creative commons
COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

의학박사 학위논문

전장유전체연관분석연구를 이용한 비결핵항산균 폐질환자의 유전적 소인 규명

2021년 8월

서울대학교 대학원

의학과 내과학 전공

조 재 영

전장유전체연관분석연구를 이용한 비결핵항산균 폐질환자의 유전적 소인 규명

지도교수 임 재 준

이 논문을 의학박사 학위논문으로 제출함
2021년 4월

서울대학교 대학원
의학과 내과학 전공
조 재 영

조재영의 의학박사 학위논문을 인준함
2021년 7월

위 원 장 _____ 김 종 일

부위원장 _____ 임 재 준

위 원 _____ 김 남 중

위 원 _____ 성 문 우

위 원 _____ 강 영 애

초 록

전장유전체연관분석연구를 이용한 비결핵항산균 폐질환자의 유전적 소인 규명

조 재 영
의학과 내과학 전공
서울대학교 대학원

서론: 비결핵항산균은 토양과 강물 등의 자연환경에 널리 분포하고 있으므로 거의 모든 사람들이 종종 비결핵항산균을 흡입하게 되는데, 이들 중 일부에서만 비결핵항산균 폐질환이 발생한다는 점은 이 질환에 대한 유전적 소인이 존재한다는 점을 시사한다. 본 연구의 목표는 최근 급격히 증가하고 있는 비결핵항산균 폐질환자들의 유전적 결함 존재 여부를 전장유전체연관분석연구를 통해 확인하는 것이다.

방법: 비결핵항산균 폐질환자와 쌍둥이 코호트에 참여한 건강인을 대조군으로 삼아 전장유전체연관분석연구를 시행했다. 쌍둥이는 무작위로 배정된 1명만 대조군으로 참여할 수 있었다. 전장유전체연관분석연구에서 후보로 선정된 유전적 변이에 대해서는 독립적인 비결핵항산균 폐질환자-대조군 코호트에서 연관성이 재현되는지 확인했다. Genotype-Tissue Expression (GTEx) 데이터베이스를 이용해서 발현 양적 형질 유전자리(expression quantitative trait loci, eQTL)를 식별하고, 멘델 무작위배정(Mendelian randomization) 분석을 시행했다.

결과: 403명의 비결핵항산균 폐질환자와 306명의 대조군을 대상으로 한 전장유전체연관분석연구를 통해서 7번 염색체 단완(7p13)에 위치한 유

전적 변이인 rs849177이 비결핵항산균 폐질환에 대한 감수성과 관련된 유전자 변이로 추정되었다 (OR, 2.34; 95% CI, 1.71 to 3.21; $p=1.36\times10^{-7}$). 이 연관성은 독립적인 비결핵항산균 폐질환자-대조군 코호트에서 재현되었고, 두 코호트를 함께 분석했을 때의 유의확률은 4.92×10^{-8}였다. eQTL 분석을 통해 rs849177의 위험 대립유전자(risk allele)가 세포자멸사를 유발하는 유전자인 *STK17A*의 발현 감소와 연관되어 있다는 것을 밝혔다. 멘델 무작위배정 분석을 통해 *STK17A*가 비결핵항산균 폐질환 발생에 인과관계가 있다는 것을 확인했다 (β, −4.627; 95% CI, −8.768 ~ −0.486; p=0.029).

결론: 7번 염색체 단완(7p13)에 위치한 유전적 변이가 한국인의 비결핵항산균 폐질환에 대한 감수성과 연관되어 있었다. 비결핵항산균 폐질환자는 이 유전적 변이의 위험 대립유전자를 가지고 있을 가능성이 높은데, 이 위험 대립유전자는 세포자멸사를 유발하는 유전자인 *STK17A*의 발현 감소와 연관되어 있다. 따라서 비결핵항산균에 감염된 대식세포의 세포자멸사가 정상적으로 진행되지 않아 비결핵항산균 폐질환에 대한 감수성이 증가할 것으로 추정된다.

주요어 : 비결핵항산균; 폐질환; 유전적 감수성; 전장유전체연관분석연구; 세포자멸사

학 번 : 2016-30558

본 학위논문의 내용은 영국 호흡기학회에서 발간하는 *Thorax*에 출판되었음 (Cho J, Park K, Choi SM, Lee J, Lee CH, Lee JK, Heo EY, Kim DK, Lee YJ, Park JS, Cho YJ, Yoon HI, Lee JH, Lee CT, Kim N, Choi KY, Lee KH, Sung J, Won S, Yim JJ. Genome-wide association study of non-tuberculous mycobacterial pulmonary disease. Thorax 2021;76:169-177).

목 차

초록 i

목차 iii

표 목차 vi

그림 목차 vii

I. 서 론 1

1. 비결핵항산균 폐질환의 세계적인 증가 추세 1

2. 비결핵항산균 폐질환 치료의 어려움 6

3. 비결핵항산균 폐질환으로 인한 사망률 및 사회·경제적 비용의 증가 6

4. 비결핵항산균 폐질환에 대한 유전적 소인 8

5. 최근의 엑솜 염기서열 분석 연구와 제한점 10

6. 연구의 필요성 14

II. 연구 방법 15

1. 발굴 코호트(discovery cohort) 15

2. 재현 코호트(replication cohort) 16

3. 발굴 코호트에서 수행된 전장유전체연관분석연구 17

4. 재현 코호트를 통한 검증(validation) 21

5. 다차원 척도법(multidimensional scaling, MDS) 23

6. 지역 플롯(regional plot) 24

7. 전장유전체 복합 형질 분석(genome-wide complex trait analysis) 24

8. 토폴로지 연관 도메인 분석(topologically associating domain analysis) 25

9. 발현 양적 형질 유전자자리(expression quantitative trait loci, eQTL)와 PrediXcan 분석.....25

10. 2-표본 멘델 무작위배정(two-sample Mendelian randomization) 분석.....26

11. 비결핵항산균에 감염된 대식세포(macrophage)의 전사체(transcriptome) 프로파일링 분석 29

12. Gene Ontology enrichment 분석 30

13. 기존에 알려진 유전자에 대한 흔한 변이(common variant) 분석과 유전자 세트(gene set) 분석.....30

III. 결 과..... 32

1. 전장유전체연관분석으로 얻은 잠재적 후보 SNP 32

2. 잠재적 후보 SNP의 검증.....45

3. 3차원 구조 분석.....48

4. 비결핵항산균 폐질환과 연관된 후보 유전자 49

5. 후보 유전자가 비결핵항산균 폐질환 발생에 미치는 인과적 영향(causal effect).....56

6. 비결핵항산균에 감염된 대식세포에서 후보 유전자의 차별적 발현(differential expression) 58

7. 비결핵항산균에 감염된 대식세포에서 세포자멸사(apoptosis) 경로의 강화(enrichment) 60

8. 기존에 알려진 비결핵항산균 폐질환 관련 유전자에 대한 연관 연구 64

IV. 고　　찰 69
V. 결　　론 79

참고문헌 80
Abstract 96

표 목차

[표 1] 비결핵항산균 감염의 발생률과 유병률 추이..... 3

[표 2] 비결핵항산균 폐질환에 대한 엑솜 염기서열 분석 연구 11

[표 3] 발굴 및 복제 코호트 참가자의 임상적 특성..... 33

[표 4] 비결핵항산균 폐질환자의 동반질환 35

[표 5] 발굴 및 재현 코호트에서 rs849177에 대한 전장유전체연관분석 결과 40

[표 6] 발굴 코호트에서 rs849177과 *Mycobacterium avium* 복합체(complex) 폐질환 및 *Mycobacterium abscessus* 복합체 폐질환과의 연관성 44

[표 7] Fisher의 방법을 사용한 발굴 및 재현 코호트의 결과 병합 46

[표 8] Genotype-Tissue Expression (GTEx) 데이터베이스에서 rs849177에 의해 유전자 발현 수준이 크게 영향을 받는 후보 유전자와 해당 조직 50

[표 9] Gene Ontology (GO) 분석을 통해 얻은 GSE72821 및 E-MTAB-1101 데이터세트에서 모두 유의한 세포자멸사 경로 61

[표 10] 기존에 비결핵항산균 폐질환 관련 유전자로 보고된 43개의 유전자에 대한 흔한 변이(common variant) 분석과 유전자 세트(gene set) 분석 65

그림 목차

[그림 1] 정도 관리(quality control)의 흐름 모식도..... 19

[그림 2] 멘델 무작위배정(two-sample Mendelian randomization)의 가정 28

[그림 3] 다차원 척도법(multidimensional scaling, MDS) 플롯36

[그림 4] 발굴 코호트에서 전장유전체연관분석으로 얻은 잠재적 후보 유전자자리 38

[그림 5] 발굴 코호트에서 체질량지수를 보정하지 않고 수행한 로지스틱 분석 42

[그림 6] 발굴 코호트에서 rs849177에 대한 조건부 분석(conditional analysis) 43

[그림 7] rs849177 주변의 토폴로지 연관 도메인(topologically associating domain, TAD) 48

[그림 8] Genotype-Tissue Expression(GTEx) 데이터베이스의 폐 조직에서 rs849177 유전자형에 따른 *STK17A* 유전자의 발현 수준 52

[그림 9] PrediXcan을 이용해 예측한 비결핵항산균 폐질환자군과 대조군에서의 *STK17A* 유전자의 발현 수준 54

[그림 10] 멘델 무작위배정 (Mendelian randomization, MR) 분석을 통해 얻은 변이-비결핵항산균 폐질환의 회귀 계수 대 변이-*STK17A* 유전자의 발현 수준의 회귀 계수 플롯 57

[그림 11] *Mycobacterium abscessus* 또는 *Mycobacterium avium*에 감염된 또는 감염되지 않은 대식세포에서 *STK17A* (A, B), *BLVRA* (C, D), *COA1* (E)의 발현 수준..... 59

I. 서 론

비결핵항산균(non-tuberculous mycobacteria, NTM)은 토양과 강물 등의 자연환경에 널리 분포하고 있으며, 광범위한 발병력(virulence) 스펙트럼을 가지고 있다. 비결핵항산균은 폐질환, 림프절염, 피부 및 연조직 감염, 파종성 감염 등을 일으킬 수 있으나, 그 중 폐질환이 대부분을 차지한다. 현재까지 200 종(species) 이상의 비결핵항산균이 확인되었고, 우리나라에서 비결핵항산균 폐질환(NTM pulmonary disease, NTM-PD)을 일으키는 가장 흔한 원인균은 *Mycobacterium avium* 복합체(complex)이고, 두 번째로 흔한 균은 *Mycobacterium abscessus* 복합체이다.

1. 비결핵항산균 폐질환의 세계적인 증가 추세

비결핵항산균 폐질환의 발생률(incidence)과 유병률(prevalence)은 세계 여러 지역에서 증가하고 있다 (표 1).[1-12] 특히, 우리나라는 지난 10년 동안 비결핵항산균 감염의 발생률과 유병률이 모두 크게 증가했다. 국민건강보험공단 데이터베이스를 분석한 연구에 따르면, 비결핵항산균 감염의 발생률은 2003년에는 인구 10만명당 1.0명이었으나, 2016년에는 17.9명으로 크게 증가했다.[1] 유병률 또한 크게 증가하여

2003년에는 인구 10만명당 1.2명이었으나 2016년에는 33.3명으로 증가한 것으로 보고되었다.[1] 2007년부터 2016년까지의 건강보험심사평가원 데이터베이스를 분석한 연구도 같은 결과를 나타냈다.[2] 비결핵항산균 감염의 발생률은 2007년에는 인구 10만명당 6.0명이었으나, 2016년에는 19.0명으로 증가했다.[2] 비결핵항산균 감염의 유병률은 지난 10년 동안 5배 이상 증가하여, 2007년에는 인구 10만명당 6.7명이었으나, 2016년에는 39.6명으로 나타났다.[2]

동아시아에 비해 비결핵항산균 폐질환의 유병률이 상대적으로 낮은 세계 여러 지역에서도 비결핵항산균 폐질환의 발생률과 유병률이 증가하고 있다 (표 1).[3-12] 미국의 의료 청구 데이터베이스를 분석한 연구에 의하면, 비결핵항산균 폐질환의 발생률은 2008년에는 인구 10만명당 3.13 명이었으나 2015년에는 4.73명으로 증가했다.[3] 유병률은 2008년에는 인구 10만명당 6.78명이었으나, 2015년에는 11.70명으로 꾸준한 증가 추세를 보였다.[3] 특히, 65세 이상의 메디케어 수혜자를 대상으로 한 연구에서는 비결핵항산균 폐질환의 유병률이 1997년에는 인구 10만명당 20명이었는데 이후 꾸준히 증가하여 2007년에는 47명이었다.[4] 한편, 오스트레일리아의 북동부에 있는 퀸즐랜드주의 비결핵항산균 폐질환 발생률은 1999년에는 인구 10만명당 2.2명, 2005년에는 3.2명이었으나,[10] 2020년 이후 2040년까지 비결핵항산균 폐질환 발생이 3배 이상 증가할 것으로 예측되었다.[13]

표 1. 비결핵항산균 감염의 발생률과 유병률 추이

Area/Population, Country	Study Period	Incidence (per 100,000)	Prevalence (per 100,000)	Remark
North America				
Four integrated health care delivery systems, USA[7]	1994–2006		• 2004–2006: 1.4–6.6 • Age ≥60: 19.6 (1994–1996) → 26.7 (2004–2006)	
Medicare beneficiaries, USA[4]	1997–2007		20 → 47	Age ≥65
Administrative claims database, USA[3]	2008–2015	3.13 → 4.73	6.78 → 11.70	
New York, USA[14]	2000–2003	2.0		
Oregon, USA[15]	2005–2006		5.6	
Oregon, USA[12]	2007–2012	4.8 → 5.6		
British Columbia, Canada[16]	1990–2006	8.3 → 6.0		
Ontario, Canada[5]	1998–2010		4.9 → 9.8	
Ontario, Canada[6]	2003–2008		4.35 → 6.81	*Mycobacterium avium* complex

				pulmonary disease
Europe				
National medical database, Denmark[17]	1997–2008	1.66 → 1.50		
Sentinel site surveillance, France[18]	2001–2003	0.74 → 0.72		
Central Greece, Greece[8]	2004–2006	0.67 → 0.94		
Health insurance claim database, Germany[9]	2009–2014		2.3 → 3.3	
Health insurance claim database, Germany[11]	2010–2011	1.12 → 1.48		
Oceania				
Queensland, Australia[10]	1999–2005	2.2 → 3.2		
New Zealand[19]	2004	1.17		
Asia				
National Health Insurance Service database, South Korea[1]	2003–2016	1.0 → 17.9	1.2 → 33.3	
Health Insurance Review and Assessment Service database, South Korea[2]	2007–2016	6.0 → 19.0	6.7 → 39.6	
Vital Statistics database, Japan[20]	2005		33–65	Estimated prevalence from

				the mortality rate
Nationwide hospital–based survey, Japan[21]	2014	14.7		

USA: United States of America.

2. 비결핵항산균 폐질환 치료의 어려움

비결핵항산균 폐질환은 장기간의 복합 항생제 치료에도 불구하고 치료 성적이 만족스럽지 않다. *M. avium* 복합체 폐질환의 치료 성적에 대한 체계적 문헌고찰 및 메타분석 결과, 치료 성공률은 60%였고, 치료 중단이나 약제 용량 변경을 필요로 하는 부작용은 6.4%였다.[22] 여러 항생제에 내성을 보이고, 수주 혹은 수개월의 주사 치료가 필요한 *M. abscessus* 복합체 폐질환의 경우, 치료 성적이 더 낮다. 메타분석에 의하면 *M. abscessus* 복합체 폐질환의 치료 성공률은 45.6%였다.[23] *M. abscessus* 복합체의 아종(subspecies)인 *M. abscessus* subsp. *abscessus*에 의한 폐질환의 치료 성공률이 특히 낮아 33.0%에 불과했다.[23]

3. 비결핵항산균 폐질환으로 인한 사망률 및 사회·경제적 비용의 증가

비결핵항산균 폐질환자는 삶의 질이 저하되어 있고,[24] 예후가 불량하다.[1,25] 2003년부터 2016년까지의 국민건강보험공단 데이터베이스를 분석한 연구에 따르면, 비결핵항산균 감염의 진단 후 1년 사망률이 4.7%였고, 5년 사망률이 17.8%에 달했다.[1] 비결핵항산균 감염 환자의 표준화사망률은 일반인구보다 2.16배 높았다.[1] 2002년부터

2017년까지의 건강보험심사평가원 데이터베이스를 분석한 연구진도 비슷한 결과를 보고했다.[25] 비결핵항산균 감염 환자의 10년 누적 생존율은 80.8%로 일반인구 (99.1%) 및 결핵 환자 (97.0%)보다 현저히 낮았다.[25] 일본에서는 비결핵항산균 감염으로 인한 사망률의 지속적인 증가 추세가 확인되었는데, 1970년부터 2010년까지 40년 동안 비결핵항산균 감염으로 인한 조사망률이 373배 증가했다.[20] 우리나라에서 시행된 전향적 코호트 연구 결과에 따르면, 고령의 나이, 남성, 저체중, 만성 폐 아스페르길루스증(chronic pulmonary aspergillosis), 악성 종양, 만성 심장질환 또는 간질환의 존재, 영상검사에서 동공의 존재, 혈액검사에서 적혈구침강속도(erythrocyte sedimentation rate, ESR)의 증가 등이 사망률을 예측할 수 있는 인자였다.[26,27]

비결핵항산균 폐질환으로 인한 사망률 증가뿐만 아니라 사회·경제적 비용의 증가 추세도 뚜렷하다. 유럽과 캐나다에서 시행된 후향적 연구에서 비결핵항산균 폐질환 치료 관련하여 상당한 경제적 부담이 발생한다고 보고되었고,[28] 독일에서 시행된 연구에서 비결핵항산균 폐질환자는 입원이 3배 이상 증가했고,[11] 영국에서 시행된 연구에서는 의료기관 이용이 6배 증가했다.[29]

4. 비결핵항산균 폐질환에 대한 유전적 소인

비결핵항산균은 폐질환, 림프절염, 피부 및 연조직 감염, 파종성 감염 등을 일으킬 수 있으나, 그 중 폐질환이 90% 이상을 차지한다. 파종성 비결핵항산균 감염은 원발성 면역결핍[30]이나 후천적 면역손상[31] 환자에서 발생한다. 이전 연구에 따르면 파종성 비결핵항산균 감염에 대한 멘델 감수성(Mendelian susceptibility)은 인터루킨(interleukin)-12 의존성 인터페론(interferon)-γ 경로의 유전자 돌연변이와 관련이 있다.[30,32] 하지만, 파종성 비결핵항산균 감염과 다르게, 비결핵항산균 폐질환은 알려진 면역 결함이 없는 사람에서 발생한다. 비결핵항산균은 토양과 강물 등의 자연환경에 널리 분포하고 있으므로 거의 모든 사람들이 종종 비결핵항산균을 흡입하게 되는데, 이들 중 일부에서만 비결핵항산균 폐질환이 발생한다는 점은 이 질환에 대한 유전적 소인이 존재한다는 점을 시사한다.

