	2.2107E-09
LGALS3	8.324	15.677	1.883	6.4112E-09
MCM3	7.172	10.017	1.397	1.3037E-10
MCM4	4.805	7.011	1.459	4.2744E-11
MCM6	2.874	5.041	1.754	4.9299E-11
MSN	29.326	43.248	1.475	1.0518E-08
NR0B2	47.037	30.922	0.657	2.6183E-06
PACSIN3	21.695	15.337	0.707	1.6059E-05
PDGFRA	13.106	19.618	1.497	9.0775E-07
PIK3IP1	5.913	8.727	1.476	2.3685E-07
RNASE1	14.444	23.755	1.645	8.7351E-07
RNF152	25.187	16.363	0.650	1.3805E-05
SFXN2	14.903	12.139	0.815	0.00147773
SULF2	22.680	31.452	1.387	1.8812E-07
TTC36	196.064	149.142	0.761	1.1039E-06
WIPI1	8.254	11.875	1.439	3.4707E-08

- [103] 해당 유전자들은 지방증(Steatosis)에서 NASH로 진행됨에 따라 유전자의 발현량이 감소하는 7개의 유전자 CEBPD, DNAJC12, NR0B2, PACSIN3, RNF152, SFXN2, TTC36 와 발현이 증가하는 21개의 유전자 AJUBA, ANXA2, ATP8B1, CCND1, CHST9, CLDN7, DPYSL2, FAT1, GPNMB, ITGA6, KRT8, LGALS3, MCM3, MCM4, MCM6, MSN, PDGFRA, PIK3IP1, RNASE1, SULF2, WIPI1 들이다.
- [104]
- [105] **실시예 5 : 복합마커의 만성 간질환 진단능 확인**
- [106] 본 발명에서는 선별한 28종의 복합마커를 이용하여 만성 간질환에 대한 진단능을 평가하기 위해, 상기 실시예 1의 비-알코올성 지방간염인 NASH 환자를 대상으로 ROC 곡선 분석을 수행하여 진단의 정확도를 분석하였다.
- [107] ROC 곡선은 결과의 가양성률(FPR, 1-specificity)과 진양성률(TPR, sensitivity)을 나타내며, 검사의 정확도를 평가한다. 28종의 복합마커를 이용하여 지방간과 NASH 그룹 분류의 ROC를 그려보았고, 추정방법으로는 경험적(empirical) 방법, 모수적(binormal) 방법과 비모수적(non-parametric) 방법의 세가지를 사용하였다.
- [108] 그 결과, 도 7에 나타난 바와 같이, AUC 값은 모두 1로 매우 정확한 만성 간질환 진단능을 보이는 것을 확인하였다.
- [109]
- [110] **실시예 6 : 복합마커의 만성 간질환 진단능 검증**
- [111] 본 발명에서는 지방증(Steatosis) 환자 및 비-알코올성 지방간염인 NASH 환자를 대상으로 하여 선별한 28종의 복합마커의 만성 간질환 진단능을 검증하였으며, 28종의 마커 중에서 일부에 대해, qPCR을 이용하여 검증(validation)을 수행하였다.
- [112] 실시예 2의 환자 조직으로부터 수득한 RNA를 역전사 시켜 cDNA를 제조한 다음, 하기 표 4의 프라이머 서열을 이용하여 PCR을 수행하였다.
- [113] RNA는 10 uM 정방향 프라이머 1 μ l, 10 uM 역방향 프라이머 1 μ l, 증류수 1 μ l, SYBR Green I(Roche사 제품) 5 μ l 및 1/5로 희석된 cDNA 2 μ l를 혼합한 다음, LightCycler 96과 SYBR Green I(Roche사 제품)을 이용하여 PCR을 수행하였다.
- [114] PCR은 95 °C에서 5분 동안 반응(pre-incubation) -> 95 °C에서 10초, 60 °C에서 10초, 및 72 °C에서 10초씩 45 사이클(amplification) 반복 -> 95 °C에서 5초 -> 65 °C에서 1분(melting curve) -> 40 °C 쿨링 과정으로 수행하였다.
- [115] PCR 완료 후, 18s rRNA값으로 정규화(normalize) 한 후 Δ 2-Ct 값을 구하여 t-test를 수행하였다.
- [116]

[117] [표3]
프라이머 서열

Gene	프라이머 서열 (5' -> 3')		서열번호
AJUBA	정방향	TCTGCAGGCCTGGCAGAAGA	서열번호 1
	역방향	CCTAAGCTCAGACTGCACACA TG	서열번호 2
ANXA2	정방향	TCGGACACATCTGGTGACTTCC	서열번호 3
	역방향	CCTCTTCACTCCAGCGTCATAG	서열번호 4
ATP8B1	정방향	GTCTTGGACAGAGTCACTTC	서열번호 5
	역방향	CGTCTTATCAGAGAAGATATA AT	서열번호 6
CEBPD	정방향	TCCGGCAGTTCTTCAAGCAGCT	서열번호 7
	역방향	GAGGTATGGGTCGTTGCTGAG T	서열번호 8
GPNMB	정방향	CTGGAGCTGAGTAGGATTCCT G	서열번호 9
	역방향	AGGACGTCTGTCATCTGGATG A	서열번호 10
KRT8	정방향	GCTGACCGACGAGATCAACTT C	서열번호 11
	역방향	TATCCTCGTACTGTGCCTTGAC	서열번호 12
LGALS3	정방향	CCATCTTCTGGACAGCCAAGT G	서열번호 13
	역방향	TATCAGCATGCGAGGCACCAC T	서열번호 14
MCM6	정방향	GACAACAGGAGAAGGGACCTC T	서열번호 15
	역방향	GGACGCTTTACCACTGGTGTA G	서열번호 16
MSN	정방향	CACCGGGAAGCAGCTATTTGA	서열번호 17

