UniversitätsSpital Zürich

Group Prof. Dr. Adriano Aguzzi
Institute of Neuropathology

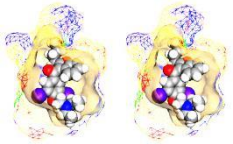
# Computational modelling: Complementary Tools to Animals in Science

Andra Chincisan

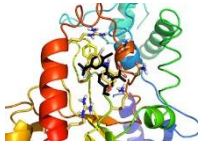
Journal Club: Special series on Laboratory Animal Science

5 September 2017

# What are computational models?



**Computational models** are **mathematical** models that are **simulated** using computation to study complex systems. In biology, one example is the use of a computational model to study an outbreak of an infectious disease such as influenza. The **parameters** of the mathematical model are adjusted using **computer simulation** to study different **possible outcomes**.



Modeling can expedite research by allowing scientists to conduct thousands of simulated experiments by computer in order to identify the actual physical experiments that are most likely to help the researcher find the solution to the problem being studied. Simulations of models can reveal hidden patterns and/or counterintuitive mechanisms in complex systems.

*https://www.nibib.nih.gov/science-education/science-topics/computational-modeling*

# Computational Modeling, Formal Analysis, and Tools for Systems Biology

As the amount of **biological data** in the public domain **grows**, so does the range of modeling and analysis techniques employed in systems biology.

The growing interest in systems biology in executable models and their analysis has necessitated the borrowing of terms and methods from computer science, such as formal analysis, model checking, static analysis, and runtime verification.

| Tool | Main case studies |
|---|---|
| BAM [92] | LDL degradation pathway [93] |
| BetaWB [12] | The MAPK biochemical pathway [12], cell-cycle [11] |
| BIOCHAM [30] | Mammalian cell cycle control [34], G protein-coupled receptor kinases [31] |
| BioDivine [69,78] | Genetic regulatory networks [79] |
| BioNetGen [22] + BioLab [33] | HMGB1 signal pathway [35]<br>Analysis of T-cell receptor signaling pathway [33] |
| Bio-PEPA WB [13] | Plant circadian clock [14] |
| BoolNet [52] | Genetic networks [52] |
| BMA [56] | Biological signaling networks [55] |
| BNS [53] | Cell cycle sequence of fission yeast [54] |
| Breach [80] | Collagen proteolysis [115], Cellular iron homeostasis network [81] |
| CompuCell3D [103] | Vertebrate segmentation and somite formation [104] |
| COPASI [8] | Biochemical networks [9] |
| dReach [82] | Cardiac cell hybrid models [83] |
| FLAME [27] | Sperm behavior [97] |
| GINsim [49] | Diversity and plasticity of Th cell types [50] |
| | MAPK network on cancer cell fate decision [51] |
| GreatSPN [43] | Signal transduction pathways for angiogenesis [44] |
| IBM Rational Rhapsody [58] | T-cell activation with statecharts [59] |
| KaSim [32] | EGFR signaling [36] |
| Mathworks Simulink [76] | Heart model for pacemaker verification [77] |
| Pathway Logic [45] | Sporulation initiation in *B. subtilis* [46]<br>MAPK signaling network [46]<br>EGF stimulation network [45] |
| PRISM [6] | Biological signaling pathways [6,143,144], bone pathologies [107] |
| Rovergene [71] | Synthetic transcription cascade [71], myocyte excitability [66], bone remodeling [107] |
| Snoopy [26] + MARCIE [42] | Systems and synthetic biology [41] |
| SPiM [15] | Modeling of the EGFR network [16], MHC class I peptide optimization [17] |
| S-TaLiRo [84] | Modeling of the insulin-glucose regulatory system [149] |
| REPAST [28] | Bone remodeling [98] |

doi:10.1371/journal.pcbi.1004591.t001

*Example of computational models in systems biology*

# Modelling methods

**Data driving modeling**

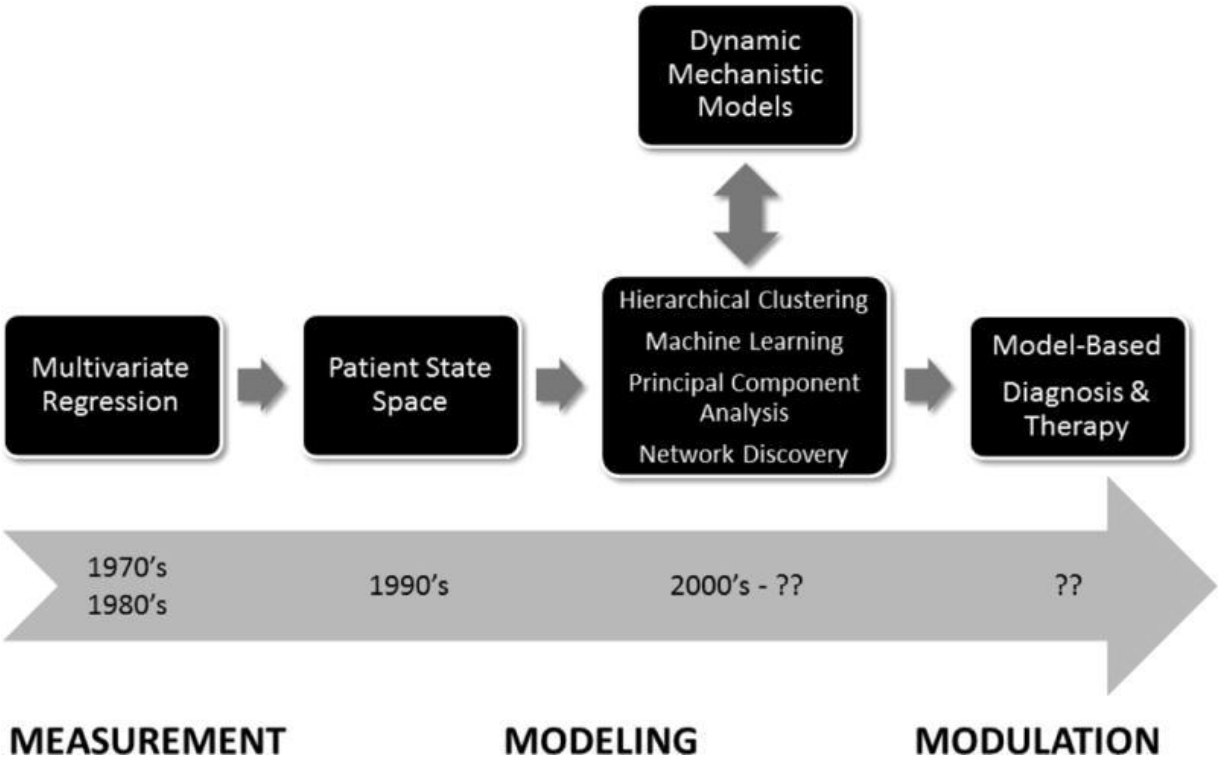
- Multivariate regression
- Hierarchical clustering
- Principal Component analysis
- Machine learning
- Network discovery

**Mechanistic modeling**

- Equation base models
- Agent-base models
- Rule-based models

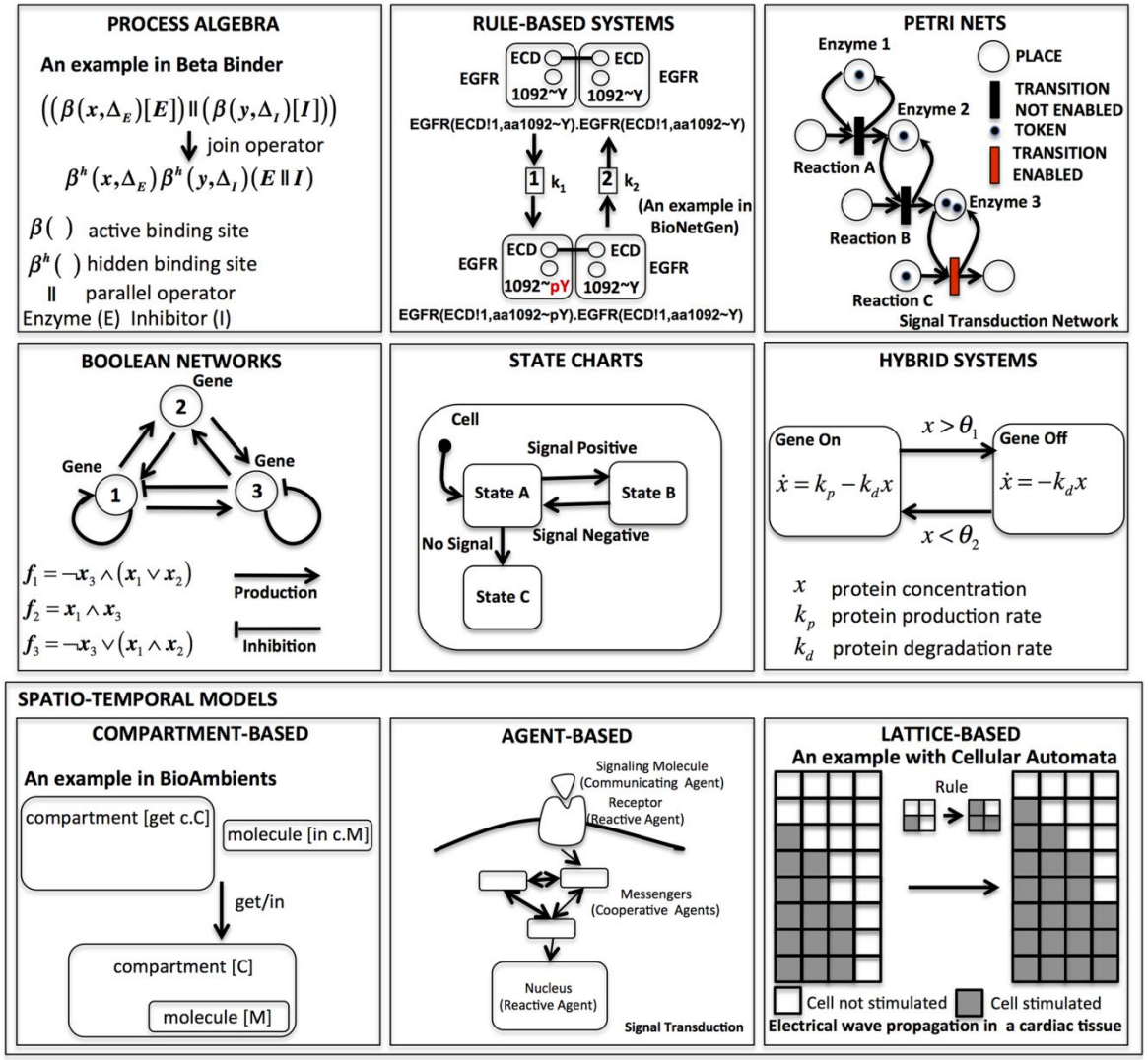
**Alternative models - concurrent system**

- Petri net (graph based)
- Process algebra




*Vodovotz, Yoram, and Timothy R. Billiar. "In Silico Modeling: Methods and Applications toTrauma and Sepsis." Critical care medicine 41.8 (2013): 2008.*

# Computational modeling approaches




*Bartocci, Ezio, and Pietro Lió. "Computational modeling, formal analysis, and tools for systems biology." PLoS computational biology 12.1 (2016): e1004591.*

# Papers

**Mechanistic modeling**
- Equation base models
- Agent-base models
- Rule-based models

**Alternative models - concurrent system**
- Petri net (graph based)
- Process algebra

**Data driving modeling**
- Multivariate regression
- Hierarchical clustering
- Principal Component analysis
- Machine learning
- Network discovery


**Petri Net computational modelling of Langerhans cell Interferon Regulatory Factor Network predicts their role in T cell activation**

Marta E. Polak, Chuin Ying Ung, Joanna Masapust, Tom C. Freeman & Michael R. Ardern-Jones

Scientific Reports. 2017;4.
doi:10.1038/s41598-017-00651-5


**Machine learning workflow to enhance predictions of Adverse Drug Reactions (ADRs) through drug-gene interactions: application to drugs for cutaneous diseases**

Kalpana Raja, Matthew Patrick, James T. Elder & Lam C. Tsoi

Scientific Reports. 2017;6.
doi:doi:10.1038/s41598-017-03914-3

SCIENTIFIC REPORTS
nature publishing group npg
Scientific Reports

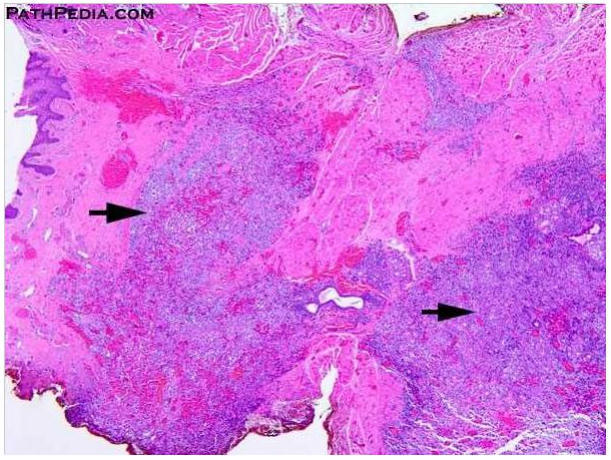
# Petri Net computational modelling of Langerhans cell Interferon Regulatory Factor Network predicts their role in T cell activation

Marta E. Polak, Chuin Ying Ung, Joanna Masapust, Tom C. Freeman & Michael R. Ardern-Jones

Scientific Reports. 2017;4.
doi:10.1038/s41598-017-00651-5

# Introduction

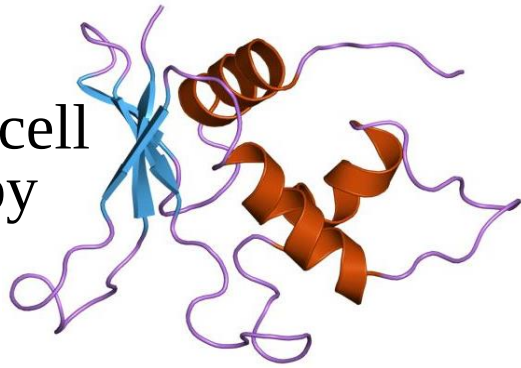
- **Langerhans cells** (LCs) are able to orchestrate adaptive immune responses in the skin by interpreting the microenvironmental context in which they encounter foreign substances, but the **regulatory basis** for this has **not** been **established**.