비결핵항산균 폐질환에 대한 유전적 소인이 존재한다는 이론적 근거는 다음과 같다. 첫째, 비결핵항산균 폐질환의 비교적 뚜렷한 표현형이 존재한다. 비결핵항산균 폐질환은 주로 체형이 마른 폐경기 여성에서 발생하며, 오목가슴, 척추측만증 또는 승모판탈출이 동반되기도 한다.[33]

둘째, 비결핵항산균 폐질환에 대한 다양한 역학연구에서 인종적 차이를 확인할 수 있다. 표 1에서 알 수 있듯이 비결핵항산균 폐질환은

아시아인에서 유병률이 가장 높고, 유럽인에서 유병률이 가장 낮다. 이민자가 많은 하와이에서 2005년부터 2013년까지 비결핵항산균 분리(isolation)의 연간 평균 유병률을 조사한 결과, 일본인, 중국인, 베트남인, 한국인에서 10만명당 34명으로 가장 높았고, 필리핀인이나 백인(Caucasians)에서 10만명당 18명이었으며, 하와이 원주민에서 10만명당 6명으로 가장 낮았다.[34] 또한, 미국 기관지확장증 레지스트리를 분석한 연구에 따르면, 아프리카계 미국인(African-American)에서 상대적으로 비결핵항산균 감염이 드물었다 (relative risk, 0.29; 95% CI, 0.15 ~ 0.54).[35]

셋째, 비결핵항산균 폐질환의 가족 클러스터링(familial clustering)을 보고한 이전 연구도 비결핵항산균 폐질환에 대한 유전적 소인의 존재를 뒷받침한다.[36] 미국에서 109명의 비결핵항산균 폐질환 계보발단자(proband)와 2,285명의 일촌 및 이촌 친척을 조사한 결과, 14명(12.8%)의 계보발단자는 일촌 및 이촌 친척 중 비결핵항산균 폐질환자가 있었다.[36] 또한 13명(11.9%)의 계보발단자는 기관지확장증, 21명(19.2%)의 계보발단자는 척추측만증, 7명(6.4%)의 계보발단자는 승모판탈출, 6명(5.5%)의 계보발단자는 오목가슴이 일촌 및 이촌 친척 중에 있다고 응답했다.[36]

5. 최근의 엑솜 염기서열 분석 연구와 제한점

아직까지 비결핵항산균 폐질환에 대한 감수성과 관련된 유전적 소인은 충분히 밝혀지지 않았다. 최근 비결핵항산균 폐질환을 가진 서양인을 대상으로 한 엑솜 염기서열 분석(whole-exome sequencing, WES) 연구가 보고되었다 (표 2). 첫 번째 연구는 69명의 비결핵항산균 폐질환자와 18명의 비결핵항산균 폐질환이 없는 가족 구성원에 대해 엑솜 염기서열 분석을 시행하고 1000 Genomes Project 참가자를 대조군으로 삼은 연구이다.[37] 후보 유전자들을 면역에 관여하는 단백질 코딩 유전자, *CFTR* (cystic fibrosis transmembrane conductance regulator) 유전자, 섬모(cilia) 관련 유전자, 그리고 결합조직(connective tissue) 관련 유전자의 4개의 카테고리로 나누어 유전자 세트(gene set) 분석을 시행했다. 그 결과는 비결핵항산균 폐질환자는 비결핵항산균 폐질환이 없는 가족 구성원 또는 대조군에 비해 모든 유전자 카테고리에서 유전적 변이(genetic variant)를 가지고 있을 가능성이 더 높았다. 저자들은 비결핵항산균 폐질환은 복합 유전성(multigenic) 질환이며, 여러 시스템에 속한 여러 유전자들의 변이(variations)와 그 조합에 의해 발생한다고 결론 지었다.[37] 그러나, 이 연구는 비결핵항산균 폐질환에 대한 감수성과 관련된 특정 유전적 변이를 밝혀내지 못했다.

표 2. 비결핵항산균 폐질환에 대한 엑솜 염기서열 분석 연구

Author	Country	Gene	Population	Findings
Szymanski et al.[37]	USA	Exome-wide	69 patients with NTM-PD, 18 unaffected family members	A significantly higher number of patients with NTM-PD than the other individuals had low-frequency, protein-affecting variants in immune, *CFTR*, cilia, and connective tissue categories (35, 26, 90, and 90%, respectively).
Becker et al.[38]	USA	*MST1R*	11 patients (2 related) with NTM-PD and pectus excavatum and scoliosis (Lady Windermere syndrome)	Rare missense variant (p.V900M) in *MST1R* was found in both related patients and 2 patients with sporadic disease.
Chen et al.[39]	USA	Exome-wide	9 families with NTM-PD (16 affected, 20 unaffected), 57 patients with sporadic NTM-PD	The genome-wide variant-level analysis identified a region on chromosome 6q12-6q16; gene-level linkage analysis of chromosome 6 found suggestive linkage of the *TTK* gene as well as *TPBG, ORC3* and *ANKRD6*; further genome-wide gene-level association analysis identified variants in *MAP2K4,*

				RCOR3, KRT83, IFNLR1 and *SLC29A1* associated with NTM-PD.

ANKRD6: ankyrin repeat domain 6; *CFTR*: cystic fibrosis transmembrane conductance regulator; *IFNLR1*: interferon lambda receptor 1; *KRT83*: keratin 83; *MAP2K4*: mitogen-activated protein kinase kinase 4; *MST1R*: macrophage stimulating 1 receptor; NTM-PD: non-tuberculous mycobacterial pulmonary disease; *ORC3*: origin recognition complex subunit 3; *TPBG*: trophoblast glycoprotein; *TTK*: TTK protein kinase; *RCOR3*: REST corepressor 3; *SLC29A1*: solute carrier family 29 member 1; USA: United States of America.

두 번째 엑솜 염기서열 분석 연구는 동일한 코호트에서 같은 연구자 그룹에서 의해 수행되었으며, 비결핵항산균 폐질환자가 있는 아홉 가족에서 전장유전체 연관분석(linkage analysis)을 시행했다.[39] 저자들은 6번 염색체 장완(6q12-6q16)에 위치한 *TTK* (TTK protein kinase)[39] 유전자를 비결핵항산균 폐질환에 대한 감수성 유전자로 특정했다.

세 번째 엑솜 염기서열 분석 연구는 Lady Windermere 증후군이라고도 불리는 비결핵항산균 폐질환과 오목가슴, 척추측만증이 있는 11명의 환자에 대한 연구이다. 저자들은 *MST1R* (macrophage stimulating 1 receptor)[38] 유전자의 드문 변이가 Lady Windermere 증후군에 대한 감수성에 기여할 것이라고 제시했다.

그러나, 비결핵항산균 폐질환자 중 극히 일부만이 이러한 유전적 변이들로 이 질병의 발생이 설명될 수 있으며, 이러한 연관성은 독립적인 다른 인구 집단에서 재현되지 않았다는 점에서 상기 기술한 엑솜 염기서열 분석 연구들은 한계점을 가지고 있다.[32]

6. 연구의 필요성

우리나라를 포함하여 세계적으로 비결핵항산균 폐질환이 크게 증가하고 있으나, 여전히 비결핵항산균 폐질환의 치료 성적은 만족스럽지 못하고, 이 질환으로 인한 사망과 사회·경제적 비용은 증가하고 있다. 질병의 발생에 유전과 환경, 그리고 그 상호작용이 복합적으로 작용할 것으로 생각되나, 아직 비결핵항산균 폐질환에 대한 감수성을 결정하는 유전적 소인에 대한 연구가 부족하다. 비결핵항산균 폐질환의 비교적 뚜렷한 표현형이 존재한다는 점과, 인종적 차이를 보고한 역학연구, 그리고 비결핵항산균 폐질환의 가족 클러스터링을 보고한 연구 등은 여전히 이 질환에 대한 유전적 소인이 존재한다는 점을 시사한다. 이에 본 연구에서는, 비결핵항산균 폐질환이 급증하고 있는 한국인에서 비결핵항산균 폐질환에 대한 유전적 소인을 규명하기 위해 전장유전체연관분석연구(genome-wide association study, GWAS)를 수행했다.

II. 연구 방법

1. 발굴 코호트(discovery cohort)

2011년 7월 1일부터 2017년 2월 28일까지 서울대학교병원의 전향적 비결핵항산균 폐질환 코호트[40]에 참여한 비결핵항산균 폐질환자 412명이 발굴 코호트(discovery cohort)에 포함되었다. 비결핵항산균 폐질환의 진단은 2007년 미국흉부학회(American Thoracic Society)/미국감염병학회(Infectious Diseases Society of America)의 비결핵항산균 질환 가이드라인[41]에 따랐다. 이 가이드라인에 의하면,[41] 다음 기준을 만족하는 경우 비결핵항산균 폐질환으로 진단할 수 있다: 1) 호흡기 증상과, 공동 또는 다발성 소결절을 동반한 다병소의 기관지확장증을 보이는 흉부 X-ray 또는 전산화단층촬영(computed tomography, CT) 소견, 2) 다른 진단의 적절한 배제, 3) 객담 검사에서 최소한 2회 이상 비결핵항산균 배양 양성이거나, 기관지세척액에서 1회 이상 배양 양성인 경우, 또는 폐생검에서 육아종 등 항산균 감염의 병리학적 증거가 있고 조직 배양이 양성이거나, 또는 폐생검에서 항산균 감염의 병리학적 증거가 있으면서 1회 이상 객담 또는 기관지세척액에서 배양 양성.

대조군은 우리나라 성인 쌍둥이와 그 가족으로 구성된 쌍둥이

코호트(Healthy Twin Study, Korea)[42]에서 선택되었다. 쌍둥이 코호트는 한국인유전체역학조사사업(Korean Genome and Epidemiology Study, KoGES)의 세부 코호트 중 하나이다. 쌍둥이 코호트는 364 가족 중 총 1,056명의 참가자에 대해 유전자형(genotype)이 이미 분석되어 있는데, 유전자형 분석에는 발굴 코호트의 비결핵항산균 폐질환자의 유전자형 분석에 사용된 것과 동일한 플랫폼이 사용되었다. 대조군으로 쌍둥이 또는 그들의 가족 중 494명이 무작위로 선택되었는데, 쌍둥이의 경우 무작위로 선택된 1명만 대조군에 포함될 수 있었다.

2. 재현 코호트(replication cohort)

발굴 코호트에서 발견된 전장 유전체(genome-wide) 차원의 유의미한 단일뉴클레오타이드다형성(single-nucleotide polymorphism, SNP)의 재현 여부를 검증하기 위해, 서울대학교병원 및 분당서울대학교병원, 보라매병원에서 비결핵항산균 폐질환을 진단받은 187명의 환자를 2017년 1월 16일부터 2017년 11월 20일까지 전향적으로 등록하여 재현 코호트(replication cohort)를 구성했다. 이 코호트의 대조군은 분당서울대학교병원에서 위암을 진단받은 환자의 가족 구성원 500명과 광주치매코호트연구단(Gwangju Alzheimer's disease and Related Dementia Cohort Research Center)[43]이 구축한

코호트의 참가자 중 건강하고 인지기능이 정상인 1,195명이다.

본 연구는 서울대학교병원 (H-1605-111-763), 분당서울대학교병원 (B-1610/366-303), 조선대학교 (CHOSUN 2013-12-018-066)의 의학연구윤리심의위원회(Institutional Review Board)로부터 승인을 받았다. 모든 참가자는 서면 동의를 제공했으며 본 연구는 헬싱키 선언에 근거한 윤리규정 등에 따라 수행되었다.

3. 발굴 코호트에서 수행된 전장유전체연관분석연구

발굴 코호트의 환자군과 대조군은 298,930개의 변이(variant)에 대해 Illumina HumanCoreExome-12 v1.0A를 통해 유전자형이 결정되었다. 첫 번째 정도 관리(quality control, QC)는 환자군과 대조군을 통합(pooling)하기 전에 환자군과 대조군 각각에서 독립적으로 수행되었다. 첫 번째 정도 관리는 크게 두 단계로 이루어지는데, 참가자 정도 관리와 변이 정도 관리가 그것이다. 자세한 과정은 그림 1에 기술하였다. 간단히 설명하면, 먼저 다음 조건을 만족하는 참가자를 제거했다: 1) 성별이 실제 임상정보와 다르게 예측, 2) 유전자형 분석 성공률(call rate)이 97% 미만, 3) 이형접합률(heterozygosity rate)이 그 평균으로부터 표준편차의 3배보다 벗어남, 4) 다른 참가자들과의 identity-by-descent 추정치가 0.185보다 크고, 짝을 이룬(paired) 참가자들 중 유전자형 분석

성공률이 가장 작음 (일촌 혹은 이촌 관계를 가진 참가자를 제거하기 위한 목적). 이후 다음 조건을 만족하는 변이를 제거했다: 1) 소수 대립유전자 빈도(minor allele frequency, MAF)가 1% 미만, 2) 유전자형 분석 성공률이 95% 미만. 첫 번째 정도 관리 후 환자군과 대조군을 통합하였는데, 환자군과 대조군 사이에 모호하지 않고, 중첩되는 상염색제 변이(autosomal variant)만 병합되었다.

환자군과 대조군을 통합한 후, 두 번째 정도 관리를 수행하였는데, 이 과정은 Anderson 등이 제안한 방법을 따랐다.[44] 두 번째 정도 관리도 첫 번째 정도 관리와 마찬가지로 두 단계로 이루어졌다. 먼저, 다음 조건을 만족하는 참가자를 제거했다: 1) 유전자형 분석 성공률이 97% 미만, 2) 이형접합률이 그 평균으로부터 표준편차의 3배보다 벗어남, 3) 다른 참가자들과의 identity-by-descent 추정치가 0.185보다 크고, 짝을 이룬 참가자들 중 유전자형 분석 성공률이 가장 작음 가장 큼 (일촌 혹은 이촌 관계를 가진 참가자를 제거하기 위한 목적). 이후 다음 조건을 만족하는 변이를 제거했다: 1) 환자군과 대조군에서 유전자형 분석 성공률이 유의하게 차이가 남 ($p < 0.00001$), 2) 소수 대립유전자 빈도가 5% 미만, 3) 유전자형 분석 성공률이 97% 미만, 4) Hardy-Weinberg 평형을 만족하지 않음 (p for exact test <0.001).

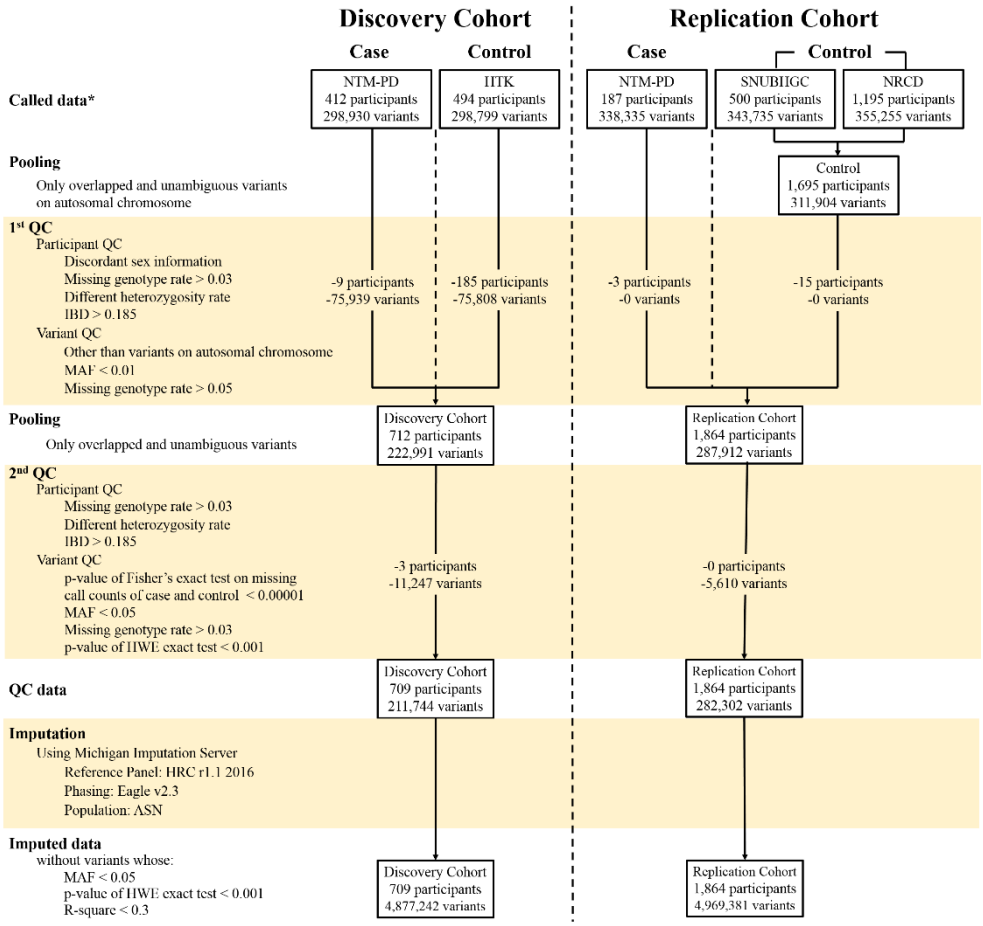

그림 1. 정도 관리(quality control)의 흐름 모식도

발굴 코호트(discovery cohort) 및 재현 코호트(replication cohort)에서 모두 동일한 정도 관리(quality control, QC) 기준이 적용되었다. 참가자에 대한 정도 관리 단계(participant QC)에서 이형접합률(heterozygosity rate)이 다르다는 것은, 이형접합률이 그 평균으로부터 표준편차의 3배보다 벗어난 것을 의미한다 (단, 재현 코호트의 마지막 정도 관리 단계에서는 4배를 기준으로 함). *재현 코호트에서 유전자형 분석(calling) 과정은 희귀한 변이(rare variant)를 제거하는 과정을 포함하고 있으며, 소수 대립유전자 빈도(minor allele frequency, MAF)가 5% 미만인 변이는 제거되었다. 이 과정에서 비결핵항산균 폐질환자군, 분당서울대병원 대조군(SNUBHGC), 광주치매코호트연구단(NRCD)의 대조군에서 각각 459,607, 450,525, 425,869개의

변이가 제거되었다. 따라서 재현 코호트의 “called data”에는 유전자형 분석(calling) 과정에서 정도 관리를 통과한 변이만 포함되어 있다.

HTK: Healthy Twin Study, Korea; HWE: Hardy-Weinberg equilibrium; IBD: identity-by-descent; MAF: minor allele frequency; NRCD: National Research Center for Dementia; NTM-PD: non-tuberculous mycobacterial pulmonary disease; SNUBHGC: Seoul National University Bundang Hospital Gastric Cancer; QC: quality control.

