precisionlife

CASE STUDY:

Patient Stratification of T2 vs Non-T2 Asthma

Identification of Novel Disease Mechanisms in Non-T2 Populations

CONTACT
E: steve@rowanalytics.com
W: precisionlife.com

precisionlife® Case Study

Genetic Underpinnings of T2 (eosinophilic) / non-T2 Asthma

EXECUTIVE SUMMARY

precisionlife **uses unique AI methods and data analytics to enable faster and deeper analysis of large multi-omic datasets to reveals new insights into the multi-factorial causes of complex diseases and predicts patients' different experience of disease and response to therapies.**

Asthma patients can broadly be categorised into two molecular phenotypes; those with high T-helper cell type 2 (T2/eosinophil) expression and those without (non-T2). Asthma patients with a T2 phenotype currently have a range of targeted biologic treatment options available to them, however non-T2 patients lack such targeted drugs and often have to rely on conventional symptomatic control therapies (such as bronchodilators and inhaled corticosteroids), which do little to combat the underlying disease pathology.

We performed a comparative study using a genotype dataset derived from UK Biobank to compare T2/non-T2 asthma patient populations. Using precisionlife **MARKERS with UK Biobank data, the study took less than 3 days from conception to assembly of data to analysis and interpretation.**

The study identified clear differences in disease associated genetic pathways between the T2 and non-T2 asthma cohorts. These differences, which cannot be detected using existing GWAS methods, offer significant potential for better patient stratification, diagnosis and new treatment options.

Whilst most of the significant 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 which affects 1 in 13 people. 5.4 million people are currently receiving asthma treatment in the UK[1]. A subset of these patients suffer 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 even older adults and pediatric patients. T2 asthma equally affects males and females[2].

The etiology of T2 asthma involves activation of the Type 2 T-Helper (Th2) cells, which result 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 1/3 of asthma patients do not have high blood eosinophil counts and present as nonallergic cases[3]. While there have been many GWAS studies on asthma to date, prior studies have not focused on or found the genetic differences between T2 and non-T2 forms of asthma[4]. Understanding of the genetics of T2 asthma has been largely based on studies of the Th2-cytokine pathways. One earlier GWAS study focused on genetic defects affecting eosinophil counts, but this study did not draw strong conclusions about the genetic architecture of T2 asthma[5].

Therefore, we analyzed UK Biobank data from asthma cases and healthy controls with an aim to identify significant genotypic differences between T2 and non-T2 asthma. Identifying the differences between combinatorial patterns of SNPs associated with T2/non-T2 patients (disease signatures) enables downstream interpretation of the genes, pathways and mechanisms involved in the different disease presentations.

METHODS

precisionlife **MARKERS** is a massively scalable multi-omics association platform that enables the detection of high order epistatic interactions at genome-wide study scale. It can find and statistically validate combinations of five or more SNP genotypes that are found in many cases and few controls for a range of phenotypes (including in this case eosinophil levels).

In this study, we identified asthma cases in UK Biobank using a slightly modified version of the case selection criteria presented by Ferreira et al[6]. Using these criteria, we identified a total of 42,205 total asthma cases. We randomly selected 90,034 age- and gender-matched subjects from the same database to serve as controls.

We then generated 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 good indicator of eosinophilia in the airways[7]. 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 less 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 more 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 data set (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. Using **precision**life **MARKERS**, we analyzed the full asthma data set 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