*Langerhans cell histiocytosis, tongue, www.Pathpedia.Com*

- Aim:
  - Utilising systems immunology approaches combining ***in silico*** **modelling** of a reconstructed gene regulatory network (GRN) with ***in vitro*** **validation** of the predictions, we sought to determine the **mechanisms** of **regulation** of **immune responses** in human primary **LCs.**

# Introduction

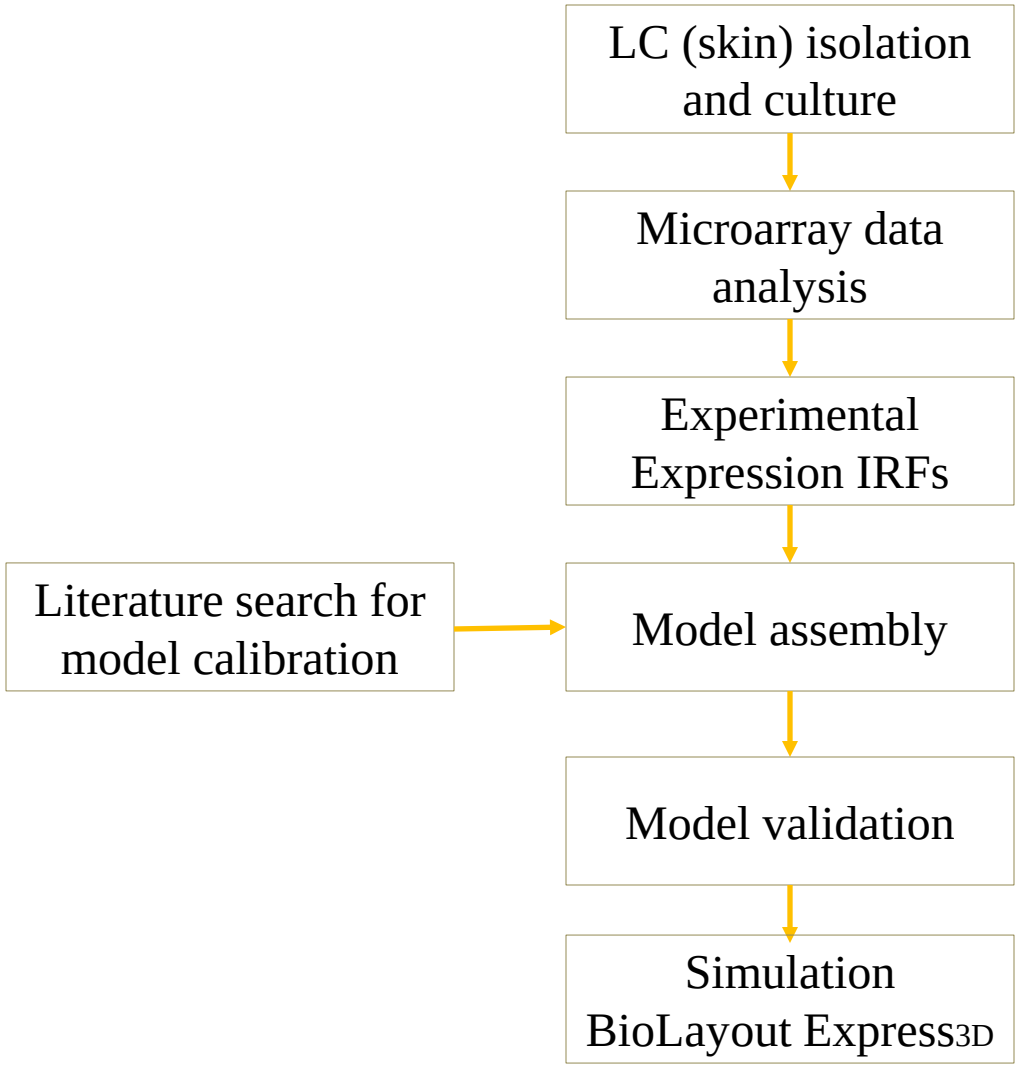


- The key role of **Interferon regulatory factors** (IRFs) as controllers of the human Langerhans cell response to epidermal cytokines was revealed by whole transcriptome analysis.

- Applying Boolean logic we assembled a **Petri net-based model of the IRF-GRN** which provides molecular **pathway predictions** for the induction of different transcriptional programmes in LCs.

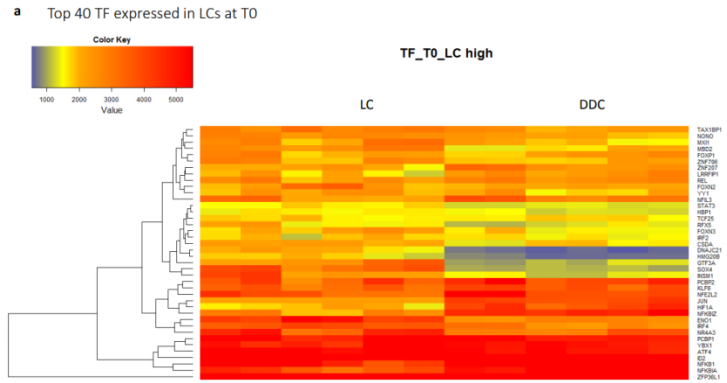
- Characterise the differential effect of **key epidermal cytokines, TNFα and TSLP**, on the ability of LCs to cross-present viral antigens to cytotoxic T cells, and to propose a transcriptional mechanism regulating this process.

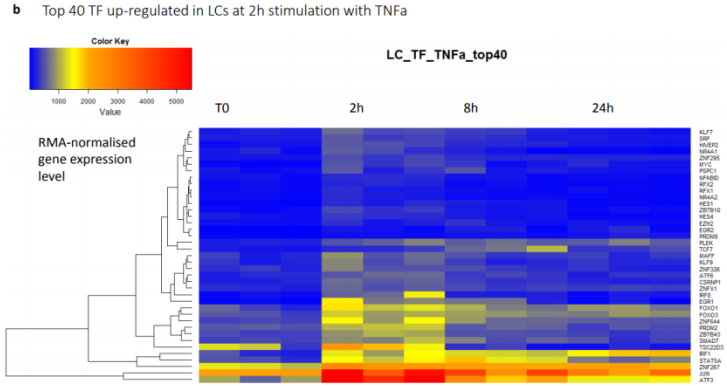
# System overview

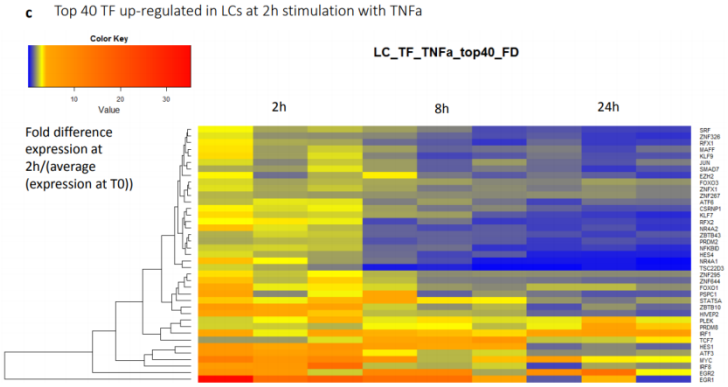


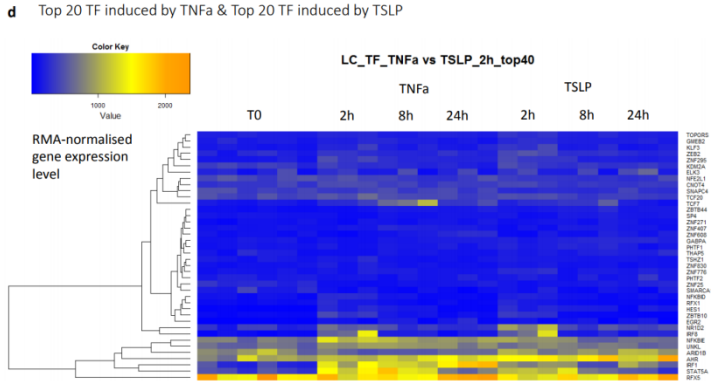
# Microarray data analysis

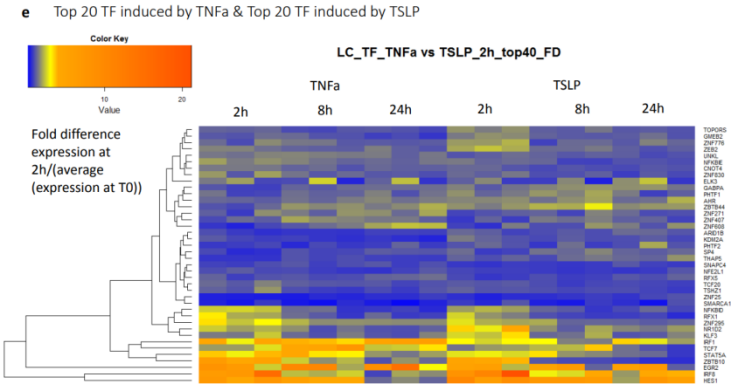
Top transcription factors

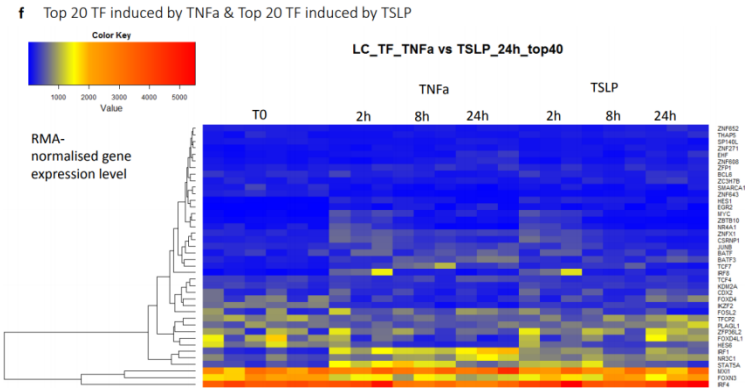


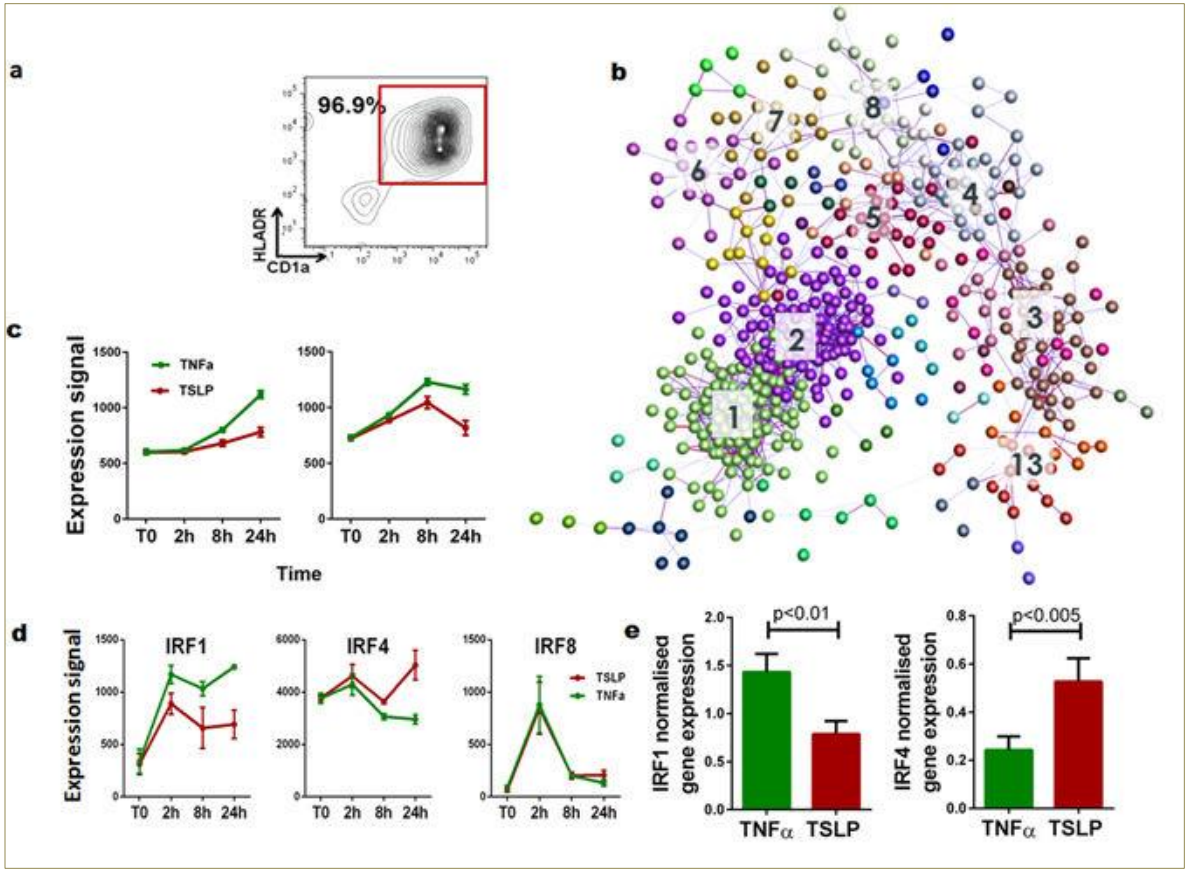



# Changes in Langerhans cell core transcriptional network

Epidermal cytokines, TNFα and TSLP, differentially regulate the expression of Interferon Regulatory Factors in human migratory LC

a) Human LCs
b) Core transcriptomic networks of human LCs
c) Expression profiles of clusters during 24 h simulations
d) Expression changes of IRF1, IRF4 and IRF8 i n LC
e) Differential induction of IRF1 and IRF4 mRNA by TNFα and TSLP