(출처: Cho J, Park K, et al. Thorax 2021;76:169-177)

두 번째 정도 관리를 마친 후 709명의 참가자로부터 211,744개의 변이가 남았고, Michigan Imputation Server (v1.0.2) 및 서버에 확립된 파이프라인을 이용해 결측치 추정(imputation)을 시행했다. 결측치 추정을 위한 참조(reference) 패널은 Haplotype Reference Consortium (HRC, r1.1 2016), phasing 프로그램은 Eagle (v2.3), 정도 관리 집단은 아시아인으로 선택했다.[45] HRC 패널은 UK10K 및 1000 Genomes Project를 포함한 20개의 서로 다른 연구에서 얻은 64,976개의 haplotype과 32,611개의 샘플로 구성되어 있다.[46] 결측치 추정 후, 소수 대립유전자 빈도가 5% 미만이거나, Hardy-Weinberg 평형을 만족하지 않거나 (p for exact test <0.001), R^2가 0.3보다 작은 변이는 제외했다. 최종 세트에는 총 709명 (비결핵항산균 폐질환자 403명, 대조군 306명), 4,877,242개의 변이가 포함되었다. 2명의 비결핵항산균 폐질환자는 체질량지수를 알 수 없어

체질량지수 결측치를 missForest 패키지를 사용해 추정했다 (R-package: missForest, v1.4).[47]

연관성 분석을 위해 로지스틱 회귀분석(logistic regression)을 수행했고, 나이, 나이^2, 성별, 체질량지수를 보정했다. 집단의 하위 구조를 보정하기 위해, 최종 세트에 포함된 변이들의 주성분(principal component, PC)을 계산했고, 가장 큰 고유값 10개에 해당하는 주성분 점수 10개가 공변량으로 포함되었다.[48] PLINK (v1.90b3.44)[49] 또는 R (v3.4.4)을 이용해 두 번째 정도 관리 및 연관성 분석을 시행했다.

4. 재현 코호트를 통한 검증(validation)

재현 코호트의 샘플은 한국인칩(Korean Chip, K-CHIP) 컨소시엄을 통해 제공되는 Affymetrix Axiom K-CHIP을 사용하여 유전자형을 분석했다. 질병관리청 국립보건연구원에서 2015년 한국인칩사업을 통해 한국인칩을 개발하였다 (4845-301, 3000-3031). 한국인칩은 827,783개의 변이를 포함하고 있다. 발굴 코호트와 마찬가지로, 재현 코호트에서도 첫 번째 정도 관리는 환자군과 대조군을 통합하기 전에 환자군과 대조군 각각에서 독립적으로 수행되었다 (그림 1). 재현 코호트의 데이터는 CEL 파일 형식으로 수집되었고, Affymetrix Power Tools (v1.18)과 Axiom Analysis Suite (v3.1.51)를 사용하여 유전자형을 분석(calling)하고, 이를 PLINK 파일

형식으로 변환했다. 유전자형 분석은 비결핵항산균 폐질환자군과 두 개의 대조군에서 각각 개별적으로 수행되었다. 비결핵항산균 폐질환자군의 유전자형 분석 단계에서, 다음 조건을 만족하는 참가자를 제거했다: 1) dish QC 값이 82% 미만, 2) 유전자형 분석 성공률이 97% 미만, 3) 플레이트의 평균 분석 성공률(call rate)과 합격률(pass rate)이 각각 98%, 95% 미만. 이후 다음 조건을 만족하는 변이를 제거했다: 1) 유전자형 분석 성공률이 95% 미만, 2) 소수 대립유전자 빈도가 5% 미만, 3) Fisher’s linear discriminant가 3.6 미만, 4) 이형 접합 강도 오프셋(heterozygous strength offset)이 −0.1 미만 (이외 기타 조건은 프로그램의 기본 옵션 설정을 적용함). 대조군의 유전자형 분석 단계도 위와 거의 동일한 단계를 따랐다. 다른 점은 플레이트의 평균 분석 성공률 99% 미만인 참가자를 제거했고, 유전자형 분석 성공률이 97% 미만인 변이를 제거했으며, 이외 추가로 플레이트 간의 유전자형 분석 성공률이 유의하게 차이가 나는 경우 ($p < 0.00001$) 제거했다. 이후 두 개의 대조군을 통합하였는데, 두 대조군 사이에 모호하지 않고, 중첩되는 변이만 병합되었다. 이후 비결핵항산균 폐질환자군과 대조군은 발굴 코호트와 동일한 정도 관리 단계를 거쳤다. 환자군과 대조군을 통합할 때는 환자군과 대조군 사이에 모호하지 않고, 중첩되는 상염색체 변이만 병합되었다.

환자군과 대조군을 통합한 후, 발굴 코호트에서 사용된 것과 거의 동일한 두 번째 정도 관리를 시행했다. 다른 점은 이형접합률이 그

평균으로부터 표준편차의 4배보다 벗어나는 참가자를 제거한 점이다. 정도 관리를 마친 후 864명의 참가자로부터 282,302개의 변이가 남았고, 결측치 추정(imputation)을 시행했다. 결측치 추정 후, 발굴 코호트에서 시행한 것과 동일하게, 소수 대립유전자 빈도가 5% 미만이거나, Hardy-Weinberg 평형을 만족하지 않거나 (p for exact test <0.001), R^2가 0.3보다 작은 변이는 제외했다 (그림 1). 최종 세트에는 총 1,864명 (비결핵항산균 폐질환자 184명, 대조군 1,680명), 4,969,381개의 변이가 포함되었다. 연관성 분석을 위해 로지스틱 회귀분석을 수행했고, 발굴 코호트에서와 같은 공변량을 보정했으나, 체질량지수는 재현 코호트의 대조군에 대해서는 정보가 없어서 보정하지 못했다.

5. 다차원 척도법(multidimensional scaling, MDS)

다차원 척도법(multidimensional scaling, MDS) 분석을 수행했으며, 결측치 추정(imputation) 후 참가자 간의 원시 해밍(Hamming) 거리를 기반으로 플롯을 만들었다. 발굴 및 복제 코호트 참가자와 1000 Genomes Phase 3 참가자 간의 유전적 거리를 나타내기 위해 상위 2개 차원을 사용했다.

6. 지역 플롯(regional plot)

발굴 코호트에서 수행한 전장유전체연관분석연구 결과를 기반으로 LocusZoom을 사용하여 가장 유의한 변이 주변에 대한 지역 플롯을 그렸다 (그 결과 7번 염색체의 43.6Mb에서 44.0Mb까지). 설정은 대부분 LocusZoom의 기본 옵션을 적용했고, 다만, 연관 불균형 (linkage disequilibrium, LD) 집단은 1000 Genomes (2014년 11월 기준)의 동아시아인으로 지정했다.

7. 전장유전체 복합 형질 분석(genome-wide complex trait analysis)

전장유전체 복합 형질 분석(genome-wide complex trait analysis, GCTA) (v1.26.0)을 이용해서 비결핵항산균 폐질환에 대한 SNP-유전성(heritability)을 추정했다.[50] 유전체 기반의 제한된 최대우도 추정법(restricted maximum likelihood)을 사용했고, 최종 세트에 포함된 변이 중 소수 대립유전자 빈도가 0.1보다 큰 변이만 추정에 사용되었다. 상위 10개의 주성분 점수를 제외한 GWAS에 사용된 모든 공변량을 포함시켰다. SNP-유전성을 추정하기 위해서는 우리나라 비결핵항산균 폐질환의 평생 유병률을 알아야 하므로, 이를 계산했는데, 알려진 비결핵항산균 폐질환의 시점 유병률[2] (2016년 기준

인구 100,000명당 39.6건)에 평균 수명 (70년)을 곱해서 구했다.[51] 이렇게 계산된 비결핵항산균 폐질환의 평생 유병률 추정치는 0.028였다.

8. 토폴로지 연관 도메인 분석(topologically associating domain analysis)

전장유전체연관분석연구를 통해 유의할 것으로 추정된 SNP 주변의 크로마틴 상호작용을 알아보기 위해 3D Genome Browser (http://promoter.bx.psu.edu/hi-c/)를 이용했다. 기본 옵션으로 종은 인간, 어셈블리는 hg19, Hi-C 데이터 소스는 Donor-LG1-raw 타입의 폐 조직을 선택했다. 유의할 것으로 추정된 SNP 주변의 토폴로지 연관 도메인(topologically associating domain, TAD)을 확인하고, TAD에 속하는 단백질 코딩 유전자를 조사했다.

9. 발현 양적 형질 유전자자리(expression quantitative trait loci, eQTL)와 PrediXcan 분석

관심 있는 SNP와 관련된 유전자와 그 발현 수준을 알아 보기 위해 Genotype-Tissue Expression (GTEx) Portal v7 (https://www.gtexportal.org/) 데이터세트를 이용하여 발현 양적 형질 유전자자리(expression quantitative trait loci, eQTL) 분석을 수행했다.

PrediXcan 분석은 GTEx v7 PredictDB를 사용했고, 본 연구진의 결측치 추정(imputation) 후의 유전자형 데이터를 적용하여 폐 조직에서 관심 있는 유전자의 발현 수준을 예측했다.[52] 선형 회귀를 통해 비결핵항산균 폐질환자군과 대조군 사이의 유전자 발현 수준을 비교했고, 전장유전체연관분석연구에 사용된 공변량을 모두 보정했다.

10. 2-표본 멘델 무작위배정(two-sample Mendelian randomization) 분석

일련의 과정에서 후보 유전자로 선택된 유전자의 발현 정도가 비결핵항산균 폐질환의 발생 위험에 미치는 인과적 관계를 확인하기 위해 2-표본 멘델 무작위배정(two-sample Mendelian randomization, MR) 분석을 수행했다 (그림 2). 먼저, GTEx v7 데이터세트의 폐 조직에서 후보 유전자의 발현 수준과 유의한 관계가 있는 유전적 변이를 찾았다. 찾은 변이들에 대해 본 연구진의 발굴 코호트에서 다른 변이와의 쌍(pairwise) 비교를 통해 상관 관계가 0.2보다 큰 변이는 제거했다. 남은 변이를 도구 변수(instrument variable, IV)로 삼고, 발굴 코호트 및 GTEx 데이터세트 분석의 요약 통계(summary statistics)를 적용하여 도구 변수들 간의 상관 관계를 보정한 MR-Egger 회귀를 수행했다.[53] 변이와 비결핵항산균 폐질환의 연관 추정치는 본 연구진의 발굴 코호트에서 얻었고, 변이와 후보 유전자와의 연관

추정치는 GTEx 데이터세트에서 얻었다. 후보 유전자의 발현 수준에 대한 상대적으로 작은 비율의 분산을 설명하는 약한 도구 변수를 감지하기 위해[54] 동일한 도구 변수 및 요약 통계를 사용해서 MR-pleiotropy residual sum and outlier (MR-PRESSO) 테스트를 적용했다.[55] 도구 변수들 간의 상관 관계를 보정하기 위해, 비결핵항산균 폐질환의 위험에 대한 후보 유전자의 발현 수준의 인과적 영향(causal effect)을 나타내는 계수 β의 신뢰구간 및 유의확률(p value)은 100,000번의 부트스트랩(bootstrap) 복제에서 추정되었다. 각 부트스트랩 복제는 GTEx v7 데이터세트의 폐 조직에서 후보 유전자의 발현 수준과 유의한 관계가 있어서 선택된 (상관 관계가 높은 변이를 제거하기 전) 유전적 변이 세트에서 복원 무작위 표본 추출(random sampling with replacement) 방법으로 만들었다. 각각의 MR-PRESSO 테스트에도 동일한 수의 도구 변수가 사용되었다.

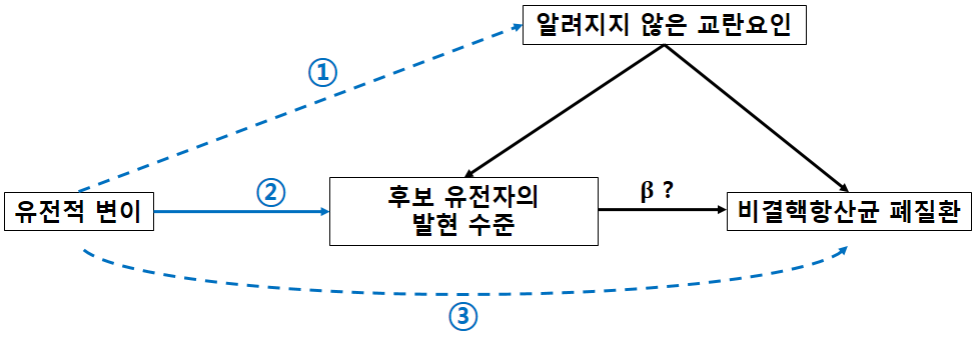

그림 2. 멘델 무작위배정(two-sample Mendelian randomization)의 가정

가정 ①: 유전적 변이(genetic variant)는 알려지지 않은 교란요인으로부터 독립적이며, 가정 ②: 유전적 변이는 후보 유전자(candidate gene)의 발현 수준과 관련이 있고, 가정 ③: 유전적 변이는 비결핵항산균 폐질환의 위험과는 독립적이며, 비결핵항산균 폐질환의 위험은 알려지지 않은 교란요인 및 후보 유전자의 발현 수준의 영향을 받는다. 계수 β는 비결핵항산균 폐질환의 위험에 대한 후보 유전자의 발현 수준의 인과적 영향(causal effect)을 나타낸다. 점선은 가정에 대한 잠재적인 위반을 의미한다.

(출처: Cho J, Park K, et al. Thorax 2021;76:169-177)

11. 비결핵항산균에 감염된 대식세포(macrophage)의 전사체(transcriptome) 프로파일링 분석

공개적으로 이용 가능한 데이터세트인 GSE72821[56] 및 E-MTAB-1101[57]을 사용하여 전사체(transcriptome) 프로파일링 분석을 수행했다. GSE72821은 THP-1 유래 대식세포(macrophage)에 Smooth 및 Rough 형태의 *Mycobacterium abscessus*를 감염시킨 후 1, 4, 24시간 후에 각각 감염된 대식세포에서의 전사 이벤트를 보여주는 mRNA 염기서열 분석 데이터세트이다.[56] *M. abscessus*에 감염된 대식세포 18개 샘플을 사례군, 감염되지 않은 대식세포 9개 샘플을 대조군으로 삼았다. 발현 수준 합계가 전체 샘플 크기의 10분의 1보다 작은 유전자를 제거한 후, DESeq2 (v1.14.1)[58]를 사용하여 사례군과 대조군 사이에 차별발현유전자(differentially expressed gene, DEG)를 확인했다.

E-MTAB-1101은 인간 유래의 대식세포에 *Mycobacterium avium*을 감염시킨 후 전사 반응을 보여주는 발현 마이크로어레이 데이터세트이다.[57] 6명의 공여자로부터 얻은 CD14+세포를 5일 동안 배양하여 대식세포로 성숙되게 한 후, 두 개의 *M. avium* subsp. *hominissuis* 분리주 및 하나의 *M. avium* subsp. *avium* 분리주로 24시간 동안 감염시킨다. 감염된 대식세포 9개 샘플을 사례군, 감염되지 않은 대식세포 3개 샘플을 대조군으로 삼았다. 정규화(normalization)

후, 데이터 분석에는 limma (v3.36.2)를 사용했다.[59]

12. Gene Ontology enrichment 분석

GSE72821 및 E-MTAB-1101의 유전자 발현 분석에서 false discovery rate (FDR)가 0.05 미만인 유전자를 DEG로 정의했다. 이러한 DEG를 Gene Ontology (GO) 컨소시엄 웹사이트 (https://geneontology.org/)로 가져와 Fisher의 정확한 검정을 수행했다.[60] 2020년 2월 21일 데이터베이스를 기반으로 분석 결과를 보고했다.

13. 기존에 알려진 유전자에 대한 흔한 변이(common variant) 분석과 유전자 세트(gene set) 분석

기존의 연구에서 비결핵항산균 폐질환과의 연관성을 보고한 유전자에 대해 본 연구진의 발굴 코호트에서 그 연관성이 재현되는지 확인하기로 했다. 기존의 연구에서 비결핵항산균 폐질환과 관련이 있다고 보고된 43개의 유전자를 선택했다.[32,37-39,61,62] University of California Santa Cruz 테이블 브라우저 (https://genome.ucsc.edu/cgi-bin/hgTables)를 이용해 43개의 유전자 각각의 전사체와 0.1MB 측면 영역을 확인했다. 비결핵항산균

폐질환과의 연관성이 가장 강한 흔한 변이(common variant)를 해당 유전자마다 하나씩 보고했다. 또한 최적의 염기서열 커널 연관성 검사(optimal sequence kernel association test)를 이용해 유전자 세트(gene set) 분석을 수행했다.[63] 비결핵항산균 폐질환과의 연관성이 가장 강한 흔한 변이에 대해 GTEx v8 데이터세트를 이용해 eQTL 분석을 수행했다.

III. 결 과

1. 전장유전체연관분석으로 얻은 잠재적 후보 SNP

비결핵항산균 폐질환의 질병 감수성 유전자자리(locus)를 확인하기 위해, 발굴 코호트의 비결핵항산균 폐질환자 403명과 대조군 306명의 4,877,242개의 변이에 대해 전장유전체연관분석연구를 수행했다. 발굴 코호트의 환자군과 대조군의 임상적 특징은 표 3과 4에 나와 있다. MDS 분석을 통해 발굴 코호트와 1000 Genomes Phase 3 데이터의 집단간 유전적 이질성(genetic heterogeneity)이 존재하는 것으로 나타났지만 (그림 3A), 발굴 코호트의 환자군과 대조군 사이의 유전적 이질성은 없었다 (그림 3B).

표 3. 발굴 및 복제 코호트 참가자의 임상적 특성

	Discovery cohort		Replication cohort	
	Patients with NTM-PD (n=403)	Control (n=306)	Patients with NTM-PD (n=184)	Control (n=1,680)
Age, mean (SD), yr	63.5 (10.1)	54.5 (11.6)	65.8 (9.9)	67.3 (11.7)
Women, n (%)	265 (65.8)	175 (57.2)	123 (66.8)	1009 (60.1)
BMI, mean (SD), kg/m^2*	20.7 (2.6)	24.4 (3.1)	20.8 (2.6)	ND
Causative organism, n (%)†				
Mycobacterium avium complex	304 (75.8)	–	146 (79.8)	–
Mycobacterium abscessus complex	80 (20.0)	–	32 (17.5)	–
Others	17 (4.2)	–	5 (2.7)	–
Radiological types, n (%)				
Nodular bronchiectatic	333 (82.6)	–	145 (78.8)	–
Fibrocavitary	66 (16.4)	–	37 (20.1)	–
Unclassifiable	4 (1.0)	–	2 (1.1)	–
History of NTM treatment, n (%)ǂ	190 (47.1)	–	101 (54.9)	–
No. of variants after imputation	4,877,242		4,969,381	

*2명의 비결핵항산균 폐질환자는 체질량지수를 알 수 없어 결측치 추정(imputation)을 시행했다.

†비결핵항산균 폐질환자 중 발굴 코호트에서 2명과 재현 코호트에서 1명은 비결핵항산균의 종이 확인되지 않았다. 한 명의 환자에서 *Mycobacterium avium* 복합체와 *Mycobacterium abscessus* 복합체가 모두 분리된 적이 있는 환자는 발굴 코호트에서 64명(16.0%), 복제 코호트에서 25명 (13.7%) 있었다. 이런 환자들은 *M. avium* 복합체에 해당하는 종이 *M. abscessus* 복합체에 해당하는 종보다 더 자주 분리된 경우, *M. avium* 복합체 폐질환으로 분류되었다. 그 반대의 경우는 *M. abscessus* 복합체 폐질환으로 분류되었다. *M. avium* 복합체에 해당하는 종과 *M. abscessus* 복합체에 해당하는 종이 분리된 횟수가 같으면, 첫 번째 분리된 종을 기준으로 분류했다.

