Allergology International 65 (2016) 123–126

Contents lists available at ScienceDirect

Allergology International

journal homepage: http://www.elsevier.com/locate/alit

Invited review article

Does epigenetics play a role in human asthma?

Donata Vercelli

Arizona Respiratory Center, Department of Cellular and Molecular Medicine, and Arizona Center for the Biology of Complex Diseases (ABCD), University of Arizona, The BIO5 Institute, Rm. 339, 1657 E. Helen Street, Tucson, AZ 85721, USA

ARTICLE INFO

Article history:
Received 1 December 2015
Received in revised form
4 December 2015
Accepted 7 December 2015
Available online 8 January 2016

Keywords:
Allergy
Birth cohort
Childhood asthma
DNA methylation
Epigenetics

Abbreviations:
GWAS genome-wide association study

ABSTRACT

Asthma and other allergic diseases are among the most prevalent chronic non-communicable diseases of childhood. According to the World Health Organization, asthma affects >7.0 million children under 18 in the United States, with an economic burden that is estimated to exceed that of tuberculosis and HIV/AIDS combined. Despite much research, the natural history of asthma and its pathogenesis are still in many ways elusive. This review discusses our current understanding of the role epigenetic processes play in asthma pathogenesis, focusing on genome-wide, population-based studies.

Introduction

Asthma and other allergic diseases are among the most prevalent chronic non-communicable diseases of childhood.[1,2] According to the World Health Organization, asthma affects >7.0 million children under 18 in the United States, with an economic burden that is estimated to exceed that of tuberculosis and HIV/AIDS combined.[3]

Despite much research, the natural history of asthma and its pathogenesis are still in many ways elusive. While epidemiologic studies indicate that the disease begins in the pre-school years even when chronic symptoms appear in early adulthood,[4,5] firm diagnostic criteria to distinguish children who will wheeze transiently during early life lower respiratory illnesses from children who will wheeze persistently, and then develop asthma, are still lacking. Yet, such criteria are urgently needed, because at least at present asthma can be treated but not cured, and therefore the focus must be on prevention. A number of asthma predictive algorithms have been proposed and refined over time,[6–8] but their sensitivity, specificity and predictive value remain suboptimal. Interestingly, these tools rely on family history and the child's clinical characteristics in early life but do not incorporate variables that can be measured already at birth.

Allergy and asthma epigenetics comes of age

It is in this context that the potential role of epigenetics in regulating the susceptibility to and the severity of asthma and allergic disease is drawing more and more attention, as shown by a continuous and steep rise in the number of publications.[9] Remarkably, certain years saw reviews outnumber primary research papers – a pattern that points to an unusual level of expectation and interest. Such interest conceivably stems from multiple considerations. Epigenetic processes, environmental stimuli and developmental programs are connected by strong functional links. That asthma has a strong environmental component was clearly illustrated by seminal epidemiologic studies that revealed major differences in asthma prevalence among countries with more or less Westernized life styles,[2,10] while the critical role of early life exposures as determinants of asthma during adulthood has been repeatedly emphasized.[2] To the extent that epigenetic mechanisms faithfully and sensitively transduce environmental signals and preside over the time-dependent unfolding of developmental differentiation programs, their involvement in asthma and allergy is both possible and probable. Another critical reason behind the rising interest in asthma and allergy epigenetics is the

E-mail address: donata@email.arizona.edu.
Peer review under responsibility of Japanese Society of Allergology.

http://dx.doi.org/10.1016/j.alit.2015.12.001

inability of genome-wide association studies (GWAS) to account for more than a limited proportion of the total phenotypic variance in asthma[11] even though the disease is well known to have a strong genetic component.[2] Current studies on rare variants (as opposed to the common ones that GWAS typically interrogate) will probably improve the situation only marginally. The realization that genetics cannot "explain" asthma and allergic disease has shifted interest towards potential alternative sources of phenotypic variance, primarily the environment and development, both of which are intertwined with epigenetic events. At the same time, interrogating the epigenome, particularly the methylome, has become increasingly feasible and can be pursued as a tool to assess the contribution of epigenetic mechanisms to allergic disease pathogenesis.

Studying epigenetics

Epigenetics studies heritable changes in gene activity that are independent of alterations in the underlying DNA sequence.[12–14] Among epigenetic modifications, studies in human asthma and allergic diseases have primarily focused on DNA methylation, a process with intimate albeit complex connections with the regulation of gene expression. DNA methylation is a robust epigenetic mark, and user-friendly, quantitative methods to extensively survey the methylome are now widely available and are replacing candidate gene studies. These genome-wide methods rely on straightforward assays and streamlined analytical pipelines, and require DNA rather than chromatin isolation procedures.[15] Therefore DNA methylation studies, unlike the more challenging analyses of post-translational histone modifications, are flourishing and virtually represent the totality of the epigenetic studies performed in human populations with asthma and allergy.