# Model Assembly

Search strategy to identify components of the IRF GRN network (PubMed)

| search term | number of publications |
|---|---|
| "Interferon regulatory factor" or IRF and antigen presentation | 71 |
| "Interferon regulatory factor" or IRF and dendritic cell and T cell stimulation | 22 |
| "Interferon regulatory factor" or IRF1 or IRF4 or IRF8 and *transcripton partner* as per the transcription partner list | 510 |
| Interferon regulatory factor or IRF1 or IRF4 or IRF8 and ChIP-seq | 15 |

→ **82** unique papers

# Model Assembly

**82** unique papers

Identify

-Stimulus
-Cell type
-Biological process controlled by IRF Interaction partner
-DNA binding sequence

Categorize

Input node
Transmission node
Output node
Mode of interaction

**Interaction database**

| | | | **Stimulus** | **Interaction partner A** | **Interaction** | **Interaction partner B** | **DNA seq** | **Gene transcription/ biological process** |
|---|---|---|---|---|---|---|---|---|
| **Citation** | **pubmed id** | **cell type** | **Stimulus** | **partner A** | **interaction** | **Partner B** | **DNA sequence** | **outcome** |
| Hildner Science 2008 | 19008445 | DC (mouse) | | BATF3 | essential | | | cross-presentation |
| Hildner Science 2008 | 19008445 | DC (mouse) | | BATF3 | essential | | | anti-viral responses |
| | | | | BATF3 | | IRF4/8??? | AICE? | cross-presentation and CD8 responses |
| Ma JBC 1997 | 9099678 | nucleic acid level | IFNg priming for LPS | ETS2 | | ? | ETS2 - site, complex F1 | IL12p40 |
| | | | | ETS2 | | ? | | IL12p40+>Th1 |
| Roy JI 2015 | 25957166 | macrophages | IFNg | IRF1 | synergy | BATF2 | IRF1 binding | Th1 |
| Marecki JI 2001 | 11359842 | fibroblasts (transf) | | IRF1 | synergy | IRF4/PU.1 | ISRE/EICE | IL1B |
| Marecki JI 2001 | 11359842 | fibroblasts (transf) | | IRF1 | synergy | IRF8/PU.1 | ISRE/EICE | IL1B |
| Shi Gene 2001 | 21803131 | monocytes | | IRF1 | | | | antigen processing to class I |
| Gabrielle J Leukocyt Biol 2006 | 16966383 | DC (mouse) | | IRF1 | | inhibits | | immunological tolerance |

# Boolean gates

**DB of interactions**

↓

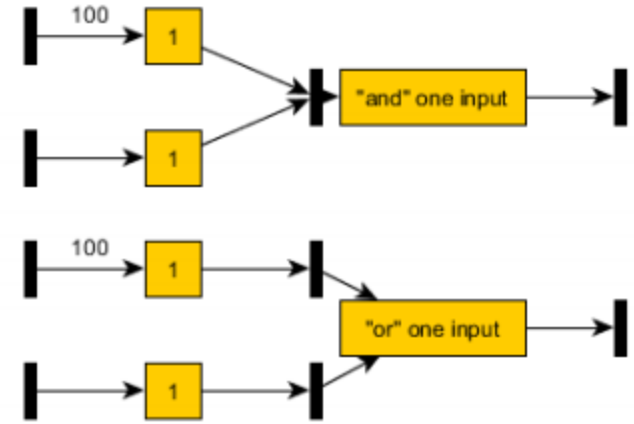
**Boolean network**

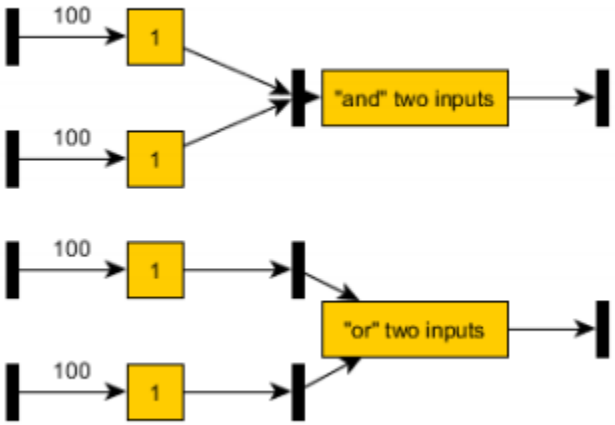
→ **Gate AND** (both essential)
**Gate OR** (one essential)
**Gate INHIBITON**

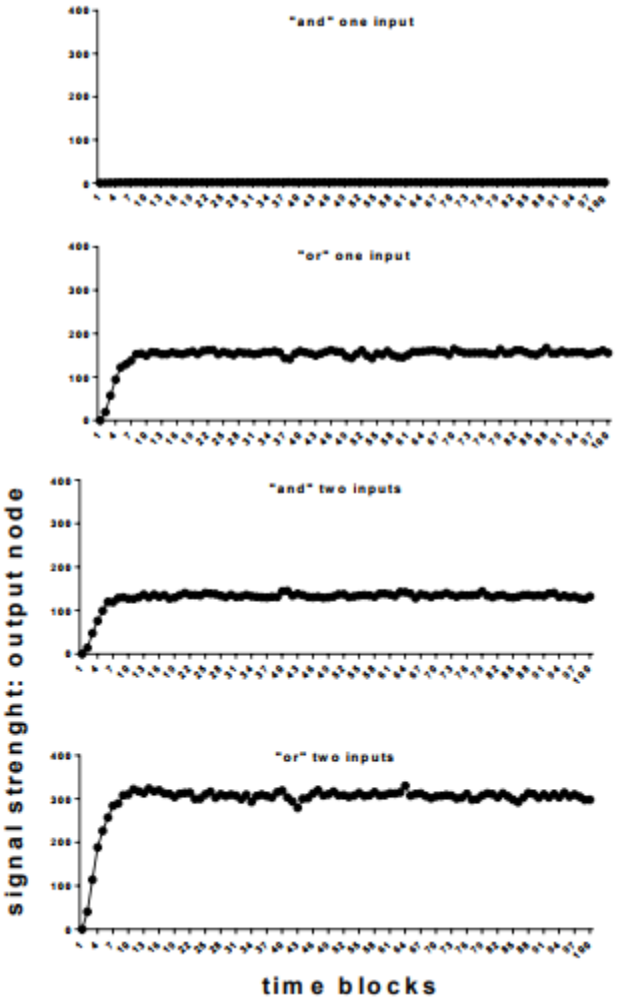
Boolean gates

| Interaction partner 1 | GATE | Interaction partner 2 | interaction | binding site | outcome | GATE |
|---|---|---|---|---|---|---|
| IRF1 | and | IRF1 | induction | ISRE | TH1/CD8 | |
| IRF1 | inhibition | IRF4 | inhibition | ISRE | TH1/CD8 | |
| IRF1 | and | IRF8 | induction | ISRE | TH1/CD8 | |
| IRF1 | not reported | AP1 | | | | |
| IRF1 | not reported | ETS | | | | |
| IRF4 | and | IRF4 | inhibition | ISRE | TH1/CD8 | |
| IRF8 | and | IRF8 | inhibition | ISRE | TH1/CD8 | OR |
| IRF4 | and | AP1 | induction | AICE | TH17 | |
| IRF4 | and | ETS | induction | EICE | TH2 | OR |
| IRF8 | and | ETS | induction | EICE | CD4 | |
| IRF8 | and | AP1 | induction | AICE | CD8 | OR |
| PRDM1 | inhibition | IRF4 | | EICE | CD4 | |
| PRDM1 | inhibition | IRF8 | | EICE | CD4 | OR |

# Boolean gates



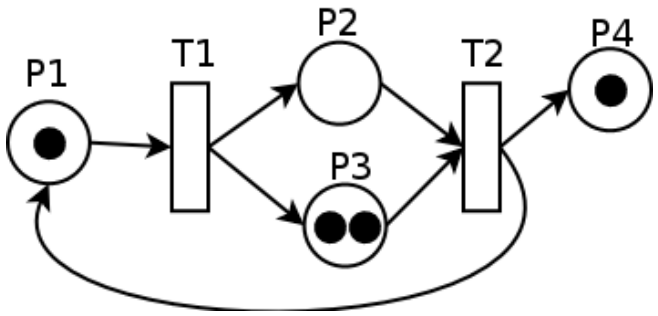
# Why Petri Net?

*Qualitative vs Quantitative methods / Petri Net vs ODES*

- Petri nets are **qualitative** logic-based models do not depend on **quantitative** data but rather on the structure of the network along with a set of logical constraints.
- Built from local **experimental observations** or knowledge-based information without kinetic information due to their finite states.
- **Quantitative** methods such as ordinary differential equations (ODEs), model the rate of change of each component in the network and require comprehensive knowledge of **kinetic** parameters, which are unknown for most networks, and therefore their applicability is limited.
- ODEs can be used for modelling **small scale** GRN.

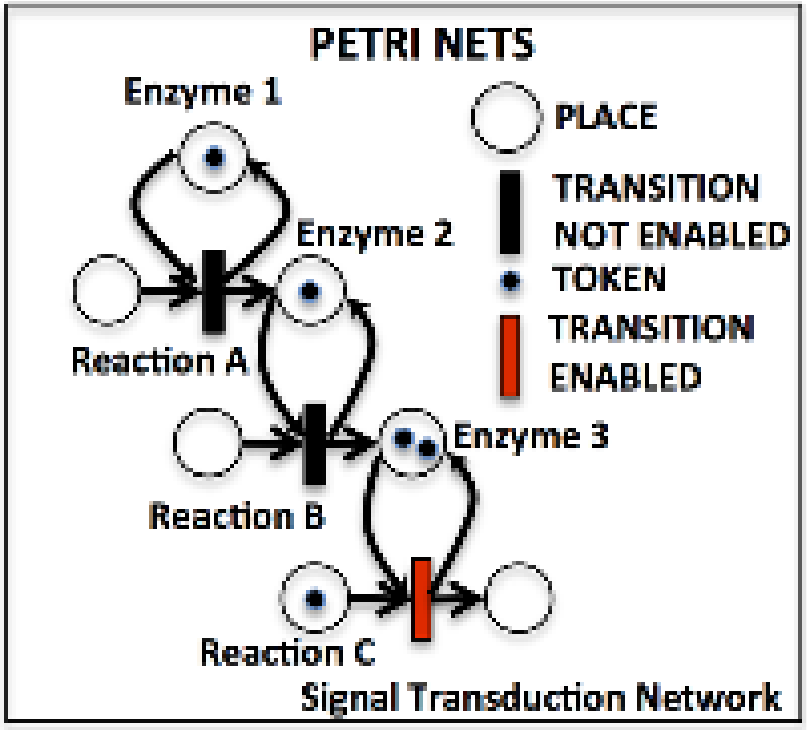
# Petri Net

- Originate from *C.A. Petri's* PhD thesis (1962), technique for the description and analysis of concurrent behavior in **natural processes - distributed systems**
- Advantages
  - based on a few **simple concepts**
  - **simple graphical** format and a precise operational semantics, attractive option for modeling the static and dynamic aspects of processes
  - many techniques & many **extensions** and variants



# Petri net

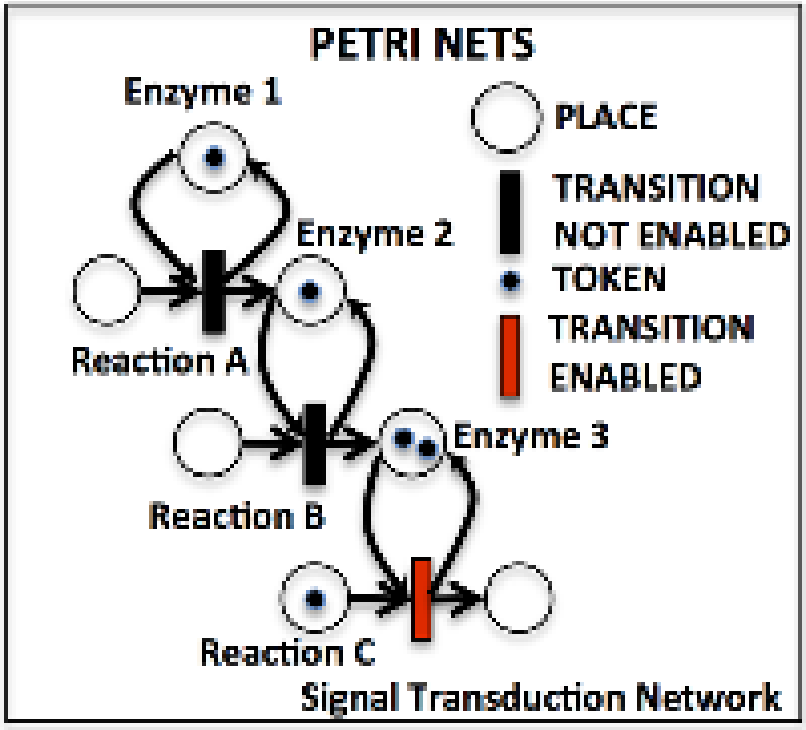
- Qualitative / quantitative
- Stochastic/ deterministic
- Limitation: model spatial information
- Model: **Stochastic Petri Net**




