Disease Study

Genetic Underpinnings of T2 (Eosinophilic) Versus Non-T2 (Non-Eosinophilic) Asthma
Genetic Underpinnings of T2 (Eosinophilic) Versus Non-T2 (Non-Eosinophilic) Asthma

Executive Summary

The PrecisionLife® platform uses a unique mathematical framework and data analytics to enable faster and deeper analysis of large, multi-omic patient datasets, bringing together genomic, epidemiological, clinical, and phenotypic data. PrecisionLife reveals new insights into the multifactorial causes of complex diseases, and predicts patients’ different experiences of disease and response to therapies.

Asthma patients can be broadly categorized into two molecular phenotypes: those with high type 2 T-helper cell expression (T2), and those with low type 2 T-helper cell expression (non-T2). Asthma patients with a T2 phenotype currently have a range of targeted biologic treatment options available to them. Non-T2 patients, however, lack personalized therapy, and often have to rely on conventional symptomatic control therapies (such as bronchodilators and inhaled corticosteroids) that do little to combat the underlying disease pathology. We performed a comparative study using a genotype dataset derived from UK Biobank to compare T2 and non-T2 asthma patient populations. Using PrecisionLife with UK Biobank data, the study took less than three days from conception to assembly of data, to analysis and interpretation.

The study identified clear differences in genetic pathways between the T2 and non-T2 asthma cohorts that hold significant potential for better patient stratification and diagnosis biomarkers, and new treatment options. While most of the significant disease-associated genes in the T2 cohort related to immune pathways and interleukins characteristic of Th2-driven allergic asthma, many of the genes that were significant in non-T2 asthmatic patients corresponded to metabolic and neuronal pathways.

Background

Asthma is a debilitating disease that affects 1 in 13 people. Some 5.4 million people are currently receiving asthma treatment in the UK. A subset of these patients suffers from a severe form of asthma now called T2 (eosinophilic) asthma. The exact prevalence of T2 asthma is unknown; however, approximately 10% of all asthma is categorized as severe. T2 asthma is most commonly diagnosed in adults 35–50 years old, although it is sometimes seen in older adults and pediatric patients. T2 asthma affects males and females equally.

The etiology of T2 asthma involves activation of the type 2 T-helper (Th2) cells, which results in the release of cytokines such as IL-5 and IL-13. In turn, these cytokines recruit eosinophils to the affected tissue to counter the antigen(s) that triggered the Th2 system.

However, as many as a third of asthma patients do not have high blood eosinophil counts and present as non-allergic cases. While there have been many Genome-Wide Association Studies (GWAS) on asthma to date, prior studies have not focused on the genetic differences between T2 and non-T2 forms of asthma. Our understanding of the genetics of T2 asthma is largely based on studies of the Th2 cytokine pathways. One earlier GWAS focused on genetic defects affecting eosinophil counts, but this study did not draw strong conclusions about the genetic architecture of T2 asthma.

Therefore, we analyzed UK Biobank data from asthma cases and healthy controls with an aim to identify significant genotype differences between T2 and non-T2 asthma.

Methods

PrecisionLife is a massively scalable, multi-omics association platform that enables the detection of high-order epistatic interactions at genome-wide study scale. PrecisionLife can find and statistically validate combinations of five or more single nucleotide polymorphism (SNP) genotypes that are found in many cases and very few controls, and associate those combinations with specific phenotypes.

In this study, we first identified asthma cases in UK Biobank using a slightly modified version of the case selection criteria presented by Ferreira et al. Using these criteria, we identified a total of 42,205 asthma cases. We randomly selected 90,034 age- and gender-matched subjects from the same database to serve as controls.

We also selected two different subsets of cases to study the genomic differences between T2 and non-T2 asthma. While UK Biobank does not have data from sputum samples, blood eosinophil counts are considered a reasonable indicator of eosinophilia in the airways. Therefore, we used blood eosinophil counts from the UK Biobank database to separate cases into T2 vs. non-T2 cohorts. Where a case had multiple blood eosinophil counts, we used the value from the earliest test. We selected a total of 15,071 cases with serum
eosinophil counts of 0.15 (1,500 cells/mm$^3$) or lower as the non-T2 cohort, and a total of 7,094 cases with serum eosinophil counts of 0.35 (3,500 cells/mm$^3$) or higher as the T2 cohort. As some asthma cases did not have any eosinophil counts recorded, we excluded them from either group. In order to reduce errors due to misclassification, we also excluded a large group of cases with eosinophil counts between 0.15 and 0.35, which we considered to be moderate or borderline values. Finally, we selected an age- and gender-matched control cohort of 21,688 subjects without asthma or similar respiratory disease. After quality control filtering, the genotype dataset included 547,147 SNPs for each case and control subject.

**Genetic Architecture of Asthma**

First, we conducted an initial study analyzing the full asthma dataset (42,205 cases versus 90,034 controls). As asthma is a fairly heterogeneous disease, we expected to find significant variation in genetic defect patterns underlying asthma subtypes. We analyzed the full asthma dataset to identify statistically significant combinations of up to five SNP genotypes. Each of these combinations (or networks) of SNP genotypes is represented by a “critical” SNP genotype based on its prominence among all networks. We then used a hierarchical clustering method on critical SNP genotypes vs. case samples to reveal the genetic organization of asthma (see Figure 1).

Our first attempt was a k-means clustering with k=4. We noted a block of SNP genotypes represented in virtually all asthma cases, while four other case blocks appeared to represent specific subpopulations of disease. A pathway enrichment analysis revealed that this block of SNPs all related to genes in pathways involved in DNA methylation and histone modification, indicating that there is a significant element of epigenetic modification involved in driving asthma pathogenesis, regardless of patient subtype.

