Biochemical and Biophysical Research Communications 452 (2014) 276–286

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Review

Integrative genomics of chronic obstructive pulmonary disease

Brian D. Hobbs, Craig P. Hersh *

Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA, United States
Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Boston, MA, United States
Harvard Medical School, Boston, MA, United States

ARTICLE INFO

Article history:
Received 20 June 2014
Available online 29 July 2014

Keywords:
Chronic obstructive pulmonary disease
Genomewide association study
Genomics
Network medicine

ABSTRACT

Chronic obstructive pulmonary disease (COPD) is a complex disease with both environmental and genetic determinants, the most important of which is cigarette smoking. There is marked heterogeneity in the development of COPD among persons with similar cigarette smoking histories, which is likely partially explained by genetic variation. Genomic approaches such as genomewide association studies and gene expression studies have been used to discover genes and molecular pathways involved in COPD pathogenesis; however, these "first generation" omics studies have limitations. Integrative genomic studies are emerging which can combine genomic datasets to further examine the molecular underpinnings of COPD. Future research in COPD genetics will likely use network-based approaches to integrate multiple genomic data types in order to model the complex molecular interactions involved in COPD pathogenesis. This article reviews the genomic research to date and offers a vision for the future of integrative genomic research in COPD.

Contents

* Corresponding author at: Channing Division of Network Medicine, 181 Longwood Ave., Boston, MA 02115, United States. Fax: +1 617 525 0958.
E-mail address: craig.hersh@channing.harvard.edu (C.P. Hersh).

http://dx.doi.org/10.1016/j.bbrc.2014.07.086
0006-291X/

1. Introduction

Chronic obstructive pulmonary disease (COPD) is the third leading cause of death worldwide, accounting for an estimated 3 million deaths in 2010 [1]. COPD is a complex disease with both genetic and environmental determinants, the most important of which is cigarette smoking. However, only a minority of smokers develop COPD, and there is significant variability in lung function across smokers with similar cigarette exposure histories [2]. Some of this heterogeneity in the development of COPD is likely a result of genetic variation. A known genetic risk factor for COPD is severe α_1-antitrypsin (AAT) deficiency, which is the result of mutations in the SERPINA1 (serpin peptidase inhibitor A1) gene [3,4]. AAT deficiency accounts for approximately 1% of COPD cases and thus is an insufficient genetic factor to account for the heterogeneity in COPD [4].

The complex genetic component of COPD was elucidated through linkage analysis and familial aggregation studies [5]. In previous decades, a series of candidate gene association studies tested genes felt to be important in the COPD pathogenesis. As candidate gene studies by design focused on genes with known potential relationship with COPD pathogenesis, they had no ability to identify novel COPD susceptibility regions. The advent of genome wide association studies (GWAS) allowed for examinations of hundreds of thousands to millions of single nucleotide polymorphisms (SNPs) across the genome. Since 2009, GWAS have revealed replicable statistical associations for COPD. GWAS have also discovered loci related to lung function, emphysema, and other COPD phenotypes (Table 1). Despite the success of GWAS, a large portion of the heritability – the phenotypic variability that is attributable to genetics – remains unexplained [6,7]. "Missing heritability" is not a unique feature for COPD and is a problem across all complex diseases [8]. There are many proposed explanation for the missing heritability in COPD and other diseases, and additional genomic research, including the integration of distinct sources of genomic data, is required.

Silverman and Loscalzo describe three generations of genomics studies (Fig. 1) [9]. First generation studies examine the association between genetic variants and disease phenotype. This concept can be broadened to include any association studies that use a single omics technology, e.g., microarray expression profiling. Second generation studies examine associations between two different types of omics data, such as GWAS and gene expression profiling, and aim to correlate the results to disease. Third generation studies utilize multiple omics data types, combined using network methods, and address not only a disease as a whole, but also consider disease subtypes. This article will review the COPD genomics studies to date, which have largely fallen into the first generation category. We will discuss the early progress in integrative genomics studies, which represent the second generation. Future second generation and eventually third generation network medicine studies have the potential to unveil the underlying biology of

Table 1
Phenotypes for COPD genomics studies.

- COPD diagnosis
- Pulmonary function tests
  - Spirometry
    - Forced expiratory volume in 1 s (FEV_1)
    - Ratio of FEV_1 to forced vital capacity
    - Decline in lung function
    - Global initiative for chronic obstructive lung disease (GOLD) stages [114]
      - Stages 1–4, from least to most severe, based on FEV_1% predicted
  - Lung volumes
  - Diffusing capacity for carbon monoxide
- Emphysema on chest computed tomography (CT) scans
  - Visual assessment
  - Quantitative image analysis
  - Emphysema pattern based on local histogram analysis
- Airway disease on chest CT scans
- Symptoms
  - Chronic bronchitis
  - Acute exacerbations
- Physiologic impairment
  - Blood oxygen levels
  - Exercise capacity
- Systemic effects
  - Body mass index
  - Co-morbidities
- COPD subtypes
  - COPD-asthma overlap syndrome
  - Machine learning subtypes

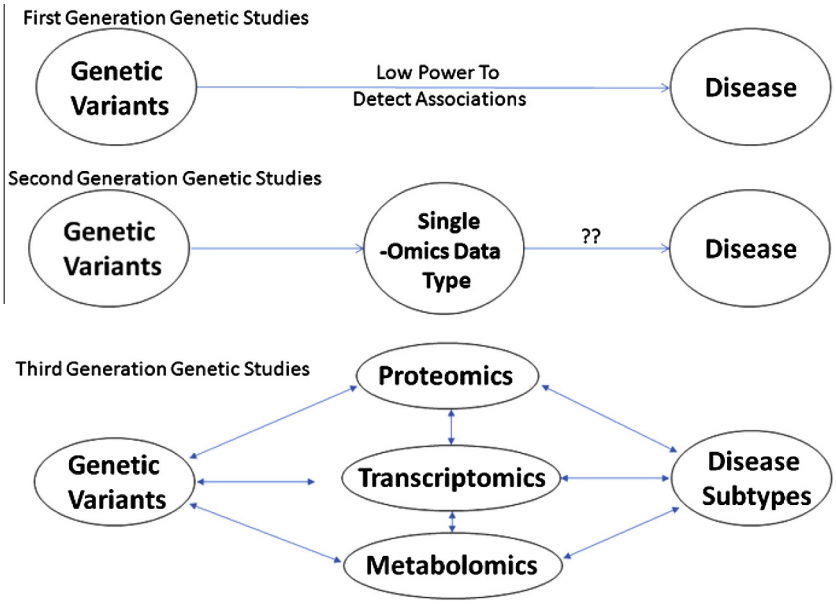

Fig. 1. Three generations of genomic studies, described by Silverman and Loscalzo [9]. Figure reproduced with permission.