†비결핵항산균 폐질환의 치료 여부는 발굴 코호트의 경우 2017년 5월 18일, 재현 코호트의 경우 2017년 11월 29일에 조사되었다.

BMI: body mass index; ND: not determined; NTM: non-tuberculous mycobacteria; NTM-PD: NTM pulmonary disease.

(출처: Cho J, Park K, et al. Thorax 2021;76:169-177)

표 4. 비결핵항산균 폐질환자의 동반질환

	Discovery cohort (n=403)	Replication cohort (n=184)
Past medical history		
Pulmonary tuberculosis	148/390 (38.0%)	77/183 (42.1%)
Measles	79/272 (29.0%)	38/150 (25.3%)
Pertussis	16/238 (6.7%)	10/134 (7.5%)
Comorbidities		
Diabetes mellitus	27/403 (6.7%)	10/184 (5.4%)
Malignancy	37/403 (9.2%)	15/184 (8.2%)
Asthma	6/403 (1.5%)	6/184 (3.3%)
Chronic obstructive pulmonary disease *	59/333 (17.7%)	27/115 (23.5%)
Bronchiectasis	383/403 (95.0%)	171/184 (92.9%)
Liver cirrhosis	3/403 (0.7%)	2/184 (1.1%)
Chronic kidney disease	2/403 (0.5%)	2/184 (1.1%)
Solid organ transplantation	1/403 (0.3%)	0/184 (0.0%)

*만성폐쇄성폐질환(chronic obstructive pulmonary disease, COPD)의 정의: 기관지확장제 흡입 후 1초간 강제날숨량(forced expiratory volume in 1 s)/강제폐활량(forced vital capacity) <0.7

(출처: Cho J, Park K, et al. Thorax 2021;76:169–177)

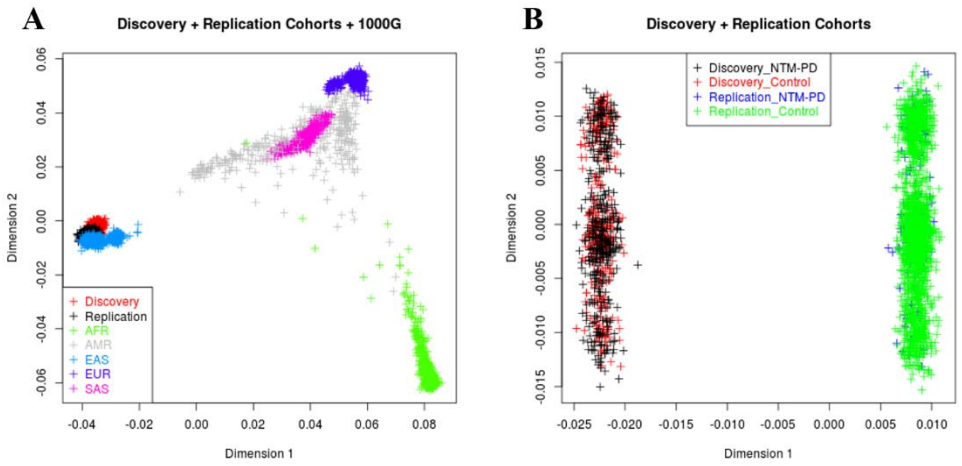

그림 3. 다차원 척도법(multidimensional scaling, MDS) 플롯

(A) 발굴 및 재현 코호트와 1000 Genomes Phase 3, (B) 발굴 및 재현 코호트

AFR: African; AMR: ad mixed American; EAS: East Asian; EUR: European; NTM-PD: non-tuberculous mycobacterial pulmonary disease; SAS: South Asian.

(출처: Cho J, Park K, et al. Thorax 2021;76:169-177)

7번 염색체의 43.6 ~ 43.8MB 범위에서 1×10^{-7}의 유의수준(significance level)에 근접한 하나의 피크가 관찰되었다 (그림 4A). 이 영역에 위치한 변이 중 rs849177 (염색체 7p13; chr7:43783788)은 비결핵항산균 폐질환과 가장 유의하게 연관되어 있었다 (OR, 2.34; 95 % CI, 1.71 ~ 3.21; $p=1.36\times10^{-7}$; 표 5). 이 연관성은 체질량지수를 보정하지 않아도 일관되게 재현되었다 ($p=2.85\times10^{-7}$; 그림 5). rs849177의 위험 대립유전자(C)의 빈도는 환자군의 경우 0.70, 대조군의 경우 0.57이었다. rs849177의 유전자형

빈도는 환자군의 경우 0.49 (CC; 197명), 0.42 (CT; 168명), 0.09 (TT; 38명), 대조군의 경우 0.33 (CC; 100명), 0.49 (CT; 149명), 0.19 (TT; 57명)이었다 (표 5). 분위수 대 분위수 플롯(quantile-quantile plot)에서 genomic inflation factor가 1.03으로, 1에 매우 가까웠다 (그림 4B). 따라서 1종 오류는 잘 제어되었다는 것을 알 수 있다. 그림 4C는 rs849177 근처의 다른 변이들이 rs849177과 높은 상관 관계가 있음을 보여준다. rs849177에 대한 조건부 분석(conditional analysis)에서 rs849177 이외의 다른 유의한 SNP는 없었다 (그림 6). 그리고 비결핵항산균 폐질환에 대한 SNP-유전성은 0.225 (p=0.266)였다.

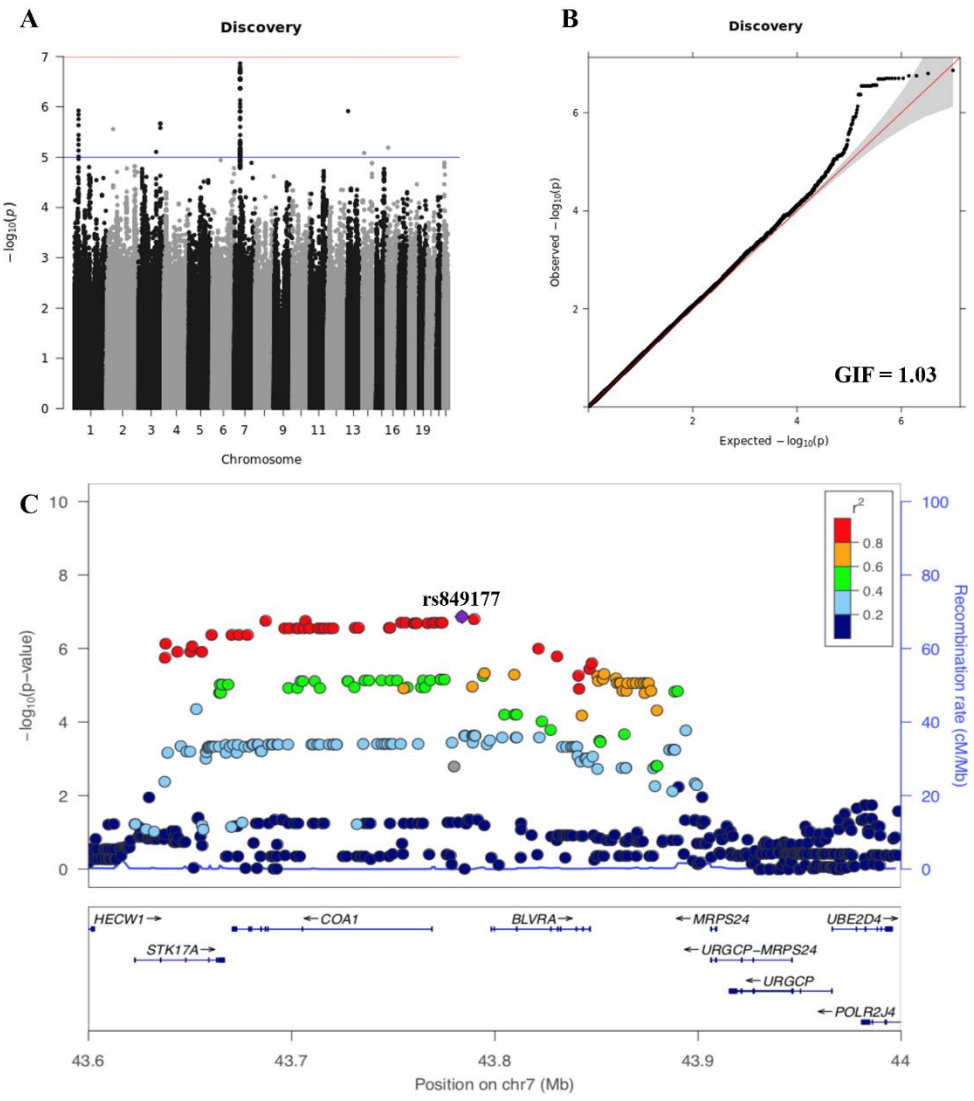

그림 4. 발굴 코호트에서 전장유전체연관분석으로 얻은 잠재적 후보 유전자자리

(A) 맨하튼 플롯(manhattan plot). 빨간색과 파란색 선은 각각 1.00×10^{-7}, 1.00×10^{-5}의 유의수준(significance level)을 나타낸다. (B) 분위수 대 분위수 플롯(quantile−quantile plot). 추정된 유의확률(p value)에 대한 95% 신뢰 구간을 회색으로 표시했다. (C) rs849177 주변의 지역 플롯(regional plot). 보라색은 rs849177을 나타내고, 다른 색상은 rs849177과 해당 변이 간의 상관 정도를 나타낸다.

Chr: chromosome; GIF: genomic inflation factor.

(출처: Cho J, Park K, et al. Thorax 2021;76:169–177)

표 5. 발굴 및 재현 코호트에서 rs849177에 대한 전장유전체연관분석 결과

Variants	Chr	Base pair	Risk/ Protective alleles	Discovery cohort				Replication cohort				Combined p value
				p value	OR (95% CI)	Risk allele frequency ‡		p value †	OR (95% CI)	Risk allele frequency ‡		
						NTM-PD	Control			NTM-PD	Control	
rs849177*	7	43783788	C/T	1.36×10^{-7}	2.34 (1.71 to 3.21)	0.70 (0.49 [197]/ 0.42 [168]/ 0.09 [38])	0.57 (0.33 [100]/ 0.49 [149]/ 0.19 [57])	0.017	1.29 (1.02 to 1.62)	0.68 (0.48 [88]/ 0.40 [73]/ 0.13 [23])	0.63 (0.39 [653]/ 0.47 [794]/ 0.14 [233])	4.92×10^{-8}

*rs849177은 결측치 추정(imputation)을 통해 얻은 SNP이다. 결측치 추정을 통해 얻은 유전자형과 실제 유전자형 간의 추정된 상관계수의 제곱값은 발굴 및 재현 코호트에서 각각 0.989 및 0.902이다.

†단측 검정 유의확률(p value).

‡위험 대립유전자(risk allele) 동형 접합 (CC), 이형 접합 (CT), 방어 대립유전자(protective allele) 동형 접합 (TT) 순서로, 유전자형 빈도는 괄호 안에, 정확한 유전자형 수는 대괄호 안에 있다.

Chr: chromosome; NTM-PD: non-tuberculous mycobacterial pulmonary disease.

(출처: Cho J, Park K, et al. Thorax 2021;76:169–177)

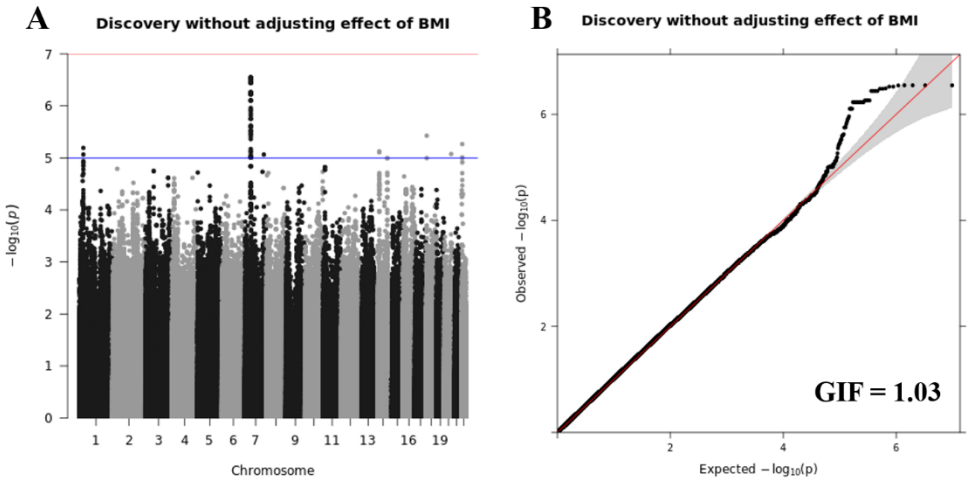

그림 5. 발굴 코호트에서 체질량지수를 보정하지 않고 수행한 로지스틱 분석

(A) 맨하튼 플롯(manhattan plot). 빨간색과 파란색 선은 각각 1.00×10−7, 1.00×10−5의 유의수준(significance level)을 나타낸다. (B) 분위수 대 분위수 플롯(quantile−quantile plot). 추정된 유의확률(p value)에 대한 95% 신뢰 구간을 회색으로 표시했다.

BMI: body mass index; GIF: genomic inflation factor.

(출처: Cho J, Park K, et al. Thorax 2021;76:169–177)

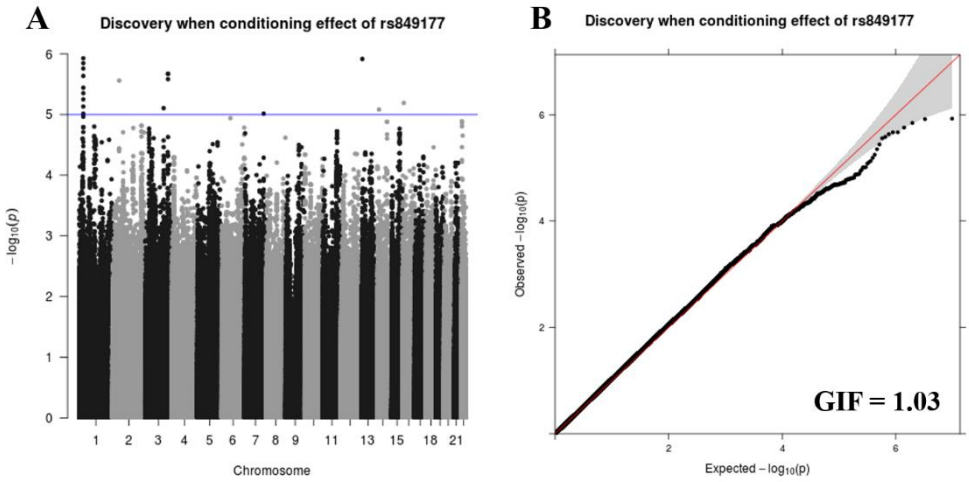

그림 6. 발굴 코호트에서 rs849177에 대한 조건부 분석(conditional analysis)

(A) 맨하튼 플롯(manhattan plot). 파란색 선은 1.00×10^{-5}의 유의수준(significance level)을 나타낸다. (B) 분위수 대 분위수 플롯(quantile−quantile plot). 추정된 유의확률(p value)에 대한 95% 신뢰 구간을 회색으로 표시했다.

GIF: genomic inflation factor.

(출처: Cho J, Park K, et al. Thorax 2021;76:169-177)

rs849177과 비결핵항산균 폐질환의 연관성은 *M. avium* 복합체 폐질환자 (OR, 2.40; 95% CI, 1.71 ~ 3.37; $p=4.30\times10^{-7}$)와 *M. abscessus* 복합체 폐질환자에서 일관되게 나타났다 (OR, 2.14; 95% CI, 1.27 ~ 3.61; $p=4.21\times10^{-3}$; 표 6). 유의확률(p value)의 차이는 주로 표본 크기의 차이와 관련이 있다고 추정되었고, 두 그룹에서 추정된 rs849177과 각 비결핵항산균 폐질환의 연관성에는 유의한 차이가 없었다 (p=0.769).

표 6. 발굴 코호트에서 rs849177과 *Mycobacterium avium* 복합체 폐질환 및 *Mycobacterium abscessus* 복합체 폐질환과의 연관성

Variants	*M. avium* complex pulmonary disease (n=304)			*M. abscessus* complex pulmonary disease (n=80)		
	OR (95% CI)	p value	Risk allele frequency*	OR (95% CI)	p value	Risk allele frequency*
rs849177	2.40 (1.71 to 3.37)	4.30×10^{-7}	0.70 (0.49/0.41/ 0.10)	2.14 (1.27 to 3.61)	4.21×10^{-3}	0.68 (0.46/0.44/ 0.10)

*rs849177의 위험 대립유전자(risk allele)는 C이다. 유전자형 빈도는 괄호 안에, 위험 대립유전자 동형 접합 (CC), 이형 접합 (CT), 방어 대립유전자(protective allele) 동형 접합 (TT) 순서로 기록했다.

(출처: Cho J, Park K, et al. Thorax 2021;76:169-177)

2. 잠재적 후보 SNP의 검증

다음으로 비결핵항산균 폐질환자 184명과 대조군 1,680명의 재현 코호트를 이용하여 전장유전체연관분석 결과를 검증했다. 재현 코호트의 환자군과 대조군의 임상적 특징은 표 3과 4에 나와 있다. MDS 분석은 재현 코호트에서 환자군과 대조군 사이에 전반적으로 유전적 이질성이 없음을 보여주었다 (그림 3B). 재현 코호트에서 rs849177의 공칭 유의확률(nominal p value)은 0.017 (OR, 1.29; 95% CI, 1.02 ~ 1.62; 표 5)였다. 그리고 표 5에 나타난 바와 같이, rs849177이 비결핵항산균 폐질환에 미치는 영향(effect)의 방향은 발굴 및 재현 코호트에서 같았다. Fisher의 방법을 사용하여 발굴 및 재현 코호트의 결과를 병합했을 때, rs849177의 유의확률(combined p value)은 4.92×10^{-8}였다. Fisher의 방법을 사용하여 발굴 및 재현 코호트의 결과를 병합했을 때, rs849177 이외의 다른 유의한 변이는 없었다 (표 7).

표 7. Fisher의 방법을 사용한 발굴 및 재현 코호트의 결과 병합

Chr	Region* (Mb)	Variant	Location (BP)	Combined p value	Risk/ Protective alleles	Risk allele frequency	
						Discovery (NTM-PD /Control)	Replication (NTM-PD /Control)
1	239.1-239.1	rs9428778	239093672	7.05×10^{-6}	G/A	0.10/0.06	0.12/0.07
3	-	rs12487736	47459679	9.62×10^{-6}	C/T	0.56/0.46	0.56/0.49
3	167.1-167.1	rs2699943	108875914	3.44×10^{-6}	C/T	0.07/0.03	0.11/0.06
7	**43.6-43.9**	**rs849177**	**43783788**	$\mathbf{4.92\times10^{-8}}$	**C/T**	**0.70/0.57**	**0.68/0.63**
8	-	rs28709344	143047788	5.95×10^{-6}	G/A	0.38/0.30	0.46/0.38
12	-	rs7301800	106481036	8.23×10^{-6}	G/A	0.60/0.57	0.77/0.67
13	-	rs9511572	25564326	7.62×10^{-6}	C/T	0.52/0.42	0.49/0.48
13	29.0-29.0	rs9508036	29015877	3.50×10^{-6}	G/A	0.42/0.39	0.46/0.35
15	25.1-25.1	rs72693656	25106554	7.77×10^{-6}	T/G	0.38/0.34	0.38/0.48
16	-	rs77953657	23825724	9.85×10^{-6}	C/T	0.83/0.73	0.84/0.76
21	-	rs45593631	45854740	2.11×10^{-6}	T/C	0.25/0.19	0.15/0.25

*지역(region)은 유의확률(p value) $<1.00\times10^{-5}$이고 연관 불균형(linkage disequilibrium, LD) $\gamma^2>0.1$인 변이들의 위치이며, 가장 유의하게 연관된 변이의 위치를 base pair로 나타냈다. 유의하게 연관된 변이가 하나인 경우, 지역(region)은 대시로 대치했다.