Next, we tried a different k-means clustering approach with k=6. We projected the resulting six clusters onto a hierarchical clustering plot of SNP genotypes vs. SNP genotypes. Since the resulting plot does not include individual cases, it cannot be mapped onto phenotype clusters. However, the new plot (see Figure 2) illustrates network connections between SNP genotypes that underlie asthma, and thus provides our first glimpse into pathways that might be implicated in asthma causation.

Subsequent pathway enrichment analysis of all the SNP genotypes in these clusters revealed that many of the most critical SNP clusters (critical SNPs which co-occur frequently in patients) are related to immune system pathways, as well as antigen presentation and processing.

**Figure 1** Hierarchical k-means clustering of 42,205 asthma cases based on critical SNP genotypes. Individual SNP genotypes are in the columns, and individual cases are in the rows. A small block of SNP genotypes is represented in virtually all cases, depicted as a narrow darker stripe along the right edge of the cluster diagram. In addition, four significant case clusters are visible (indicated along the left-hand side of the cluster diagram in blue, red, pink, and green).

**Figure 2** Hierarchical k-means clustering of 42,205 asthma cases into six clusters. Both columns and rows indicate individual SNP genotypes. Six different clusters are visible, indicated with pink (cluster 1), green (cluster 2), blue (cluster 3), orange (cluster 4), red (cluster 5), and cyan (cluster 6).
Exploring the Genetic Differences Between T2 and Non-T2 Asthma Subtypes

Our next study aimed to compare the genetic underpinnings of T2 versus non-T2 asthma using the approach described above in the Methods section. We performed several studies comparing the T2 cohort to the non-T2 cohort, and either cohort to healthy controls.

First, we compared the lists of “critical” SNPs with the lowest p-values from two of the studies: T2 vs. controls, and non-T2 vs. controls. We expected this comparison to identify three sets of critical SNP genotypes: those that are significantly present in T2 asthma; those that are significantly present in non-T2 asthma; and those that are common to both subtypes. Figure 3 illustrates the numbers of critical SNP genotypes that are significant in each of these categories, indicating clear differences in SNPs between T2 and non-T2 cases.

The unique SNPs identified in the replication study (that is, those that show up in both cohorts as statistically significant minor alleles) follow a striking pattern. When prioritized by p-value, we see a large number of SNPs that relate to immune system disorders and asthma—a result that confirms our hypothesis that we are finding biologically relevant high-order combinations of genotypic features.

We then mapped these SNPs into genes within +/-1 kilobase and plotted the corresponding genes in a network diagram to illustrate the genetic differentiation of the two subtypes of asthma at the level of genes (see Figure 4).

Figure 3 Critical SNP genotypes that are significantly represented in non-T2 asthma (2,064) vs. T2 asthma (1,456) vs. those that are common to both subtypes (332)

Figure 4 Cytoscape plot illustrating major genes involved in asthma disease architecture. Genes in red are prominently involved in T2 asthma, genes in green are prominently involved in non-T2 asthma, and genes in blue are involved in both subtypes.
Next, we mapped the prominent SNP genotypes in T2 and non-T2 asthma to gene sets and performed a pathway enrichment analysis (see Figure 5). As expected, the pathway enrichment analysis showed that T2 and non-T2 asthma are dramatically different diseases that share a common symptomatology, but little else.

While many of the most significant genes we identified in the T2 asthma population corresponded to classic T2-driven immune pathways, we identified a range of different non-immune pathways that were significant in the non-T2 cohort, including metabolic and neuronal mechanisms.

Several of the most significant genes in the non-T2 population encode enzymes that are involved in key stages of fatty acid synthesis and oxidation pathways. Although all of the genes we identified represent novel asthma targets, both of these overall pathways have been implicated in driving asthma pathogenesis.\(^8,9\) We also identified targets that are involved in the promotion of low-density lipoprotein (LDL) oxidation. Increased oxidized LDLs are hypothesized to increase bronchial inflammation through recruitment and degranulation of neutrophils,\(^10\) and inhibitors of this pathway are already of interest for several pharmaceuticals as potential new asthma therapies. Furthermore, we found a range of genes that modulate several different neuronal pathways, including regulation of GABAergic transmission, purinergic receptor activation, and glutamate signaling. This implies that non-allergic asthma is driven by a variety of different mechanisms that are not directly related to the immune system.

To summarize, we observed key differences in pathway enrichment between the two asthma subtypes. This clear segregation of pathways demonstrates that they are two distinct diseases, with different underlying mechanisms of action. We have identified over 20 novel genes that are significant only in the non-T2 population with strong, testable hypotheses for their mechanism of action. These represent promising opportunities for the development of personalized therapies for patients presenting with non-allergic asthma.

Figure 5 Pathway enrichment results for T2 vs. non-T2 asthma (calculated using the ClusterProfiler R package)
Conclusions

The study described above confirms the key genetic differences between T2 and non-T2 asthma, and further differentiates these two major asthma subtypes as radically different diseases based on genetic architecture. Our findings are well-aligned with the common understanding that cytokine regulation (especially IL-5 and IL-13) plays a key role in T2 asthma, and provide some promising novel indications of the mechanisms underpinning the pathogenesis of non-T2 asthma. While there are a number of effective biologic drugs targeting the T2 subset, long-term treatment of non-T2 asthma is still a clinical challenge. We are hopeful that the identification of the several key pathways that are enriched in the non-T2 subset will contribute to the development of novel treatments for patients that make up the vast majority of asthma cases worldwide.

Our future studies will involve further exploration of the genetic architecture of asthma in order to map genetic clusters to clinical clusters that may be easily differentiated using readily available biomarkers.

Notes and References


