

I COPD 전사체 연구의 최신 동향

윤정현^{1,2}, 서민석²

¹Division of Pulmonary and Critical Care Medicine, ² Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA

COPD is a complex and heterogeneous disease that involves a variable degree of parenchymal destruction and large and small airway disease. Genome wide gene expression profiling or transcriptomics, provides a global assessment of cellular function and has provided insights to COPD pathogenesis, heterogeneity and treatment response. In this review we will review the current state of the art transcriptomic studies and discuss the limitations and future directions for transcriptomics in COPD.

Key Words: COPD, Transcriptomics, Gene expression, Precision medicine

Corresponding author: Jeong Hyun Yun, M.D., M.P.H.

Channing Division of Network Medicine, Brigham and Women's Hospital, 181 Longwood Ave. Boston, MA 02115, USA

Tel: +1-617-525-2136, Fax: +1-617-264-6304, Email: jeong.yun@channing.harvard.edu

1. 서론

만성폐쇄성폐질환(COPD)은 폐조직의 파괴로 일어나는 폐기종(emphysema)과 만성기관지염(chronic bronchitis)을 포괄하며 환자 간의 임상 증상과 진행 향상의 개인차가 큰 복합적인 질환이다. COPD의 진단 기준으로 사용되는 폐 기능 검사로는 이렇게 상이한 질병 형태를 구분할 수 없기 때문에 유전체와 전사체 연구를 통하여 질병 형태에 따른 분자생물학적 기전을 밝히고자 하는 시도가 활발히 이루어지고 있다. 전사체(transcriptomics) 연구는 다양한 질병 형질(phenotype)을 유전자 발현 양상과 연결함으로써 특정 질병 형질에 대한 분자 생물학적 해석의 실마리를 제공하며, 궁극적으로 개개인에게 맞춤형 치료를 가능케 하는 개인 맞춤 의학(precision medicine)의 수단으로 응용될 수 있다¹. 가장 흔한 형태의 COPD 전사체 연구는 환자군과 대조군을 비교하여 차등 발현되는 유전자를 찾는 연구이다. 지금까지 이러한 연구를 통하여 COPD에 기여하는 다양한 생물학적 기전이 밝혀졌다. 하지만, 폐 기능을 바탕으로 한 진단명으로는 다양한 임상 양상과 형질을 분류할 수 없기 때문에 여기에서는 COPD 세부 형질과 치료 반응을 분류하는 연구에 대해서 소개할 것이다. 또한, 기존 전사체 연구의 한계점과 이를 개선하기 위한 새로운 방법론에 대해서 소개하고자 한다.

2. 전사체 연구란 무엇인가

과거의 전사체에 대한 정의는 세포와 조직²에서 발현되어 단백질로 번역되는 mRNA만을 지칭하였으나, 최근 전사체 연구에서는 miRNA, lncRNA 등 다양한 종류의 비번역 RNA를 포함한다. 전사체는 발달 단계, 병태생리학적 변화, 환경적 요인에 대한 각 세포/조직의 전반적인 유전자 발현 변화를 반영하기 때문에 유전체 연구와

더불어 가장 먼저 시작된 오믹스(omics) 연구 분야 중 하나이다. 전사체 연구는 1990년대에 일부 mRNA 조각을 시퀀싱하는 serial analysis of gene expression (SAGE)를 바탕으로 활성화되었으며, 다량의 전사체 발현량을 동시에 측정할 수 있는 방법은 1995년 마이크로어레이 기술이 개발된 이후 본격적으로 시작되었다³. 마이크로어레이는 알려진 전사체의 염기서열의 상보 서열(cDNA)을 독립적으로 결합시킬 수 있는 프로브를 담고 있는 전자 칩을 응용하여 형광 물질로 염색시킨 전사체 검체의 발현량을 레이저 반사 파장을 스캐너로 인식함으로써 측정할 수 있다. 2000년대부터 사용되고 있는 RNA 시퀀싱(RNA Sequencing) 방법은 전사체의 염기서열 자체를 인식하는 기술로 지놈 프로젝트(Genome project)를 통해 만들어진 참조 유전체 서열과의 매핑을 통해 모든 전사체의 발현량을 상대적으로 정량화할 수 있다. 마이크로어레이의 경우 전자 칩 상의 프로브 디자인을 통해 측정할 수 있는 상보적 서열의 한계가 정해진 반면, RNA 시퀀싱 기술은 보다 적은 양의 RNA로 훨씬 넓은 범위의 발현 정도를 측정할 수 있다는 장점이 있다. 일례로 마이크로어레이가 수백 배의 유전자 발현 차이를 측정할 수 있다면 RNA-Seq은 8,000배 이상의 차이를 감지해낼 수 있다⁴. 또한, 기존에 알려지지 않은 이형질성 구조를 지닌 전사체를 새롭게 발견할 수 있는 시퀀싱의 장점을 가지고 있다. 이러한 장점들과 더불어 최근 시퀀싱 기술의 비용이 저렴해짐에 따라 RNA 시퀀싱 방법은 전사체 연구에서 가장 흔히 쓰이는 방법이 되었다.