BP: base pair; Chr: chromosome; NTM-PD: non-tuberculous mycobacterial

pulmonary disease.

(출처: Cho J, Park K, et al. Thorax 2021;76:169–177)

3. 3차원 구조 분석

STK17A (serine/threonine kinase 17a), *COA1* (cytochrome c oxidase assembly factor 1 homolog), *BLVRA* (biliverdin reductase A) 유전자가 rs849177과 동일한 TAD에 있다 (그림 7).

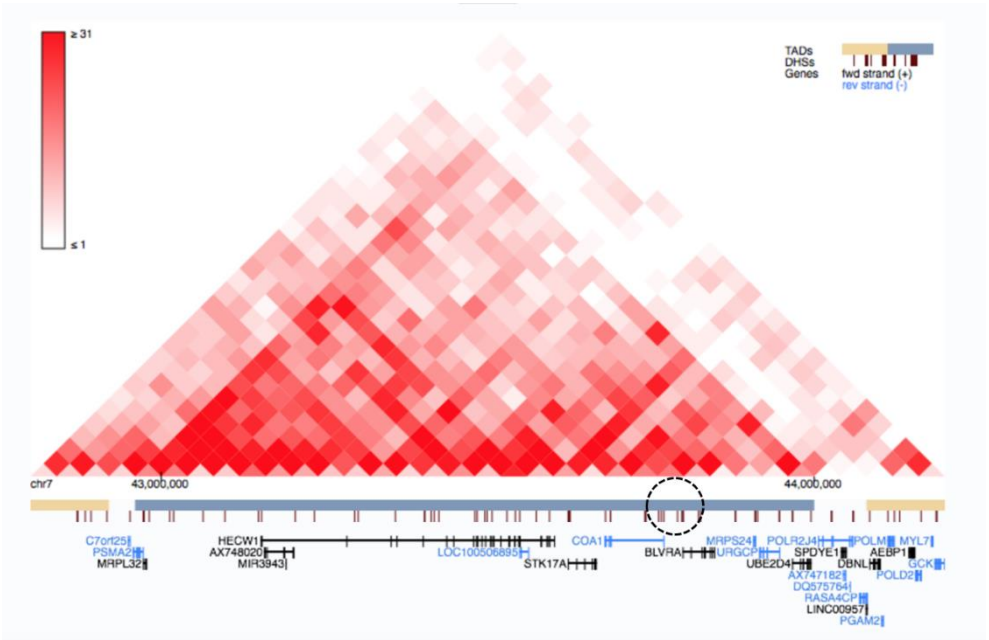

그림 7. rs849177 주변의 토폴로지 연관 도메인(topologically associating domain, TAD)

STK17A, COA1, BLVRA 유전자를 포함한 몇 개의 유전자가 rs849177 (검은 색 점선 원)과 동일한 TAD (파란색 블록)에 있다.

(출처: Cho J, Park K, et al. Thorax 2021;76:169-177)

4. 비결핵항산균 폐질환과 연관된 후보 유전자

rs849177은 GTEx 폐 조직에서 *STK17A* 유전자의 발현에 유의한 영향을 미쳤다 ($p=7.1\times10^{-19}$; 표 8). 폐 조직 이외의 GTEx 조직에서, *STK17A*, *COA1, BLVRA* 유전자의 발현 수준은 rs849177에 의해 영향을 받았다 (표 8). eQTL 분석 결과, rs849177에 위험 대립유전자(C)를 가진 사람은 GTEx 폐 조직에서 *STK17A* 유전자의 발현 수준이 낮았다 (그림 8).

표 8. Genotype-Tissue Expression (GTEx) 데이터베이스에서 rs849177에 의해 유전자 발현 수준이 크게 영향을 받는 후보 유전자와 해당 조직

Gene	p value	Tissue
STK17A	3.1×10^{-19}	Heart – Left Ventricle
	7.1×10^{-19}	Lung
	2.4×10^{-16}	Thyroid
	1.3×10^{-9}	Heart – Atrial Appendage
	1.2×10^{-5}	Skin – Not Sun Exposed
	1.4×10^{-5}	Skin – Sun Exposed
	2.0×10^{-5}	Artery – Aorta
COA1	6.8×10^{-20}	Skin – Sun Exposed
	5.8×10^{-19}	Artery – Tibial
	4.8×10^{-18}	Skin – Not Sun Exposed
	6.1×10^{-15}	Esophagus – Muscularis
	1.6×10^{-14}	Adipose – Subcutaneous
	7.2×10^{-14}	Nerve – Tibial
	2.5×10^{-13}	Artery – Aorta
	4.0×10^{-12}	Esophagus – Mucosa
	4.8×10^{-12}	Thyroid
	2.5×10^{-10}	Pancreas
	4.4×10^{-10}	Colon – Sigmoid
	6.4×10^{-10}	Colon – Transverse
	5.0×10^{-6}	Esophagus – Gastroesophageal Junction
	5.9×10^{-6}	Artery – Coronary
	7.4×10^{-6}	Adipose – Visceral
	3.2×10^{-5}	Muscle – Skeletal
	4.5×10^{-5}	Cells – Transformed fibroblasts
BLVRA	7.7×10^{-10}	Whole Blood

	1.7×10^{-8}	Esophagus – Mucosa
	1.1×10^{-6}	Artery – Tibial
POLR2J4	3.3×10^{-6}	Artery – Aorta
AC004951.5	1.1×10^{-8}	Esophagus – Mucosa
AC004985.12	1.4×10^{-6}	Esophagus – Mucosa

(출처: Cho J, Park K, et al. Thorax 2021;76:169–177)

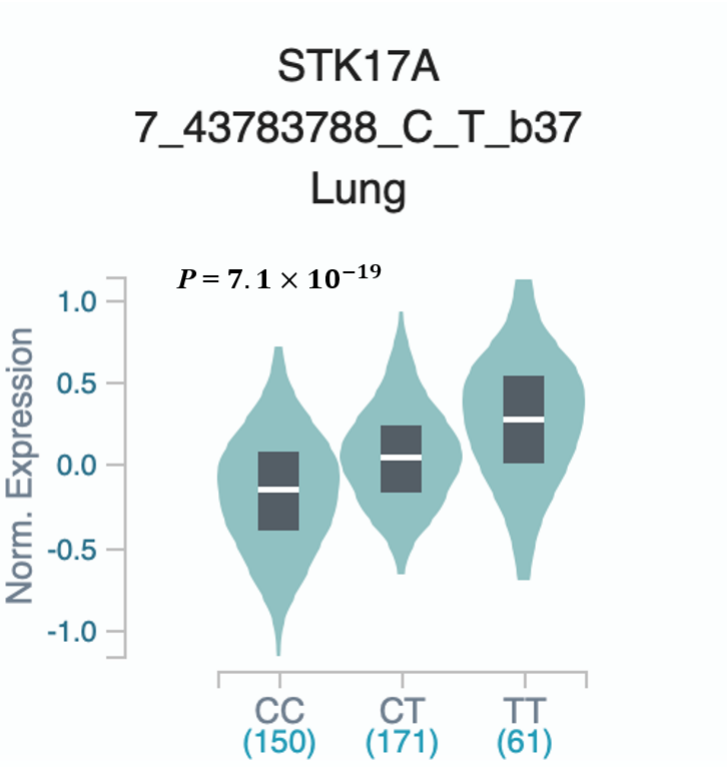

그림 8. Genotype-Tissue Expression(GTEx) 데이터베이스의 폐 조직에서 rs849177 유전자형에 따른 *STK17A* 유전자의 발현 수준

rs849177의 위험 대립유전자(risk allele)는 C이고, 방어 대립유전자(protective allele)는 T이다. 바이올린 플롯의 너비는 샘플 밀도를 나타내고, 안의 상자는 사분위수 범위를 나타내며 중심선이 중앙값이다.

CC, risk allele homozygotes; CT, risk allele heterozygotes; TT, protective allele homozygotes; *STK17A*, serine/threonine kinase 17a.

(출처: Cho J, Park K, et al. Thorax 2021;76:169-177)

다음으로, PrediXcan을 사용하여 본 연구진의 유전자형 데이터를 이용해 폐 조직 특이적인 후보 유전자의 발현 수준을 예측했다. 그 결과, *STK17A* 유전자가 발굴 코호트의 비결핵항산균 폐질환자와 대조군 사이에서 가장 유의하게 차별적으로 발현되는 유전자였고 ($p=5.00\times10^{-7}$, $FDR=2.26\times10^{-4}$; 그림 9A), 이는 재현 코호트에서 검증되었다 ($p=0.048$; 그림 9B). 또한, *STK17A* 유전자 발현 수준은 *M. avium* 복합체 폐질환과 *M. abscessus* 복합체 폐질환에서 서로 다르지 않았는데, 발굴 코호트 ($p=0.829$; 그림 9C) 및 재현 코호트 ($p=0.634$; 그림 9D)에서 마찬가지였다.

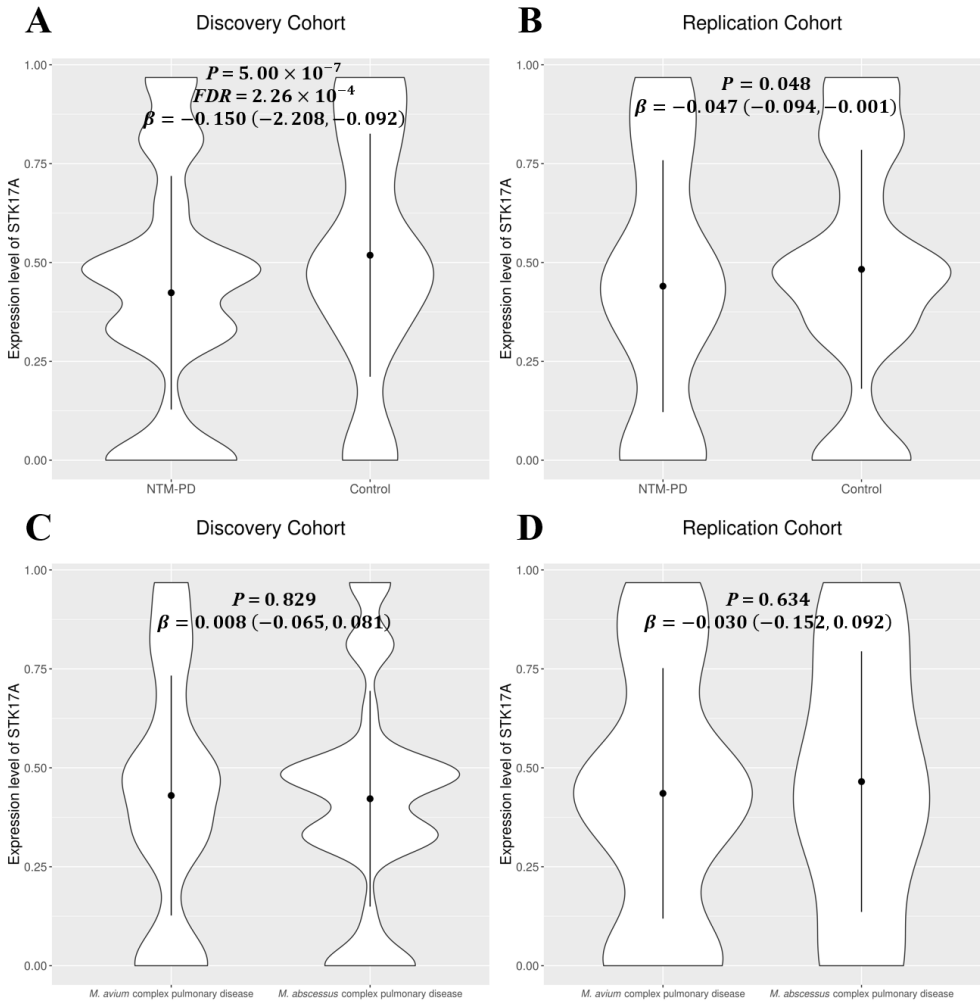

그림 9. PrediXcan을 이용해 예측한 비결핵항산균 폐질환자군과 대조군에서의 *STK17A* 유전자의 발현 수준

(A) 발굴 코호트의 비결핵항산균 폐질환자군과 대조군에서의 *STK17A* 유전자의 발현 수준, (B) 재현 코호트의 비결핵항산균 폐질환자군과 대조군에서의 *STK17A* 유전자의 발현 수준, (C) 발굴 코호트의 *Mycobacterium avium* 복합체 폐질환자군과 *Mycobacterium abscessus* 복합체 폐질환자군에서의 *STK17A* 유전자의 발현 수준, (D) 재현 코호트의 *M. avium* 복합체 폐질환자군과 *M. abscessus* 복합체 폐질환자군에서의 *STK17A* 유전자의 발현 수준

바이올린 플롯의 너비는 샘플 밀도를 나타내고, 그 안의 검은 색 점과 수직선은 각각 평균과 1 표준 편차(1 SD) 구간을 나타낸다.

FDR, false discovery rate; NTM-PD, non-tuberculous mycobacterial pulmonary disease; *STK17A*, serine/threonine kinase 17a.

(출처: Cho J, Park K, et al. Thorax 2021;76:169-177)

5. 후보 유전자가 비결핵항산균 폐질환 발생에 미치는 인과적 영향(causal effect)

MR 분석을 위해 GTEx 데이터세트의 폐 조직에서 *STK17A* 유전자의 발현 수준과 관련된 5개의 변이가 도구 변수로 선택되었다: rs7790685 (chr7:43597398), rs1100407 (chr7:43756898), rs757859 (chr7:43885849), rs929380 (chr7:43982139), rs59505066 (chr7:44108863). 선택된 변이들이 *STK17A* 유전자의 발현 수준에 미치는 영향은 이러한 변이들이 비결핵항산균 폐질환 발생에 미치는 영향과 유의한 음의 상관관계가 있었고, 이는 MR-Egger 테스트(β, −4.627; 95% CI, −8.768 ~ −0.486; p=0.029; 그림 10)나 MR-PRESSO 테스트 (β, −1.780; 95% CI, −2.974 ~ −1.618; p=0.049)에서 모두 마찬가지였다.

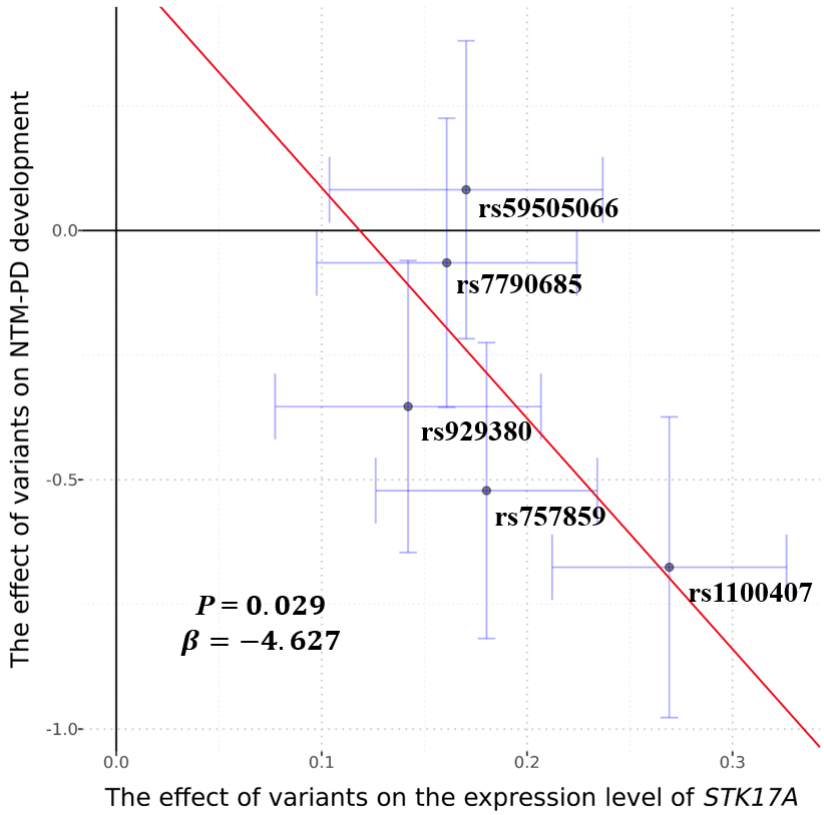

그림 10. 멘델 무작위배정 (Mendelian randomization, MR) 분석을 통해 얻은 변이-비결핵항산균 폐질환의 회귀 계수 대 변이-*STK17A* 유전자의 발현 수준의 회귀 계수 플롯

X 축은 변이가 *STK17A* 유전자의 발현 수준에 미치는 영향을 나타내고, Y 축은 변이가 비결핵항산균 폐질환 발생에 미치는 영향을 나타낸다. MR-Egger 회귀 추정치(β)는 빨간색 선의 기울기로 표시했다.

STK17A, serine/threonine kinase 17a.

(출처: Cho J, Park K, et al. Thorax 2021;76:169-177)

6. 비결핵항산균에 감염된 대식세포에서 후보 유전자의 차별적 발현(differential expression)

대식세포가 비결핵항산균에 의해 감염될 경우 *STK17A, BLVRA, COA1* 유전자의 발현 수준에 변화가 생기는지 확인하기 위해 공개적으로 사용 가능한 두 데이터세트(GSE72821 및 E-MTAB-1101)를 사용하여 대식세포 전사체 프로파일링 분석을 수행했다. *STK17A* 유전자의 발현 수준은 감염되지 않은 대식세포와 비교하여 *M. abscessus*에 감염된 대식세포 ($p=4.24\times10^{-5}$, $FDR=7.94\times10^{-4}$; 그림 11A) 및 *M. avium*에 감염된 대식세포 ($p=9.50\times10^{-3}$, $FDR=3.76\times10^{-1}$; 그림 11B)에서 모두 유의하게 증가했다 (공칭 유의수준 [nominal significance level] 0.05 기준). 그러나, *BLVRA* 유전자의 발현 수준은 두 데이터세트(GSE72821 및 E-MTAB-1101)에서 모두 감염된 대식세포와 감염되지 않은 대식세포에서 차이가 없었다 (그림 11C와 D). *COA1* 유전자의 발현 수준도 mRNA 염기서열 분석 데이터세트(GSE72821)에서 *M. abscessus*에 감염된 대식세포와 감염되지 않은 대식세포에서 차이가 없었다 (그림 11E). 발현 마이크로어레이 데이터세트(E-MTAB-1101)에 *COA1* 유전자가 포함되어 있지 않았기 때문에, *M. avium*에 감염된 대식세포에서의 *COA1* 유전자의 차별적 발현 수준은 알 수 없었다.

