# The AsthmaMap: advanced machine- and human-readable representations of mechanisms underlying asthma using domain expert knowledge

Alexander Mazein, PhD1,2,3, Olga Ivanova, MSc1,4,5, Irina Balaur, PhD1, Valeriya Berzhitskaya, PhD6, Tatiana Serebriyskaya, PhD6, Thomas Ligon, PhD7, Jan Hasenauer, PhD8, Bertrand De Meulder, PhD1, Richard G. Knowles, PhD9, Craig E. Wheelock, PhD10, Sven-Erik Dahlen, MD, PhD11, Kian Fan Chung, MD12, Ian M. Adcock, PhD12, Graham Roberts, DM, FRCPCH13,14, Ratko Djukanovic, MD, FRCP, FERS15,16, Johann Pellet, MSc1, Piotr Gawron, PhD3, Marek Ostaszewski, PhD3, Reinhard Schneider, PhD3, Rudi Balling, PhD3, Anke H. Maitland-van der Zee, PharmD, PhD5, Peter J. Sterk, MD5, Charles Auffray, PhD1

1 European Institute for Systems Biology and Medicine, CIRI UMR5308, CNRS-ENS-UCBL- INSERM, Université de Lyon, 50 Avenue Tony Garnier, 69007 Lyon, France

2 Institute of Cell Biophysics, Russian Academy of Sciences, 3 Institutskaya Street, Pushchino, Moscow Region, 142290, Russia

3 Luxembourg Centre for Systems Biomedicine, University of Luxembourg, 6 Avenue du Swing, L-4367 Belvaux, Luxembourg

4 Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands 5 Respiratory Medicine, Amsterdam University Medical Centers, University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands

6 Moscow Institute of Physics and Technology, 9 Institutskiy Pereulok, Dolgoprudny, Moscow Region, 141700, Russia

7 Faculty of Physics and Center for NanoScience (CeNS), Ludwig-Maximilians-Universität, 80539 München, Germany

8 Institute of Computational Biology, Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764, Neuherberg, Germany

9 Knowles Consulting, Stevenage Bioscience Catalyst, Gunnels Wood Road, Stevenage SG1 2FX, UK

10 Division of Physiological Chemistry 2, Department of Medical Biochemistry and Biophysics, Karolinska Institute, Stockholm, 17177, Sweden

11 Institute of Environmental Medicine (IMM), C6, Experimental asthma and allergy research, Imm/Eaaf, Karolinska Institutet, Solnavägen 9, Kvarter 5b 171 65 Solna, Sweden

12 National Heart & Lung Institute, Imperial College London, London SW3 6LY, UK

13 University of Southampton, Southampton, UK

14 NIHR Southampton Biomedical Research Centre, Southampton, UK

15 Clinical and Experimental Sciences, University of Southampton Faculty of Medicine, Southampton University Hospital, Southampton, UK

16 NIHR Southampton Respiratory Biomedical Research Unit, Southampton University Hospital, Southampton, UK

# Corresponding author

Alexander Mazein, PhD

Luxembourg Centre for Systems Biomedicine, University of Luxembourg 6 Avenue du Swing, L-4367 Belvaux, Luxembourg

+33 612900142, [a.mazein@gmail.com](mailto:a.mazein@gmail.com)

# Word count

Total: 5297 words, not including the abstract, figure legends and references Abstract: 250/250 words

Capsule summary: 30/30 words

# Competing interests

The authors declare no conflict of interest.

# Funding

Funded in part by the Innovative Medicines Initiative funded by the European Union and the European Federation of Pharmaceutical Industry Associations through U-BIOPRED (Unbiased Biomarkers for the Prediction of Respiratory Disease Outcomes, n°115010, https://[www.europeanlung.org/en/projects-and-research/projects/u-biopred)](http://www.europeanlung.org/en/projects-and-research/projects/u-biopred)) and eTRIKS (European Translational Information and Knowledge Management Services, IMI n°115446, https://www.etriks.org), the French National Center for Scientific Research (CNRS), the European Institute for Systems Biology and Medicine and the Luxembourg Center for Systems Biomedicine.

# Abstract

**Background:** A dedicated resource is needed for a systematic overview of existing knowledge about asthma mechanisms and for making sense of newly-generated data.

**Objective:** The goal of this work is the integration of prior knowledge on asthma into a single conceptual model to make possible advanced data interpretation, hypothesis generation and validation for many related projects.

**Methods:** The resource was developed following and further improving the best practices and guidelines of the Disease Maps Community (https://disease-maps.org/guidelines).

**Results:** The AsthmaMap (http://asthma-map.org) is a pathway-based representation of asthma mechanisms. It is designed as a living and evolving resource for research in the field of asthma, lung and allergic diseases. The AsthmaMap project is an open community effort driven by an international team of respiratory clinicians and scientists, systems biologists, experimental biologists, modellers and software developers. The information on disease mechanisms is offered on three interconnected layers of granularity: overview-type cellular interactions (overview), activity flows (compact), and process descriptions (most detailed). We discuss the semi-automatic assembly of disease maps using pathway modules, which renders the resource highly flexible, less dependent on manual work, and easily updated when new information on disease mechanisms becomes available.

**Conclusion:** We envision that the AsthmaMap will catalyse personalised clinical and primary care applications. Combined with methods from bioinformatics and systems biomedicine, the AsthmaMap will augment our ability to explain molecular mechanisms of asthma, and thus improve diagnostics, phenotyping and treatment strategies. This may further contribute to the development of clinical decision support systems toward personalised medicine of asthma.

# Key messages

* We present the AsthmaMap, a literature-based evolving conceptual model that depicts asthma mechanism on several layers of granularity.
* The representation of disease mechanisms is built by a community of domain experts in a systematic and reproducible way.
* The AsthmaMap is an open resource to be used for advanced data interpretation in multiple translational medicine projects.

# Capsule summary

We present the AsthmaMap, an expert-driven online, open, reusable and updatable resource that systematically integrates knowledge on asthma on the level of pathway maps united into a single virtual network.

# Keywords

asthma, allergic diseases, disease mechanisms, disease model, disease maps, knowledge representation, knowledge base, data interpretation, systems biology, translational medicine

# Abbreviations

AF, Activity Flow; BioPAX, Biological Pathway Exchange format; eTRIKS, European Translational Information and Knowledge Management Services; ILC, innate lymphoid cells; ILC2, type 2 innate lymphoid cells; ILC3, type 3 innate lymphoid cells; PD, Process Description; SBGN, Systems Biology Graphical Notation; SBGN-ML, Systems Biology Graphical Notation Markup Language; SBML, Systems Biology Markup Language; Th0, type 0 T helper cells; Th1, type 1 T helper cells; Th2, type 2 T helper cells; Th17, type 17 T helper cells; Treg, regulatory T cells; U-BIOPRED, Unbiased Biomarkers for the Prediction of Respiratory Disease Outcomes

# Introduction

Asthma is a chronic disease that affects all ages - 300 million people worldwide. The persistent symptoms with acute exacerbations are shortness of breath, wheezing and cough that vary in their occurrence, frequency and intensity. The disease can be managed to help a person to lead an active life, often with the use of inhaled corticosteroids. No cure is currently available. Hypotheses on asthma aetiology include reduced microbiome exposure, childhood infections and growing air pollution1.