3. 전사체 연구를 통한 COPD 형질 분류

전사체를 통해 COPD 형질을 분류하고 이해하기 위해 연구의 대상이 되는 형질이 명확하게 정의되어 있어야 하며, 다양한 형질의 분류가 가능토록 검체의 수가 충분해야 한다. 폐조직을 이용한 전사체 연구는 질병 조직을 직접 연구할 수 있다는 장점이 있는 반면, 침습적 채취에 의존하기 때문에 검체 수가 적은 경우가 많다. 그러나 검체 수가 적어도 같은 환자에서 서로 다른 정도의 병변을 채취함으로써 유용한 정보를 얻을 수도 있다. Campbell 등은 6명의 환자와 2명의 대조군에서 각각 폐기종의 정도가 다른 8곳의 폐조직을 채취하여 64개 검체의 마이크로어레이 데이터와 micro CT를 통해 측정된 mean linear intercept를 연계분석하여 폐기종 정도와 관련된 유전자 127개를 보고하였다. B cell receptor signaling pathway와 연관된 유전자의 발현은 증가되어 있었으며 TGF β pathway 유전자는 발현이 감소되어 있었다⁵. 또한 유전자 리스트를 기존의 therapeutic compendium 데이터와 비교함으로써 tripeptide GHK가 폐기종으로 인한 유전자 발현 변화를 정상화시킬 수 있음을 보고하였다⁵. 최근 폐조직을 이용한 연구에서 충분한 양의 검체 수를 기반으로 COPD 형질과의 관계를 조사하는 연구도 이루어지고 있다. Faner 등은 COPD 환자 70명에서 얻은 폐조직 전사체 연구에서 각각의 환자를 CT 영상과 폐 기능검사에 근거해 폐기종과 모세 기관지염(bronchiolitis)의 정도에 따라 분류한 후, 각 임상 형질마다 나타나는 유전자 발현을 여러 가지로 비교하여 보고하였다⁶. 폐기종과 모세 기관지염을 비교하였을 때, 폐기종에서 면역 관련 유전자 특히 B cell response 관련 유전자의 발현이 증가되어 있었으며 remodeling and scarring과 관련된 유전자의 발현이 감소되어 있었다⁶. 폐 기능 저하 정도가 같은 환자들을 대상으로 폐기종과 모세 기관지염을 비교하였을 때는 폐기종에서 B cell activation과 lymphoid follicle formation 관련 유전자의 발현이 증가되어 있었다⁶. 또한, 폐기종 환자만을 대상으로 중증과 경증을 비교하였을 때 역시 B cell activation과 lymphoid follicle formation과 관련된 유전자가 중증의 폐기종과 연관되어 있었다⁶. 이 연구는 폐 기능 저하의 정도와 독립적으로 폐기종과 모세 기관지염이 서로 다른 유전자의 발현, 즉 분자생물학적 기전을 가지고 있음을 보여주었으며, 폐기종에서 B cell과 lymphoid follicle formation이 중요하다는 점을 시사하였다. Morrow 등은 COPD 환자 111명과 흡연자 대조군 40명에서 얻은 폐조직에서 상호 발현되는 유전자들에 대한 네트워크 분석(weighted gene co-expression network, WGCNA)을 통해 CT로 측정된 폐기종 및 FEV1 percent predicted 형질들이 B cell과 관련된 기전과 연관되어 있음을 보고하였다⁷. Faner 연구에서는 연구자가 정의한 형질과 전사체 발현 간 상관 분석 방법을 사용하였으며, Morrow 등의 연구에서는 전사체 발현 정보에 대한 클러스터링 분석을 통해 공동 발현되는 유전자 모듈을 찾은 후 이를 형질과 상관분석하였다. 이 두 연구에서 서로 다른