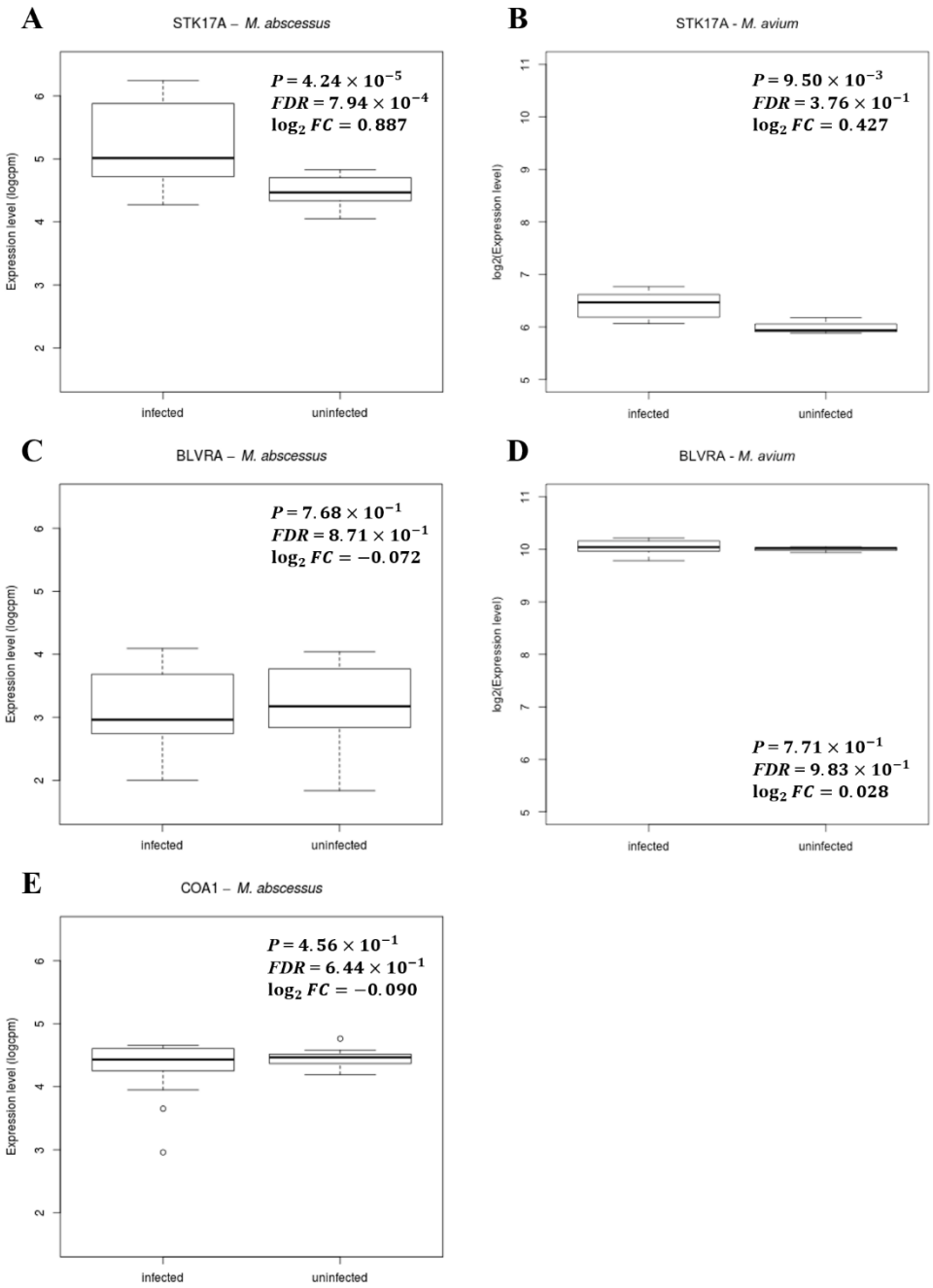

그림 11. *Mycobacterium abscessus* 또는 *Mycobacterium avium*에 감염된 또는 감염되지 않은 대식세포에서 *STK17A* (A, B), *BLVRA* (C, D), *COA1* (E)의 발현 수준

A, C, E: GSE72821 (*M. abscessus*에 감염된 대식세포)

B, D: E-MTAB-1101 (*M. avium*에 감염된 대식세포)

X 축은 비결핵항산균에 감염된 또는 감염되지 않은 대식세포를 나타내고, Y 축은 유전자의 발현 수준을 나타낸다 (log counts per million [logcpm] 값 또는 log2 값으로 표현).

BLVRA, biliverdin reductase A; FC, fold change; FDR, false discovery rate; *STK17A*, serine/threonine kinase 17a.

(출처: Cho J, Park K, et al. Thorax 2021;76:169-177)

7. 비결핵항산균에 감염된 대식세포에서 세포자멸사(apoptosis) 경로의 강화(enrichment)

GSE72821에서 유의한 DEG의 수는 2,939개였고, E-MTAB-1101에서 유의한 DEG의 수는 143개였다. GO 분석 결과, 비결핵항산균에 감염된 대식세포에서 세포자멸사 신호 전달 경로의 현저한 강화를 보여 주었다 (표 9).

표 9. Gene Ontology (GO) 분석을 통해 얻은 GSE72821 및 E-MTAB-1101 데이터세트에서 모두 유의한 세포자멸사 경로

GO biological process	GSE72821		E-MTAB-1101	
	p value	FDR	p value	FDR
Regulation of apoptotic process (GO:0042981)	1.25×10^{-17}	2.81×10^{-15}	1.28×10^{-9}	3.93×10^{-7}
Positive regulation of apoptotic process (GO:0043065)	8.80×10^{-14}	1.20×10^{-11}	3.87×10^{-6}	2.88×10^{-4}
Regulation of apoptotic signaling pathway (GO:2001233)	2.20×10^{-11}	2.24×10^{-9}	7.65×10^{-8}	9.90×10^{-6}
Negative regulation of apoptotic process (GO:0043066)	1.55×10^{-10}	1.44×10^{-8}	1.02×10^{-9}	3.31×10^{-7}
Apoptotic process (GO:0006915)	1.27×10^{-9}	1.05×10^{-7}	2.98×10^{-8}	4.52×10^{-6}
Regulation of extrinsic apoptotic signaling pathway (GO:2001236)	1.31×10^{-7}	7.82×10^{-6}	3.55×10^{-7}	3.65×10^{-5}
Regulation of intrinsic apoptotic signaling pathway (GO:2001242)	1.64×10^{-6}	7.89×10^{-5}	3.42×10^{-4}	1.23×10^{-2}

Apoptotic signaling pathway (GO:0097190)	2.26×10^{-6}	1.06×10^{-4}	1.17×10^{-3}	3.24×10^{-2}
Regulation of cysteine-type endopeptidase activity involved in apoptotic process (GO:0043281)	2.56×10^{-6}	1.18×10^{-4}	3.89×10^{-5}	2.15×10^{-3}
Negative regulation of apoptotic signaling pathway (GO:2001234)	4.26×10^{-6}	1.86×10^{-4}	9.56×10^{-8}	1.19×10^{-5}
Extrinsic apoptotic signaling pathway (GO:0097191)	4.32×10^{-5}	1.44×10^{-3}	2.60×10^{-4}	9.86×10^{-3}
Negative regulation of extrinsic apoptotic signaling pathway (GO:2001237)	1.29×10^{-4}	3.79×10^{-3}	3.39×10^{-5}	1.92×10^{-3}
Negative regulation of intrinsic apoptotic signaling pathway (GO:2001243)	2.11×10^{-4}	5.87×10^{-3}	2.34×10^{-5}	1.41×10^{-3}
Negative regulation of cysteine-type endopeptidase activity involved in apoptotic process (GO:0043154)	2.65×10^{-4}	7.14×10^{-3}	1.45×10^{-4}	6.32×10^{-3}

FDR: false discovery rate.

(출처: Cho J, Park K, et al. Thorax 2021;76:169–177)

8. 기존에 알려진 비결핵항산균 폐질환 관련 유전자에 대한 연관 연구

기존에 알려진 비결핵항산균 폐질환 관련 유전자에 대한 흔한 변이(common variant) 분석 결과, *RCOR3* (repressor element-1-silencing transcription factor [REST] corepressor 3) 유전자의 rs3111229 (chr1:211363628)가 비결핵항산균 폐질환과 유의한 연관이 있었다 (family-wise error rate 0.05 기준). 기존에 알려진 비결핵항산균 폐질환 관련 유전자에 대한 유전자 세트(gene set) 분석 결과에서는 *NFATC2* (nuclear factor of activated T cells 2) 유전자가 비결핵항산균 폐질환과 유의한 연관이 있었다 (FDR 0.05 기준) (표 10).

표 10에 나열된 변이들에 대해 GTEx 데이터세트를 이용해 eQTL 분석을 수행했다. *RCOR3* 유전자에서 비결핵항산균 폐질환과의 연관성이 가장 강한 변이(rs3111229)는 eQTL 분석에서 *RCOR3* 유전자의 발현 수준과 연관성이 없었다. *NFATC2* 유전자에서 비결핵항산균 폐질환과의 연관성이 가장 강한 변이(rs111933191) 또한 eQTL 분석에서 해당 유전자의 발현 수준과 연관성이 없었다.

표 10. 기존에 비결핵항산균 폐질환 관련 유전자로 보고된 43개의 유전자에 대한 흔한 변이(common variant) 분석과 유전자 세트(gene set) 분석

Gene	Common variant analysis			Gene set analysis	
	Locus of the most significant variant (hg19)	p value	FWER	p value	FDR
RCOR3	**chr1:211363628**	**5.66×10^{-5}**	**0.030**	0.684	0.865
NFATC2	chr20:51549716	1.53×10^{-4}	0.093	**7.47×10^{-4}**	**0.032**
SLC29A1	chr6:44212688	2.88×10^{-4}	0.110	0.072	0.621
ISG15	chr1:920503	4.93×10^{-4}	0.164	0.022	0.481
IL12RB1	chr19:17967816	1.03×10^{-3}	0.445	0.250	0.718
ANKRD6	chr6:89624219	1.07×10^{-3}	0.628	0.052	0.564
TARP	chr7:38332854	1.43×10^{-3}	0.896	0.161	0.709
IFNGR2	chr21:33437648	1.83×10^{-3}	0.938	0.115	0.709
MUC12	chr7:101046298	2.65×10^{-3}	1	0.388	0.782
SLC11A1	chr2:218438843	4.88×10^{-3}	1	0.439	0.782
SAMD3	chr6:130310686	5.11×10^{-3}	1	1	1
CRTAM	chr11:122811024	7.94×10^{-3}	1	0.477	0.782
MAP2K4	chr17:12195571	8.22×10^{-3}	1	0.477	0.782
TTK	chr6:80117150	9.52×10^{-3}	1	0.731	0.899
CFTR	chr7:117550950	0.0101	1	0.127	0.709

TPBG	chr6:82408766	0.0120	1	0.315	0.753
NELL2	chr12:44966063	0.0120	1	1	1
PSPH	chr7:56096567	0.0134	1	0.450	0.782
KRT83	chr12:52257245	0.0138	1	0.052	0.564
IRF8	chr16:85967604	0.0158	1	0.545	0.782
IL2RB	chr22:37143875	0.0206	1	0.569	0.789
GATA2	chr3:128419054	0.0212	1	0.232	0.713
LDHB	chr12:21536230	0.0214	1	0.168	0.709
GZMK	chr5:55127417	0.0219	1	0.546	0.782
AK5	chr1:77312414	0.0221	1	0.657	0.857
IFNLR1	chr1:24054800	0.0222	1	0.861	1
IFNG	chr12:68249467	0.0230	1	0.185	0.709
STAT1	chr2:190968863	0.0239	1	0.601	0.808
TIGIT	chr3:114233728	0.0285	1	0.502	0.782
PMS2P1	chr7:100347079	0.0296	1	0.153	0.709
ORC3	chr6:87621339	0.0298	1	0.220	0.713
XCL1	chr1:168647094	0.0331	1	0.198	0.709
FLJ45825	chr6:37576163	0.0337	1	1	1
XCL2	chr1:168624717	0.0430	1	0.285	0.753
IFNGR1	chr6:137253162	0.0478	1	1	1
PZP	chr12:9172646	0.0631	1	0.442	0.782
A2M	chr12:9172646	0.0631	1	0.514	0.782
FAHD2A	chr2:95485567	0.0653	1	0.305	0.753

PPIH	chr1:42576537	0.0814	1	0.397	0.782
FCRL3	chr1:157694755	0.1410	1	1	1
MPEG1	chr11:59249666	0.1443	1	0.390	0.782
MST1R	chr3:49979727	0.1536	1	0.763	0.912
GUSBP14	chr5:70309998	0.3003	1	1	1

AK5: adenylate kinase 5; *A2M*: alpha-2-macroglobulin; *ANKRD6*: ankyrin repeat domain 6; *CFTR*: cystic fibrosis transmembrane conductance regulator; *CRTAM*: cytotoxic and regulatory T cell molecule; *FAHD2A*: fumarylacetoacetate hydrolase domain containing 2A; *FCRL3*: Fc receptor like 3; FDR: false discovery rate; *FLJ45825*: uncharacterized LOC100505530; FWER: family-wise error rate; *GATA2*: GATA binding protein 2; *GUSBP14*: GUSB pseudogene 14; *GZMK*: granzyme K; *IFNG*: interferon gamma; *IFNGR1*: interferon gamma receptor 1; *IFNGR2*: interferon gamma receptor 2; *IFNLR1*: interferon lambda receptor 1; *IL2RB*: interleukin 2 receptor subunit beta; *IL12RB1*: interleukin 12 receptor subunit beta 1; *IRF8*: interferon regulatory factor 8; *ISG15*: ISG15 ubiquitin like modifier; *KRT83*: keratin 83; *LDHB*: lactate dehydrogenase B; *MAP2K4*: mitogen-activated protein kinase kinase 4; *MPEG1*: macrophage expressed 1; *MST1R*: macrophage stimulating 1 receptor; *MUC12*: mucin 12, cell surface associated; *NELL2*: neural EGFL like 2; *NFATC2*: nuclear factor of activated T cells 2; *ORC3*: origin recognition complex subunit 3; *PMS2P1*: postmeiotic segregation increased 2 pseudogene 1; *PPIH*: peptidylprolyl isomerase H; *PSPH*: phosphoserine phosphatase; *PZP*: PZP alpha-2-macroglobulin like; *TARP*: TCR gamma alternate reading frame protein; *TIGIT*: T cell immunoreceptor with Ig and ITIM domains; *TPBG*: trophoblast glycoprotein; *TTK*: TTK protein kinase; *RCOR3*: REST corepressor 3; *SAMD3*: sterile alpha motif domain containing 3; *SLC11A1*: solute carrier family 11 member 1; *SLC29A1*: solute

carrier family 29 member 1; *STAT1*: signal transducer and activator of transcription 1; *XCL1*: X–C motif chemokine ligand 1; *XCL2*: X–C motif chemokine ligand 2.

(출처: Cho J, Park K, et al. Thorax 2021;76:169–177)

IV. 고 찰

전장유전체연관분석연구를 통해 7번 염색체 단완(7p13)에 위치한 SNP인 rs849177이 한국인의 비결핵항산균 폐질환에 대한 감수성과 연관되어 있다는 것을 발견하였다. 이 연관성은 독립적인 비결핵항산균 폐질환자-대조군 코호트에서 재현되었다. GTEx 폐 조직에서 rs849177 유전자형은 세포자멸사 촉진 유전자인 *STK17A*의 발현 수준과 강한 상관 관계가 있었다. 비결핵항산균 폐질환자는 rs849177의 위험 대립유전자를 가지고 있을 가능성이 높은데, 이 위험 대립유전자는 세포자멸사를 유발하는 유전자인 *STK17A*의 발현 감소와 연관되어 있었다. 또한 MR 분석 결과도 *STK17A* 유전자가 비결핵항산균 폐질환 발생 위험과 인과적으로 관련되어 있다는 본 연구진의 가설을 뒷받침했다. 비결핵항산균 감염 시 대식세포에서 *STK17A* 유전자의 발현이 증가하는 것을 보여준 전사체 데이터 분석 결과도 본 연구진의 가설을 뒷받침했다. 요약하면, 본 연구진의 연구결과는 *STK17A* 유전자의 낮은 발현과 관련된 rs849177의 위험 대립유전자를 가지고 사람은 비결핵항산균에 감염된 대식세포의 세포자멸사가 정상적으로 진행되지 않아 비결핵항산균 폐질환에 대한 감수성이 증가될 수 있음을 시사한다.

활성화된 대식세포는 포식체 성숙(phagosomal maturation),[64]

자가포식(autophagy) 유도[65] 및 세포자멸사[66]를 통해 세포 내 마이코박테리아를 죽일 수 있다. 마이코박테리아의 발병력(virulence)은 주로 숙주의 면역 세포, 특히 폐포 대식세포 내에서 생존할 수 있는 능력에 기인한다.[67] 선천 면역 반응(innate immune response)의 일부로서의 세포자멸사는 마이코박테리아 생존력을 감소시키는 데 중요하다. 또한 세포자멸사는 후천 면역 반응(acquired immune response)의 일부로서도 중요한데, 마이코박테리아 항원을 포함하는 세포사멸체(apoptotic body)가 수상돌기세포(dendritic cell)에 의해 포식(phagocytosis)되어 항원의 제시(presentation)와 이후 T 세포 활성화로 이어지기 때문이다.[68] 마이코박테리아가 세포자멸사를 억제하는 능력은 주로 *Mycobacterium tuberculosis* 균주를 이용한 연구를 통해 전부터 알려졌다.[69,70] 이전 연구에 의하면, 치명적인(virulent) *M. tuberculosis* 균주는 인간 폐포 대식세포의 세포자멸사를 억제하는 반면, 그렇지 않은(avirulent) *M. tuberculosis* 균주는 세포자멸사를 억제하지 못했다.[69]

STK17A 유전자는 serine/threonine-protein kinase 17A를 코딩하는데, serine/threonine-protein kinase 17A는 death-associated protein kinase-related apoptosis-inducing protein kinase (DRAK) 1으로도 알려져 있다.[71] DRAK1은 핵에 위치하며 세포자멸사를 촉진시키는 조절인자(regulator) 역할을 한다.[72] 이전에 발표된 발현 마이크로어레이 데이터 연구에 따르면 *STK17A* 유전자는 *M.*

*tuberculosis*에 감염된 인간의 대식세포[73]와 *M. avium*에 감염된 소의 대식세포에서 그 발현이 증가한다.[74] 본 연구진 또한 공개된 대식세포 전사체 데이터세트를 분석해서 *M. avium* 또는 *M. abscessus*에 감염된 인간 대식세포에서 *STK17A* 유전자의 발현이 증가한다는 것을 보여줬다. 본 연구진은 비결핵항산균 폐질환자에서 rs849177의 대립 위험유전자(C)의 빈도가 높다는 것을 보여줬는데, 이 대립 위험유전자는 *STK17A* 유전자의 발현 감소와 연관되어 있었다. 마이코박테리아에 감염된 폐포 대식세포의 세포자멸사가 인간의 중요한 면역 방어 기전이라는 점을 감안할 때, *STK17A* 유전자의 낮은 발현의 결과로 폐포 대식세포의 세포자멸사가 저해되어 비결핵항산균이 대식세포 내에서 생존하고 지속적으로 복제되는 것은 비결핵항산균 폐질환에 대한 감수정 증가에 기여할 것이다.