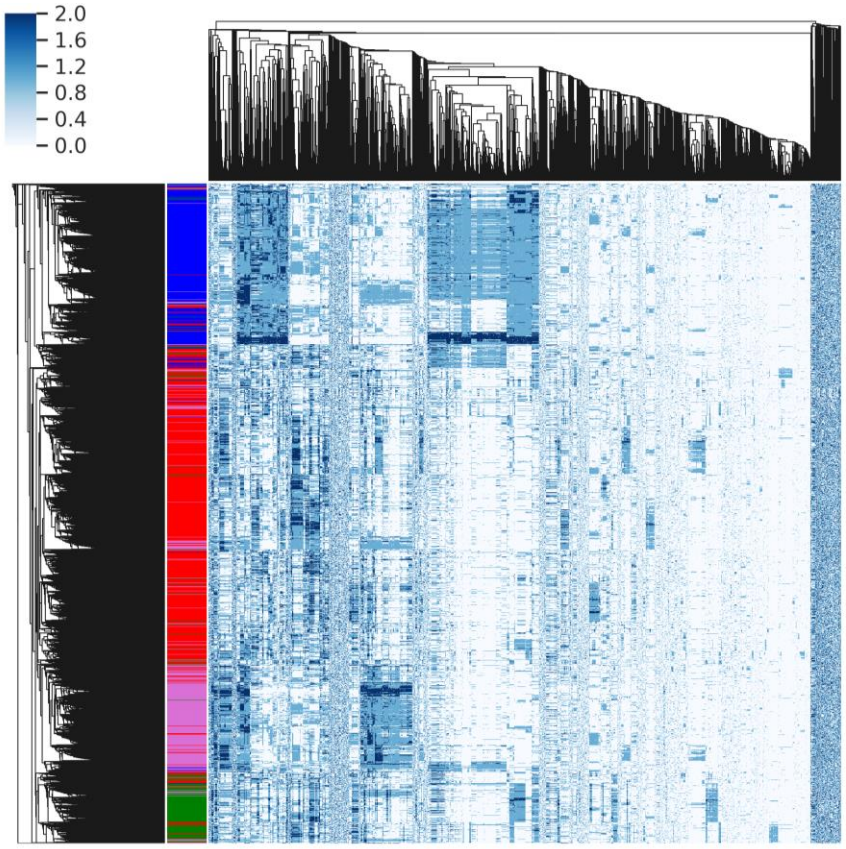

Figure 1. *Hierarchical k-means clustering of 42,205 asthma cases based on critical SNP genotypes. Critical SNP genotypes are in the columns and individual cases are in the rows. A small block of SNP genotypes is found in virtually all cases, shown as the narrow darker stripe along the right edge of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

based on its prominence among all networks. We then used a hierarchical clustering method on critical SNP genotypes versus case samples to reveal the genetic organization of asthma (Figure 1).

Using 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 (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.

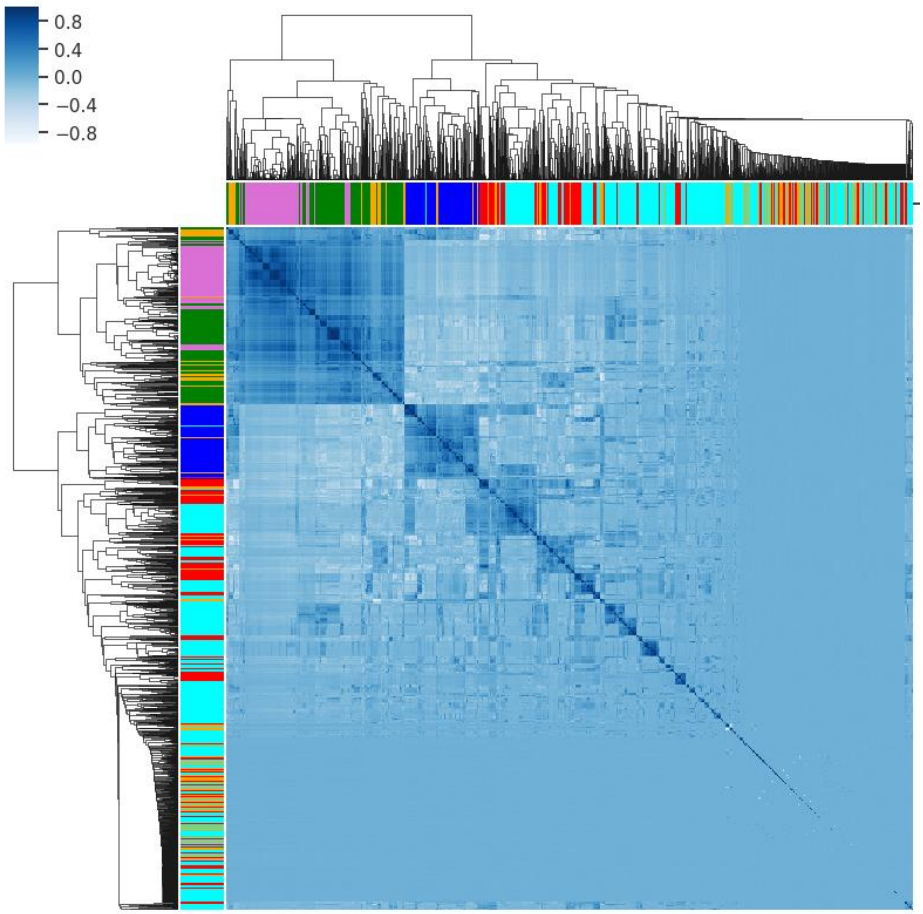

Figure 2. *Hierarchical k-means clustering of 42,205 asthma cases into 6 clusters. Both columns and rows indicate individual SNP genotypes. Six different clusters are visible (indicated with pink (cluster 1), green (cluster 2), blue (cluster 3), orange (cluster 4), red (cluster 5), and cyan (cluster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

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

EXPLORING THE GENETIC DIFFERENCES BETWEEN T2 AND NON-T2 ASTHMA SUBTYPES

The main purpose of our study was to compare the genetic underpinnings of T2 versus non-T2 asthma. Using **precision**life **MARKERS**, 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: 1) those that are significantly present in T2 asthma; 2) those that are significantly present in non-T2 asthma; 3) those that are common to both subtypes. Figure 3 illustrates the numbers of these critical SNP genotypes, indicating clear differences in SNPs between T2 and non-T2 cases.