*http://hmi.ewi.utwente.nl/verslagen/capita-selecta/CS-Bos-Wim.pdf*

# Petri net

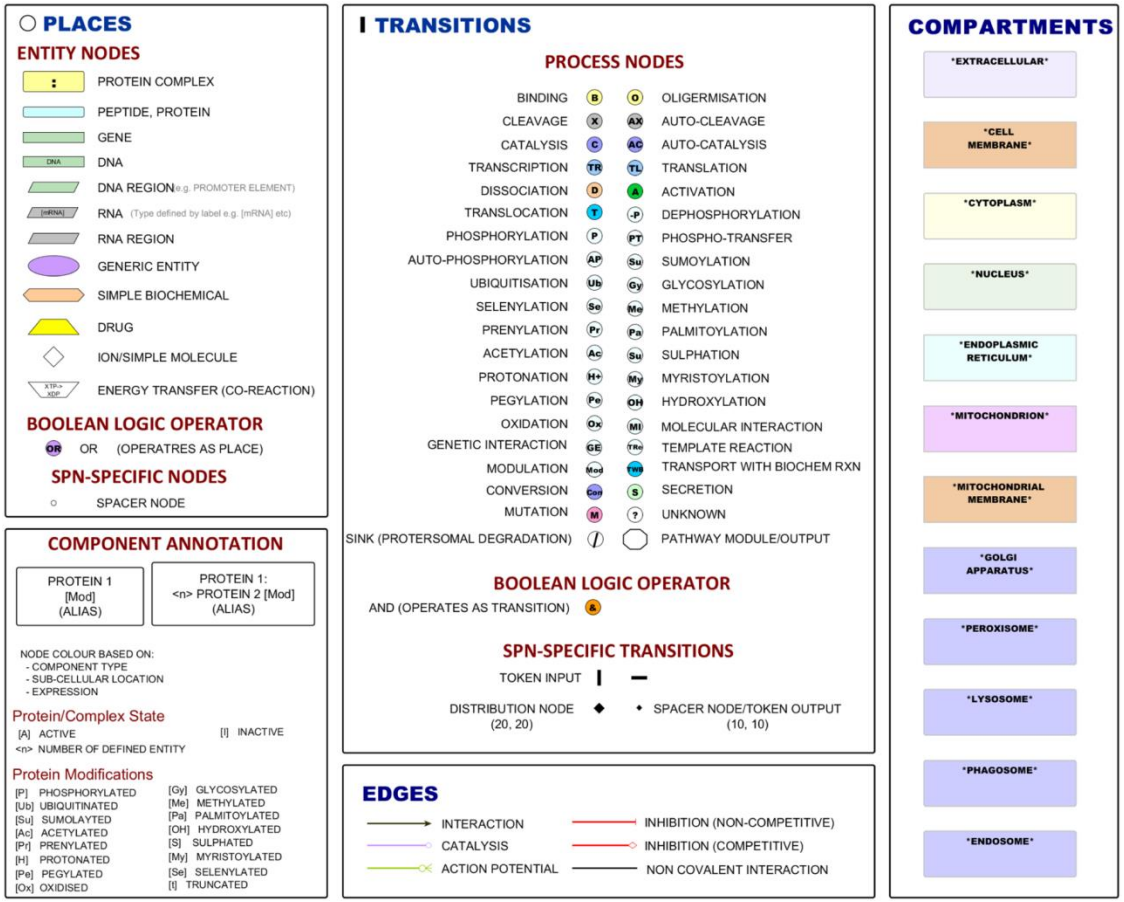
- **Tokens** A place can contain zero or more tokens.
- **Places** represent intermediate states that may exist during the operation of a process.
- **Transitions** correspond to the activities or events of which the process is made up.
- **Arcs** connect places and transitions in a way that places can only be connected to transitions and vice -versa




*http://hmi.ewi.utwente.nl/verslagen/capita-selecta/CS-Bos-Wim.pdf*

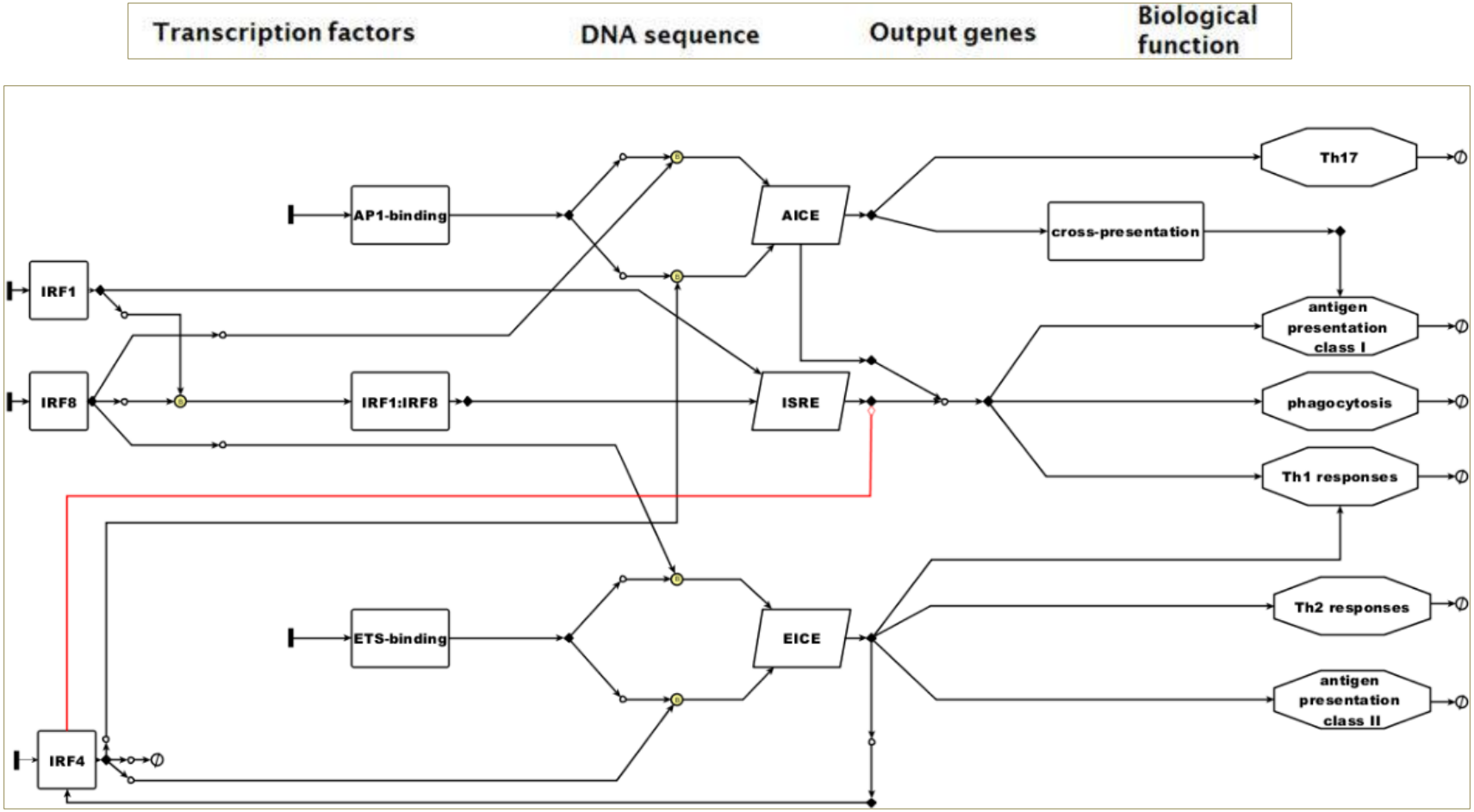
# IRF GRN model parametrization and design

- Token → GRN entry transitions → levels of expression of the transcript (level or on/off)




mEPN notation, allowing computational modelling of concurrent systems.

# IRF GRN model



# SPN model simulations

Experimentally measured expression values at

| *Time (in vitro)* | *Time (in silico)* |
|---|---|
| 0h | 0-8 time block |
| 2h | 9-32 time block |
| 8h | 33-75 time block |
| 24h | 76-100 time block |

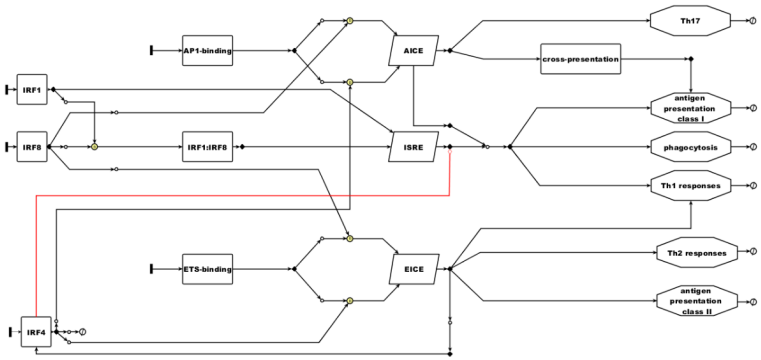
| | LC TNFa | LC TSLP |
|---|---|---|
| IRF1 | 0-8,325;9-32,1267;33-75,1209;76-100,1782 | 0-8,293;9-32,841;33-75,585;76-100,796 |
| IRF8 | 0-8,89;9-32,879;33-75,200;76-100,131 | 0-8,63;9-32,847;33-75,203;76-100,206 |
| IRF4 | 0-8,3762;9-32,4296;33-75,3067;76-100,2961 | 0-8,3773;9-32,4618;33-75,3638;76-100,5034 |
| cJUN | 0-8,2206;9-32,4831;33-75,3571;76-100,2797 | 0-8,2204;9-32,4798;33-75,3147;76-100,2207 |
| cFOS | 0-8,1072;9-32,811;33-75,153;76-100,34 | 0-8,1125;9-32,783;33-75,109;76-100,43 |
| BATF | 0-8,259;9-32,490;33-75,393;76-100,513 | 0-8,259;9-32,449;33-75,290;76-100,276 |
| BATF3 | 0-8,174;9-32,299;33-75,511;76-100,697 | 0-8,174;9-32,270;33-75,325;76-100,469 |
| ELF1 | 0-8,650;9-32,1112;33-75,724;76-100,521 | 0-8,669;9-32,1234;33-75,692;76-100,457 |
| ELF4 | 0-8,159;9-32,244;33-75,204;76-100,198 | 0-8,155;9-32,238;33-75,181;76-100,163 |
| ELK1 | 0-8,182;9-32,172;33-75,200;76-100,175 | 0-8,170;9-32,182;33-75,168;76-100,176 |
| ELK3 | 0-8,194;9-32,273;33-75,249;76-100,261 | 0-8,272;9-32,317;33-75,248;76-100,423 |
| ETS1 | 0-8,775;9-32,868;33-75,883;76-100,972 | 0-8,849;9-32,935;33-75,900;76-100,1292 |
| ETS2 | 0-8,404;9-32,413;33-75,225;76-100,118 | 0-8,389;9-32,463;33-75,250;76-100,130 |
| EHF | 0-8,92;9-32,133;33-75,117;76-100,295 | 0-8,105;9-32,146;33-75,112;76-100,229 |
| ELF2 | 0-8,234;9-32,306;33-75,209;76-100,252 | 0-8,241;9-32,341;33-75,231;76-100,243 |
| ETV3 | 0-8,956;9-32,889;33-75,544;76-100,785 | 0-8,884;9-32,843;33-75,563;76-100,749 |
| ETV6 | 0-8,558;9-32,412;33-75,363;76-100,392 | 0-8,527;9-32,489;33-75,438;76-100,448 |
| GABPa | 0-8,141;9-32,121;33-75,167;76-100,184 | 0-8,144;9-32,190;33-75,243;76-100,234 |

# Model validation

- Stochastic simulation of a logic-based diagram of the IRF gene regulatory network with Petri Nets => correctly **re-capitulates** the **observations** from multiple experimental systems
- Qualitative: Input/Output

**Model calibration**

IRF GRN model




Theoretical quantities

0 – lack of transcription

100 – expressed gene

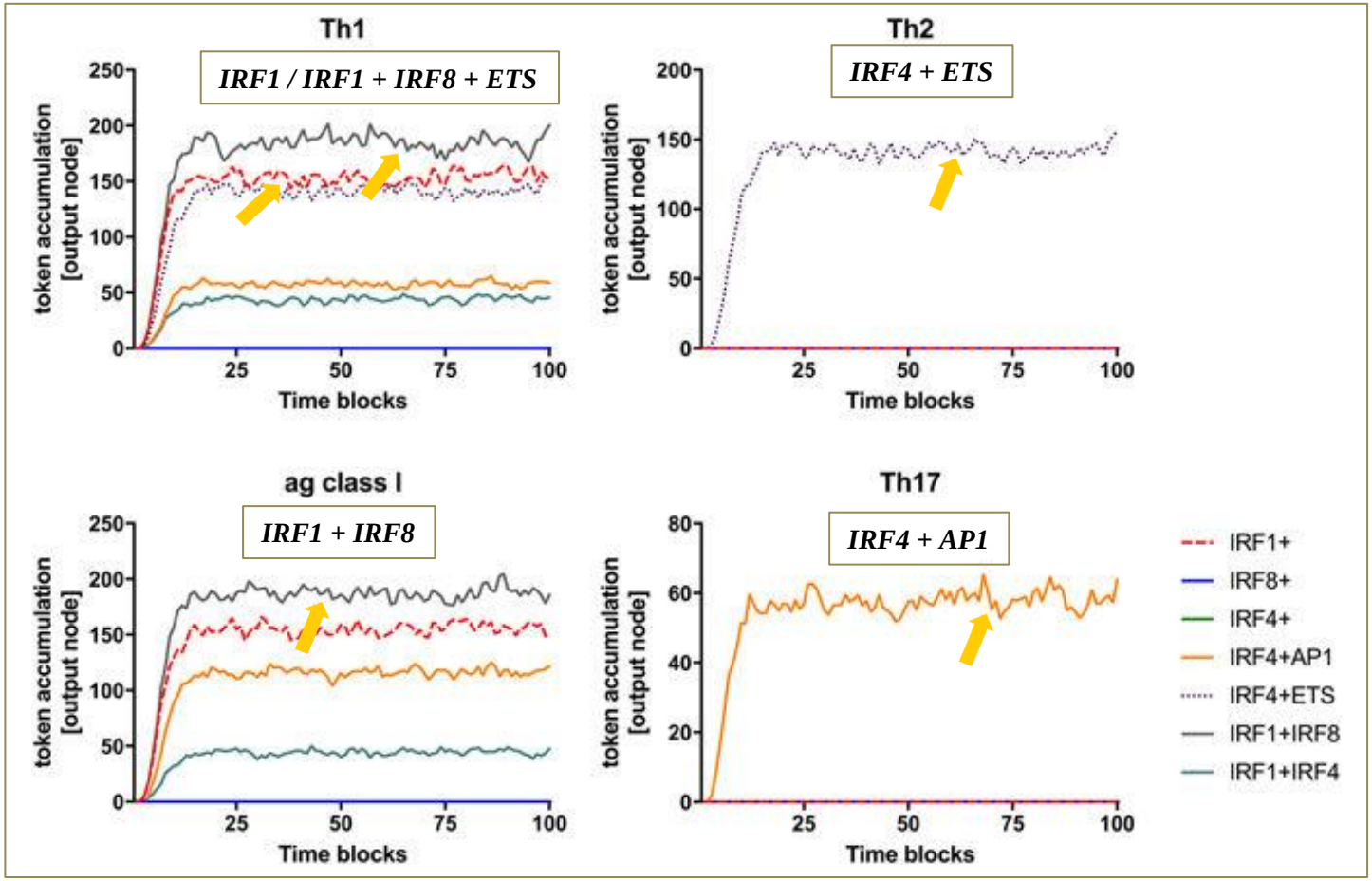
Entry nodes

IRF1
IRF4
IRF8
AP1-binding
ETS-binding

# Model validation

Model of IRF-GRN assembled based on a systematic literature review have been simulated with Signalling Petri Nets in BioLayout *Express* $^{3D}$.