	역방향	AGAACTTGGCACGGAAGTTAA	서열번호 18
	정방향	CAGTGAACACCTTTGTGCACG AG	서열번호 19
RNASE1	역방향	TGCTGGAGTTGCTCTTGTAGCA	서열번호 20
	정방향	CGGCTTTGGTGACTCTAGAT	서열번호 21
18s rRNA	정방향	CGGCTTTGGTGACTCTAGAT	서열번호 21
	역방향	GCGACTACCATCGAAAGTTG	서열번호 22

[118] [표4]

검증 결과

Gene	환자수 Steatosis/NASH	p-value (t-test)
AJUBA	n=10 / n= 20	0.198
ANXA2	n=20 / n= 24	0.178
ATP8B1	n=20 / n= 21	0.051
CEBPD	n=20 / n= 23	0.048
GPNMB	n=10 / n= 20	0.010
KRT8	n=10 / n= 20	0.032
LGALS3	n=11 / n= 20	0.105
MCM6	n=12 / n= 21	0.149
MSN	n=11 / n= 21	0.118
RNASE1	n=10 / n= 20	0.014

[119] 그 결과, ATP8B1, CEBPD, GPNMB, KRT8 및 RNASE1 마커를 포함하는 일부 마커들이 통계적으로 유의한 결과를 보여주는 것을 확인하였다. 즉, 본 발명의 복합마커는 만성 간질환 진단에 적합한 것을 확인하였으며, 28종의 마커 중에서 ATP8B1, CEBPD, GPNMB, KRT8 및 RNASE1 마커는 단일마커로도 만성 간질환 진단에 활용될 수 있음을 확인하였다.

[120]

산업상 이용가능성

[121] 본 발명에서 선별한 28종의 복합마커 또는 ATP8B1(ATPase phospholipid transporting 8B1) 마커를 이용하는 경우 만성 간질환 진단능이 현저하게 향상된 것을 확인하였으므로, 본 발명의 복합마커는 만성 간질환 진단을 위한 바이오마커로 적용할 수 있는 것을 확인하였다.

청구범위

- [청구항 1] AJUBA(ajuba LIM protein), ANXA2(annexin A2), ATP8B1(ATPase phospholipid transporting 8B1), CCND1(cyclin D1), CEBPD(CCAAT enhancer binding protein), CHST9(carbohydrate sulfotransferase 9), CLDN7(claudin 7), DNAJC12(DnaJ heat shock protein family (Hsp40) member C12), DPYSL2(dihydropyrimidinase like 2), FAT1(FAT atypical cadherin 1), GPNMB(glycoprotein nmb), ITGA6(integrin subunit alpha 6), KRT8(keratin 8), LGALS3(galectin 3), MCM3(minichromosome maintenance complex component 3), MCM4(minichromosome maintenance complex component 4), MCM6(minichromosome maintenance complex component 6), MSN(moesin), NR0B2(nuclear receptor subfamily 0 group B member 2), PACSIN3(protein kinase C and casein kinase substrate in neurons 3), PDGFRA(platelet derived growth factor receptor alpha), PIK3IP1(phosphoinositide-3-kinase interacting protein 1), RNASE1(ribonuclease A family member 1, pancreatic), RNF152(ring finger protein 152), SFXN2(sideroflexin 2), SULF2(sulfatase 2), TTC36(tetratricopeptide repeat domain 36) 및 WIPI1(WD repeat domain, phosphoinositide interacting 1)을 포함하는 복합마커; 또는 ATP8B1(ATPase phospholipid transporting 8B1) 마커를 포함하는 만성 간질환 진단 마커 조성물.
- [청구항 2] 제1항에 있어서, 상기 만성 간질환은 지방증, 비-알코올성 지방간질환(non-alcoholic fatty liver disease; NAFLD), 비-알코올성 지방간염((non-alcoholic steatohepatitis; NASH) 또는 간경화증인 것을 특징으로 하는 만성 간질환 진단 마커 조성물.
- [청구항 3] AJUBA(ajuba LIM protein), ANXA2(annexin A2), ATP8B1(ATPase phospholipid transporting 8B1), CCND1(cyclin D1), CEBPD(CCAAT enhancer binding protein), CHST9(carbohydrate sulfotransferase 9), CLDN7(claudin 7), DNAJC12(DnaJ heat shock protein family (Hsp40) member C12), DPYSL2(dihydropyrimidinase like 2), FAT1(FAT atypical cadherin 1), GPNMB(glycoprotein nmb), ITGA6(integrin subunit alpha 6), KRT8(keratin 8), LGALS3(galectin 3), MCM3(minichromosome maintenance complex component 3), MCM4(minichromosome maintenance complex component 4), MCM6(minichromosome maintenance complex component 6), MSN(moesin), NR0B2(nuclear receptor subfamily 0 group B member 2), PACSIN3(protein kinase C and casein kinase substrate in neurons 3), PDGFRA(platelet derived growth factor receptor alpha), PIK3IP1(phosphoinositide-3-kinase interacting protein 1),