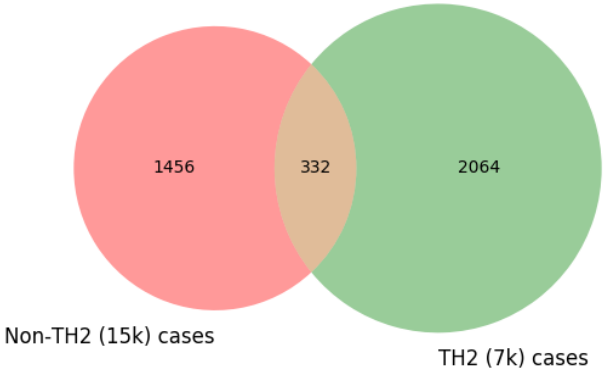

Figure 3. *Critical SNP genotypes that are significantly represented in non-T2 asthma (2,064) versus T2 asthma (1,456) versus those that are common to both subtypes (332).*

The unique SNPs identified in the replication study (that is, those that show up in both cohorts as statically 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 finding that confirms our hypothesis that we're finding biologically relevant high-order combinations of genotypic features.

We then mapped these SNPs into genes within +/- 1 KB 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 (Figure 4).

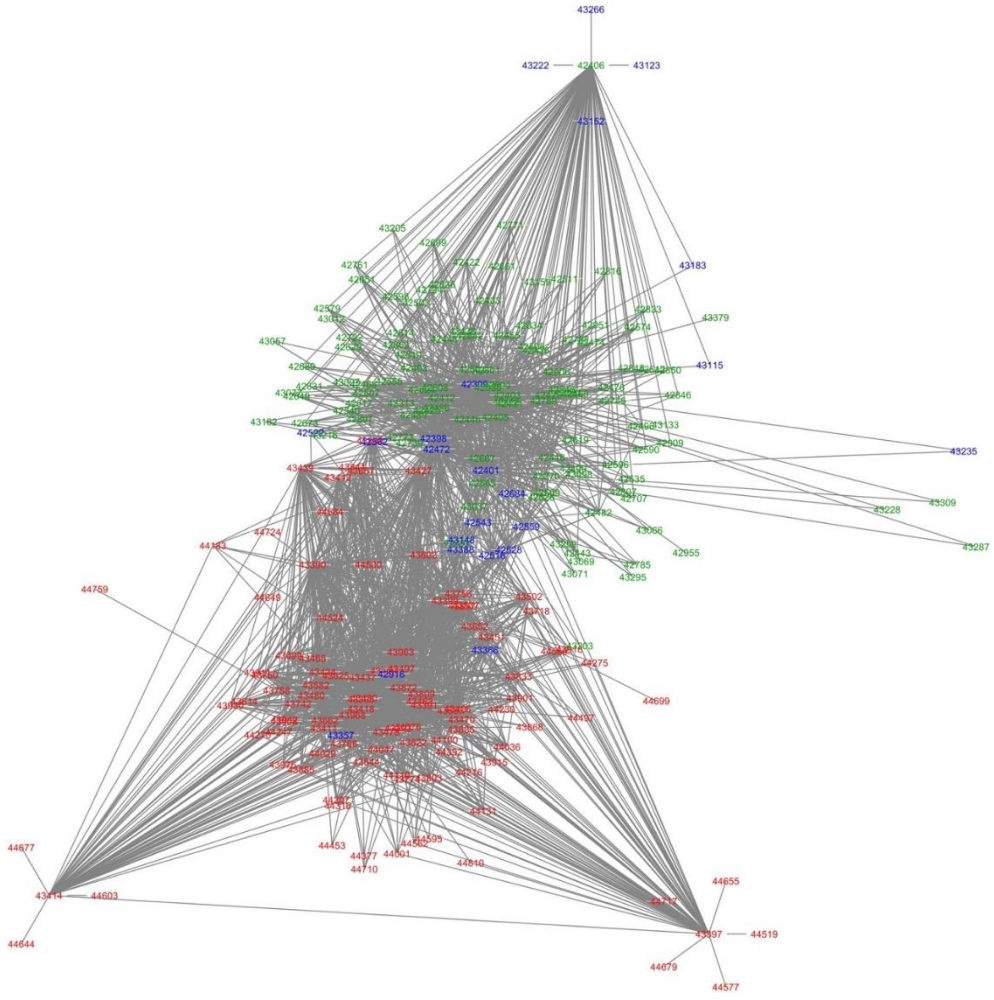

Figure 4. *Cytoscape plot illustrating major genes involved in asthma disease architecture. Genes in red are prominently involved in T2 asthma, genes in green are prominently involved in non-T2 asthma, and genes in blue are involved in both subtypes.*

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 shows that T2 and non-T2 asthma are dramatically different diseases that share a common symptomatology but little else. The eosinophilic subtype is, as expected, primarily enriched for:

- Type 2 immune response
- Positive regulation of IL-5 and IL-13 production
- Regulation of chemokine production
- Immunoglobulin production and regulation
- Lymphocyte differentiation

Common enrichment patterns (albeit at a slightly weaker association level) for both major subtypes of asthma include:

- Regulation of leukocyte-mediated immunity
- Regulation of leukocyte activation
- Secondary immune response based on somatic recombination of immune receptors

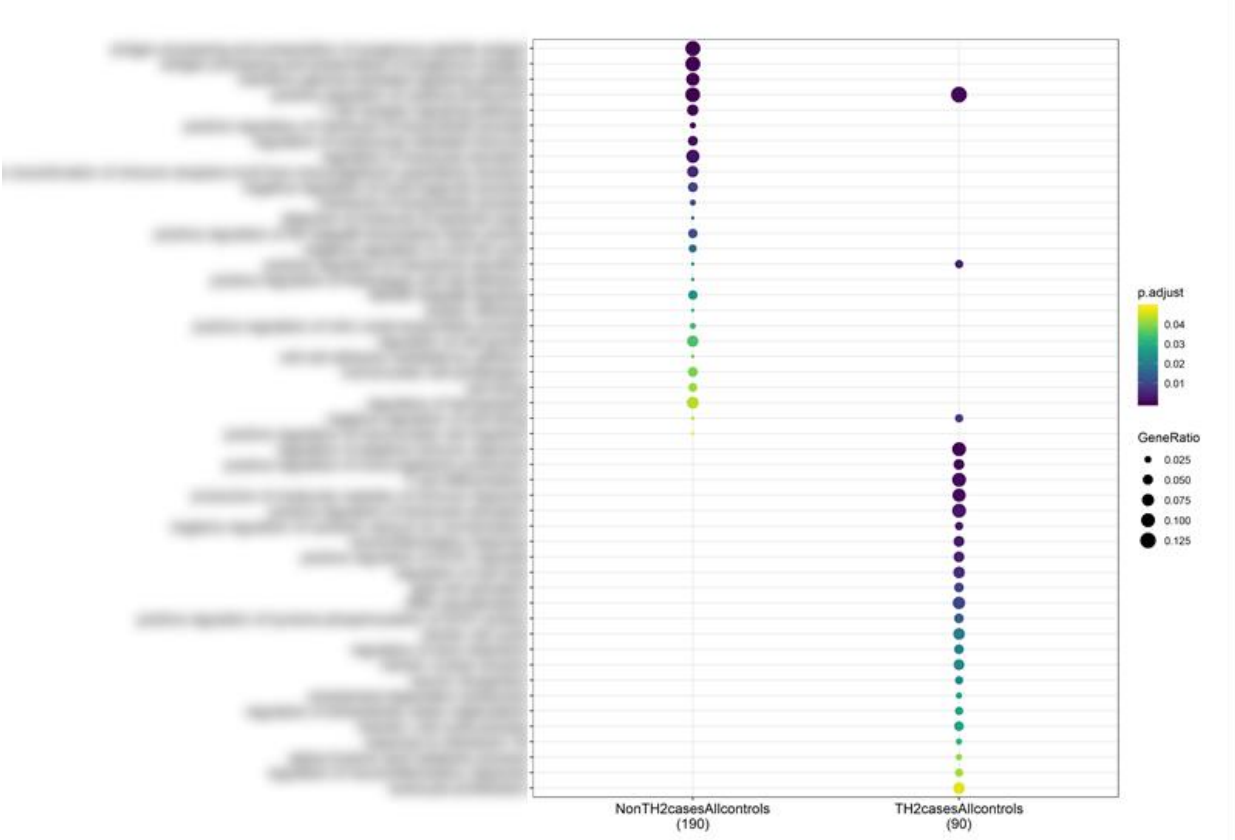

Figure 5. *Pathway enrichment results for T2 vs non-T2 asthma (calculated using the ClusterProfiler R package).*

Whilst 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 LDL oxidation. Increased oxdLDLs are hypothesized to increase bronchial inflammation through recruitment and degranulation of neutrophils[10], and inhibitors of this pathway are already of interest to several pharmaceutical companies 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 at least some forms of non-allergic asthma are driven by a variety of different mechanisms that are not directly related to the immune system.