독립적인 방법론을 사용하였지만 같은 결론에 도달하였기 때문에 폐기종에서 B cell의 중요성에 대한 결과가 높은 신뢰성이 있다고 평가된다.

폐조직 대신 접근이 상대적으로 용이한 조직을 이용한 전사체 연구도 활발하다. 기관지 내시경으로 채취한 기관지 상피세포, 비상피 세포(nasal epithelial cell), 객담, 혈액이 대표적이다. Steiling 등은 87명의 COPD 환자와 151명의 대조군의 기관지 상피세포 전사체 분석을 통해 98개의 유전자가 COPD, FEV1 percent predicted, FEV1/FVC와 연관되어 있음을 보여주었다⁸. Christenson 등은 COPD와 천식 환자의 기관지 상피세포 전사체를 비교하여 천식-COPD 중복 증후군(asthma-COPD overlap)에서 존재하는 T helper type 2 cell의 분자생물학적 특징을 밝혔으며, 흡입성 스테로이드에 대한 치료 반응을 예측할 수 있음을 보였다⁹. 이는 천식-COPD 중복 증후군의 임상적 특성들(기관지 확장제 반응, 호산구 증가, 흡입성 스테로이드에 의한 hyperinflation 감소 등) 혹은 천식 병력이 없는 환자에서도 천식의 분자생물학적 기전과의 공통점이 있다는 것을 뒷받침하며, 전사체 연구를 통해 COPD의 형질을 세분화할 수 있는 좋은 예라고 할 수 있다. COPD 환자에서 기관지 상피세포와 비상피 세포 전사체가 유사하다는 것이 보고되었으며^{10,11}, 대규모 COPD 코호트인 COPDGene 환자를 대상으로 수행되는 비상피 세포 RNA 시퀀싱 프로젝트가 예정되어 있어 비상피 세포 전사체를 이용한 COPD 형질 연구도 곧 이루어질 전망이다.

객담과 혈액은 검체 획득이 상대적으로 용이하기 때문에 연구뿐 아니라 임상적으로 응용할 수 있는 가능성이 가장 높은 조직이다. 객담과 혈액에서 유래된 전사체를 분석하여 COPD 격화(exacerbation)가 잦은 환자군과 없는 환자군을 구별하려는 시도가 있었으며, 유전자 발현을 기반으로 예측한 정확도는 임상 병력(90%)과 비슷한 민감도(91%)를 보였다¹². 혈액을 이용한 전사체 연구의 경우, 충분한 검체 수를 바탕으로 다양한 통계 방법을 활용한 연구가 다양하게 이루어지고 있다. Morrow 등은 혈액 전사체의 네트워크 분석을 이용하여 빈번한 격화와 관련된 유전자 그룹을 보고하였으며¹³, Obeidat 등도 COPD 환자 혈액 전사체를 이용한 네트워크 분석을 통해 FEV1과 연관되어 있는 유전자들을 보고하였다¹⁴. 북미의 대규모 COPD코호트인 SPIROMICS와 COPDGene에서 혈액으로부터 유래된 전사체 데이터가 수집되고 있으며, 이러한 노력들을 통해 기존에 알려지지 못한 COPD 형질과의 관계에 대한 다양한 연구결과가 나올 전망이다.