To design prevention and treatment strategies for complex diseases such as asthma, it is necessary to understand its mechanisms and causes. Results from in-depth studies of asthma- related research are scattered across thousands of publications and new findings are constantly reported. In order to empower this research, we require a dedicated resource where asthma mechanisms are assembled into a single repository, represented in both human- and machine-readable formats, verified by domain experts and used for advanced data interpretation and hypothesis generation beyond the current functional analysis approaches.

Recent advances in systems biology enabled the creation of representations of mechanisms for multiple diseases including cancer, rheumatoid arthritis, type A influenza, Parkinson’s and Alzheimer’s diseases2–6. These maps are used for modelling, hypothesis generation and validation in the relevant clinical context6–10. Such resources are referred to as “disease maps” defined as conceptual models of the corresponding diseases, collections of interconnected signalling, metabolic and gene regulatory processes11. The disease maps are literature-based resources, with their backbone established over well-defined hallmarks of a given disease. Often the development of the maps is supported by adapting high-quality pathway information available in Reactome12, PANTHER13 and related databases. The information is stored in a format compatible with the Systems Biology Graphical Notation (SBGN)14; convertible to the Biological Pathway Exchange format (BioPAX)15 for sharing pathway information and the Systems Biology Markup Language (SBML)16 for modelling and simulation. It is important to highlight that disease maps are designed to allow continuous updates and improvements because the conceptual models and the clinical definition of a disease may evolve.

Despite a substantial body of literature on the topic, knowledge of asthma mechanisms is insufficient. On the one hand, we know a lot about molecules and processes involved and are able to capture asthma complexity and its multiple forms in individual patients. This includes mechanisms connected to genetics, allergen exposure and sensitisation, viruses, bacteria, pollution, diet, hormones, stress and psychological factors. On the other hand, we do not know enough to provide effective prevention strategies for asthma and to be able to address the unmet clinical needs17. Until recently, no dedicated resource has been available that embraces a systematic approach to describing asthma mechanisms.

A collective effort, initiated within the U-BIOPRED project and further developed within the eTRIKS project, led to shaping the AsthmaMap resource, a map for computational description of the mechanisms relevant to asthma development, progression and treatment18. The AsthmaMap project progresses as part of the large-scale Disease Maps collaborative community11,19. The development is supported by experts in the field of asthma from the Amsterdam University Medical Centers, University of Amsterdam, the National Heart & Lung Institute of Imperial College London, the Karolinska Institute and the University of Southampton [(http://a](http://asthma-map.org/team))s[thma-map.org/team).](http://asthma-map.org/team))

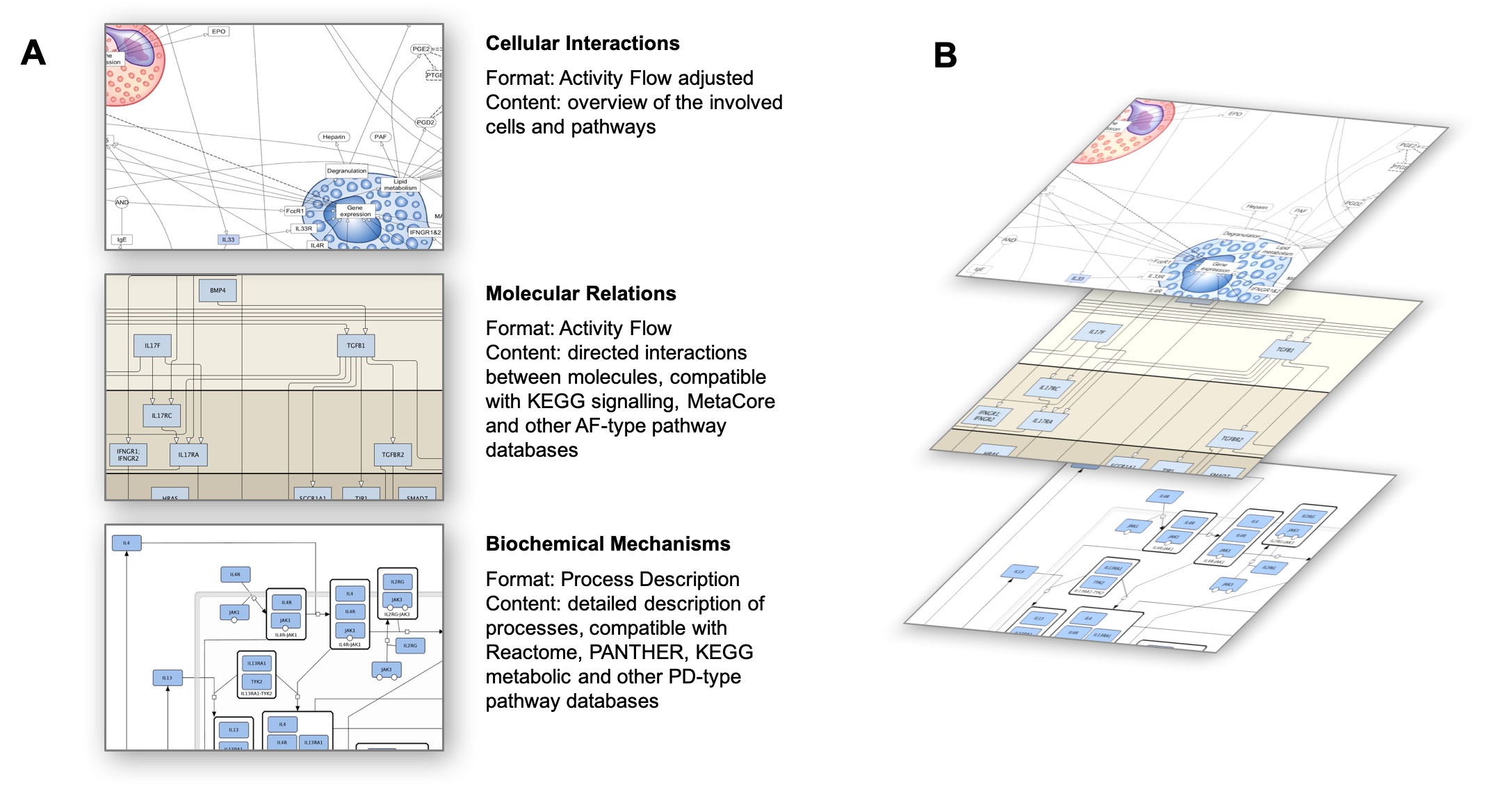
We present the AsthmaMap resource (http://asthma-map.org) and introduce strategies for semi-automatic management of its content. We describe challenges identified during the development of the AsthmaMap and the solutions implemented for some of them, outline the ongoing efforts and the expected future developments.

# Resource architecture

The ambitious aim of representing asthma mechanisms should be supported by efficient reuse of information from the existing databases and resources. However, those resources store information in different formats and levels of granularity. We thus decided to embrace separate levels of resolution and make it possible to integrate high-quality sources of network-based information in the AsthmaMap, while keeping its content compliant with the existing systems biology standards.

The AsthmaMap consists of three interconnected layers of granularity (Figure 1): **Cellular Interactions** - a machine- and human-readable overview diagram of the involved cell types and the corresponding pathways represented by cytokines and receptors; **Molecular**

**Relations** - a collection of diagrams split according to cell types that can be seen as a single virtual map with an intermediate level of details; **Biochemical Mechanisms** - the most detailed layer where information is shown on the level of molecular processes. Each higher-level representation is reflected at a lower level. This enables semantic zooming when the type of representation and the level of digitalisation changes with the zoom: the layers of the AsthmaMap represent the same systems, but with more or fewer details shown, depending on the zoom level.