COA1 유전자가 코딩하는 cytochrome c oxidase assembly factor 1 homolog[75]과 *BLVRA* 유전자가 코딩하는 biliverdin reductase A[76] 또한 세포자멸사에 기여한다고 알려져 있다. 그러나, *STK17A* 유전자와 달리 GTEx 폐 조직에서 *COA1* 및 *BLVRA* 유전자의 발현 수준은 rs849177 유전자형과 관련이 없었다. 또한, 공개된 대식세포 전사체 데이터세트에 대한 분석 결과 *COA1* 및 *BLVRA* 유전자는 비결핵항산균에 감염된 대식세포에서 차별적으로 발현되지 않았다. 이는 *COA1* 및 *BLVRA* 유전자는 7번 염색체 단완(7p13)에 위치한 유전적 변이가 관련된 비결핵항산균 폐질환에 대한 감수성 유전자가 아님을

시사한다.

이전 연구[39]의 유전자 수준(gene-level) 분석에서 비결핵항산균 폐질환과 관련 있다고 보고된 *RCOR3* 유전자도 본 연구에서 비결핵항산균 폐질환과 유의한 관련이 있었다. RCOR3는 REST corepressor를 구성하는 단백질인데, 비결핵항산균 폐질환의 병인과 관련해서 어떤 역할을 하는지에 대해서는 알려져 있지 않아, 추가적인 연구가 필요하다. 그리고 *NFATC2* 유전자는 본 연구진이 수행한 유전자 세트(gene set) 분석에서 비결핵항산균 폐질환과 유의한 연관이 있었다. 비결핵항산균 폐질환자의 혈액을 이용한 유전자 발현 마이크로어레이 분석 결과 *NFATC2* 유전자의 발현이 감소되어 있다는 것을 보고한 기존 연구가 있었고,[61] NFATc2가 인터페론-γ[77] 및 종양괴사 인자(tumor necrosis factor)-α[78]의 발현을 조절한다는 것을 고려하면 NFATc2의 억제는 비결핵항산균 폐질환의 발생과 관련이 있을 수 있다.

한편, 본 연구결과가 학술지에 발표되고 난 후, 일본에서 *M. avium* 복합체 폐질환자 475명에 대한 전장유전체연관분석연구의 결과가 보고되었다.[79] 간단히 요약하면, 16번 염색체 단완(16p21)에 위치한 유전적 변이(rs109592)가 일본인의 *M. avium* 복합체 폐질환과 가장 유의하게 연관되어 있었다. 이 연관성은 한국과 유럽의 *M. avium* 복합체 폐질환 코호트에서 각각 재현되었다. GTEx 데이터세트를 이용한 eQTL 분석에서 이 유전적 변이와 *CHP2* (calcineurin like EF-hand

protein 2) 유전자의 연관성이 확인되었다. rs109592에 위험 대립유전자(C)를 가진 사람은 GTEx 폐 조직에서 *CHP2* 유전자의 발현 수준이 낮았다. CHP2는 칼슘 결합 단백질로 세포막에 존재하는 나트륨 이온/수소 이온 교환체(plasma membrane-type Na+/H+ exchanger, NHE)의 보조인자(cofactor)로 작용하여 그 활성을 조절한다.[80] NHE는 세포 pH 항상성 유지에 매우 중요한데, *CHP2* 유전자의 발현이 감소되어 NHE를 통한 이온 항상성이 깨질 경우, 비결핵항산균 감염에 취약해진다고 설명하고 있다.

이전 연구에 따르면 폐결핵의 SNP-유전성은 0.15이다.[51] 기존 폐결핵 연구[51]에서 사용된 방법과 동일한 방법으로 비결핵항산균 폐질환의 평생 유병률을 추정했을 때, 본 연구에서 추정된 비결핵항산균 폐질환의 SNP-유전성은 0.225로, 폐결핵의 유전성보다 더 컸다. 이는 본 연구진이 예상했던 바이기도 하다.

본 연구의 결과를 적절하게 평가하려면 그 한계 또한 제대로 인식하고 있어야 한다. 첫째, 본 연구는 대상환자수가 많지 않아 검정력이 부족했을 가능성이 있고, 이것이 rs849177과 비결핵항산균 폐질환의 연관성이 발굴 코호트에서 기존의 유전체 차원의 유의수준(significance level)인 5×10^{-8}에 도달하지 못한 원인일 수 있다. 하지만, 이 연관성은 독립적인 비결핵항산균 폐질환자-대조군 코호트인 재현 코호트에서 검증되었고, Fisher의 방법을 사용하여 발굴 및 재현 코호트의 결과를 병합했을 때, rs849177의 유의확률(combined

p value)은 4.92×10^{-8}였고 이는 기존 유의수준을 만족하는 결과이다. 한편, 7번 염색체 위치한 이 유전적 변이와 비결핵항산균 폐질환의 연관성이 한국인 재현 코호트에서 검증되기는 했으나, 이 연관성이 다른 인종이나 인구 집단에서 공통적으로 재현되는지 여부는 알 수 없다. 비결핵항산균 폐질환은 인종에 따라 역학에 차이가 있으므로[34] 후속 연구를 통해 이 유전적 소인이 한국인에 특이적인지 또는 한국인 외 다른 아시아인이나 유럽인 등에 공통적인지 확인이 필요하다. 이를 위해 다인종-다국가로 구성된 국제 비결핵항산균 폐질환 컨소시엄을 구성한다면, 전장유전체연관분석연구에 필수적인 충분한 환자수를 확보하여 통계적 검정력을 높이고, 드문 유전적 변이에 대한 연관성 확인도 가능할 것으로 보인다. 이러한 일련의 노력을 통해 비결핵항산균 폐질환의 병인에 대한 이해를 높이고, 치료 면에서도 새로운 시각을 제공할 수 있을 것으로 기대된다.

둘째, 발굴 및 재현 코호트의 대조군에서 rs849177의 대립유전자 빈도가 달랐다. 발굴 코호트에서 rs849177의 위험 대립유전자(C)의 빈도는 환자군의 경우 0.70, 대조군의 경우 0.57이었다. 재현 코호트에서 rs849177의 위험 대립유전자(C)의 빈도는 환자군의 경우 0.68, 대조군의 경우 0.63이었다. 즉, 대조군의 rs849177의 위험 대립유전자(C)의 빈도가 발굴 코호트에서 0.57, 재현 코호트에서 0.63으로 약간 차이가 있었다. 다차원 척도법 분석 결과인 그림 3B를 보면 발굴 및 재현 코호트의 참가자들이 각각 클러스터링

되는 경향을 보이는데, 두 코호트 참가자들 사이에 유전적 이질성이 존재한다고 보기는 어렵고, 이는 각 코호트 참가자에 대한 유전자형 분석에 각각 다른 플랫폼이 사용되었기 때문으로 생각된다. 발굴 코호트에서는 Illumina 사의 칩이 사용되었고, 재현 코호트에서는 Affymetrix 사의 칩이 사용되었다. 그림 3B에서 보이는 것처럼 각 코호트 내에서는 환자군과 대조군 사이에 전반적인 유전적 이질성은 없었기 때문에 본 연구의 결과를 해석하는데 큰 영향이 없을 것으로 생각된다. 한편, 질병관리청에서 2012년부터 현재까지 한국인 1722명의 염기서열 분석(whole genome sequencing)한 데이터베이스인 한국인 참조유전체 데이터베이스(Korean Reference Genome Database, KRGDB)[81]에서 rs849177의 위험 대립유전자(C)의 빈도는 0.64였다 (1722명 통합 데이터베이스 기준). 이는 재현 코호트 대조군의 rs849177의 위험 대립유전자(C)의 빈도인 0.63와는 비슷했지만, 발굴 코호트 대조군의 rs849177의 위험 대립유전자(C)의 빈도인 0.57과는 약간 차이가 있었다. 본 연구진은 rs849177과 비결핵항산균 폐질환의 연관성이 다른 대조군을 설정했을 때도 재현되는지 확인하기 위해 임상유전체생명정보시스템(Clinical & Omics Data Archive, CODA)의 데이터베이스를 이용해 추가 분석을 시행했다 (http://coda.nih.go.kr/coda/). 임상유전체생명정보시스템은 보건복지부 지원으로 생산된 임상연구데이터와 유전체데이터의 통합축적 및 유통을 담당하며, 한국인칩을 이용해 유전자형이 분석된 한국인 66,407명의 데이터가 축적되어 있다. 재현 코호트의 비결핵항산균 폐질환자의

유전자형 분석을 한국인칩을 이용해 수행했기 때문에, 재현 코호트의 비결핵항산균 폐질환자 184명과 임상유전체생명정보시스템의 데이터베이스에서 추출한 대조군 1840명을 분석했다. 임상유전체생명정보시스템의 대조군의 규모가 재현 코호트의 비결핵항산균 폐질환자의 환자수보다 매우 컸기 때문에, 비결핵항산균 폐질환자와 나이와 성별을 매칭해서 10배수를 임상유전체생명정보시스템의 데이터베이스에서 추출했다. 분석 결과, rs849177과 비결핵항산균 폐질환의 연관성은 재현되었다 (OR, 1.32; 95% CI, 1.05 ~ 1.66; p=0.009).

셋째, rs849177과 관련된 유전자와 그 발현 수준을 알아 보기 위해 GTEx 데이터세트를 이용해 eQTL 분석을 수행할 때, 관심 있는 조직을 폐로 한정했으나, 폐 이외의 조직, 특히 혈액에서의 특정 유전자 발현 수준이 비결핵항산균 폐질환의 발생에 영향을 미칠 수 있다. 본 연구에서 GTEx 데이터세트의 전혈(whole bood)에서 *BLVRA* 유전자의 발현 수준이 rs849177에 의해 영향을 받는 것으로 나타났다. 즉, rs849177에 위험 대립유전자(C)를 가진 사람은 GTEx 혈액에서 *BLVRA* 유전자의 발현 수준이 낮았다. Biliverdin reductase A는 NADPH 또는 NADH의 존재 하에 biliverdin을 bilirubin으로 환원시키는 역할을 한다.[76] 세포 내 bilirubin이 부족해질 경우, 반응산소종(reactive oxygen species)이 증가하고, 세포자멸사가 일어날 수 있다고 보고된 바 있고, 따라서 biliverdin reductase A가 세포의

산화-환원 주기에 중요한 역할을 담당할 수 있다.[76,82] 비결핵항산균 폐질환의 병인론 측면에서 *BLVRA* 유전자의 역할에 대해 추가적인 연구가 필요하다. 한편, 관심 있는 조직을 폐로 한정하여 eQTL 분석을 시행한 후, 본 연구진은 특히 대식세포의 역할에 주목하였고, 비결핵항산균을 감염시킨 대식세포에서 *STK17A* 유전자의 발현이 증가하는 것을 보여준 전사체 데이터 분석 결과도 본 연구진의 가설을 뒷받침하기는 했지만, 실제 GTEx 데이터세트 폐 조직에서 대식세포의 비율은 매우 적다. 또한 비결핵항산균 감염에 대항한 면역 반응에서 대식세포 외에도 T 세포, NK 세포, B 세포, 수상돌기세포를 포함한 면역 세포와 상피 세포의 역할도 고려되어야 한다. 그리고 같은 조직 내에서도 세포 유형에 따라 eQTL은 특이적일 수 있기 때문에, 향후 비결핵항산균 폐질환자의 말초 혈액 단핵세포(peripheral blood mononuclear cell)를 이용해 단일 세포 eQTL 분석을 수행한다면,[83] 본 연구를 통해 발견한 유전적 변이가 각 세포 유형별로 유전자 발현 및 조절에 미치는 영향을 구체화할 수 있을 것으로 기대된다.

마지막으로, 본 연구진이 GTEx 데이터세트를 이용한 유전자 발현 분석 및 질병에 대한 인과적 영향 평가를 수행하고, 연구의 결과를 뒷받침하는 대식세포 전사체 데이터세트의 분석 결과를 확인하였지만, 비결핵항산균 감염에 대한 인간 면역 반응에서 rs849177 및 *STK17A* 유전자의 실제 역할을 실험적이며 직접적인 방법으로 밝힌 것은 아니다. 특히, 분석 과정에서 많이 사용된 GTEx 데이터세트에는 아시아인

기증자가 대략 1%에 불과하다. 후속 연구를 통해서 실제 비결핵항산균 폐질환자에서 건강인과 비교해 혈액이나 폐 조직에서 *STK17A* 유전자의 발현이 감소되어 있는지 확인이 필요하다. 그리고 보다 근본적으로 *STK17A* 유전자를 조작(manipulation)했을 때, 세포자멸사의 정도에 변화가 생기는지, 비결핵항산균 감염에 취약해지는지 등에 대한 세포 또는 동물 수준의 실험실 연구가 필요하다.

IV. 결　　론

GWAS를 통해 7번 염색체 단완에 위치한 유전적 변이가 세포자멸사를 촉진하는 유전자의 발현을 변화시키고, 한국인의 비결핵항산균 폐질환에 대한 감수성과 관련이 있다는 것을 확인했다. 이러한 발견은 비결핵항산균 폐질환의 병인에 대한 통찰력을 제공하고 새로운 치료제를 찾는 데 도움이 될 수 있을 것이다.

참고 문헌

1. Park SC, Kang MJ, Han CH, Lee SM, Kim CJ, Lee JM, Kang YA. Prevalence, incidence, and mortality of nontuberculous mycobacterial infection in Korea: a nationwide population-based study. BMC Pulm Med 2019;19:140.

2. Lee H, Myung W, Koh WJ, Moon SM, Jhun BW. Epidemiology of Nontuberculous Mycobacterial Infection, South Korea, 2007-2016. Emerg Infect Dis 2019;25:569-72.

3. Winthrop KL, Marras TK, Adjemian J, Zhang H, Wang P, Zhang Q. Incidence and Prevalence of Nontuberculous Mycobacterial Lung Disease in a Large U.S. Managed Care Health Plan, 2008-2015. Ann Am Thorac Soc 2020;17:178-85.

4. Adjemian J, Olivier KN, Seitz AE, Holland SM, Prevots DR. Prevalence of nontuberculous mycobacterial lung disease in U.S. Medicare beneficiaries. Am J Respir Crit Care Med 2012;185:881-6.

5. Marras TK, Mendelson D, Marchand-Austin A, May K, Jamieson FB. Pulmonary nontuberculous mycobacterial disease, Ontario, Canada, 1998-2010. Emerg Infect Dis 2013;19:1889-91.

6. Al-Houqani M, Jamieson F, Mehta M, Chedore P, May K,

Marras TK. Aging, COPD, and other risk factors do not explain the increased prevalence of pulmonary Mycobacterium avium complex in Ontario. Chest 2012;141:190-7.

7. Prevots DR, Shaw PA, Strickland D, Jackson LA, Raebel MA, Blosky MA, Montes de Oca R, Shea YR, Seitz AE, Holland SM, Olivier KN. Nontuberculous mycobacterial lung disease prevalence at four integrated health care delivery systems. Am J Respir Crit Care Med 2010;182:970-6.

8. Gerogianni I, Papala M, Kostikas K, Petinaki E, Gourgoulianis KI. Epidemiology and clinical significance of mycobacterial respiratory infections in Central Greece. Int J Tuberc Lung Dis 2008;12:807-12.

9. Ringshausen FC, Wagner D, de Roux A, Diel R, Hohmann D, Hickstein L, Welte T, Rademacher J. Prevalence of Nontuberculous Mycobacterial Pulmonary Disease, Germany, 2009-2014. Emerg Infect Dis 2016;22:1102-5.

10. Thomson RM. Changing epidemiology of pulmonary nontuberculous mycobacteria infections. Emerg Infect Dis 2010;16:1576-83.

11. Diel R, Jacob J, Lampenius N, Loebinger M, Nienhaus A, Rabe KF, Ringshausen FC. Burden of non-tuberculous mycobacterial pulmonary disease in Germany. Eur Respir J 2017;49.

12. Henkle E, Hedberg K, Schafer S, Novosad S, Winthrop KL. Population-based Incidence of Pulmonary Nontuberculous Mycobacterial Disease in Oregon 2007 to 2012. Ann Am Thorac Soc 2015;12:642-7.

13. Ratnatunga CN, Lutzky VP, Kupz A, Doolan DL, Reid DW, Field M, Bell SC, Thomson RM, Miles JJ. The Rise of Non-Tuberculosis Mycobacterial Lung Disease. Front Immunol 2020;11:303.

14. Bodle EE, Cunningham JA, Della-Latta P, Schluger NW, Saiman L. Epidemiology of nontuberculous mycobacteria in patients without HIV infection, New York City. Emerg Infect Dis 2008;14:390-6.

15. Cassidy PM, Hedberg K, Saulson A, McNelly E, Winthrop KL. Nontuberculous mycobacterial disease prevalence and risk factors: a changing epidemiology. Clin Infect Dis 2009;49:e124-9.

16. Hernandez-Garduno E, Rodrigues M, Elwood RK. The incidence of pulmonary non-tuberculous mycobacteria in British Columbia, Canada. Int J Tuberc Lung Dis 2009;13:1086-93.

17. Andrejak C, Thomsen VO, Johansen IS, Riis A, Benfield TL, Duhaut P, Sorensen HT, Lescure FX, Thomsen RW. Nontuberculous pulmonary mycobacteriosis in Denmark: incidence and prognostic factors. Am J Respir Crit Care Med 2010;181:514-21.

18. Dailloux M, Abalain ML, Laurain C, Lebrun L, Loos-Ayav C, Lozniewski A, Maugein J. Respiratory infections associated with nontuberculous mycobacteria in non-HIV patients. Eur Respir J 2006;28:1211-5.

19. Freeman J, Morris A, Blackmore T, Hammer D, Munroe S, McKnight L. Incidence of nontuberculous mycobacterial disease in New Zealand, 2004. N Z Med J 2007;120:U2580.

20. Morimoto K, Iwai K, Uchimura K, Okumura M, Yoshiyama T, Yoshimori K, Ogata H, Kurashima A, Gemma A, Kudoh S. A steady increase in nontuberculous mycobacteriosis mortality and estimated prevalence in Japan. Ann Am Thorac Soc 2014;11:1-8.

21. Namkoong H, Kurashima A, Morimoto K, Hoshino Y, Hasegawa N, Ato M, Mitarai S. Epidemiology of Pulmonary Nontuberculous Mycobacterial Disease, Japan(1). Emerg Infect Dis 2016;22:1116-7.

22. Kwak N, Park J, Kim E, Lee CH, Han SK, Yim JJ. Treatment Outcomes of Mycobacterium avium Complex Lung Disease: A Systematic Review and Meta-analysis. Clin Infect Dis 2017.

23. Kwak N, Dalcolmo MP, Daley CL, Eather G, Gayoso R, Hasegawa N, Jhun BW, Koh WJ, Namkoong H, Park J, Thomson R, van Ingen J, Zweijpfenning SMH, Yim JJ. M ycobacterium abscessus pulmonary disease: individual patient data meta-analysis. Eur Respir

J 2019;54.