4. 전사체 연구를 통한 COPD 치료 반응 분류

약물에 대한 유전자 발현 연구는 치료 기전을 이해하고 치료 반응을 예측하는 데 응용될 수 있다. 한 예로, 스테로이드는 유전자 전사에 직접적으로 영향을 주기 때문에 스테로이드로 인한 유전자 발현은 일관적이고 강하게 나타난다¹⁵. van den Berge 등은 기존에 스테로이드에 노출되지 않은 COPD 환자를 흡입성 스테로이드와 위약군으로 나누어 추적한 GLUCOLD 연구를 통해 흡입성 스테로이드로 인한 기관지 상피세포의 전사체 발현 변화를 보고하였다¹⁶. 이 연구에서 치료 전, fluticasone +/- salmeterol 치료 후 6개월, 30개월의 기관지 상피세포 전사체를 분석하여 278개의 유전자가 fluticasone +/- salmeterol 치료와 연관되어 있음을 보였다¹⁶. 특히 epithelial cell signaling, oxidative stress, remodeling에 관련된 유전자 발현이 하향 조절되어 있었으며, 하향 조절된 유전자 발현에 따라 FEV1 감소 폭이 적었다¹⁶.

최근에 발표된 Christenson 등의 연구는 전사체 분석을 통해 특정 임상 형질이 치료 반응과 연관이 있음을 보여주었다¹⁷. COPD 환자군과 대조군의 기관지 상피세포 전사체분석을 통해 얻어진 IL-17A 반응 유전자들이 독립적으로 얻어진 SPIROMICS 코호트에서 기도폐색(airway obstruction)과 CT상 소기관지 질환과 연관되어 있고, GLUCOLD 환자군에서는 호산구 수나 type 2 면역 반응과 상관없이 스테로이드에 대한 반응이 적은 특징을 지닌다는 것을 밝혔다¹⁷.

5. 기존 전사체 연구의 한계와 발전 방향

대부분 전사체 발현량 측정방법은 세포 전체의 평균적인 유전자 발현을 측정하는 벌크-세포 발현량 측정 기법을 기반으로 수행된다. 하지만, 폐처럼 다양한 종류의 세포로 구성되어 있는 조직에서는 조직 내 세포의 구성 차이에 의해 유전자 발현의 차이가 발생할 수 있다. 앞서 기술한 폐기종과 B cell 유전자 발현 연구의 경우, 폐기종의 B cell 유전자 발현이 달라진 것일 수도 있지만 단순히 B cell의 숫자가 증가했을 가능성이 있다. 그러므로 벌크-세포 발현량 측정 기법으로는 이와 같은 차이를 구별해 낼 수 없다는 한계점이 존재한다. 최근, 이를 해결하기 위한 단일 세포 전사체 발현량 측정 기법(single cell transcriptomics) 기반 전사체 연구가 보편화되는 추세이다. 특히, 단일 세포 RNA 시퀀싱 기술은 전사체 서열 정보와 함께 각각의 세포에 대한 발현량 정보를 동시에 얻을 수 있기 때문에 복잡한 조직에서 세포의 종류와 구성, 세포별 유전자 발현 상태와 차이에 대한 분석을 수행할 수 있다. 수천 개에서 수만 개의 세포에 대한 단일 세포 전사체에 대한 탐구가 가능해짐에 따라 폐섬유화증 등의 질환에서 연구 결과가 보고되었으며^{18,19}, 낭포성 섬유증의 원인이 되는 CFTR 유전자가 발현되어 있는 새로운 종류의 세포도 밝혀졌다^{20,21}. 아직까지 COPD 관련 전사체 연구 분야에서 단일 세포 시퀀싱 기술을 활용한 연구 결과가 발표된 적은 없다. 하지만, 필자가 속한 연구실을 비롯하여 COPD에서 보다 높은 분자 생물학적 해상력을 지닌 단일 세포 RNA 시퀀싱 기술을 활용한 연구가 진행되고 있다.

두 번째로 유전자 발현에 대한 상관 분석에서 특정 생물 현상에 대한 인과관계를 구분할 수 없는 한계가 있다. 전사체 통계 분석을 통해 도출된 차등 발현 유전자가 COPD 특정 형질에 대한 원인 또는 결과임을 구분하기 위한 추가적인 실험을 수행해야만 한다. 또한, 질병의 경중에 따른 차이, 시간에 따른 차이를 보는, 보다 더 복잡한 실험 디자인을 통해 이와 같은 한계를 보완할 수 있다.