**Figure 1.** The interconnected layers of granularity in the AsthmaMap. **A.** Fragments of maps of the Cellular Interactions, Molecular Relations and Biochemical Mechanisms representations. **B.** The three representations visualised as layers. The three representations reflect the same systems but use different scales, levels of details and conceptual models.

*AsthmaMap Cellular Interactions*

The overview diagram (Figure 1) is a comprehensive summary of the complexity of underlying mechanisms in connection to the cell types and tissues involved. It facilitates the discussions with domain experts, and the creation and curation of the AsthmaMap conceptual shape: modules, cell types, receptors, cytokines and other mediators, connected causes and effects.

The Cellular Interaction layer can be seen as a rule-based graphical review of mechanisms underlying asthma. Information is stored in a format compatible with the SBGN Activity Flow (AF) language with additional background images for better readability.

*AsthmaMap Molecular Relations*

SBGN AF diagrams encode simple logical relations, e.g., biochemical species A activates or inhibits biochemical species B. This is close to the traditional representation of signalling pathways and most convenient for building logical models20.

The Molecular Relations layer of the AsthmaMap (<http://asthma-map.org/mr)>is represented by diagrams for 16 cell types, including the following: dendritic, airway epithelial, macrophage,

neutrophil, eosinophil, eosinophil precursor, mast, airway smooth muscle, fibroblast, B, Th0, Th2, ILC precursor, ILC2, goblet and Treg. The planned development includes adding Th1, Th17 and ILC3 cell types (discussed in section “Missing and unknown relationships”).

The diagrams represent a mix of signalling, regulatory and metabolic networks; they contain proteins, metabolites, regulatory elements and transcription factors, and submaps. The smallest diagram (ILC2) includes 35 nodes and 50 edges. The largest one (Th0) - 351 edges and 647 nodes. Altogether the Molecular Relations layer has 831 unique entities.

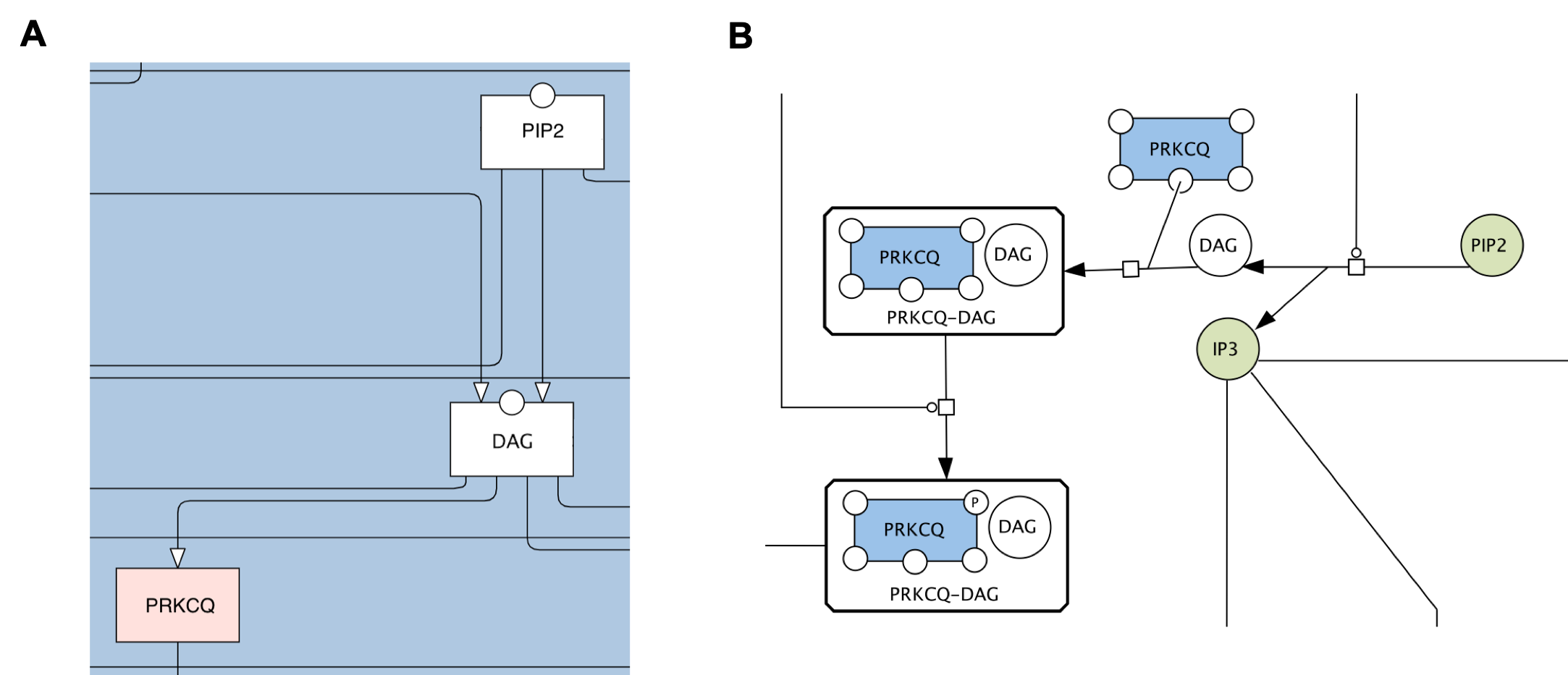
Each of the 16 diagrams was created by merging pathways in SBGN-AF-compatible format. The SBGN AF pathways were assembled using information from the AsthmaMap Cellular Interactions layer and the MetaCore pathway database (https://portal.genego.com) that we consider a high-quality resource for making network analysis disease-specific. Information on selected relevant interactions was then reduced and adjusted to the standard AF format, and the compatibility of the resulting AF modules was ensured for further processing and merging.

It is also possible to employ freely available resources for building Activity Flow type of representations. The corresponding AF-compatible databases and approaches are discussed in section “Challenges specific for the AsthmaMap Molecular Relations layer”.

As the SBGN AF representation is well-suited for developing logical models6,20, we anticipate that the AsthmaMap Molecular Relations layer will constitute a robust basis for the development of further mathematical modelling of asthma.

*AsthmaMap Biochemical Mechanisms*

The Biochemical Mechanisms layer provides detailed definitions for biochemical reactions, essentially encoding the stoichiometric matrix of the network. In contrast to the AF language, where the relations between entities are shown via direct links or via logical operators, in the Process Description (PD) language21, the attention is on the molecular processes that describe the involved reactants, products and modifiers, for example, enzymes or inhibitors (Figure 2).