24. Kwak N, Kim SA, Choi SM, Lee J, Lee CH, Yim JJ. Longitudinal changes in health-related quality of life according to clinical course among patients with non-tuberculous mycobacterial pulmonary disease: a prospective cohort study. BMC Pulm Med 2020;20:126.

25. Lee H, Myung W, Lee EM, Kim H, Jhun BW. Mortality and Prognostic Factors of Nontuberculous Mycobacterial Infection in Korea: A Population-based Comparative Study. Clin Infect Dis 2021;72:e610-e9.

26. Kim HJ, Kwak N, Hong H, Kang N, Im Y, Jhun BW, Yim JJ. BACES Score for Predicting Mortality in Nontuberculous Mycobacterial Pulmonary Disease. Am J Respir Crit Care Med 2021;203:230-6.

27. Jhun BW, Moon SM, Jeon K, Kwon OJ, Yoo H, Carriere KC, Huh HJ, Lee NY, Shin SJ, Daley CL, Koh WJ. Prognostic factors associated with long-term mortality in 1445 patients with nontuberculous mycobacterial pulmonary disease: a 15-year follow-up study. Eur Respir J 2020;55.

28. Goring SM, Wilson JB, Risebrough NR, Gallagher J, Carroll S, Heap KJ, Obradovic M, Loebinger MR, Diel R. The cost of Mycobacterium avium complex lung disease in Canada, France,

Germany, and the United Kingdom: a nationally representative observational study. BMC Health Serv Res 2018;18:700.

29. Axson EL, Bual N, Bloom CI, Quint JK. Risk factors and secondary care utilisation in a primary care population with non-tuberculous mycobacterial disease in the UK. Eur J Clin Microbiol Infect Dis 2019;38:117-24.

30. Filipe-Santos O, Bustamante J, Chapgier A, Vogt G, de Beaucoudrey L, Feinberg J, Jouanguy E, Boisson-Dupuis S, Fieschi C, Picard C, Casanova JL. Inborn errors of IL-12/23- and IFN-gamma-mediated immunity: molecular, cellular, and clinical features. Semin Immunol 2006;18:347-61.

31. Ristola MA, von Reyn CF, Arbeit RD, Soini H, Lumio J, Ranki A, Buhler S, Waddell R, Tosteson AN, Falkinham JO, 3rd, Sox CH. High rates of disseminated infection due to non-tuberculous mycobacteria among AIDS patients in Finland. J Infect 1999;39:61-7.

32. Wu UI, Holland SM. Host susceptibility to non-tuberculous mycobacterial infections. Lancet Infect Dis 2015;15:968-80.

33. Kim RD, Greenberg DE, Ehrmantraut ME, Guide SV, Ding L, Shea Y, Brown MR, Chernick M, Steagall WK, Glasgow CG, Lin J, Jolley C, Sorbara L, Raffeld M, Hill S, Avila N, Sachdev V, Barnhart LA, Anderson VL, Claypool R, Hilligoss DM, Garofalo M, Fitzgerald

A, Anaya-O'Brien S, Darnell D, DeCastro R, Menning HM, Ricklefs SM, Porcella SF, Olivier KN, Moss J, Holland SM. Pulmonary nontuberculous mycobacterial disease: prospective study of a distinct preexisting syndrome. Am J Respir Crit Care Med 2008;178:1066-74.

34. Adjemian J, Frankland TB, Daida YG, Honda JR, Olivier KN, Zelazny A, Honda S, Prevots DR. Epidemiology of Nontuberculous Mycobacterial Lung Disease and Tuberculosis, Hawaii, USA. Emerg Infect Dis 2017;23:439-47.

35. Reich JM, Kim JS. Diminished Susceptibility of African-Americans to Non-tuberculous Mycobacterial Disease. Lung 2018;196:125-7.

36. Leung JM, Fowler C, Smith C, Adjemian J, Frein C, Claypool RJ, Holland SM, Prevots RD, Olivier K. A familial syndrome of pulmonary nontuberculous mycobacteria infections. Am J Respir Crit Care Med 2013;188:1373-6.

37. Szymanski EP, Leung JM, Fowler CJ, Haney C, Hsu AP, Chen F, Duggal P, Oler AJ, McCormack R, Podack E, Drummond RA, Lionakis MS, Browne SK, Prevots DR, Knowles M, Cutting G, Liu X, Devine SE, Fraser CM, Tettelin H, Olivier KN, Holland SM. Pulmonary Nontuberculous Mycobacterial Infection. A Multisystem, Multigenic Disease. Am J Respir Crit Care Med 2015;192:618-28.

38. Becker KL, Arts P, Jaeger M, Plantinga TS, Gilissen C, van Laarhoven A, van Ingen J, Veltman JA, Joosten LA, Hoischen A, Netea MG, Iseman MD, Chan ED, van de Veerdonk FL. MST1R mutation as a genetic cause of Lady Windermere syndrome. Eur Respir J 2017;49.

39. Chen F, Szymanski EP, Olivier KN, Liu X, Tettelin H, Holland SM, Duggal P. Whole-Exome Sequencing Identifies the 6q12-q16 Linkage Region and a Candidate Gene, TTK, for Pulmonary Nontuberculous Mycobacterial Disease. Am J Respir Crit Care Med 2017;196:1599-604.

40. Kim SJ, Yoon SH, Choi SM, Lee J, Lee CH, Han SK, Yim JJ. Characteristics associated with progression in patients with of nontuberculous mycobacterial lung disease : a prospective cohort study. BMC Pulm Med 2017;17:5.

41. Griffith DE, Aksamit T, Brown-Elliott BA, Catanzaro A, Daley C, Gordin F, Holland SM, Horsburgh R, Huitt G, Iademarco MF, Iseman M, Olivier K, Ruoss S, von Reyn CF, Wallace RJ, Jr., Winthrop K, Subcommittee ATSMD, American Thoracic S, Infectious Disease Society of A. An official ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases. Am J Respir Crit Care Med 2007;175:367-416.

42. Gombojav B, Song YM, Lee K, Yang S, Kho M, Hwang YC, Ko

G, Sung J. The Healthy Twin Study, Korea updates: resources for omics and genome epidemiology studies. Twin Res Hum Genet 2013;16:241-5.

43. Nijenhuis JT, Choi KY, Choi YY, Lee JJ, Seo EH, Kim H, Lee KH. Differences Between APOE Carriers and Non-APOE Carriers on Neurocognitive Tests: Jensen Effects? Am J Alzheimers Dis Other Demen 2018;33:353-61.

44. Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. Nat Protoc 2010;5:1564-73.

45. Das S, Forer L, Schonherr S, Sidore C, Locke AE, Kwong A, Vrieze SI, Chew EY, Levy S, McGue M, Schlessinger D, Stambolian D, Loh PR, Iacono WG, Swaroop A, Scott LJ, Cucca F, Kronenberg F, Boehnke M, Abecasis GR, Fuchsberger C. Next-generation genotype imputation service and methods. Nat Genet 2016;48:1284-7.

46. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, Kang HM, Fuchsberger C, Danecek P, Sharp K, Luo Y, Sidore C, Kwong A, Timpson N, Koskinen S, Vrieze S, Scott LJ, Zhang H, Mahajan A, Veldink J, Peters U, Pato C, van Duijn CM, Gillies CE, Gandin I, Mezzavilla M, Gilly A, Cocca M, Traglia M, Angius A, Barrett JC, Boomsma D, Branham K, Breen G, Brummett CM, Busonero F, Campbell H, Chan A, Chen S, Chew E, Collins FS,

Corbin LJ, Smith GD, Dedoussis G, Dorr M, Farmaki AE, Ferrucci L, Forer L, Fraser RM, Gabriel S, Levy S, Groop L, Harrison T, Hattersley A, Holmen OL, Hveem K, Kretzler M, Lee JC, McGue M, Meitinger T, Melzer D, Min JL, Mohlke KL, Vincent JB, Nauck M, Nickerson D, Palotie A, Pato M, Pirastu N, McInnis M, Richards JB, Sala C, Salomaa V, Schlessinger D, Schoenherr S, Slagboom PE, Small K, Spector T, Stambolian D, Tuke M, Tuomilehto J, Van den Berg LH, Van Rheenen W, Volker U, Wijmenga C, Toniolo D, Zeggini E, Gasparini P, Sampson MG, Wilson JF, Frayling T, de Bakker PI, Swertz MA, McCarroll S, Kooperberg C, Dekker A, Altshuler D, Willer C, Iacono W, Ripatti S, Soranzo N, Walter K, Swaroop A, Cucca F, Anderson CA, Myers RM, Boehnke M, McCarthy MI, Durbin R. A reference panel of 64,976 haplotypes for genotype imputation. Nat Genet 2016;48:1279-83.

47. Stekhoven DJ, Bühlmann P. MissForest--non-parametric missing value imputation for mixed-type data. Bioinformatics 2012;28:112-8.

48. Zhao H, Mitra N, Kanetsky PA, Nathanson KL, Rebbeck TR. A practical approach to adjusting for population stratification in genome-wide association studies: principal components and propensity scores (PCAPS). Stat Appl Genet Mol Biol 2018;17.

49. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA,

Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559-75.

50. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 2011;88:76-82.

51. Speed D, Cai N, Consortium U, Johnson MR, Nejentsev S, Balding DJ. Reevaluation of SNP heritability in complex human traits. Nat Genet 2017;49:986-92.

52. Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, Eyler AE, Denny JC, Consortium GT, Nicolae DL, Cox NJ, Im HK. A gene-based association method for mapping traits using reference transcriptome data. Nat Genet 2015;47:1091-8.

53. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015;44:512-25.

54. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. Stat Methods Med Res 2017;26:2333-55.

55. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread

horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet 2018;50:693-8.

56. Aulicino A, Dinan AM, Miranda-CasoLuengo AA, Browne JA, Rue-Albrecht K, MacHugh DE, Loftus BJ. High-throughput transcriptomics reveals common and strain-specific responses of human macrophages to infection with Mycobacterium abscessus Smooth and Rough variants. BMC Genomics 2015;16:1046.

57. Agdestein A, Jones A, Flatberg A, Johansen TB, Heffernan IA, Djonne B, Bosco A, Olsen I. Intracellular growth of Mycobacterium avium subspecies and global transcriptional responses in human macrophages after infection. BMC Genomics 2014;15:58.

58. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.

59. Smyth GK. Limma: linear models for microarray data. In: Smyth GK, Ritchie M, Thorne N, Wettenhall J, eds. Bioinformatics and computational biology solutions Using R and Bioconductor. New York, NY: Springer; 2005:397-420.

60. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res

2019;47:D419-d26.

61. Cowman SA, Jacob J, Hansell DM, Kelleher P, Wilson R, Cookson WOC, Moffatt MF, Loebinger MR. Whole-Blood Gene Expression in Pulmonary Nontuberculous Mycobacterial Infection. Am J Respir Cell Mol Biol 2018;58:510-8.

62. Koh WJ, Kwon OJ, Kim EJ, Lee KS, Ki CS, Kim JW. NRAMP1 gene polymorphism and susceptibility to nontuberculous mycobacterial lung diseases. Chest 2005;128:94-101.

63. Lee S, Wu MC, Lin X. Optimal tests for rare variant effects in sequencing association studies. Biostatistics 2012;13:762-75.

64. Vergne I, Chua J, Singh SB, Deretic V. Cell biology of mycobacterium tuberculosis phagosome. Annu Rev Cell Dev Biol 2004;20:367-94.

65. Gutierrez MG, Master SS, Singh SB, Taylor GA, Colombo MI, Deretic V. Autophagy is a defense mechanism inhibiting BCG and Mycobacterium tuberculosis survival in infected macrophages. Cell 2004;119:753-66.

66. Molloy A, Laochumroonvorapong P, Kaplan G. Apoptosis, but not necrosis, of infected monocytes is coupled with killing of intracellular bacillus Calmette-Guerin. J Exp Med 1994;180:1499-509.

67. Smith I. Mycobacterium tuberculosis pathogenesis and

molecular determinants of virulence. Clin Microbiol Rev 2003;16:463-96.

68. Briken V, Miller JL. Living on the edge: inhibition of host cell apoptosis by Mycobacterium tuberculosis. Future Microbiol 2008;3:415-22.

69. Keane J, Remold HG, Kornfeld H. Virulent Mycobacterium tuberculosis strains evade apoptosis of infected alveolar macrophages. J Immunol 2000;164:2016-20.

70. Sly LM, Hingley-Wilson SM, Reiner NE, McMaster WR. Survival of Mycobacterium tuberculosis in host macrophages involves resistance to apoptosis dependent upon induction of antiapoptotic Bcl-2 family member Mcl-1. J Immunol 2003;170:430-7.

71. Shiloh R, Bialik S, Kimchi A. The DAPK family: a structure-function analysis. Apoptosis 2014;19:286-97.

72. Sanjo H, Kawai T, Akira S. DRAKs, novel serine/threonine kinases related to death-associated protein kinase that trigger apoptosis. J Biol Chem 1998;273:29066-71.

73. Silver RF, Walrath J, Lee H, Jacobson BA, Horton H, Bowman MR, Nocka K, Sypek JP. Human alveolar macrophage gene responses to Mycobacterium tuberculosis strains H37Ra and H37Rv. Am J Respir Cell Mol Biol 2009;40:491-504.

74. Kabara E, Kloss CC, Wilson M, Tempelman RJ, Sreevatsan S, Janagama H, Coussens PM. A large-scale study of differential gene expression in monocyte-derived macrophages infected with several strains of Mycobacterium avium subspecies paratuberculosis. Brief Funct Genomics 2010;9:220-37.

75. Mick DU, Dennerlein S, Wiese H, Reinhold R, Pacheu-Grau D, Lorenzi I, Sasarman F, Weraarpachai W, Shoubridge EA, Warscheid B, Rehling P. MITRAC links mitochondrial protein translocation to respiratory-chain assembly and translational regulation. Cell 2012;151:1528-41.

76. Baranano DE, Rao M, Ferris CD, Snyder SH. Biliverdin reductase: a major physiologic cytoprotectant. Proc Natl Acad Sci U S A 2002;99:16093-8.

77. Kiani A, García-Cózar FJ, Habermann I, Laforsch S, Aebischer T, Ehninger G, Rao A. Regulation of interferon-gamma gene expression by nuclear factor of activated T cells. Blood 2001;98:1480-8.

78. Kaminuma O, Kitamura F, Kitamura N, Hiroi T, Miyoshi H, Miyawaki A, Miyatake S. Differential contribution of NFATc2 and NFATc1 to TNF-alpha gene expression in T cells. J Immunol 2008;180:319-26.

79. Namkoong H, Omae Y, Asakura T, Ishii M, Suzuki S, Morimoto

K, Kawai Y, Emoto K, Oler AJ, Szymanski EP, Yoshida M, Matsuda S, Yagi K, Hase I, Nishimura T, Sasaki Y, Asami T, Shiomi T, Matsubara H, Shimada H, Hamamoto J, Jhun BW, Kim SY, Huh HJ, Won HH, Ato M, Kosaki K, Betsuyaku T, Fukunaga K, Kurashima A, Tettelin H, Yanai H, Mahasirimongkol S, Olivier KN, Hoshino Y, Koh WJ, Holland SM, Tokunaga K, Hasegawa N. Genome-wide association study in patients with pulmonary Mycobacterium avium complex disease. Eur Respir J 2021.

80. Orlowski J, Grinstein S. Na+/H+ exchangers. Compr Physiol 2011;1:2083-100.

81. Jung KS, Hong KW, Jo HY, Choi J, Ban HJ, Cho SB, Chung M. KRGDB: the large-scale variant database of 1722 Koreans based on whole genome sequencing. Database (Oxford) 2020;2020.

82. Kim SY, Kang HT, Choi HR, Park SC. Biliverdin reductase A in the prevention of cellular senescence against oxidative stress. Exp Mol Med 2011;43:15-23.

83. van der Wijst M, de Vries DH, Groot HE, Trynka G, Hon CC, Bonder MJ, Stegle O, Nawijn MC, Idaghdour Y, van der Harst P, Ye CJ, Powell J, Theis FJ, Mahfouz A, Heinig M, Franke L. The single-cell eQTLGen consortium. Elife 2020;9.

Abstract

Genome-wide association study of non-tuberculous mycobacterial pulmonary disease

Jaeyoung Cho
Internal Medicine Major
Graduate School of Medicine
Seoul National University

Background The prevalence and incidence of non-tuberculous mycobacterial pulmonary disease (NTM-PD) are increasing in different parts of the world including South Korea. Despite ubiquitous exposure to NTM, a subset of people develop NTM-PD. Moreover, the presence of ethnic disparity, familial clustering, and the distinct phenotype of NTM-PD indicate the presence of genetic predispositions to this disease. However, the genetic factors associated with susceptibility to NTM-PD have been unclear. We aimed to find genetic variants in individuals with NTM-PD by a genome-wide association study (GWAS).

Methods We performed a GWAS with Korean patients with NTM-PD and controls from the Healthy Twin Study, Korea cohort. Candidate single-nucleotide polymorphisms (SNPs) from this discovery cohort were then validated in another Korean replication cohort. The Genotype-Tissue Expression (GTEx) data set was utilized to find expression quantitative trait loci (eQTL). In addition, we conducted the PrediXcan analysis to predict the candidate gene expression

levels using our imputed genotype data and the GTEx data set. We performed a Mendelian randomization (MR) analysis to evaluate the causal effect related to the expression level of a candidate gene on the risk of development of NTM-PD. Moreover, we performed a transcriptome profiling analysis with publicly accessible data sets to identify differentially expressed genes (DEGs) in NTM-infected macrophages.

Results Through the GWAS with 403 patients with NTM-PD and 306 controls, we found rs849177 on chromosome 7p13 as the candidate SNP associated with susceptibility to NTM-PD (odds ratio, 2.34; 95% confidence interval [CI], 1.71 ~ 3.21; $p=1.36\times10^{-7}$). Its association was replicated in the independent cohort of 184 individuals with NTM-PD and 1,680 controls. When the Fisher's method was applied, the combined p-value for both cohorts was 4.92×10^{-8}. The eQTL analysis revealed that a risk allele (C) at rs849177 was associated with decreased expression levels of *STK17A*, a proapoptotic gene. According to the PrediXcan analysis, *STK17A* was the most significant DEG between individuals with NTM-PD and controls. A causal effect of *STK17A* on the development of NTM-PD was found in the MR analysis (β, -4.627; 95% CI, -8.768 ~ -0.486; p=0.029). According to the transcriptome profiling analysis, the expression levels of *STK17A* increased significantly in macrophages infected with NTM.

Conclusions We found rs849177 on chromosome 7p13 as a susceptibility marker for NTM-PD in a Korean population. This genetic variant might be related to susceptibility to NTM-PD by modifying the expression level of a proapoptotic gene.

Keyword : non-tuberculous mycobacteria; lung diseases; disease susceptibility; genome-wide association study; apoptosis

Student Number : 2016-30558

Parts of this dissertation were previously published in *Thorax* (Cho J, Park K, Choi SM, Lee J, Lee CH, Lee JK, Heo EY, Kim DK, Lee YJ, Park JS, Cho YJ, Yoon HI, Lee JH, Lee CT, Kim N, Choi KY, Lee KH, Sung J, Won S, Yim JJ. Genome-wide association study of non-tuberculous mycobacterial pulmonary disease. Thorax 2021;76:169-177).