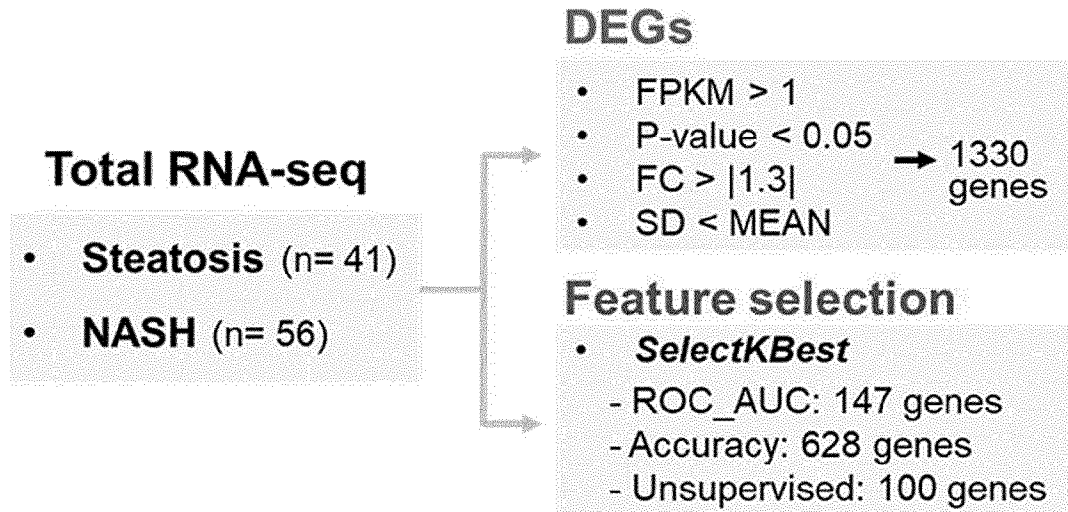
RNASE1(ribonuclease A family member 1, pancreatic), RNF152(ring finger protein 152), SFXN2(sideroflexin 2), SULF2(sulfatase 2), TTC36(tetratricopeptide repeat domain 36) 및 WIPI1(WD repeat domain, phosphoinositide interacting 1)을 포함하는 만성 간질환 진단용 복합마커; 또는 ATP8B1(ATPase phospholipid transporting 8B1)을 포함하는 만성 간질환 진단 마커의 mRNA 또는 단백질 수준을 측정하는 제제를 포함하는 만성 간질환 진단용 조성물.

- [청구항 4] 제3항에 있어서, 상기 만성 간질환은 지방증, 비-알코올성 지방간질환(non-alcoholic fatty liver disease; NAFLD), 비-알코올성 지방간염((non-alcoholic steatohepatitis; NASH) 또는 간경화증인 것을 특징으로 하는 만성 간질환 진단용 조성물.
- [청구항 5] 제3항에 있어서, 상기 mRNA 수준을 측정하는 제제는 마커에 특이적으로 결합하는 프라이머쌍, 프로브 또는 안티센스 뉴클레오타이드인 것을 특징으로 하는 만성 간질환 진단용 조성물.
- [청구항 6] 제3항에 있어서, 상기 단백질 수준을 측정하는 제제는 마커 또는 복합마커의 펩타이드 단편에 특이적으로 결합하는 항체, 상호작용 단백질, 리간드, 나노입자(nanoparticles) 또는 압타머(aptamer)인 것을 특징으로 하는 만성 간질환 진단용 조성물.
- [청구항 7] 제3항 내지 제6항 중 어느 한 항의 만성 간질환 진단용 조성물을 포함하는 만성 간질환 진단용 키트.
- [청구항 8] (a) 환자의 생물학적 시료로부터 제1항의 만성 간질환 진단 마커 조성물의 mRNA 또는 단백질 발현 수준을 측정하는 단계; 및
(b) 상기 mRNA 또는 단백질 발현 수준을 대조군 시료와 비교하는 단계를 포함하는 만성 간질환 진단을 위한 정보 제공방법.
- [청구항 9] 제8항에 있어서, 상기 만성 간질환 진단을 위한 정보 제공방법은 CEBPD(CCAAT enhancer binding protein), DNAJC12(DnaJ heat shock protein family (Hsp40) member C12), NR0B2(nuclear receptor subfamily 0 group B member 2), PACSIN3(protein kinase C and casein kinase substrate in neurons 3), RNF152(ring finger protein 152) 및 SFXN2(sideroflexin 2), TTC36(tetratricopeptide repeat domain 36) 마커의 mRNA 또는 이의 단백질 발현 수준이 대조군에 비해 감소하고, AJUBA(ajuba LIM protein), ANXA2(annexin A2), ATP8B1(ATPase phospholipid transporting 8B1), CCND1(cyclin D1), CHST9(carbohydrate sulfotransferase 9), CLDN7(claudin 7), DPYSL2(dihydropyrimidinase like 2), FAT1(FAT atypical cadherin 1), GPNMB(glycoprotein nmb), ITGA6(integrin subunit alpha 6), KRT8(keratin 8), LGALS3(galectin 3), MCM3(minichromosome maintenance complex component 3), MCM4(minichromosome maintenance complex component 4),