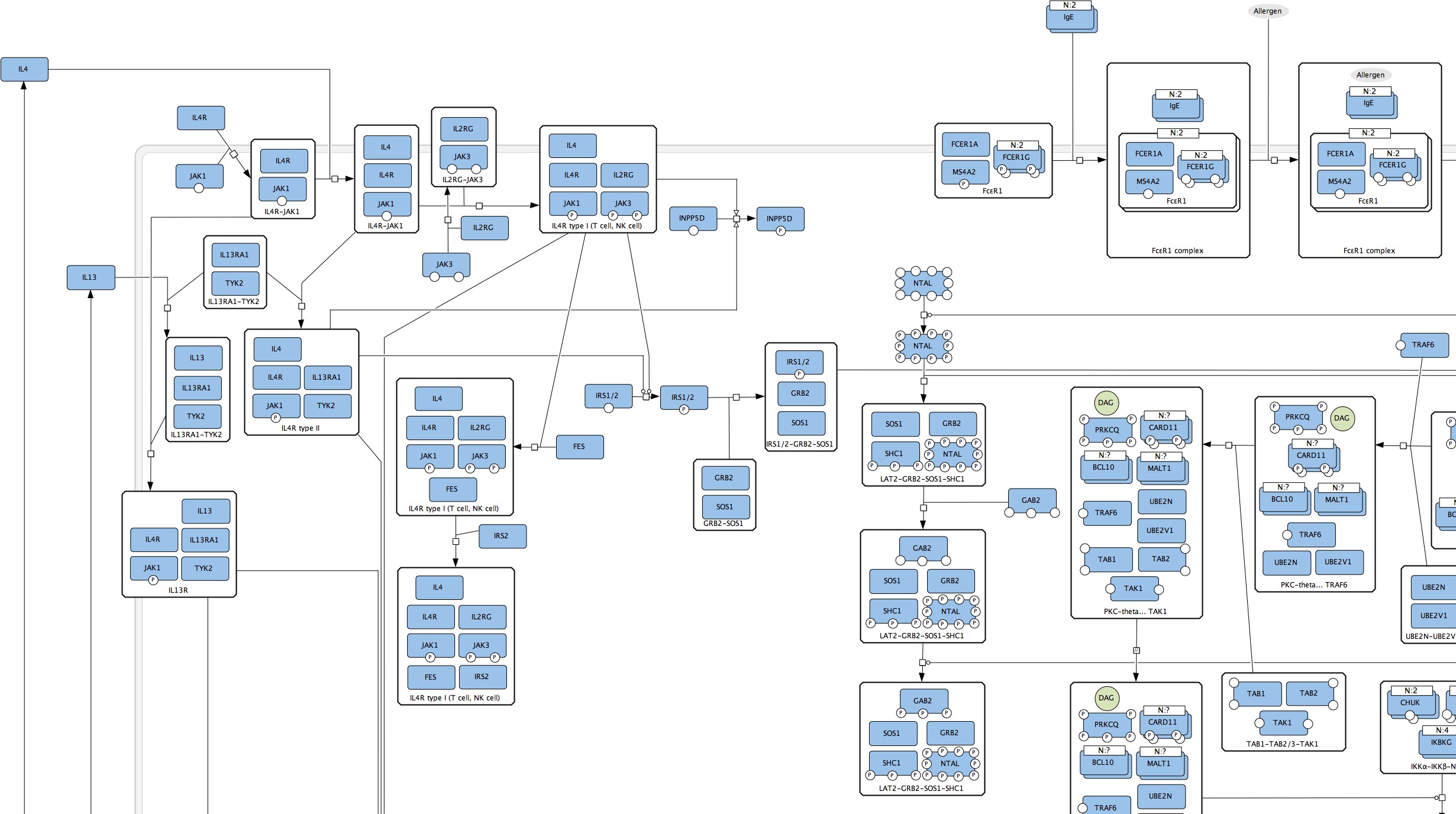
**Figure 2.** An example of representing the same biological event of PRKCQ activation in the SBGN Activity Flow language on the Molecular Relation layer and in the SBGN Process Description language on the Biochemical Mechanisms layer. **A.** The Activity Flow representation shows that DAG activates PRKCQ. **B.** The Process Description representation shows the process in detail: the formation of the DAG-PRKCQ with the following phosphorylation of PRKCQ.

This is the most detailed and most challenging layer of the AsthmaMap architecture [(http://a](http://asthma-map.org/bm))s[thma-map.org/bm).](http://asthma-map.org/bm)) The diagrams are manually designed as a result of an expert- driven curation of PubMed search of the relevant literature and are represented in the unambiguous PD language of the SBGN standard. The AsthmaMap Biochemical Mechanisms layer is currently represented by three cell-specific diagrams: mast cell, eosinophil cell and eicosanoid modules.

An example of such maps of the Biochemical Mechanisms layer is the mast cell module (a fragment of which is available in Figure 3) which consists of 342 species and 414 reactions: 79 unique metabolites, 122 unique proteins (166 if counted in different compartments), 10

genes, 10 RNAs, and 64 complexes.

The PD format is supported by the MINERVA platform22; the AsthmaMap Biochemical Mechanisms diagrams are available for browsing, search and data visualisation at https://asthma.uni.lu.



**Figure 3.** The mast cell module of the AsthmaMap Biochemical Mechanisms layer: a fragment of the map with the IL4/IL13 receptor system signalling. The map is available for browsing via the MINERVA platform at https://asthma.uni.lu.

The AsthmaMap Biochemical Mechanisms layer is suited for mechanistic modelling as it is developed using the SBGN PD language20.

*Communication between layers*

Each layer is reflected in the other layers. For example, on the Cellular Interactions layer, a pathway can be represented by a single entity, e.g. a receptor; then, on the Molecular Relations layer, more details are given and it is a pathway that starts from the receptor described. On the Biochemical Mechanisms layer, the same receptor can be shown multiple times, demonstrating its transition from one state to another.

# Map development

The curation was performed as described previously for the published comprehensive pathway maps23–25 and disease maps2–6 and according to the guidelines of the Disease Maps Community (https://disease-maps.org/guidelines). We also took into account the extensive experience of metabolic network curation described by Thiele and coauthors26.

# Semi-automatic assembly and updates

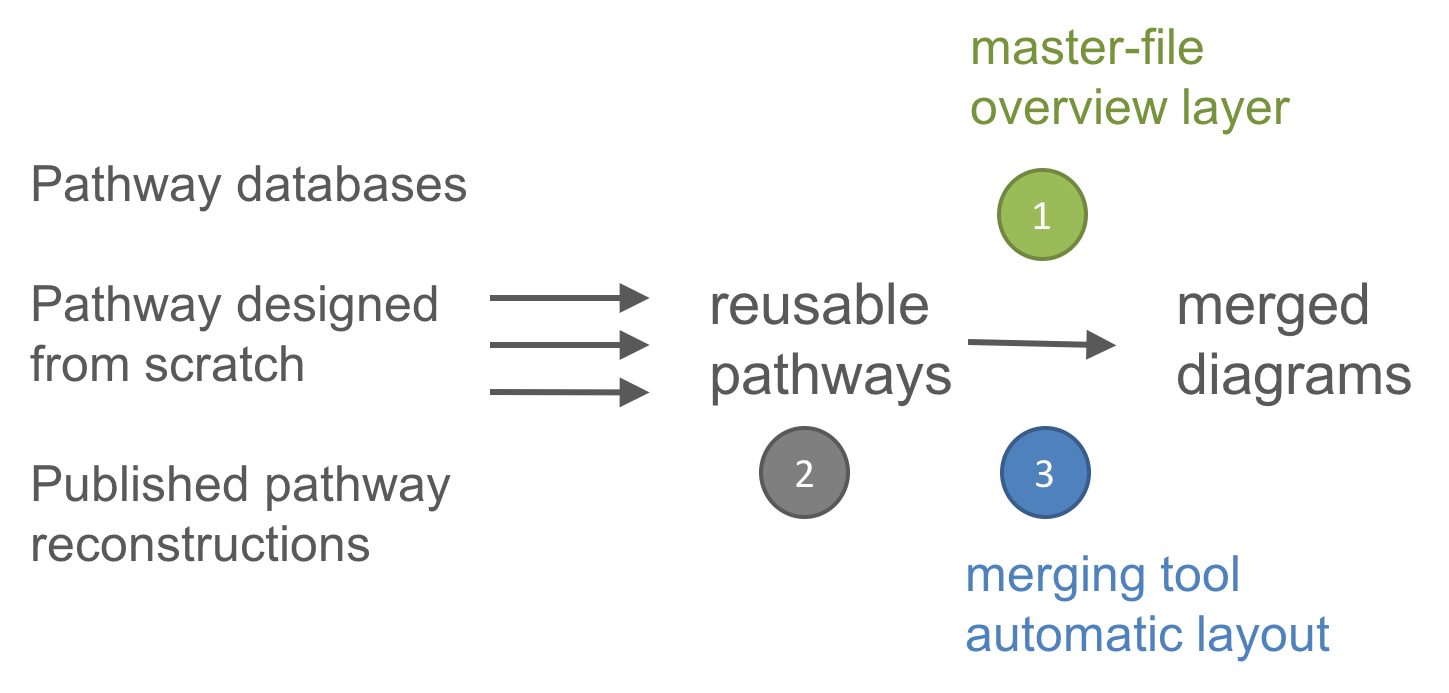
One of the powerful features of the AsthmaMap resource is the use of semi-automatic tools for creating and managing the content of the Molecular Relations layer. This approach addresses the shortcomings of manual curation of large-scale diagrams and allows avoiding redundancy and facilitating continuous updates.