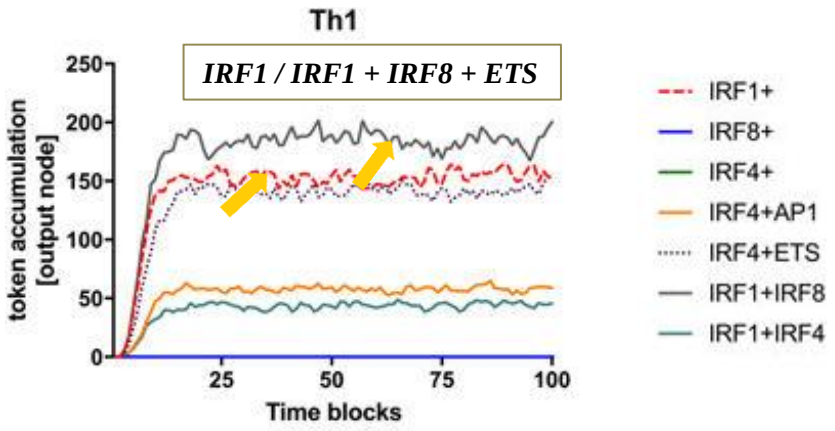
*Literature*

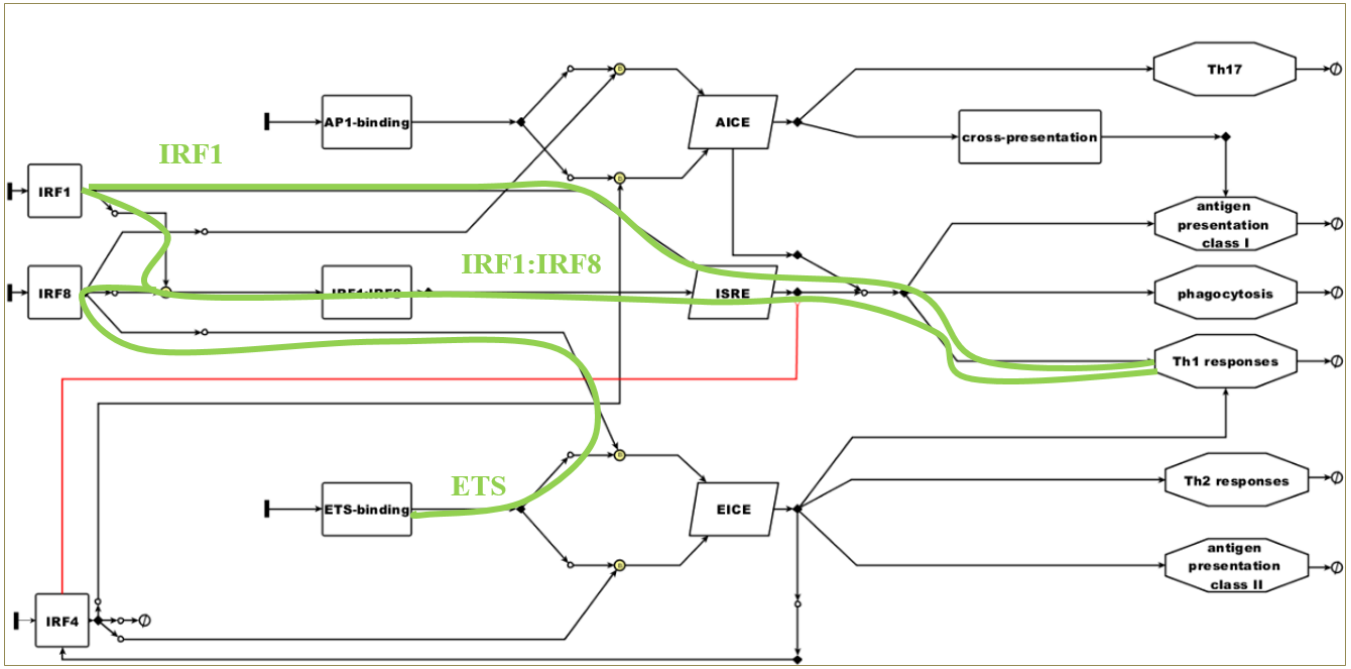



*Biological processes: Th17 responses, antigen presentation in class I, phagocytosis, Th1 responses, Th2 responses, antigen presentation in class II*

# Model validation Th1

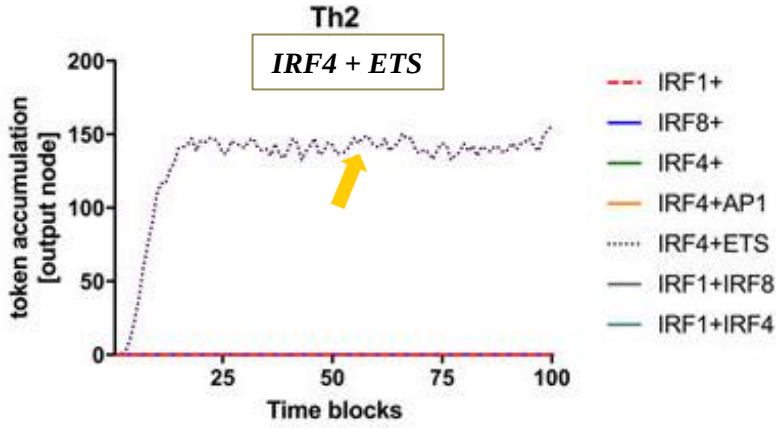
| Interaction partner 1 | GATE | Interaction partner 2 | interaction | binding site | outcome |
|---|---|---|---|---|---|
| IRF1 | and | IRF1 | induction | ISRE | TH1/CD8 |
| IRF1 | inhibition | IRF4 | inhibition | ISRE | TH1/CD8 |
| IRF1 | and | IRF8 | induction | ISRE | TH1/CD8 |
| IRF8 | and | ETS | induction | EICE | CD4 |

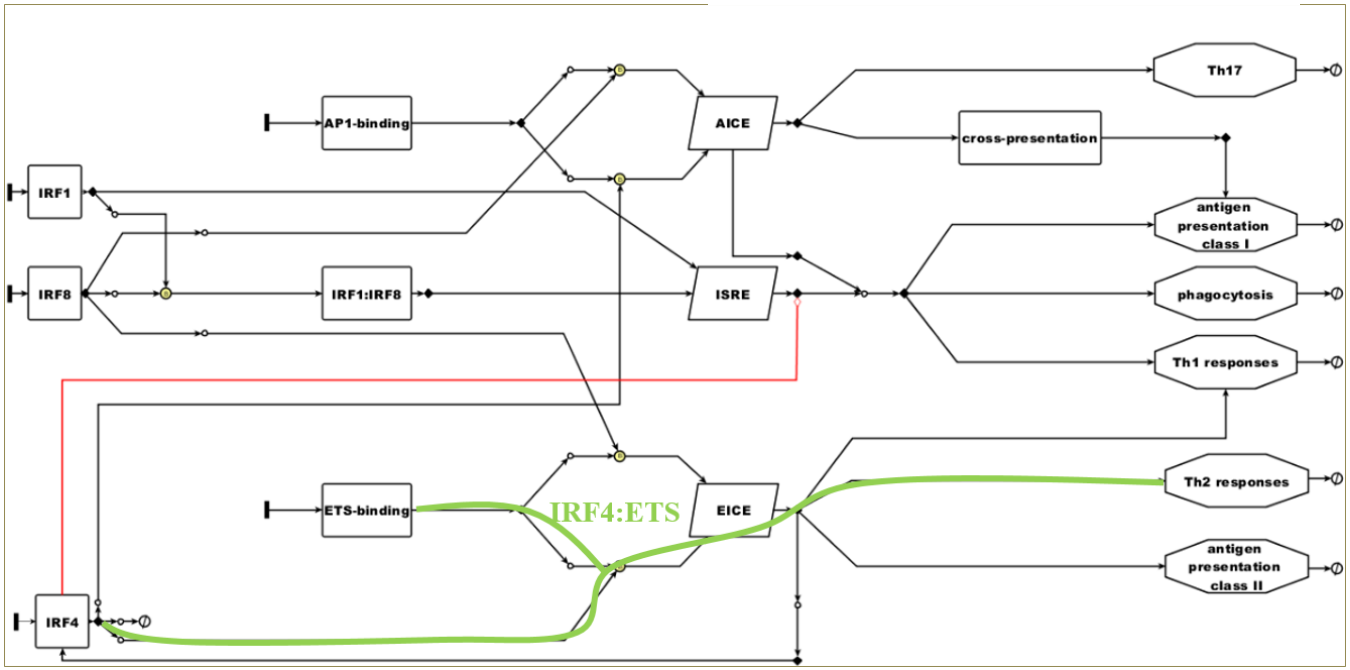


# Model validation Th2

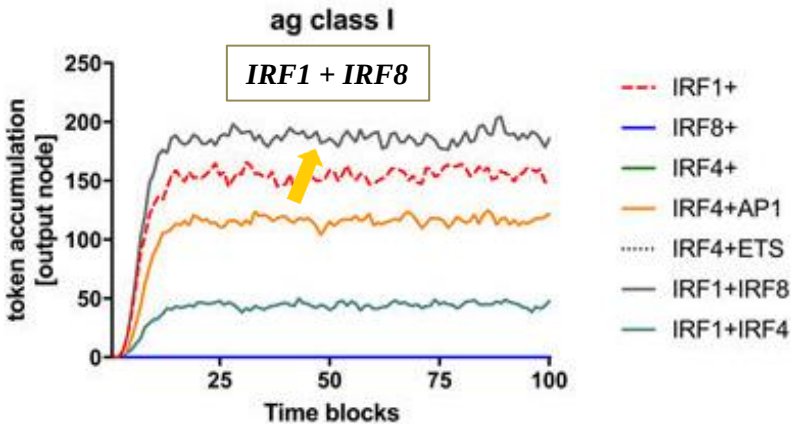
| Interaction partner 1 | GATE | Interaction partner 2 | interaction | binding site | outcome |
|---|---|---|---|---|---|
| IRF4 | and | ETS | induction | EICE | TH2 |

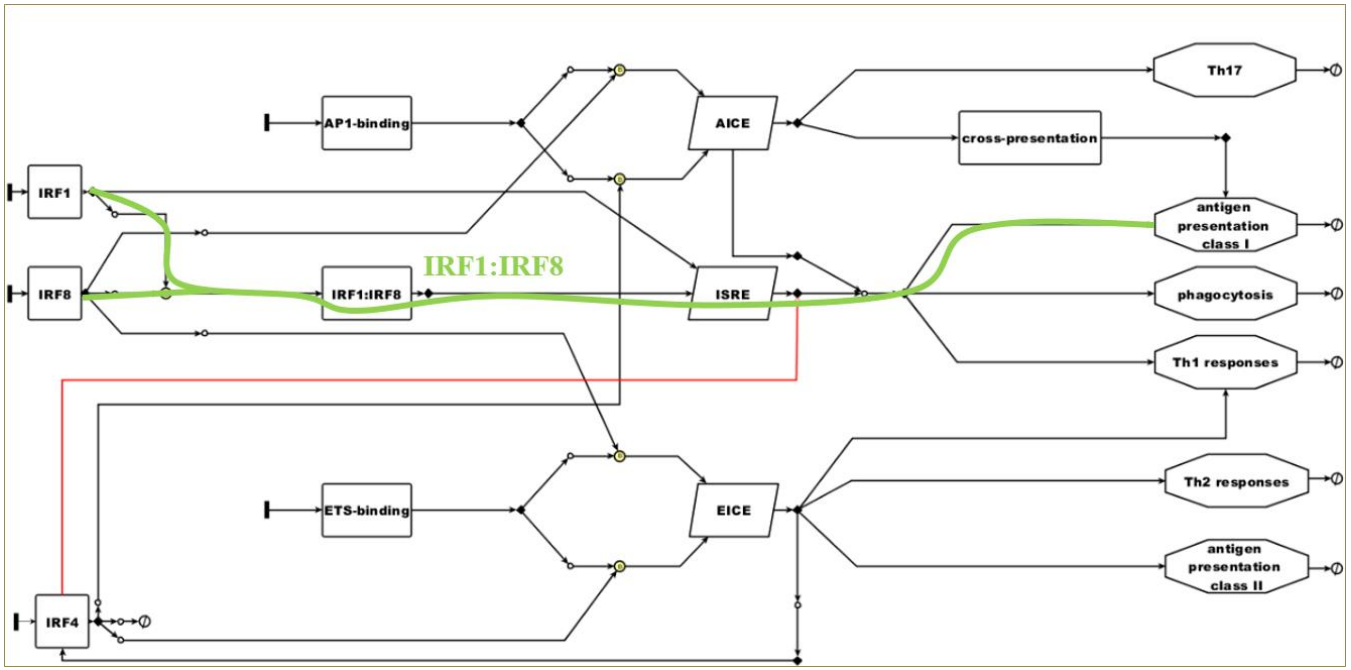


# Model validation ag class I

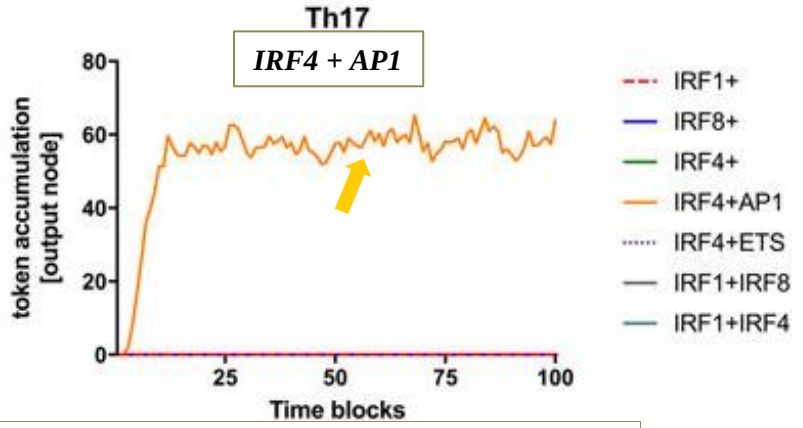
| Interaction partner 1 | GATE | Interaction partner 2 | interaction | binding site | outcome |
|---|---|---|---|---|---|
| IRF1 | and | IRF8 | induction | ISRE | TH1/CD8 |