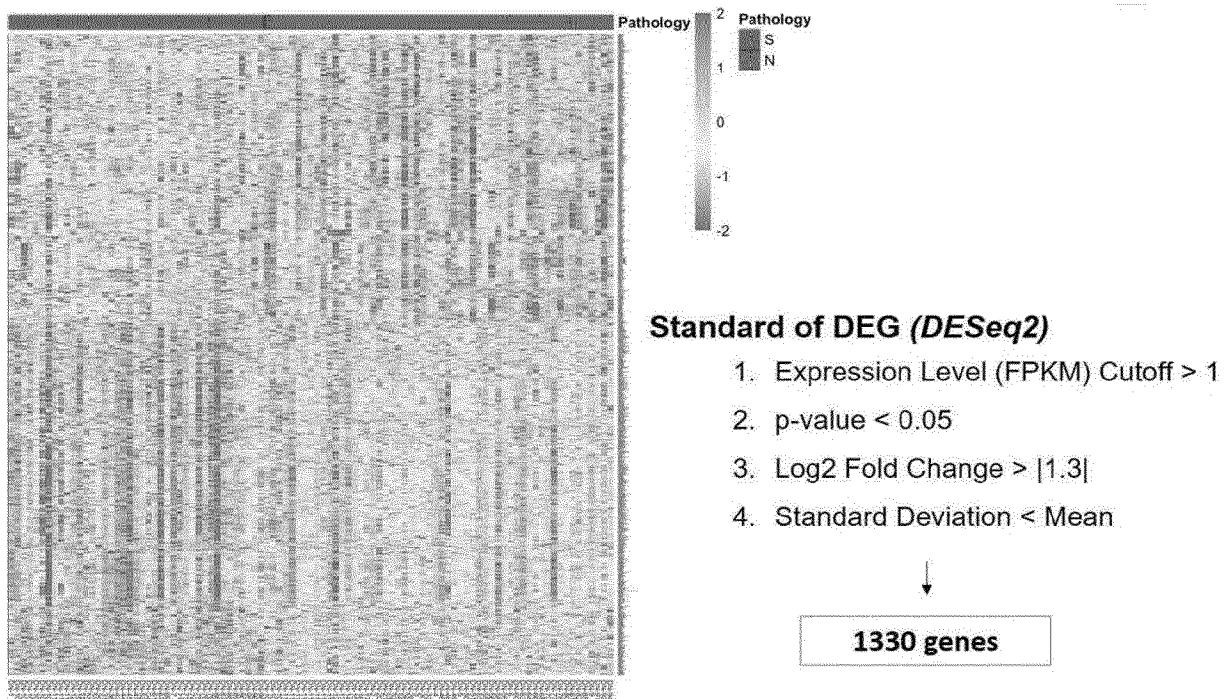
MCM6(minichromosome maintenance complex component 6), MSN(moesin), PDGFRA(platelet derived growth factor receptor alpha), PIK3IP1(phosphoinositide-3-kinase interacting protein 1), RNASE1(ribonuclease A family member 1, pancreatic), SULF2(sulfatase 2), 및 WIPI1(WD repeat domain, phosphoinositide interacting 1) 마커의 mRNA 또는 이의 단백질 발현 수준이 대조군에 비해 증가하면, 만성 간질환인 것으로 정보를 제공하는 단계를 추가로 포함하는 것을 특징으로 하는 만성 간질환 진단을 위한 정보 제공방법.

- [청구항 10] 제8항에 있어서, 상기 만성 간질환 진단을 위한 정보 제공방법은 ATP8B1(ATPase phospholipid transporting 8B1) 마커의 mRNA 또는 이의 단백질 발현 수준이 대조군에 비해 증가하면, 만성 간질환인 것으로 정보를 제공하는 단계를 추가로 포함하는 것을 특징으로 하는 만성 간질환 진단을 위한 정보 제공방법.