Avoiding redundancy: because a large set of pathways is involved in signal processing in the various cell types, the same subnetworks would be repeated and would have to be manually redrawn many times.

Facilitating continuous updates: updating maps manually would be increasingly difficult for large and dense diagrams when new information needs to be introduced. Regular updates are important from a clinical/biological point of view so new discoveries can be quickly incorporated, keeping the AsthmaMap up-to-date.

Therefore, we had to consider the semi-automatic assembly of the maps based on the content of the Cellular Interactions layer as an overview of what is to be included in the Molecular Relations layer. Similarly, the Biochemical Mechanisms layer may be further built up from the Molecular Relations layer but this is a challenging task: the SBGN PD requires mechanistic details which are often unknown.

We have developed a customised pipeline for integrating information on asthma-relevant pathways into a larger network of the Molecular Relations layer. Specifically, the input for this pipeline was represented by i) a master-file containing a set of pathways or sub-pathway modules to be considered for a particular cell type, based on review papers and discussions with asthma experts, and ii) detailed information on biological entities involved in the selected pathways (Figure 4).



**Figure 4.** The semi-automatic assembly of the Molecular Relations layer. The three key components are 1) the master-file that reflects the overview layer; 2) reusable pathway modules selected for merging for a particular cell

type; 3) merging pipeline with automatic layout employed at the end of the process. The merging pipeline uses the information from the master-file in order to select pathways for merging.

After merging, the resulting diagrams were uploaded in the yEd Graph Editor for preparing a human-readable automatic layout. Finally, the content was converted to the standard SBGN- ML format using a dedicated AF extension of the ySBGN tool (https://github.com/sbgn/ySBGN).

The experience gained through the semi-automatic generation of the Molecular Relations layer maps pointed toward the question of the quality of the merging. The quality can be assessed by comparing the content of the source pathways and the resulting map. For example, each element of the source pathways and each of the relations between objects should be present in the final map.

All of the listed elements are important for making the resource manageable and updatable so that the pipeline could be reused to ensure the sustainable development of the AsthmaMap.

# Challenges in the development

A first challenge in generating the AsthmaMap is the definition of the gold standard: namely its clinical diagnosis. It can be envisaged that rather than the clinical features currently being used for defining asthma, biomarkers and biomarker networks will be increasingly used to define this disease or more preferably its phenotypes23. This emphasizes that the map will not have a discrete position and borders in relation to other medical conditions.

The further challenges of development and use of disease maps include: representation of the disease mechanisms with correct focus and on correct scale, engagement of research community to contribute and use the map, and exploration and curation of increasingly complex content11,19. The Disease Maps Community strives to address them as common issues for a number of projects. At the same time, because of its multi-layered structure, the AsthmaMap faces specific challenges related to representation formats.

*Challenges specific for the AsthmaMap Molecular Relations layer*

One of the challenges for the development of the AsthmaMap Molecular Relations layer is the size and legibility of the diagrams. This problem, even though indicated for other disease maps, became even more pressing in the AsthmaMap due to automated content generation. The introduction of the process of diagram merging allowed us to efficiently build a large repository of molecular relations, which in turn became challenging to compartmentalise and to lay out in an understandable way. Further growth of our resource will be greatly facilitated by scalable human-like layout algorithms24 or solutions for the data-driven splitting of molecular diagrams25. Such approaches could allow progress beyond the currently represented cell types and thus extending the number of cell-specific diagrams.

Another challenge is the source of the content for the Molecular Relations layer. First, to interconnect interesting modules of the network, cell-unspecific pathways were used as parts of cell-specific diagrams. Evaluation of these interactions, especially using well-annotated single cell data, will help to further refine the most relevant structure of the map. Moreover, using multiple, well-annotated knowledge sources can help to generate proper networks. Currently, the contents of the map were built with the support of MetaCore, a subscription-

based system. In the future, other sources, like OmniPath collection of pathways26 or PathwayCommons15,27 will support the curation and evolution of the AsthmaMap.

*Challenges specific for the AsthmaMap Biochemical Mechanisms layer*

Ensuring consistency and accuracy of the Biochemical Mechanisms layer is a particularly difficult topic. Describing separate pathways within one pathway database is possible, assuming that all the pathways form a consistent non-conflicting virtual map.

This means that, for example, the MAPK cascade described in one part of the Biochemical Mechanisms layer should be compatible and comparable with a similar module in another pathway.

The compatibility issue is an interesting and important one: within the same language events can be shown in less or more details, and both compatibility and comparison might be difficult to achieve if the representation is inconsistent or built using a different quality of evidence about physiological processes. This can happen when, for instance, one of the articles does not specify the phosphorylation residue or a cell type where the process was measured. In this case, one map may show specific phosphorylation sites for proteins and another map might omit this type of information. Both maps would be valid PD diagrams but then the first one would have more information, and a sophisticated comparison tool is needed to detect the differences and similarities. In the case of merging one version should be chosen, preferably the most detailed one.

On the AsthmaMap Biochemical Mechanisms layer, the compatibility is currently achieved by careful manual curation, taking into account the content of all pathways. Eventually, this will have to be replaced with an automatic system that checks the consistency and accuracy.

A particular issue is the compatibility of entities that are described in a system: a single map or a collection of maps. Before the pieces can be combined, it is important to ensure that they use the same approach to identify entities. This issue arises when an entity is a part of a pool of entities that are also presented as an object on the same map or a different map. This may happen, for example, when a member of a protein family is represented in one case as a specific molecule and in another case as a generic representative of a group. A possible solution would involve text mining and ontology mapping approaches28 combined with manual curation.

*Automatically aligning the content of the multiple layers*

The key to aligning the content of the layers is in the choice of entities included and their relationships. In order to address this issue, we rely on the notion of pathways or subpathways in order to group entities and provide a familiar (used in most pathway databases) navigation system. For example, the same pathway could be represented differently but the set of genes and proteins involved would be similar, with some exceptions when the Molecular Relations (or the Biochemical Mechanisms) layer evolves more intensively and then the other one would have to be updated correspondingly. This is different for the Cellular Interactions layer when normally one or two entities (for example a ligand and a receptor) would represent the whole pathway available on the more detailed Molecular Relations and Biochemical Mechanisms layers.