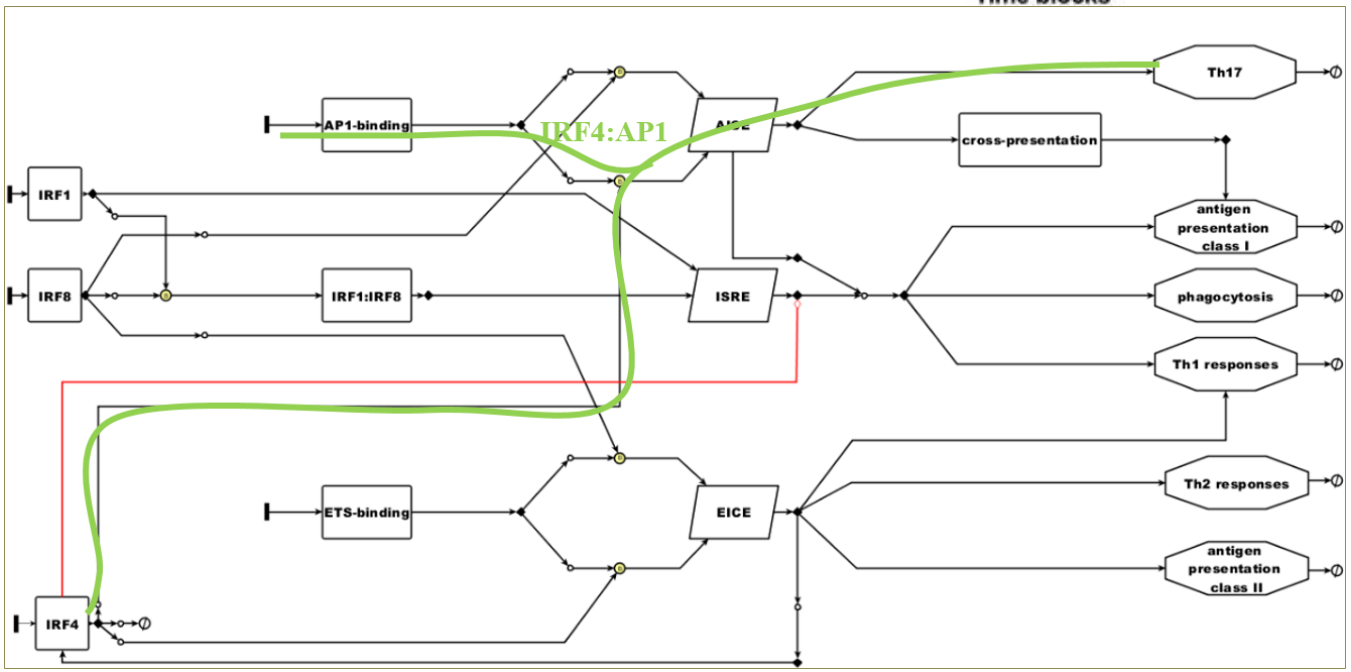


# Model validation Th17

| Interaction partner 1 | GATE | Interaction partner 2 | interaction | binding site | outcome |
|---|---|---|---|---|---|
| IRF4 | and | AP1 | induction | AICE | TH17 |

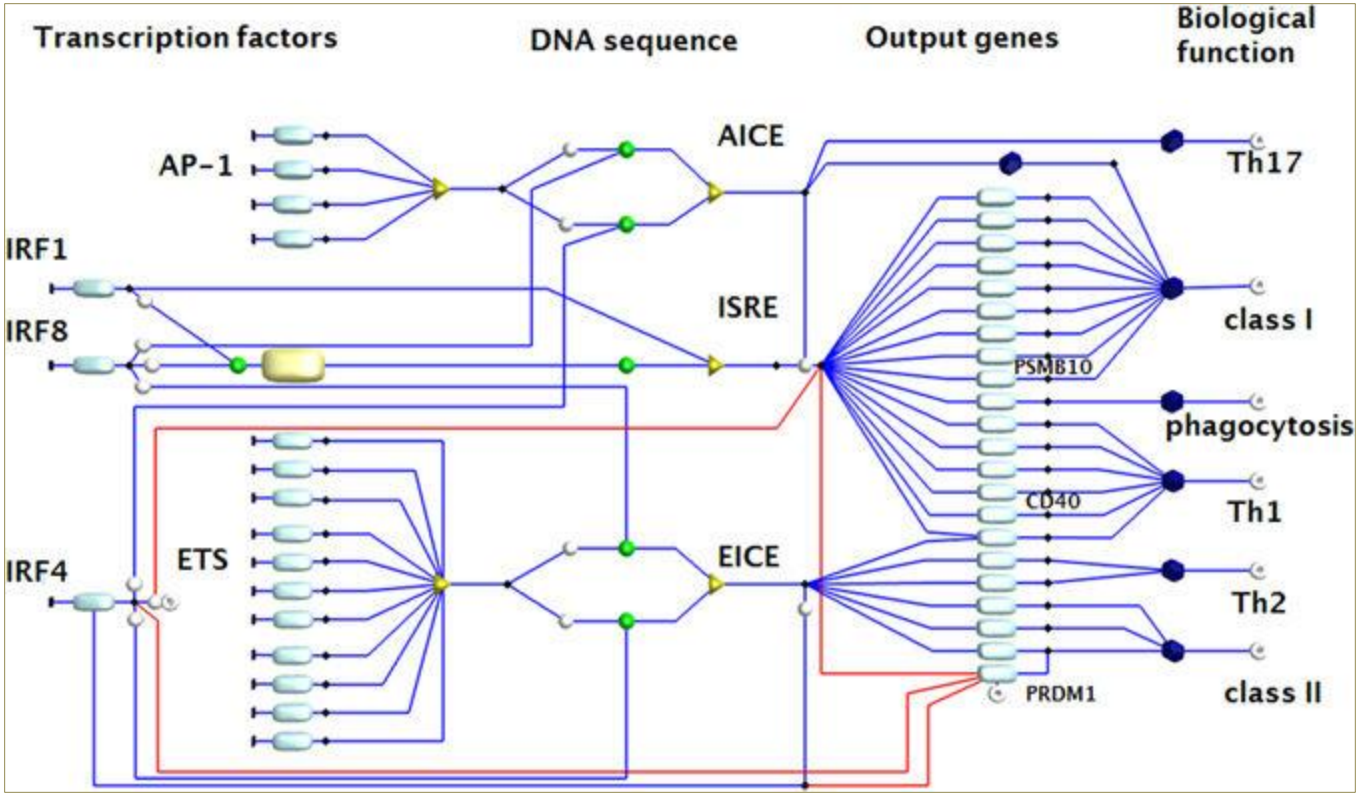


# Extended IRF GRN model parametrization and *in silico* simulations

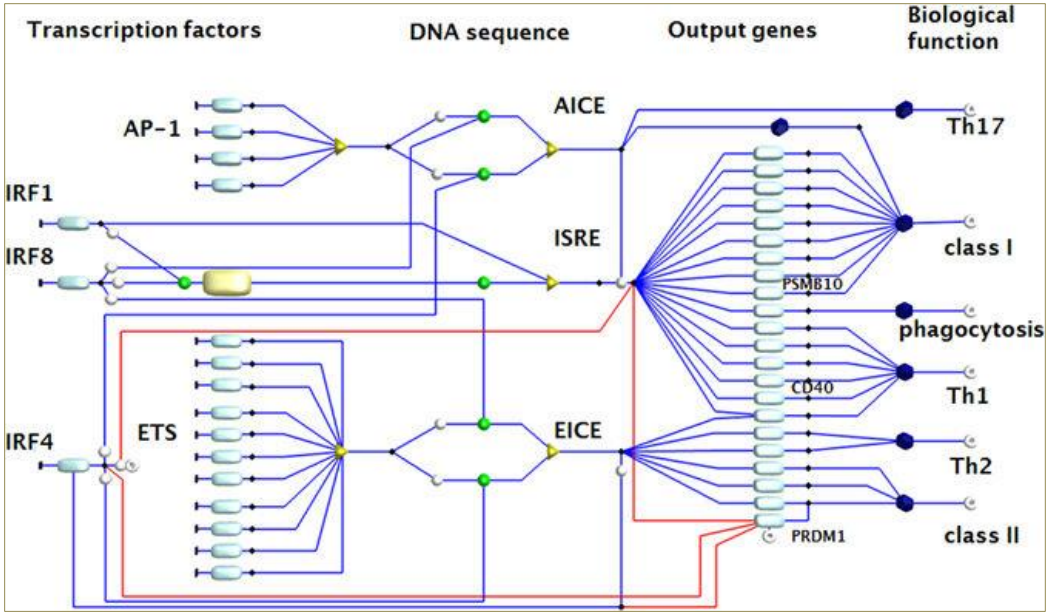
The modulation by epidermal cytokines of LC ability to activate antigen-specific CD8 T cell responses is predicted by *in silico* modelling of IRF-GRN parametrised with experimental data.

All members of ETS & AP1 expressed in human LC



# Extended IRF GRN model parametrization and *in silico* simulations

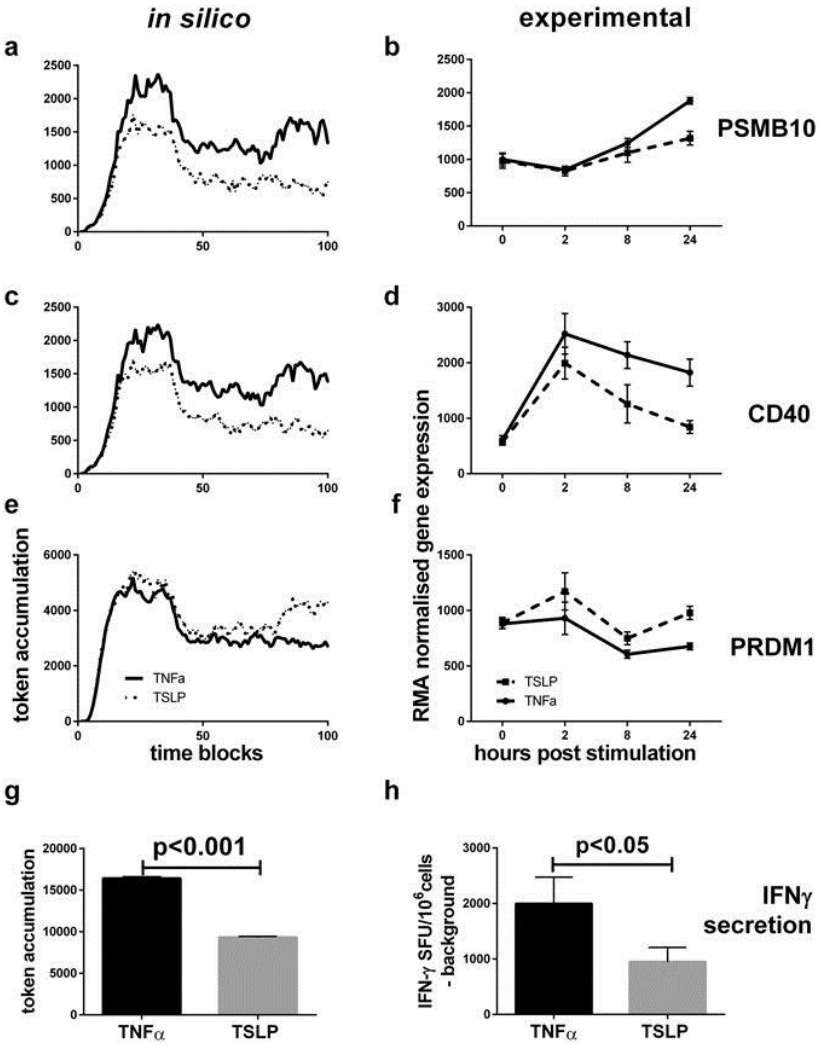
- No. of **tokens** in the network entry nodes: 0 /100
- SPN Stochastic Petri Nets  has been set as per the **levels of expression** from microarray data analysis
- **BioLayout** *Express*$^{3D}$, **100 time blocks**, **500 runs**.