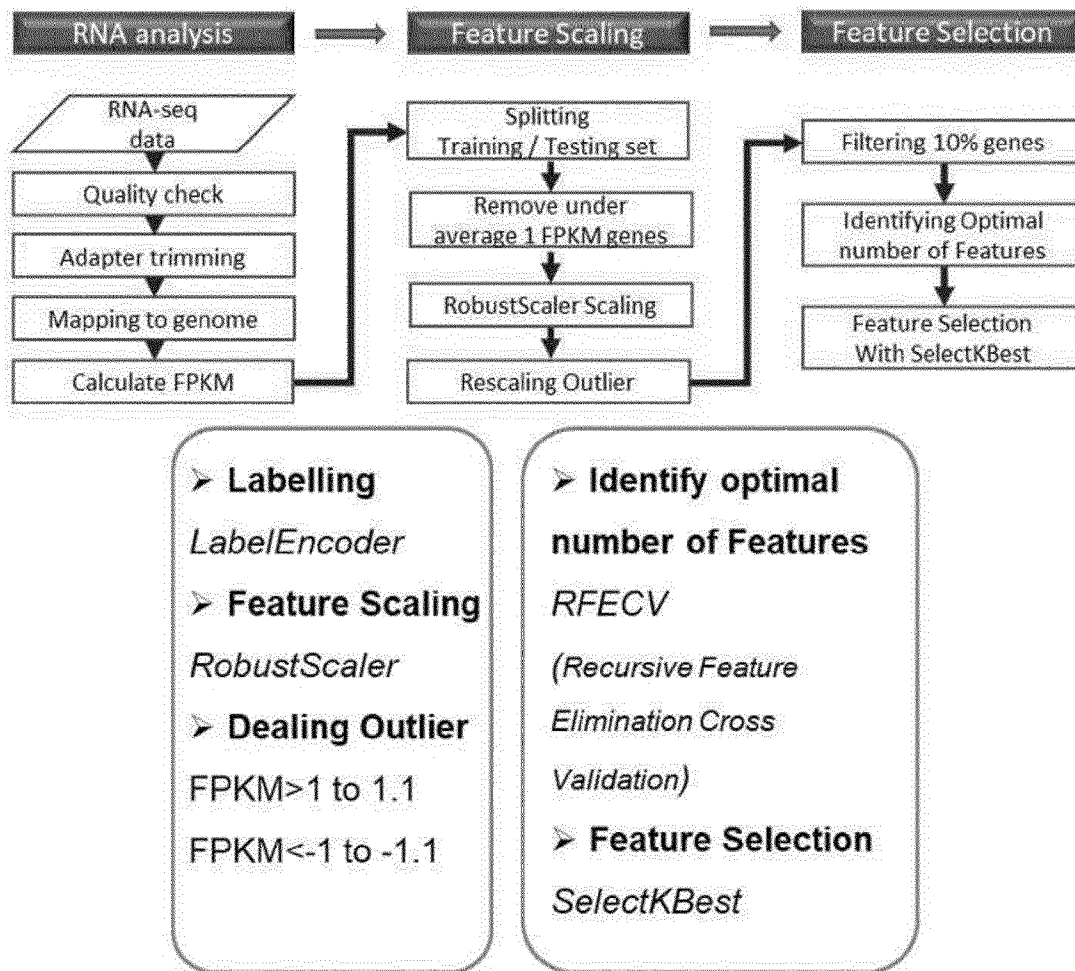
[도1]



[도2]



[도3]



[도4]

Cases of Features selected for each modeling purpose

Case1. Supervised machine learning model scored by ROC_AUC	147 genes – CAPG
Case2. Supervised machine learning model scored by Accuracy	628 genes – CAPG
Case3. Unsupervised machine learning model	100 genes – MCM6

Process

Identification of Optimal Feature number

RFECV (Recursive Feature Elimination Cross Validation)

Feature Selection in the 3 Cases

SelectKBest

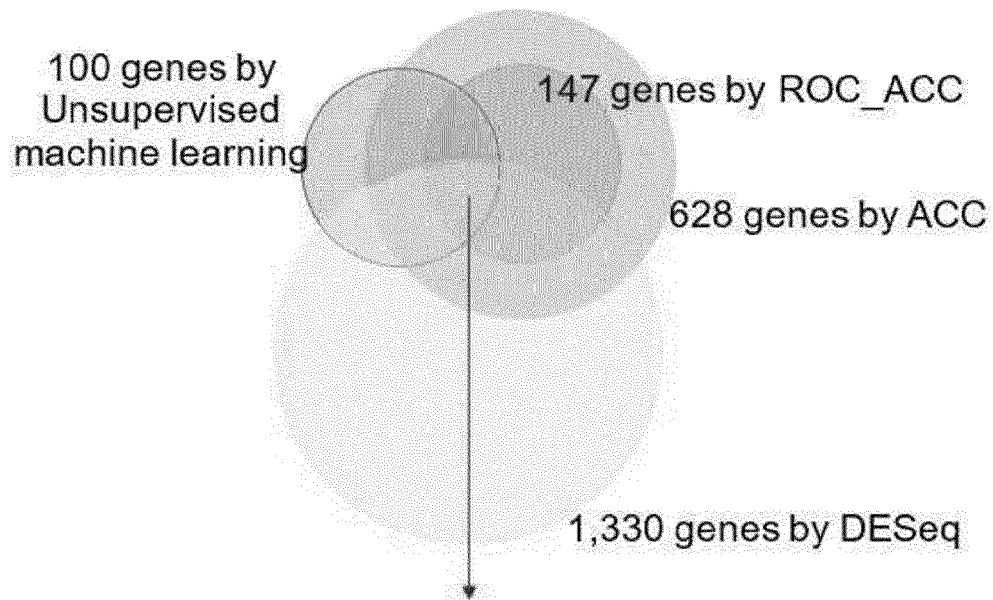
Ranking Feature Importance

SelectKBest – Feature Importance Algorithm

Evaluation of the selected features

Comparison of Real FPKM value in Steatosis with NASH
 (Heatmaps sorted by CAPG or MCM6)

