Patient Stratification of UK Biobank Asthma Population Reveals Distinct Genetic & Mechanistic Differences between T2-High and T2-Low Asthma

K. Taylor, S. Das, A. Malinowski, C. Stubberfield
PrecisionLife, Oxford, United Kingdom

Scan to find out more:

Introduction

PrecisionLife is a computational biology company focusing on precision medicine analytics in complex chronic diseases. Our mechanistic patient stratification identifies subgroups of patients who share causal drivers of disease and treatment response, generating biomarkers that inform and de-risk drug discovery and development.

Asthma can be split into two major endotypes (T2-high and T2-low) characterised by distinct molecular drivers. The genetic risk factors and underlying pathology of T2-low asthma are more poorly understood than the T2-driven endotype and there is no clear consensus on the clinical definition, although patients with low eosinophil counts often display a greater proportion of neutrophilic inflammation.

PrecisionLife aimed to identify genetic differences between the two cohorts and further characterise the biological drivers of T2-low asthma.

Methods

DATASET:
PrecisionLife analysed 2 genetic datasets constructed from the UK Biobank using blood eosinophil data:

T2-LOW CASES (n = 14,996):
ICD-10 code, J45.x
Serum eosinophil counts of 0.15 (1,500 cells/mm3) or fewer[1]
T2-HIGH CASES (n = 7,072):
ICD-10 code, J45.x
Serum eosinophil counts of 0.35 (3,500 cells/mm3) or more[1]
HEALTHY CONTROLS (n = 21,566)
No evidence of chronic respiratory disease (inc. asthma, allergic rhinitis)

COMBINATORIAL ANALYSIS:
Each case-control dataset was analysed in the PrecisionLife platform to identify combinations of SNP genotypes that when observed together in a patient are strongly associated with each endotype of asthma.

SNP combinations that have high odds ratios, low *p*-values and high prevalence in cases are prioritized. This process undergoes 1,000 cycles of fully randomized permutations and combinations must meet a specified FDR threshold.

SNPs are scored using a Random Forest algorithm in a 5-fold cross-validation framework and prioritized based on their ability to differentiate cases and controls.

The highest scoring SNPs are then mapped to genes and clustered by the patients they co-occur in to generate a disease architecture.

The results from each dataset (T2-high and T2-low asthma) were compared to identify significant genetic differences between the two.

Results

Figure 1. Combinatorial analysis of genomic data vs GWAS

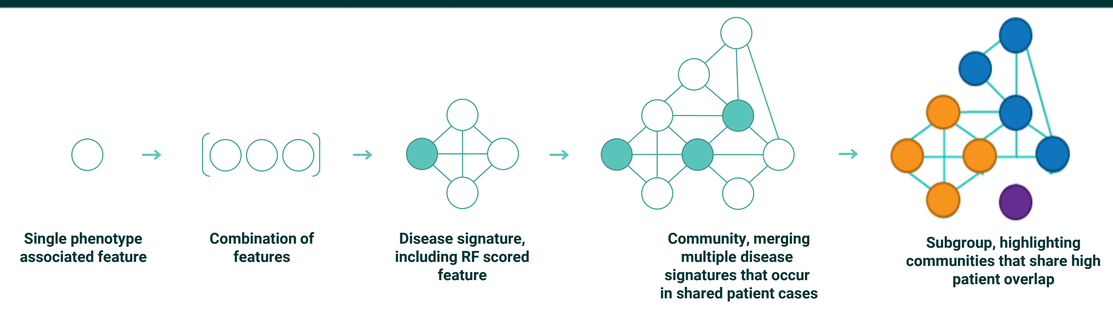

Figure 1. Conceptual representation of features, combinations, disease signatures and communities used to build up the disease architecture in the PrecisionLife combinatorial methodology.

GWAS	Combinatorial Analysis
Single SNP associations must be significant across large groups of patients	Specific combinations of variants associated with each patient subgroup serve as a genetic stratification biomarker
Limited insights unless disease is likely to be caused by a small number of rare variants with large effect sizes (often in gene coding regions affecting protein 3D structure)	Patient subgroups with different causes of disease or even incorrect diagnoses can be distinguished (stratified) by different mechanistic aetiology
Does not account for the effects of interactions between SNPs, genes and metabolic networks	Captures epistatic and non-linear additive effects of all interactions between SNPs, genes, environmental factors and metabolic networks

Table 1. Key results from PL analysis of T2-high and T2-low datasets

	T2-Low Asthma	T2-High Asthma
Validated disease signatures (SNP combinations)	2,380	4,449
SNPs associated with endotype	1,182	2,483
Significant (RF-scored) genes associated with endotype	200	200
Penetrance* (%)	100	100

* Proportion of cases represented by at least one validated disease signature

Table 1. Summary of the results of PrecisionLife's combinatorial analysis of the two asthma endotypes from UK Biobank genetic data.

Figure 2. Overlap analysis of SNPs identified in T2-high and T2-low asthma

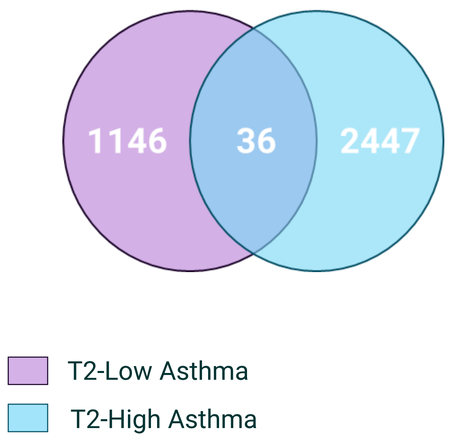

Figure 2. Venn diagram demonstrating the overlap in the RF-scored SNPs identified in the T2-low and T2-high asthma studies.