세 번째로 COPD같은 복합질환에서 단일 전사체로는 질병 설명력의 한계가 존재하며²² 유전체학, 후성 유전체학, 단백질체학, 메타지노믹스 등의 다양한 오믹스 마커 사이의 관계를 설명하지 못한다. 이에 지난 10년간 다양한 오믹스 데이터를 통합 분석할 수 있는 전산학적/통계학적 분석 방법이 개발되어 왔으며, 그중, 유전체-전사체의 관계를 분석하는 eQTL (expression quantitative trait loci) 분석의 발전이 두드러졌다²³. 이 통합 분석은 유전체 변이가 특정 전사체의 발현량의 차이를 만듦에 따라 간접적으로 질병 관련 형질에 영향을 끼칠 것이라는 가설 하에 실시하는 유전체-전사체의 상관 분석 방법이다. 특히, 전사체의 프로모터에 위치한 유전체 변이가 직접적으로 인근 전사체의 발현량을 조절하는 기전인 cis-eQTL을 조사하는 연구가 COPD에서 진행되어 이미 알려진 유전체 변이들에 대한 기능적 해석을 하는데 성공하였다²⁴. 최근에 발표된 환경-특이적 eQTL (conditional-specific eQTL)은 COPD 환자에 대한 산소치료의 효능을 알아보는 LOTT trial 데이터를 기반으로, 산소 요법 여부에 따라 eQTL의 방향성이 달라질 수 있음을 처음으로 제시하였다²⁵. 이 외에도 후성유전체-전사체의 관계를 살펴보는 meQTL (methylation-expression quantitative trait loci)²⁶, 단백질 발현량-유전체의 관계를 살펴보는 pQTL (protein quantitative trait loci)²⁷ 등이 연구 중이며, 가까운 미래에 COPD 관련 연구에도 적용될 것으로 기대된다. 이 같은 전사체와 다른 오믹스 데이터 간 다중 분석을 통해 기존에 볼 수 없었던 생물학적 마커들 사이의 관계를 밝히고, 더 나아가 COPD에 대한 이해를 높일 것으로 기대한다.

References

1. Steiling K, Lenburg ME, Spira A. Personalized management of chronic obstructive pulmonary disease via transcriptomic profiling of the airway and lung. *Ann Am Thorac Soc* 2013;10 Suppl:S190-6.
2. Velculescu VE, Zhang L, Zhou W, Vogelstein J, Basrai MA, Bassett DE Jr, et al. Characterization of the yeast transcriptome. *Cell* 1997;88:243-51.
3. Lowe R, Shirley N, Bleackley M, Dolan S, Shafee T. Transcriptomics technologies. *PLoS Comput Biol* 2017;13:

- e1005457.
4. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 2009;10:57-63.
 5. Campbell JD, McDonough JE, Zeskind JE, Hackett TL, Pechkovsky DV, Brandsma CA, et al. A gene expression signature of emphysema-related lung destruction and its reversal by the tripeptide GHK. *Genome Med* 2012; 4:67.
 6. Faner R, Cruz T, Casserras T, López-Giraldo A, Noell G, Coca I, et al. Network analysis of lung transcriptomics reveals a distinct B-cell signature in emphysema. *Am J Respir Crit Care Med* 2016;193:1242-53.
 7. Morrow JD, Zhou X, Lao T, Jiang Z, DeMeo DL, Cho MH, et al. Functional interactors of three genome-wide association study genes are differentially expressed in severe chronic obstructive pulmonary disease lung tissue. *Sci Rep* 2017;7:44232.
 8. Steiling K, van den Berge M, Hijazi K, Florido R, Campbell J, Liu G, et al. A dynamic bronchial airway gene expression signature of chronic obstructive pulmonary disease and lung function impairment. *Am J Respir Crit Care Med* 2013;187:933-42.
 9. Christenson SA, Steiling K, van den Berge M, Hijazi K, Hiemstra PS, Postma DS, et al. Asthma-COPD overlap. Clinical relevance of genomic signatures of type 2 inflammation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2015;191:758-66.
 10. Steiling K, Lenburg ME, Spira A. Airway gene expression in chronic obstructive pulmonary disease. *Proc Am Thorac Soc* 2009;6:697-700.
 11. Boudewijn IM, Faiz A, Steiling K, van der Wiel E, Telenga ED, Hoonhorst SJM, et al. Nasal gene expression differentiates COPD from controls and overlaps bronchial gene expression. *Respir Res* 2017;18:213.
 12. Singh D, Fox SM, Tal-Singer R, Bates S, Riley JH, Celli B. Altered gene expression in blood and sputum in COPD frequent exacerbators in the ECLIPSE cohort. *PLoS One* 2014;9:e107381.
 13. Morrow JD, Qiu W, Chhabra D, Rennard SI, Belloni P, Belousov A, et al. Identifying a gene expression signature of frequent COPD exacerbations in peripheral blood using network methods. *BMC Med Genomics* 2015;8:1.
 14. Obeidat M, Nie Y, Chen V, Shannon CP, Andiappan AK, Lee B, et al. Network-based analysis reveals novel gene signatures in peripheral blood of patients with chronic obstructive pulmonary disease. *Respir Res* 2017; 18:72.
 15. Kan M, Shumyatcher M, Himes BE. Using omics approaches to understand pulmonary diseases. *Respir Res* 2017;18:149.
 16. van den Berge M, Steiling K, Timens W, Hiemstra PS, Sterk PJ, Heijink IH, et al. Airway gene expression in COPD is dynamic with inhaled corticosteroid treatment and reflects biological pathways associated with disease activity. *Thorax* 2014;69:14-23.
 17. Christenson SA, van den Berge M, Faiz A, Inkamp K, Bhakta N, Bonser LR, et al. An airway epithelial IL-17A response signature identifies a steroid-unresponsive COPD patient subgroup. *J Clin Invest* 2018. doi: 10.1172/JCI121087. [Epub ahead of print]
 18. Xu Y, Mizuno T, Sridharan A, Du Y, Guo M, Tang J, et al. Single-cell RNA sequencing identifies diverse roles of epithelial cells in idiopathic pulmonary fibrosis. *JCI Insight* 2016;1:e90558.
 19. Reyfman PA, Walter JM, Joshi N, Anekalla KR, McQuattie-Pimentel AC, Chiu S, et al. Single-cell transcriptomic analysis of human lung reveals complex multicellular changes during pulmonary fibrosis. *bioRxiv* 2018;296608. doi: 10.1101/296608. [Epub ahead of print]
 20. Montoro DT, Haber AL, Biton M, Vinarsky V, Lin B, Birket SE, et al. A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. *Nature* 2018;560:319-24.
 21. Plasschaert LW, Žilionis R, Choo-Wing R, Savova V, Knehr J, Roma G, et al. A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. *Nature* 2018;560:377-81.
 22. Castaldi PJ, Cho MH, Zhou X, Qiu W, Mcgeachie M, Celli B, et al. Genetic control of gene expression at novel and established chronic obstructive pulmonary disease loci. *Hum Mol Genet* 2015;24:1200-10.
 23. Shabalín AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* 2012;28:1353-8.

24. Kim WJ, Lee SD. Candidate genes for COPD: current evidence and research. *Int J Chron Obstruct Pulmon Dis* 2015;10:2249-55.
25. Seo M, Qiu W, Bailey W, Criner GJ, Dransfield MT, Fuhlbrigge AL, et al. Genomics and response to long-term oxygen therapy in chronic obstructive pulmonary disease. *J Mol Med (Berl)* 2018;96:1375-85.
26. Heyn H, Moran S, Hernando-Herraez I, Sayols S, Gomez A, Sandoval J, et al. DNA methylation contributes to natural human variation. *Genome Res* 2013;23:1363-72.
27. Wu L, Candille SI, Choi Y, Xie D, Jiang L, Li-Pook-Than J, et al. Variation and genetic control of protein abundance in humans. *Nature* 2013;499:79-82.