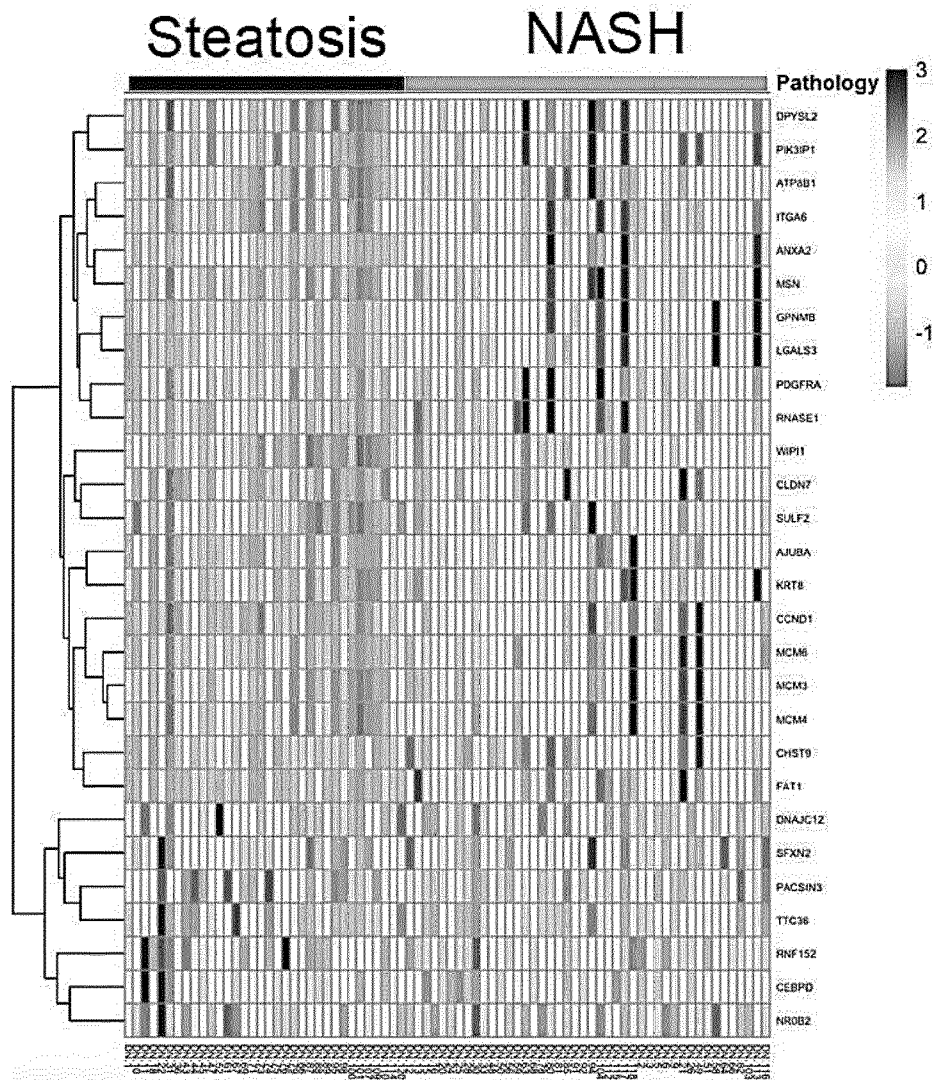
[도5]



28 intersected Genes

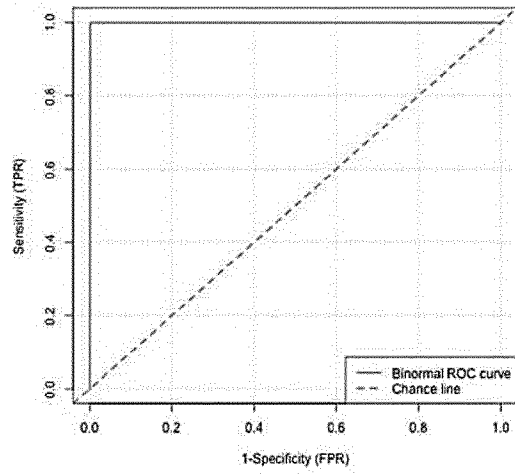
AJUBA, ANXA2, ATP8B1, CCND1, CEBPD, CHST9, CLDN7, DNAJC12, DPYSL2, FAT1, GPNMB, ITGA6, KRT8, LGALS3, MCM3, MCM4, MCM6, MSN, NR0B2, PACSIN3, PDGFRA, PIK3IP1, RNASE1, RNF152, SFXN2, SULF2, TTC36, WIP1

[도6]

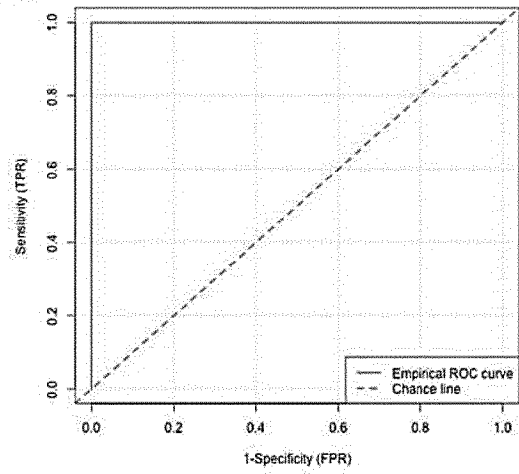


[도7]