Figure 3. Comparative pathway enrichment analysis demonstrates biological differences between the genes identified in T2-high and T2-low asthma

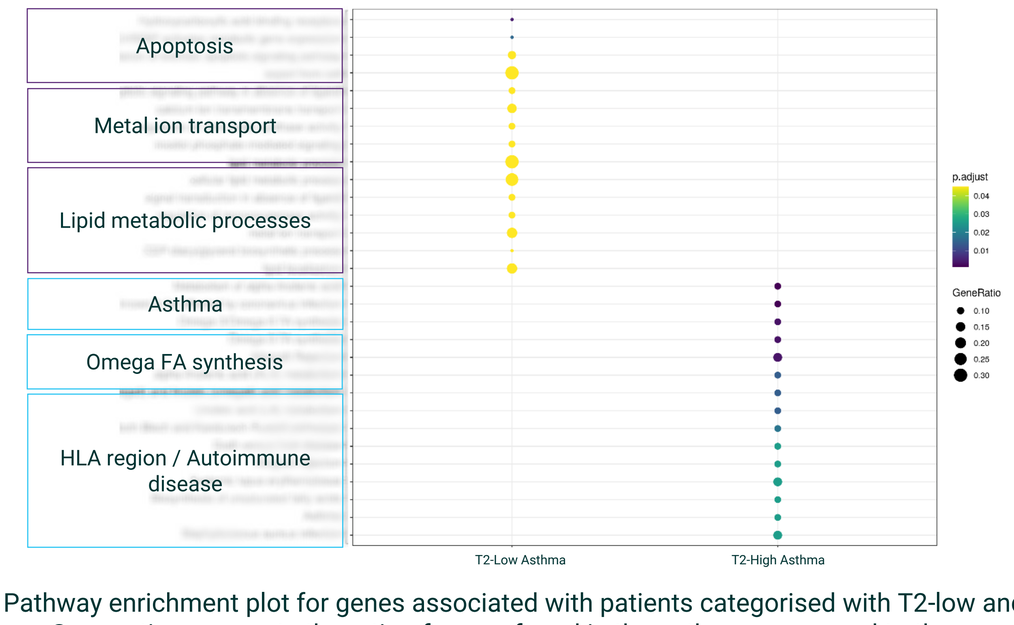

Figure 3. Pathway enrichment plot for genes associated with patients categorised with T2-low and T2-high asthma. Gene ratio represents the ratio of genes found in the pathway compared to the genes associated with a community and *p.adjust* represents the p-value adjusted for multiple testing. The dots in the plot are color-coded based on their corresponding *p.adjust* values.

Table 2. Target assessment indicates 35+ novel drug targets within T2-low asthma

TARGET	PSBs	MECHANISM OF ACTION	PREVALENCE IN ASTHMA	EXPRESSED IN LUNG	TARGET TRACTABILITY	CONFIRMED IN GASP STUDY
GENE 1	T2-Low Asthma	Lipid Metabolism Th17 inflammation			Tool Compounds	
GENE 2	T2-Low Asthma	Lipid Metabolism				✓
GENE 3	T2-Low Asthma	Metal Ion Homeostasis			Tool Compounds	✓
GENE 4	T2-Low Asthma	Lipid Metabolism Neutrophilic Inflammation			Drugs in Phase II	✓
....						
GENE 35	T2-Low Asthma	Apoptosis				✓

Table 2. Identified asthma targets are evaluated against druggability criteria to facilitate target prioritisation. Unique insights from the combinatorial analysis include estimation of case penetrance and identification of patient stratification biomarkers (PSBs) for each target.

Discussion

PATIENT STRATIFICATION
We demonstrate distinct differences in the biological pathways associated with the genes identified in each asthma endotype when analysed using the PrecisionLife platform.

T2-high:
- T2 cytokine pathways
- HLA region genes

T2-low:
- Neutrophilic inflammation
- Lipid metabolic processes
- Apoptosis
- Metal ion binding

NOVEL DRUG TARGETS
PrecisionLife have prioritised 35 of the most tractable gene targets that have strong mechanistic links to the processes identified in the pathway enrichment analysis. These can be used for the development of more targeted and efficient treatment strategies to improve asthma control in each endotype.

ANALYSIS OF GASP DATASET
PrecisionLife have confirmed 8 of the prioritised targets in an independent asthma population using genetic data derived of cases from the GASP [the Genetics of Asthma Severity and Phenotypes][2] study and UK Biobank controls. Further analysis of this dataset using the PrecisionLife platform is ongoing.

Conclusion

The results demonstrate that the PrecisionLife combinatorial analysis platform is uniquely able to stratify heterogenous patient populations with complex disease pathologies. We can use these insights to identify more effective therapeutic strategies and accompanying biomarker sets to match them to the patient subgroups that are most likely to demonstrate benefit in downstream clinical trials.

References

1. Carr TF, Zeki AA, Kraft M. Eosinophilic and Noneosinophilic Asthma. *Am J Respir Crit Care Med*. 2018;197(1):22-37. doi:10.1164/rccm.201611-2232PP

2. Shrine N, Portelli MA, John C, et al. Moderate-to-severe asthma in individuals of European ancestry: a genome-wide association study. Lancet Respir Med. 2019;7(1):20-34. doi:10.1016/S2213-2600(18)30389-8

Acknowledgements

Research described in this study has been conducted using data from the UK Biobank Resource (application number 44288).

We would like to thank the teams at SAIL, HDR UK BREATHE and Nottingham University for their provision and analysis of the GASP data.

precisionlife
improving health | **for everyone**

For more information, visit:
www.precisionlife.com