# Simulation experiments extended IRF GRN

*In silico* vs *in vitro* results

Pattern of gene expressions in LC by TNFa & TSLP stimulation

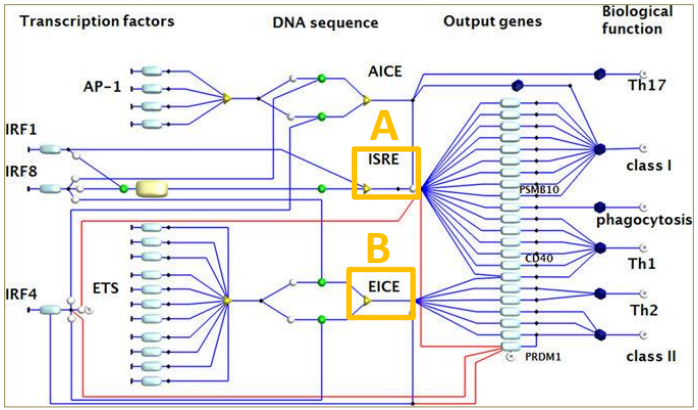


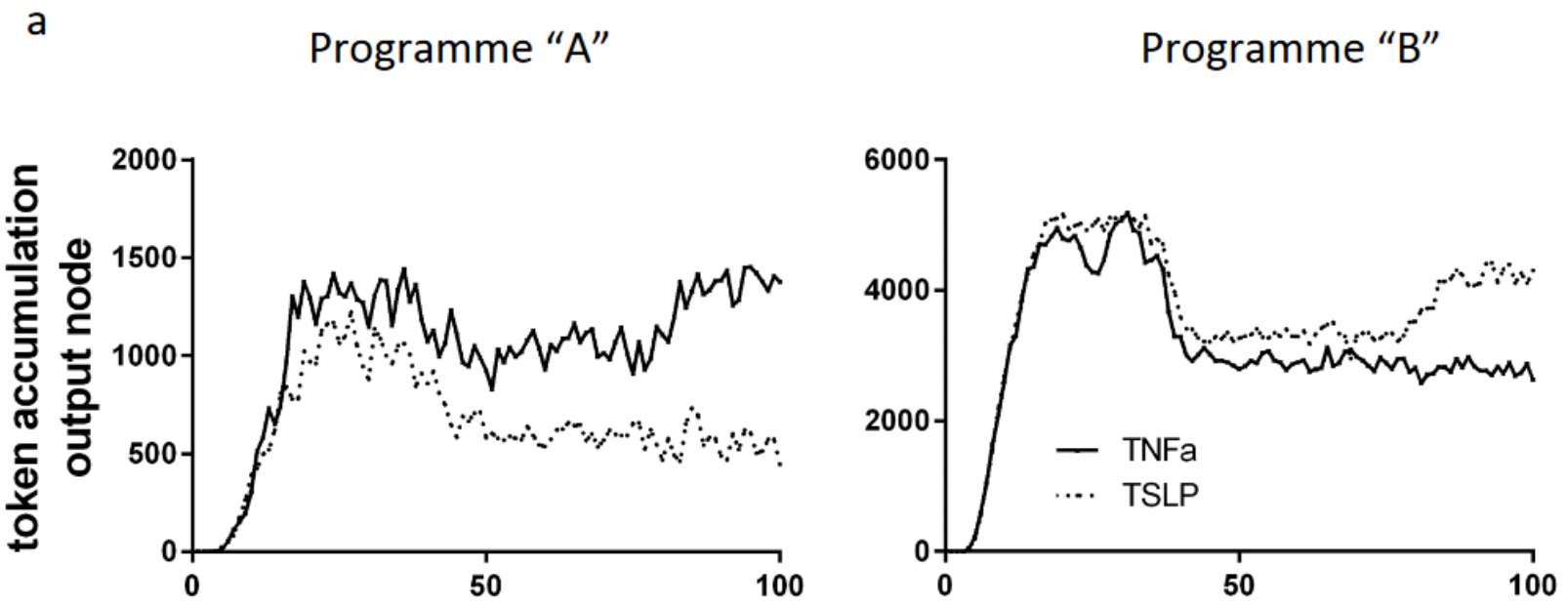
# *In silico* profiles of genes involved in programme “A” and “B”

34/50 genes

A : genes up-regulated by TNF-α

B : genes up-regulated by TNF-α and TSLP



# Summary

- Simulation experiments indicated that the ability of **LCs to present a peptide to CD8 T cells** would be **altered** by the **cytokine milieu** (TNFα/TSLP), which has not previously been reported and was not anticipate.
- *In silico* simulations performed after model parameterisation with transcription factor expression values predicted that human LC activation of antigen-specific CD8 T cells would be differentially regulated by epidermal cytokine induction of specific IRF-controlled pathways. This was **confirmed** by ***in vitro*** measurement of IFN-γ production by activated T cells.

# Summary

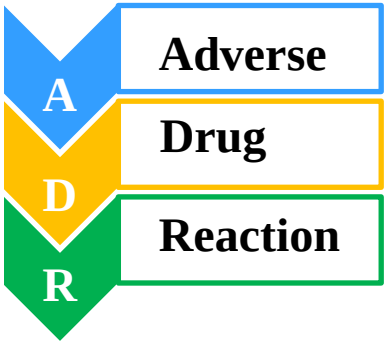
- The model demonstrate that computational modelling of a specific immune network can **predict functional outcomes** of immune responses based on experimentally data.
- Platform for many future studies of human immunity, utilising data from individual transcriptomic analyses to provide **predictions** of how **molecular interventions** may alter cellular phenotype based on the actual gene expression patterns in an individual.
- This can determine the outcome of immune responses in health and in disease, and offers the possibility of **predictive *in silico* testing** of the effectiveness of therapeutic interventions.


# Machine learning workflow to enhance predictions of Adverse Drug Reactions (ADRs) through drug-gene interactions: application to drugs for cutaneous diseases

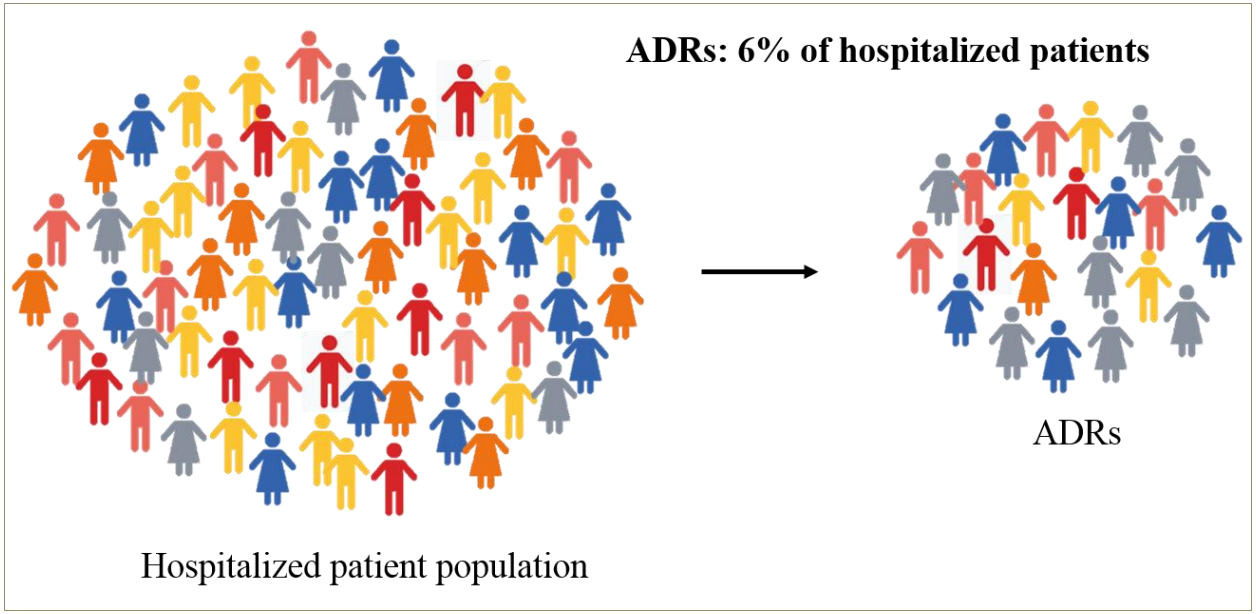
Kalpana Raja, Matthew Patrick, James T. Elder & Lam C. Tsoi

Scientific Reports. 2017;6.
doi:doi:10.1038/s41598-017-03914-3

# Introduction - Adverse drug reaction

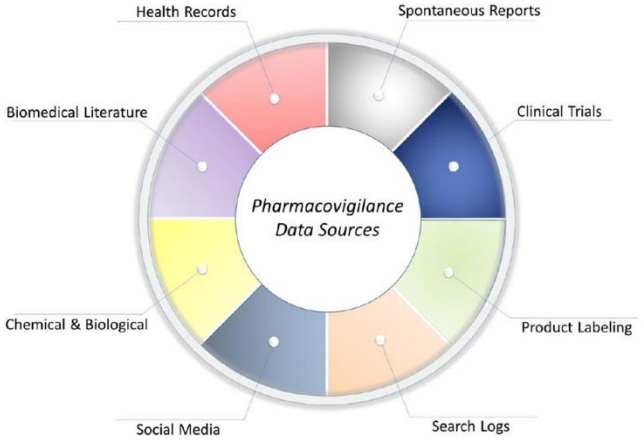


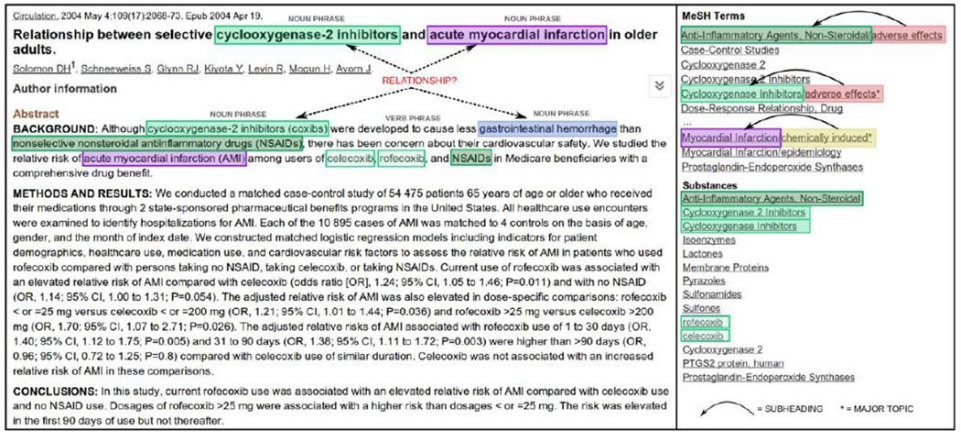

- ADRs may occur due to prolonged administration of a drug, or **combined** usage of two or more **drugs**
- ADR is the one **major reason** for **failure** in drug **clinical** trial



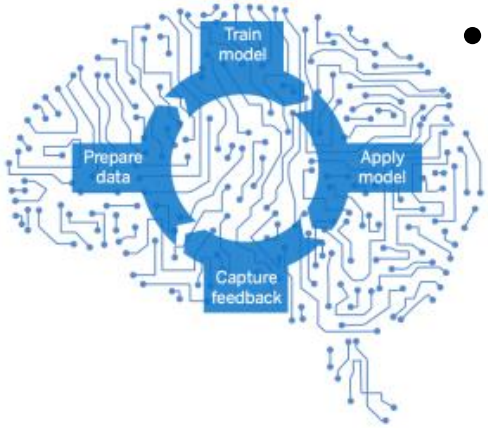
# Introduction – Biomedical sources

- A large amount of data (e.g., Medline, DrugBank, …) is available in public domain.
- It is not structured data stored in databases, but it is **free text**.
- Complexities and variability of natural language, and challenging to deal with algorithmically requires dedicated *computational* **text -mining** approaches.



# Machine learning



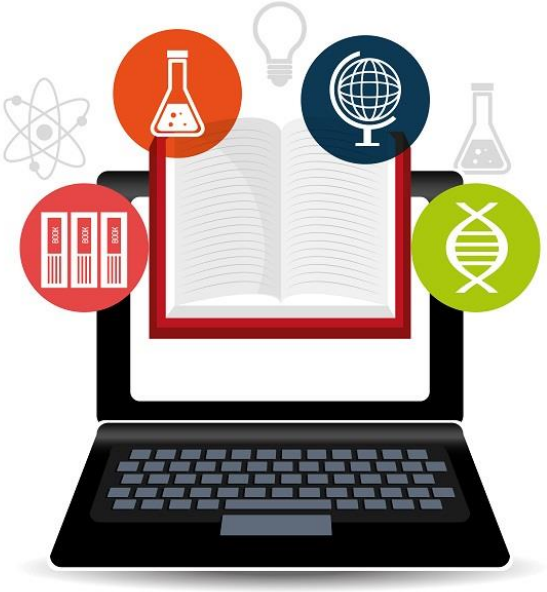

- In big data science, machine learning methods are computer algorithms that can automatically learn to recognize complex patterns based on empirical data.

- The goal of an machine learning method is to enable an algorithm to **learn from data of the past** or present and use that knowledge to **make predictions** or decisions for unknown future events.