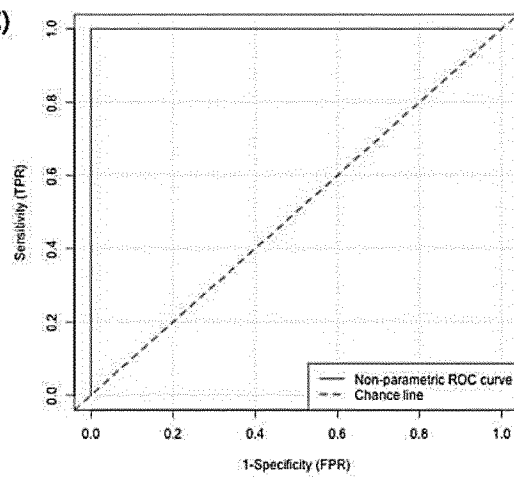
(A)



(B)



(C)



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2021/017751

A. CLASSIFICATION OF SUBJECT MATTER

C12Q 1/6883(2018.01)i; G01N 33/68(2006.01)i

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12Q 1/6883(2018.01); G01N 33/68(2006.01)

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

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

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

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

eKOMPASS (KIPO internal) & keywords: 만성 간질환(chronic liver disease), 진단(diagnosis), 마커(marker), 발현(expression), 통합전사체분석(integrative transcriptomic analysis), 지방증(steatosis), 비-알코올성 지방간염(non-alcoholic steatohepatitis, NASH)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	DENG, Bao-Cheng et al. Novel ATP8B1 mutation in an adult male with progressive familial intrahepatic cholestasis. World Journal of Gastroenterology. 28 November 2012 (online publication date), vol. 18, no. 44, pp. 6504-6509. See abstract.	1,2 3-10
A	US 2014-0303018 A1 (SOMALOGIC, INC.) 09 October 2014 (2014-10-09) See abstract; and claims 1, 2, 4, 5, 7-9, 12-14, 26, 28-34, 36, 37, 40-43 and 56.	1-10
A	WO 2020-087037 A2 (MOLECULAR STETHOSCOPE, INC.) 30 April 2020 (2020-04-30) See abstract; and claims 1-84.	1-10
A	HAYASHI, Hisamitsu et al. Assessment of ATP8B1 deficiency in pediatric patients with cholestasis using peripheral blood monocyte-derived macrophages. EBioMedicine. 07 October 2017 (online publication date), vol. 27, pp. 187-199. See pages 187-198.	1-10



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

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

“D” document cited by the applicant in the international application

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

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

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

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

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

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

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

“&” document member of the same patent family

Date of the actual completion of the international search

17 March 2022

Date of mailing of the international search report

18 March 2022

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
Government Complex-Daejeon Building 4, 189 Cheongsaro, Seo-gu, Daejeon 35208

Facsimile No. +82-42-481-8578

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2021/017751

C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MARTINEZ-FERNANDEZ, Pilar et al. Knockdown of ATP8B1 expression leads to sepcific downregulation of the bile acid sensor FXR in HepG2 cells: effect of the FXR agonist GW4064. American Journal of Physiology-Gastrointestinal and Liver Physiology. 19 February 2009 (online publication date), vol. 296, no. 5, pp. G1119-G1129. See pages G1119-G1128.	1-10

INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2021/017751

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☒ forming part of the international application as filed:
 - ☒ in the form of an Annex C/ST.25 text file.
 - ☐ on paper or in the form of an image file.
 - b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
 - c. ☐ furnished subsequent to the international filing date for the purposes of international search only:
 - ☐ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).
 - ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).
2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2021/017751

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	2014-0303018	A1	09 October 2014	BR	112015021992	A2	18 July 2017
				BR	112015021992	A8	03 December 2019
				CN	105143887	A	09 December 2015
				CN	105143887	B	28 August 2018
				CN	108957006	A	07 December 2018
				CN	108957006	B	27 July 2021
				DK	3296747	T3	30 March 2020
				EP	2972386	A2	20 January 2016
				EP	2972386	B1	10 January 2018
				EP	3296747	A1	21 March 2018
				EP	3296747	B1	08 January 2020
				EP	3660512	A1	03 June 2020
				ES	2662108	T3	05 April 2018
				ES	2773999	T3	16 July 2020
				HK	1220251	A1	28 April 2017
				MX	2015011756	A	14 January 2016
				MX	360140	B	24 October 2018
				US	10359435	B2	23 July 2019
				US	2017-0227550	A1	10 August 2017
				US	2018-0259537	A1	13 September 2018
				US	9612248	B2	04 April 2017
				US	9945875	B2	17 April 2018
				WO	2014-150198	A2	25 September 2014
				WO	2014-150198	A3	18 December 2014
				ZA	201506441	B	31 July 2019
WO	2020-087037	A2	30 April 2020	AU	2019-367010	A1	17 June 2021
				CA	3117488	A1	30 April 2020
				CN	113825864	A	21 December 2021
				EP	3870742	A2	01 September 2021
				WO	2020-087037	A3	04 June 2020