Manual or semi-automatic alignment allows us to complete the task but are inefficient and open to human error. There is a need for an automatic alignment process, which assumes producing the corresponding representation of biochemical events in a complementary layer and the use of automatic layout algorithms that would preserve the original layout as much as possible while making space for new entities and relationships24,29.

*Missing and unknown relationships*

The use of mechanistic models (compatible with PD) and logical models (compatible with AF) raises a generic issue that is very interesting and challenging but is rarely discussed: that of unknown and missing information. There might be relationships and processes we have not yet identified, and that could substantially change the behaviour of the models and reflect the underlying biology. Specifically, we could consider three categories:

1. known missing information that has not yet been included in the disease description for various reasons (incomplete coverage of the literature or planned updates);
2. identified gaps in our knowledge about a system (known unknowns): we know that there is a mechanism but do not know the details;
3. presumed missing relationships when we do not know if something is missing (unknown unknowns).

The last category in the AsthmaMap could include cell types that are not so far identified or incorporated (perhaps a regulatory cell type or sub-type), feedback mechanisms, tachyphylaxis mechanisms or mediators not so far identified or incorporated. Specific additions planned for asthma-relevant signalling in Th1, Th17 and ILC3 cell types and the involvement of these cells in intercellular communication and pathophysiological processes.

The way to address the problem of possible missing information is to continue reviewing relevant papers and regularly updating the resource using new publications, employing both text mining technologies and manual integration of knowledge. The assumption about missing relationships itself could be made, for example, in cases where the system is comprehensively described but the model still does not agree with experimental data. Then the task would be to go from unknown unknowns to known unknowns via offering hypotheses and trying to confirm or reject them, then plan experiments for learning more about the system.

*Single-cell sequencing*

It is going to be important to update and reshape the AsthmaMap resource based on the on single-cell sequencing analysis following on from the recent publication on the Human Lung Cell Atlas30. Thus single cell analysis will define new cell types, their status and cell-to-cell interactions.

*Timescales*

One important question in connection to the representation of biological systems, disease mechanisms and disease progression is that the biology can be and often is very different on different time-scales. For example, an acute allergen challenge to allergic asthmatics reveals rapid airway changes that are predominantly driven by rapid (in the range of 5-60 minutes) mast cell responses, whereas late phase airway responses (in the range of 1-12 hours) have a different biology driven by a different range of cells and mediators, and beyond that (days)

there are responses such as eosinophil influx into the lung, and the relevant mechanisms are different again. Would this mean different maps for different timescales or possible behaviour and different scenarios can be described in some other way?

The approach we would like to explore is that temporal changes on different timescales have to be addressed during the computational modelling stage. A disease map as a description of mechanisms can be seen as a “blueprint of possibilities” while possible scenarios can be simulated as a computational model. For that, different types of models can be developed based on a disease map. Another approach is building disease maps in such a way that all the building blocks are also executable modules readily available for simulation. The topic of transforming static representations into quantitative dynamic disease models is also discussed in the next section.

# Planned advances and applications

The AsthmaMap is intended to be used 1) as a source of systematically-organised information on mechanisms underlying asthma to support various related projects, e.g. Chronic Obstructive Pulmonary Disease (COPD); 2) as the basis for network analysis with or without using omics data; 3) for mathematical modelling and computational experiments; 4) as a teaching aid. The following sections outline the ongoing work and the planned development of the AsthmaMap resource and its applications.

*Effect of current medicines*

Asthmatics use inhaled corticosteroids (ICS) or inhaled corticosteroids in combination with long-acting beta2-agonists (ICS-LABA) as recommended by the latest guidelines of the Global Initiative for Asthma (GINA, https://ginasthma.org), while those with more severe asthma in addition take other medications such as leukotriene receptor antagonists or monoclonal antibody therapies e.g. those targeting the IL-5 pathway. Therefore, it would be highly desirable if the effects of these treatments could be built into the AsthmaMap. This may be challenging in some instances e.g. corticosteroids which have multiple effects on different cells and pathways; in other cases such as anti-IL-5 therapies, it should be straightforward to determine the effect of interventions within the AsthmaMap. In the even longer term, this could be applied to novel mechanisms of action of an oral prostaglandin DP2 receptor (CRTh2) antagonist31 or thymic stromal lymphopoietin (TSLP) antibody32, and combinations of pharmacological agents. This challenging and potentially very impactful work now can be approached based on the AsthmaMap resource, by adding drug action mechanisms directly to the map. In this way, it will be possible to model and simulate the effects of drugs even if it is a complex combination of possible effects. In addition, the impact of drugs on other non- asthma complex diseases may produce super-response signatures that could be applied to the AsthmaMap to indicate the impact of these therapies.

*Expanding the AsthmaMap Biochemical Mechanisms layer*

The currently available maps of the Biochemical Mechanisms layer cover a subset of asthma mechanisms. Expanding this layer will be done in the following directions: 1) taking into account the content of the Molecular Relations layer in the Biochemical Mechanisms layer, for example by adding modules for dendritic cells, macrophages, airway muscle cells, goblet cells etc. - the multiple modules as included in the Molecular Relations layer; 2) reusing high-quality

PD pathway databases, for example, the Reactome database12 from which pathways can be exported in SBGN and BioPAX formats; and 3) automatically creating Biochemical Mechanisms layer on demand.

Automated technology for this purpose does not exist yet. It assumes the extension of the pathway merging pipeline for supporting the SBGN PD format and the employment of advanced automatic layout algorithms referred to as “human-like layout algorithms”24,33.

*Applications for omics data analysis and interpretation*

A critical step in omics data analysis is the pathway-level interpretation of statistically significant results. There are many issues with the way in which pathways are defined as separate and linear sequences of reactions instead of a complex network of reactions34. Moreover, we know that signalling networks have a bow-tie structure35, which is often lost in the pathway-centric representation of biology. Additionally, pathways are usually defined for the “healthy state” and are not cell-specific. These difficulties in pathway definition are at the source of a well-known issue when interpreting omics results: the researcher ends up with a long list of over- represented pathways, sharing many common elements between them that are not relevant to the particular disease or the topic studied in the omics data analysis, thus including many false positives.

The AsthmaMap can help circumvent those issues. The molecular events are represented as a single network, the information is disease-specific and cell-specific. Combined with a map- topology enrichment analysis36, the AsthmaMap is a promising tool for interpretation of omics analysis results, as it is designed to be an easy-to-use intuitive environment for non- computational biologists.

*Transforming static representations into quantitative dynamic disease models*

The knowledge about pathway topology available via the AsthmaMap resource is currently used to develop a dynamic mechanistic model of mast cell dynamics (https://asthma.uni.lu/minerva/index.xhtml?id=m07). Mechanistic modelling will facilitate the integration of experimental data, comparison of competing hypotheses, the assessment of patient-patient variability, the prediction of drug responses, etc.37 So far, the mast cell module of the AsthmaMap Biochemical Mechanisms layer has been translated into an SBML model [(http://a](http://asthma-map.org/model))s[thma-map.org/model).](http://asthma-map.org/model)) In this process, the reaction kinetics were implemented and plausible parameter ranges were extracted from the literature. To infer the unknown parameters of the quantitative dynamical model, we started setting up a database containing quantitative experimental data from several published studies38 and implemented a parameter estimation pipeline39,40. As the model contains more than 300 biochemical species (genes, transcripts, proteins, small molecules, etc.) and 400 parameters, the parameter estimation requires substantial amounts of data and is highly non-trivial. However, initial results already helped to improve the AsthmaMap resource further, e.g., by detecting incorrect reactions.

