



Frequently Asked Questions:

1. What is AllergyGenDB?

AllergyGenDB is a comprehensive knowledge based allergic diseases genome database. It is an online tool that effectively integrates, and visualizes allergic diseases and associated genetic, pathway and network information from research articles and curated databases.

2. How do I interpret the calculated LPF values in literature mining?

Log of the Product of Frequency (LPF) is a way to measure how relevant a gene or variant is to the disease. It is calculated as follows:

$$\text{LPF (gene/SNP, disease)} = \log_{10}\left(\frac{X}{G} * \frac{X}{T}\right)$$

Where X is the number of abstracts containing both gene and disease term, G is the number of abstracts containing the gene, and T is the number of abstracts containing the disease term.

3. Which database is used for literature mining?

Diseases-associated genes, SNPs and pathways are collected from PubMed by using the pre-annotated file collection from [PubTator](#), which is text mining tool for annotating the entire PubMed articles with key biological entities (e.g. genes and diseases).

4. What databases are available in the curated databases?

AllergyGenDB supports curated association retrieval from three public data sources: GWAS Catalog, dbGaP and Monarch.

5. Why I don't see the content of the Attributes section?

Attribute selection is available when choosing only one single disease and querying genes or SNPs.

6. Why the Visualization section cannot be selected?

Cloud Tag Graph is available when selecting single disease; while Venn Diagram and Jaccard Index Matrix are available when more than one diseases are selected.

7. Can I download the result table?

Yes, the output table can be downloaded to Excel, CSV or PDF files, copied to the clipboard for further use or printed directly by clicking on the link the buttons above the table.

8. Where should I contact if I have any other question?

Please contact Dr. Tesfaye Mersha at Tesfaye.Mersha@cchmc.org.