DNA methylation occurs at cytosines in individual CpG dinucleotides and in clusters of CpG sites called CpG islands. The relationship between DNA methylation and gene expression is context-dependent. Typically, promoter methylation promotes gene silencing, whereas an unmethylated promoter is necessary but not sufficient for gene expression and the relevant gene is said to be poised for expression. However, methylation in gene bodies is often associated with high gene transcription.[15,16]

Current techniques to study DNA methylation rely on bisulfite conversion of DNA, a reaction in which unmethylated cytosines are converted to thymines and only methylated cytosines are preserved as such in the sequence. While the decreasing cost and higher quality of next-generation sequencing are improving research in many areas of genomics, whole-genome bisulfite sequencing is still challenging because the reduced complexity of the DNA sequence that results from the bisulfite conversion process makes it difficult to unambiguously map 100–200 bp-long reads to the genome. Therefore, targeted (gene- and/or pathway-specific) bisulfite sequencing or genome-wide microarrays are still widely used. Currently, the most popular platform for genome-wide DNA methylation profiling is the 450k Illumina Human Methylation BeadChip that interrogates approximately 450,000 CpG sites throughout the genome. This platform surveys individual CpG sites and its output can be intuitively understood as the percentage of DNA methylation at each site.[17] The Illumina 450k array was designed to cover 99% of RefSeq genes with a global average of 17.2 probes per gene region, 96% of CpG islands, 92% of CpG shores, and 86% of CpG shelves.[17] Differences in DNA methylation are detected using two types of Infinium probes on bisulfite-converted DNA. Infinium I probes are designed in pairs (one against the methylated locus and the other against the unmethylated locus). In contrast, Infinium II probes are designed to bind both the methylated and the unmethylated locus. The methylation state is detected upon a single base extension and a distinct signal is given from labeled nucleotides.[17] It is noteworthy that despite their popularity, these arrays are unable to discriminate 5-methylcytosine from 5-hydroxymethylcytosine,[18] an epigenetic modification that occurs primarily in the brain and embryonic stem cells.

Most genome-wide studies performed in asthma and allergy have targeted DNA methylation in peripheral blood cells, and only a few have surveyed airway cells or tissues. Indeed, although the lung is a major target organ in asthma and allergy, obtaining lung tissue can be problematic in adults and is virtually impossible in children. On the other hand, because immune alterations accompany and often precede a clinical diagnosis of allergic diseases and asthma,[2,19–22] information from peripheral blood immune cells may still be relevant to disease pathogenesis.

Epigenetic alterations in concurrent allergic disease

So far, most if not all genome-wide epigenetic studies in allergy and asthma sought to identify DNA methylation signatures in patients with *concurrent* disease. Thus a recent study surveyed associations between serum IgE concentrations and DNA methylation in 95 nuclear pedigrees from individuals in their twenties. Replicated associations between IgE and low methylation were found at 36 loci. Genes annotated to these loci encode known eosinophil products, and also implicate phospholipid inflammatory mediators, specific transcription factors and mitochondrial proteins.[23] Another recent study compared DNA methylation patterns and gene expression in 6–12 years old inner-city children with persistent atopic asthma versus healthy control subjects. Results were validated in an independent population of asthmatic patients. Eighty-one differentially methylated regions were identified. Several immune genes were hypomethylated in asthma, including IL13, RUNX3, and TIGIT. Among asthmatic patients, 11 differentially methylated regions were associated with higher serum IgE concentrations, and 16 were associated with percent predicted FEV_1. Methylation marks involved in T-cell maturation (RUNX3), T_H2 immunity (IL4), and oxidative stress (catalase) were validated in an independent asthmatic cohort of children living in the inner city.[24]

Food allergy was the phenotype analyzed in 11–15 month old children using a supervised learning approach to discover a 96-CpG signature that distinguished food-allergic and food-sensitized individuals as well as food-allergic and non-allergic infants.[25] The authors also showed that their methylation signature outperformed both egg- and peanut-specific serum IgE levels as a predictor of clinical allergy. Of note, food allergy status was correctly predicted in a replication cohort of 48 individuals with an accuracy of 79.2%.

A growing challenge with epigenetic epidemiology is that a vast amount of data is generated and new statistical techniques are necessary to make sense of it. This is because of small-n-large-p (few observations relative to the number of predictors) and because traditional methods are not optimized for identifying complex biological processes. Two recent studies relied on Random Forest, a machine learning algorithm used for classification that can handle the data issues discussed above.[26] A forest composed of classification trees is grown using randomly selected bootstrap samples of the data to form training and testing sets of study participants. At each node within each tree, the training set is partitioned into different classes with the split determined by a subset of randomly chosen predictors. These two levels of randomness, random selection of training/testing sets and random testing of predictors, allow the random forest to produce robust classification predictions. Once the forest is grown using the training sets, the observations in the testing sets are classified via the forest and misclassification rates can be used to evaluate the accuracy of the forest.[26] Of note, it is likely that methylation changes at a