A. 발명이 속하는 기술분류(국제특허분류(IPC)) C12Q 1/6883(2018.01)i; G01N 33/68(2006.01)i		
B. 조사된 분야 조사된 최소문헌(국제특허분류를 기재) C12Q 1/6883(2018.01); G01N 33/68(2006.01) 조사된 기술분야에 속하는 최소문헌 이외의 문헌 한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우)) eKOMPASS(특허청 내부 검색시스템) & 키워드: 만성 간질환(chronic liver disease), 진단(diagnosis), 마커(marker), 발현(expression), 통합전사체분석(integrative transcriptomic analysis), 지방증(steatosis), 비-알코올성 지방간염(non-alcoholic steatohepatitis, NASH)		
C. 관련 문헌		
카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
X A	DENG, BAO-CHENG 등, "Novel ATP8B1 mutation in an adult male with progressive familial intrahepatic cholestasis", World Journal of Gastroenterology, 2012년 11월 28일(온라인 공개일), 18권, 4호, 페이지 6504-6509 요약	1,2 3-10
A	US 2014-0303018 A1 (SOMALOGIC, INC.) 2014.10.09 요약; 및 청구항 1, 2, 4, 5, 7-9, 12-14, 26, 28-34, 36, 37, 40-43, 56	1-10
A	WO 2020-087037 A2 (MOLECULAR STETHOSCOPE, INC.) 2020.04.30 요약; 및 청구항 1-84	1-10
A	HAYASHI, HISAMITSU 등, "Assessment of ATP8B1 deficiency in pediatric patients with cholestasis using peripheral blood monocyte-derived macrophages", EBioMedicine, 2017년 10월 7일(온라인 공개일), 27권, 페이지 187-199 페이지 187-198	1-10
☒ 추가 문헌이 C(계속)에 기재되어 있습니다. ☒ 대응특허에 관한 별지를 참조하십시오.		
* 인용된 문헌의 특별 카테고리: "A" 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌 "D" 본 국제출원에서 출원인이 인용한 문헌 "E" 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌 "L" 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌 "O" 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌 "P" 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌 "T" 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌 "X" 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다. "Y" 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다. "&" 동일한 대응특허문헌에 속하는 문헌		
국제조사의 실제 완료일 2022년03월17일(17.03.2022)		국제조사보고서 발송일 2022년03월18일(18.03.2022)
ISA/KR의 명칭 및 우편주소 대한민국 특허청 (35208) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사) 팩스 번호 +82-42-481-8578		심사관 허주형 전화번호 +82-42-481-5373

C. 관련 문헌		
카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
A	MARTINEZ-FERNANDEZ, PILAR 등, "Knockdown of ATP8B1 expression leads to sepcific downregulation of the bile acid sensor FXR in HepG2 cells: effect of the FXR agonist GW4064", American Journal of Physiology-Gastrointestinal and Liver Physiology, 2009년 2월 19일(온라인 공개일), 296권, 5호, 페이지 G1119-G1129 페이지 G1119-G1128	1-10

제1기재란 핵산염기 및/또는 아미노산 서열(첫 번째 용지의 1.c의 계속)

1. 국제출원에 개시된 핵산염기 및/또는 아미노산 서열과 관련하여, 국제조사는 다음에 기초하여 수행되었습니다.

- a. ☒ 아래의 형태로 출원시 국제출원의 일부를 구성하는 서열목록
 - ☒ 부록 C/ST.25 텍스트 파일
 - ☐ 서면 혹은 이미지 파일
- b. ☐ PCT 규칙 13의3.1(a)에 따라 국제출원과 함께 국제조사만을 목적으로 부록 C/ST.25 텍스트 파일의 형태로 제출된 서열목록
- c. ☐ 국제조사만을 목적으로 국제출원일 이후에 아래 형태로 제출된 서열목록
 - ☐ 부록 C/ST.25 텍스트 파일 (규칙 13의3.1(a))
 - ☐ 서면 혹은 이미지 파일 (규칙 제13의3.1(b) 및 시행세칙 713).

2. ☐ 추가로 서열목록에 대하여 하나 이상의 버전이나 사본이 제출된 경우, 후속 버전 또는 추가된 사본에 기재되어 있는 정보가 출원시 출원의 일부를 구성하는 정보와 동일하거나 또는 출원시의 개시범위를 벗어나지 않는다는 진술서가 제출되었습니다.

3. 추가 의견:

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
US 2014-0303018 A1	2014/10/09	BR 112015021992 A2	2017/07/18
		BR 112015021992 A8	2019/12/03
		CN 105143887 A	2015/12/09
		CN 105143887 B	2018/08/28
		CN 108957006 A	2018/12/07
		CN 108957006 B	2021/07/27
		DK 3296747 T3	2020/03/30
		EP 2972386 A2	2016/01/20
		EP 2972386 B1	2018/01/10
		EP 3296747 A1	2018/03/21
		EP 3296747 B1	2020/01/08
		EP 3660512 A1	2020/06/03
		ES 2662108 T3	2018/04/05
		ES 2773999 T3	2020/07/16
		HK 1220251 A1	2017/04/28
		MX 2015011756 A	2016/01/14
		MX 360140 B	2018/10/24
		US 10359435 B2	2019/07/23
		US 2017-0227550 A1	2017/08/10
		US 2018-0259537 A1	2018/09/13
		US 9612248 B2	2017/04/04
		US 9945875 B2	2018/04/17
		WO 2014-150198 A2	2014/09/25
		WO 2014-150198 A3	2014/12/18
		ZA 201506441 B	2019/07/31
WO 2020-087037 A2	2020/04/30	AU 2019-367010 A1	2021/06/17
		CA 3117488 A1	2020/04/30
		CN 113825864 A	2021/12/21
		EP 3870742 A2	2021/09/01
		WO 2020-087037 A3	2020/06/04