To summarize, we observed key differences both in patient stratification biomarkers and in the pathway enrichment between the two asthma subtypes. This clear segregation of markers and pathways demonstrates that they are two distinct diseases, with different underlying mechanisms of action. We have identified (and validated with multiple KOLs) 12 novel, druggable genes that are significant only in the non-T2 population. Many of these genes are targets of approved on-market compounds, offering either off-the-shelf tool compounds for novel drug discovery or potential repurposing opportunities.

These targets pass the 5Rs heuristics for in-silico validation with strong genetic explanation, testable hypotheses for their mechanisms of action, differentiated tissue expression, safe and bioavailable starting compounds, clear patient stratification biomarkers and good commercial potential with unmet medical need and good commercial/IP landscape. These represent promising opportunities for the development of personalized therapies for patients presenting with nonallergic asthma.

CONCLUSIONS

The study described above is wholly dependent on being able to identify high-order combinations of SNPs that together are associated with different patient phenotypes. The significant majority of the SNPs making up the disease signatures would not pass the standard GWAS significance threshold when evaluated in isolation across the whole patient population. Using higher order combinations of SNPs, in this case 5-SNP combinations, enables a much more granular view of the disease and higher resolution patient stratification.

This study identified the key genetic differences between T2 and non-T2 asthma and further differentiated these two major asthma subtypes as clearly different diseases based on their genetic architectures. It provides a clear basis for the stratification of patients and selection of different therapeutic approaches for the different patient groups.

Our findings are well-aligned with the common understanding that cytokine regulation (especially IL-5 and IL-13) play a key role in T2 asthma[7] and provide some indication of the mechanisms underpinning the pathogenesis of non-T2 asthma. Whilst there are several promising biologic drugs targeting the T2 subset, long-term treatment of non-T2 asthma is still a clinical challenge[3]. 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.

REFERENCES

1. Asthma UK, https://www.asthma.org.uk/about/media/facts-and-statistics (accessed August 13 2019)
2. American Partnership for Eosinophilic Disorders, https://apfed.org/about-ead/egids/eoe/ (accessed October 8 2019)
3. Baos S., Calzada D., Cremades-Jimeno L.. et al. Nonallergic Asthma and Its Severity: Biomarkers for Its Discrimination in Peripheral Samples. Front Immunol. 2018;9:1416.
4. Kim K.W. and Ober C. Lessons Learned From GWAS of Asthma. Allergy Asthma Immunol Res. 2019 Mar;11(2):170-187.
5. Gudbjartsson DF, Bjornsdottir US, Halapi E, Helgadottir A, Sulem P, Jonsdottir GM, et al. Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. Nat Genet 2009;41:342-7.
6. Ferreira M.A.R., Mathur R., Vonk J.M. et al. Genetic Architectures of Childhood- and Adult-Onset Asthma Are Partly Distinct. American Journal of Human Genetics, 2019, Volume 104:4. 665-684.
7. Carr T.F., Zeki A.A., and Kraft M. Eosinophilic and Noneosinophilic Asthma. American Journal of Respiratory and Critical Care Medicine, 2018; 197(1):22-37.
8. Lee JY., Zhao L., Youn HS. et al. Saturated Fatty Acid Activates but Polyunsaturated Fatty Acid Inhibits Toll-like Receptor 2 Dimerized with Toll-like Receptor 6 or 1. J. Biol. Chem. 2004; 279:16971-16979.
9. Al-Khami AA., Ghonim MA, Del Valle L., et al. Fuelling the mechanisms of asthma: Increased fatty acid oxidation in inflammatory immune cells may represent a novel therapeutic target. Clin Exp Allergy. 2017;47(9):1170–1184.
10. Sedgwick JB., Hwang YS., Gerbyshak HA. et al. Oxidized Low-Density Lipoprotein Activates Migration and Degranulation of Human Granulocytes. Am J. Respir Cell Mol Biol. 2003;29(6):702-9.