*Regular updates, annotation and text-mining technologies*

In order to keep up with the newly published findings, it is necessary to employ available text- mining technologies to identify papers relevant for curation. This needs to be followed by manual curation to keep the resource reliable. A direct publication-to-network transformation

can be considered in the future if the quality of such a transformation is verified. Newly found relationships can be assimilated by the AsthmaMap via updating the disease map modules. Systems such as GitHub (https://github.com) and FAIRDOMHub (https://fairdomhub.org) may be used to track changes and to control the quality.

An important part of the update should be the annotation of nodes and connections in the AsthmaMap, so the map users are able to verify the reliability and relevant publications themselves. Annotations may also be done using text-mining technology. Further visualization of the reliability of the connections would provide help for further mathematical modelling on the basis of the map.

In the updates, the position of entities on the diagram would have to be adjusted but the main layout should be preserved in order to keep the human-readable interface as stable as possible. This will be possible to achieve by using incremental layout algorithms that take into account the existing layout29.

*Community curation*

As discussed previously, one of the main advantages of our approach is a consensus-based reconstruction of disease mechanisms supported by many experts. To continue supporting this approach, we use an environment that allows adding new information and comments not only by the map developers but also by the broader community. The MINERVA platform enables such a communication: for the Biochemical Mechanisms layer maps it is possible to select an element via right click and then choosing “Add comment” (https://asthma.uni.lu), email is requested so the moderators can request additional information or start a discussion. In a similar fashion, a section of a map can be selected and a comment added. Researchers can directly contact the developers using the contact email ([http://asthma-map.org/about/),](http://asthma-map.org/about/)) make a request or propose changes.

*Towards practical applications*

Practical applications of disease maps resources have been clearly demonstrated in the fields of cancer7–10 and neurodegenerative diseases41. We plan to build on this shared community expertise and work towards clinically applicable solutions for personalised diagnostics and treatment of asthma. Similarly to the published cases, the way to do this is by producing hypotheses via network-based data analysis and computational modelling, by designing and performing validation experiments and then by bridging findings to their use in clinics.

*Outlook*

The development of the AsthmaMap in the future will be coordinated with similar projects in the field of allergic diseases, other respiratory and immune-system-related diseases. This collaboration would lead to the distribution of shared tasks and optimisation of the work for all the participants. It is also important for the research focused on comorbidities and identification of mechanisms common for several diseases. In the end, it is possible that there will be collaborations on topics that at first sight might seem to be less likely related to asthma, e.g. diseases that might share inflammatory profiles with asthma such as rheumatoid arthritis (RA) or inflammatory bowel disease (IBD). Furthermore, in systems medicine, there is a tendency

of moving away from organ-specific diseases towards thinking about the whole-organism level and new disease ontologies.

The Disease Maps Community is the environment designed to facilitate such collaborations. Possible directions to explore are: 1) identification of other efforts focused on describing disease mechanisms or pathways of interests for many diseases and their clinical phenotypes;

2) identification and maintenance of modules shared between several diseases and co- morbidities; and 3) shaping clusters of disease maps for various topics such as inflammation, allergies, autoimmune diseases, cancers, age-related diseases etc. Development of the shared modules might increase the speed of disease map development, optimise work of many groups and stimulate collaboration between experts from related but traditionally distant fields. This will naturally contribute to shifting the focus from symptoms to mechanisms on the level of a whole organism.

# Conclusion

We present a machine- and human-readable representation of mechanisms underlying asthma, an open project driven by a community of experts in systems biomedicine, including lung diseases and computational systems biology specialists, combining human knowledge supported by machine performance at the crossroad of data and medical sciences. We offer an evolving conceptual model of disease mechanisms that can be transformed into executable dynamic models for computational hypothesis testing and predictions. This is in line with the efforts to actively maintain and reuse models42 and to make them reproducible (https://reproduciblebiomodels.org).

The multi-layer architecture of the AsthmaMap makes it compatible with the existing widely used pathway databases. For example, the Biochemical Mechanisms layer is compatible with the Reactome database format where metabolic, signalling and gene regulatory events are presented as detailed reactions and processes. That allows efficient exchange of pathway information while adapting the content to particular cell types in the asthma-specific context.

The AsthmaMap is designed to be useful for biologists working in asthma as well as clinical scientists working on defining endotypes of asthma towards personalised medicine. It is intended as a reference resource for making sense of omics datasets by integrating and analysing the prior knowledge and newly generated data. We believe this approach can move towards clinical practice via contributing to the design of clinical decision support systems and via using it for allergy- and inflammation-related research in systems medicine and systems pharmacology. It is expected that on the basis of a detailed description of relevant molecular and physiological processes it will be possible to move from ambiguous and imprecise description of asthma phenotypes to evidence-supported description of asthma endotypes via identifying and validating underlying mechanisms specific to various asthma subtypes. This will serve to deliver more precise medical treatments for this prevalent class of chronic disease.

# Availability

All relevant data are available at https://asthma-map.org. The code of the yEd-specific GraphML-SBGN converter is written in Java and available at https://github.com/sbgn/ySBGN.

# License

The MetaCore license was purchased by the CNRS-EISBM within the eTRIKS project for developing the Molecular Relations layer of the AsthmaMap. This included access to asthma- specific MetaCore pathways.

# Acknowledgements

The AsthmaMap project progresses in alliance with the efforts of the Disease Maps Community (https://disease-maps.org).

# References

1. Ivanova O, Richards LB, Vijverberg SJ, Neerincx AH, Sinha A, Sterk PJ, et al. What did we learn from multiple omics studies in asthma? Allergy. 2019;
2. Kuperstein I, Bonnet E, Nguyen H-A, Cohen D, Viara E, Grieco L, et al. Atlas of Cancer Signalling Network: a systems biology resource for integrative analysis of cancer data with Google Maps. Oncogenesis. 2015;4:e160.
3. Fujita KA, Ostaszewski M, Matsuoka Y, Ghosh S, Glaab E, Trefois C, et al. Integrating Pathways of Parkinson’s Disease in a Molecular Interaction Map. Mol Neurobiol. 2013;49:88–102.
4. Mizuno S, Iijima R, Ogishima S, Kikuchi M, Matsuoka Y, Ghosh S, et al. AlzPathway: a comprehensive map of signaling pathways of Alzheimer’s disease. BMC Syst Biol. 2012;6:52.
5. Matsuoka Y, Matsumae H, Katoh M, Eisfeld AJ, Neumann G, Hase T, et al. A comprehensive map of the influenza A virus replication cycle. BMC Syst Biol. 2013;7:97.
6. Singh V, Ostaszewski M, Kalliolias GD, Chiocchia G, Olaso R, Petit-Teixeira E, et al. Computational Systems Biology Approach for the Study of Rheumatoid Arthritis: From a Molecular Map to a Dynamical Model. Genomics Comput Biol. 2018;4.
7. Monraz Gomez LC, Kondratova M, Ravel J-M, Barillot E, Zinovyev A, Kuperstein I. Application of Atlas of Cancer Signalling Network in preclinical studies. Brief Bioinform. 2019;20:701–16.
8. Chanrion M, Kuperstein I, Barrière C, El Marjou F, Cohen D, Vignjevic D, et al. Concomitant Notch activation and p53 deletion trigger epithelial-to-mesenchymal transition and metastasis in mouse gut. Nat Commun. 2014;5:5005.
9. Jdey W, Thierry S, Russo C, Devun F, Al Abo M, Noguiez-Hellin P, et al. Drug-Driven Synthetic Lethality: Bypassing Tumor Cell Genetics with a Combination of AsiDNA and PARP Inhibitors. Clin Cancer Res Off J Am Assoc Cancer Res. 2016;
10. Kuperstein I, Grieco L, Cohen DPA, Thieffry D, Zinovyev A, Barillot E. The shortest path is not the one you know: application of biological network resources in precision oncology research. Mutagenesis. 2015;30:191–204.
11. Mazein A, Ostaszewski M, Kuperstein I, Watterson S, Le Novère N, Lefaudeux D, et al. Systems medicine disease maps: community-driven comprehensive representation of disease mechanisms. Npj Syst Biol Appl. 2018;4.
12. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, et al. The Reactome Pathway Knowledgebase. Nucleic Acids Res. 2018;46:D649–55.
13. Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, et al. PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. Nucleic Acids Res. 2017;45:D183–9.
14. Le Novère N, Hucka M, Mi H, Moodie S, Schreiber F, Sorokin A, et al. The Systems Biology Graphical Notation. Nat Biotechnol. 2009;27:735–41.
15. Demir E, Cary MP, Paley S, Fukuda K, Lemer C, Vastrik I, et al. The BioPAX community standard for pathway data sharing. Nat Biotechnol. 2010;28:935–42.
16. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, et al. The systems biology

markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinforma Oxf Engl. 2003;19:524–31.

1. Edwards MR, Saglani S, Schwarze J, Skevaki C, Smith JA, Ainsworth B, et al. Addressing unmet needs in understanding asthma mechanisms: From the European Asthma Research and Innovation Partnership (EARIP) Work Package (WP)2 collaborators. Eur Respir J. 2017;49.
2. Mazein A, Knowles RG, Adcock I, Chung KF, Wheelock CE, Maitland-van der Zee AH, et al. AsthmaMap: An expert-driven computational representation of disease mechanisms. Clin Exp Allergy. 2018;48:916–8.
3. Ostaszewski M, Gebel S, Kuperstein I, Mazein A, Zinovyev A, Dogrusoz U, et al. Community-driven roadmap for integrated disease maps. Brief Bioinform. 2018;
4. Le Novère N. Quantitative and logic modelling of molecular and gene networks. Nat Rev Genet. 2015;16:146–58.
5. Moodie S, Le Novère N, Demir E, Mi H, Villéger A. Systems Biology Graphical Notation: Process Description language Level 1 Version 1.3. J Integr Bioinforma. 2015;12:263.
6. Gawron P, Ostaszewski M, Satagopam V, Gebel S, Mazein A, Kuzma M, et al. MINERVA-a platform for visualization and curation of molecular interaction networks. NPJ Syst Biol Appl. 2016;2:16020.
7. Saglani S, Menzie-Gow AN. Approaches to Asthma Diagnosis in Children and Adults. Front Pediatr. 2019;7:148.
8. Siebenhaller M, Nielsen SS, McGee F, Balaur I, Auffray C, Mazein A. Human-like layout algorithms for signalling hypergraphs: outlining requirements. Brief Bioinform. 2018;
9. Ostaszewski M, Kieffer E, Danoy G, Schneider R, Bouvry P. Clustering approaches for visual knowledge exploration in molecular interaction networks. BMC Bioinformatics. 2018;19:308.
10. Türei D, Korcsmáros T, Saez-Rodriguez J. OmniPath: guidelines and gateway for literature-curated signaling pathway resources. Nat Methods. 2016;13:966–7.
11. Cerami EG, Gross BE, Demir E, Rodchenkov I, Babur Ö, Anwar N, et al. Pathway Commons, a web resource for biological pathway data. Nucleic Acids Res. 2011;39:D685–90.
12. Groß A, Pruski C, Rahm E. Evolution of biomedical ontologies and mappings: Overview of recent approaches. Comput Struct Biotechnol J. 2016;14:333–40.
13. Dogrusoz U, Karacelik A, Safarli I, Balci H, Dervishi L, Siper MC. Efficient methods and readily customizable libraries for managing complexity of large networks. PLOS ONE. 2018;13:e0197238.
14. Schiller HB, Montoro DT, Simon LM, Rawlins EL, Meyer KB, Strunz M, et al. The Human Lung Cell Atlas: A High-Resolution Reference Map of the Human Lung in Health and Disease. Am J Respir Cell Mol Biol. 2019;61:31–41.
15. Bateman ED, Guerreros AG, Brockhaus F, Holzhauer B, Pethe A, Kay RA, et al. Fevipiprant, an oral prostaglandin DP2 receptor (CRTh2) antagonist, in allergic asthma uncontrolled on low-dose inhaled corticosteroids. Eur Respir J. 2017;50.
16. Corren J, Parnes JR, Wang L, Mo M, Roseti SL, Griffiths JM, et al. Tezepelumab in Adults with Uncontrolled Asthma. N Engl J Med. 2017;377:936–46.
17. Kieffer S, Dwyer T, Marriott K, Wybrow M. HOLA: Human-like Orthogonal Network Layout. IEEE Trans Vis Comput Graph. 2016;22:349–58.
18. Khatri P, Sirota M, Butte AJ. Ten Years of Pathway Analysis: Current Approaches and Outstanding Challenges. Ouzounis CA, editor. PLoS Comput Biol. 2012;8:e1002375.
19. Abd-Rabbo D, Michnick SW. Delineating functional principles of the bow tie structure of a kinase-phosphatase network in the budding yeast. BMC Syst Biol. 2017;11:38.
20. Alexeyenko A, Lee W, Pernemalm M, Guegan J, Dessen P, Lazar V, et al. Network enrichment analysis: extension of gene-set enrichment analysis to gene networks. BMC Bioinformatics. 2012;13:226.
21. Wolkenhauer O. Why model? Front Physiol [Internet]. 2014 [cited 2017 Feb 10];5. Available from: <http://journal.frontiersin.org/article/10.3389/fphys.2014.00021/abstract>
22. Parravicini V, Gadina M, Kovarova M, Odom S, Gonzalez-Espinosa C, Furumoto Y, et al. Fyn kinase initiates complementary signals required for IgE-dependent mast cell

degranulation. Nat Immunol. 2002;3:741–8.

1. Fröhlich F, Kaltenbacher B, Theis FJ, Hasenauer J. Scalable Parameter Estimation for Genome-Scale Biochemical Reaction Networks. PLOS Comput Biol. 2017;13:e1005331.
2. Stapor P, Weindl D, Ballnus B, Hug S, Loos C, Fiedler A, et al. PESTO: Parameter EStimation TOolbox. Bioinforma Oxf Engl. 2018;34:705–7.
3. Satagopam V, Gu W, Eifes S, Gawron P, Ostaszewski M, Gebel S, et al. Integration and Visualization of Translational Medicine Data for Better Understanding of Human Diseases. Big Data. 2016;4:97–108.
4. Glont M, Nguyen TVN, Graesslin M, Hälke R, Ali R, Schramm J, et al. BioModels: expanding horizons to include more modelling approaches and formats. Nucleic Acids Res. 2018;46:D1248–53.