# Introduction

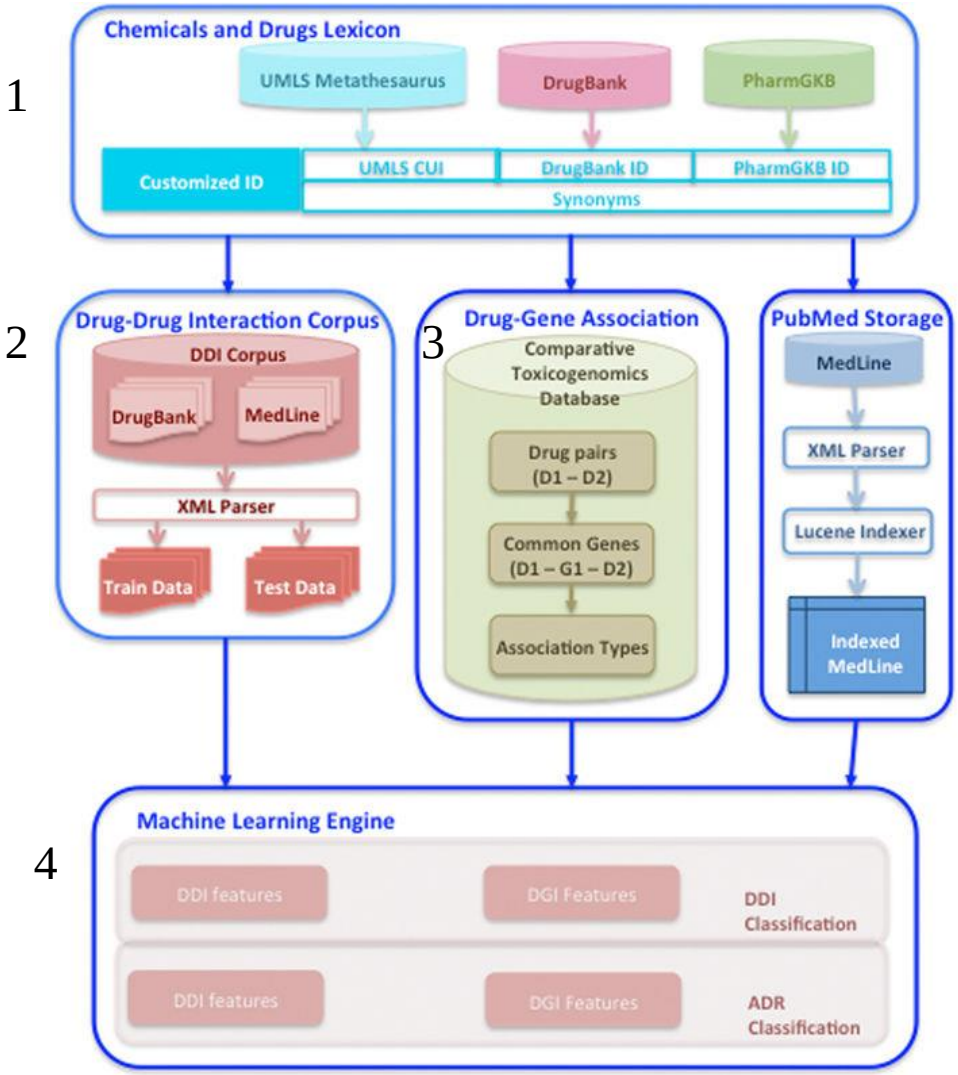
Aims:

- Develop novel and robust **literature-mining framework** for enhancing the **predictions of DDI-based ADRs by integrating DGIs.**
- Use **machine learning models** to learn syntactic and semantic information from the literature, and to claissfy of ADR types.



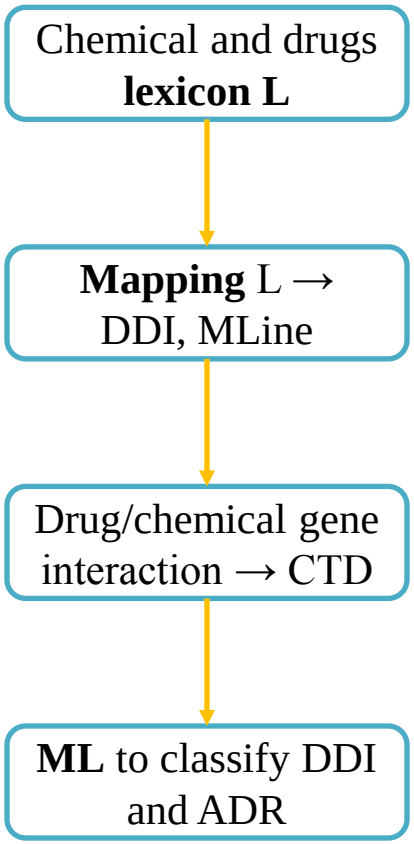
*ADR* – adverse drug reaction,
*DDI* – drug-drug interaction
*DGI* – drug-gene interaction

# System architecture



# Overview

- Predicting DDI-based ADR types consists of 4 steps




*L* - lexicon for chemicals and drugs (L)

*DDI* – drug-drug interaction

*MDLine* - MedLine abstracts

*CTD* - Comparative Toxicogenomics Database

*ML* - machine learning to classify literature sentences for DDIs and then categorize different ADR types.

# Language processing

***Corpus***
A corpus is a large collection of text used for accumulating statistics and linguistic analysis on natural language text.

***Lexicon***
A lexicon is a collection of information about the words of a language about the lexical categories to which they belong (dictionary).

# Chemicals and drugs lexicon

***L* - lexicon for chemicals and drugs**

*Sources*

*S1* 


→ concepts belonging to semantic groups **"Chemicals and Drugs"**

*S2* 


→ 8,203 drugs and 1,201 drug salts

*S3* 


→ 3,175 drugs

# DDI corpus & Medline abstract

DDI corpus

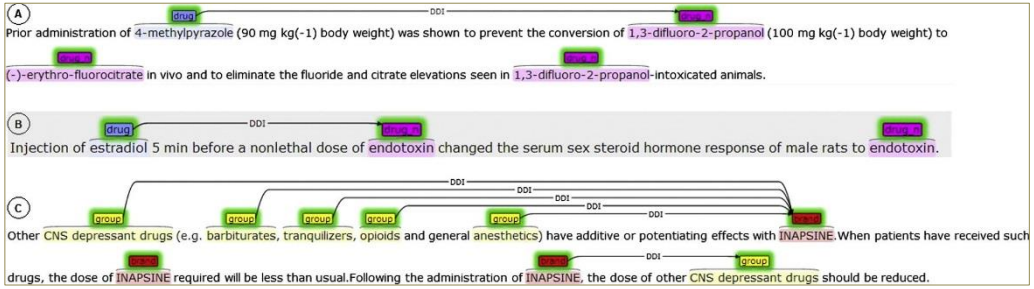
DRUGBANK

6,793 documents

**Medline docs**

PubMed.gov
US National Library of Medicine
National Institutes of Health
PubMed

1,701 documents




- **DDI annotations** (i.e. "True"/"False" between any two drugs
- "True" annotations include four DDI-derived ADR types

**4 DDI-derived ADR types**:

- adverse effect
- effect at molecular level
- effect related to pharmacokinetics
- drug interactions without known ADR

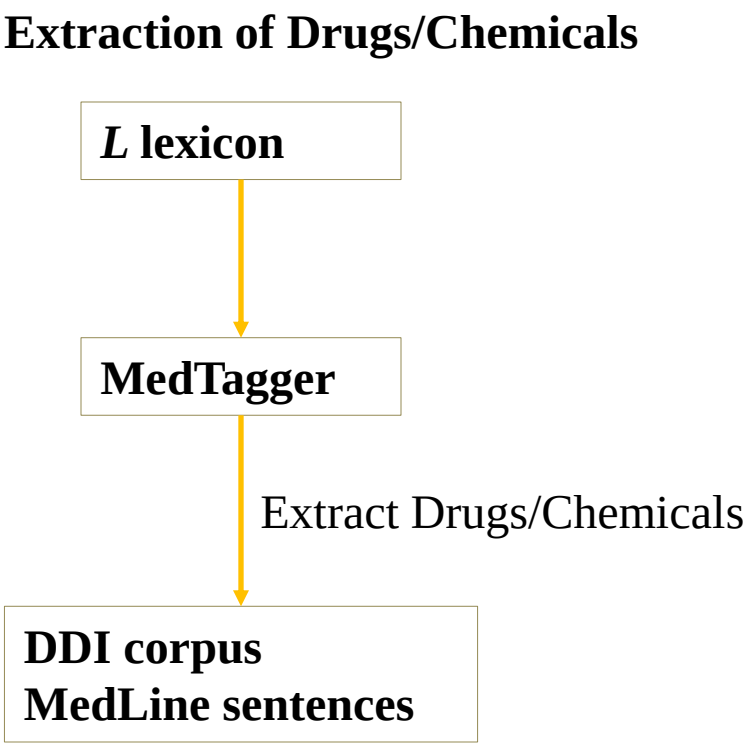
Medline

**Medline abstracts**

469,995 (i.e. >97%) citations → map 5 human genes

4,712,812 sentences from 469,995 MedLine abstracts

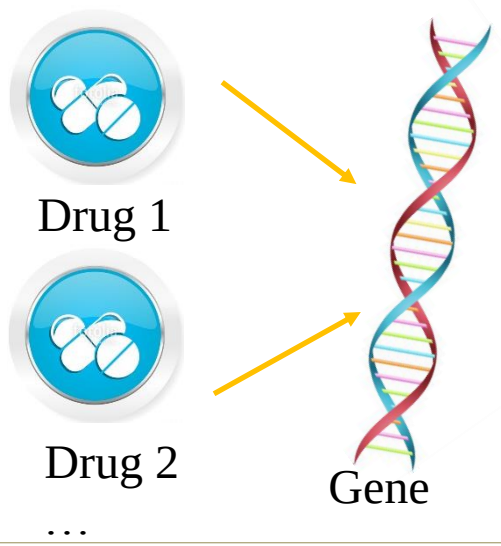
# Extraction of Drugs/Chemicals



# Extraction of Chemical/Drug-Gene interaction

*Information regarding DGIs can enhance the prediction of DDIs as well as ADR types classification by using ML approaches*

**DGI Drug-gene interaction**




*Sources:* **CTD, DGIdb**

- CTD: 500,000 DGIs from CTD, pertaining to 21,986 human genes
- Medline: 8,176 chemicals/drugs from 24,311

- Use lexicon to map DGIs in CTD
- gene(s) from the CTD database that interacts with both the drugs and retrieved the DGI associations

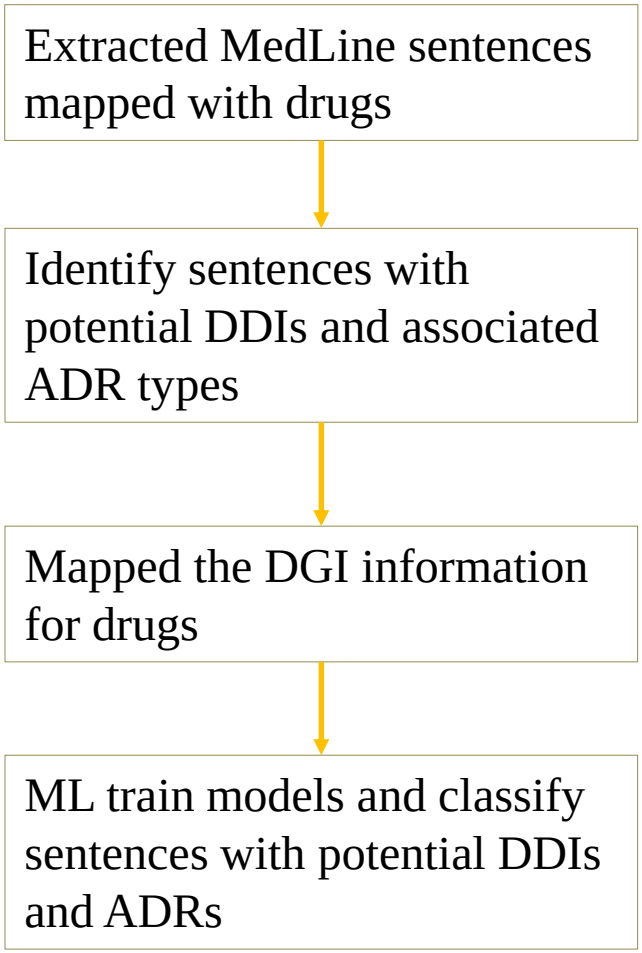
- DDI corpus: 193,294 DGIs for 5,773 drug pairs
- Medline sent: 49,188 DGIs for 935 drug pairs

**CTD** Comparative Toxicogenomics DB
**DGIdb** Drug-Gene Interaction DB

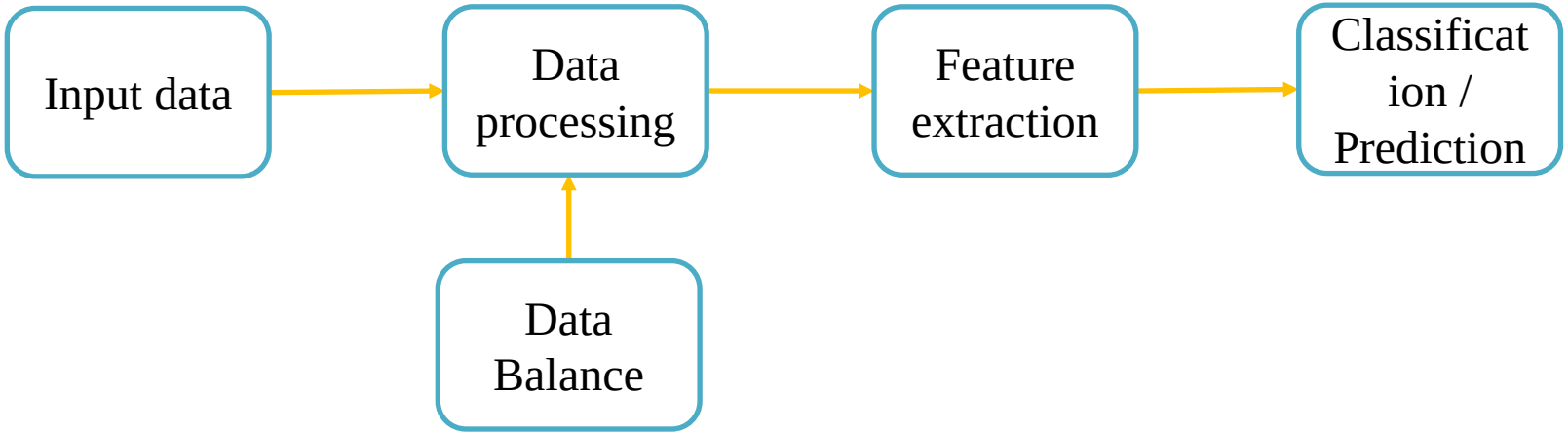
# Application on cutaneous diseases

*Identify medications for cutaneous diseases that might induce adverse reactions when taken together with other drugs*

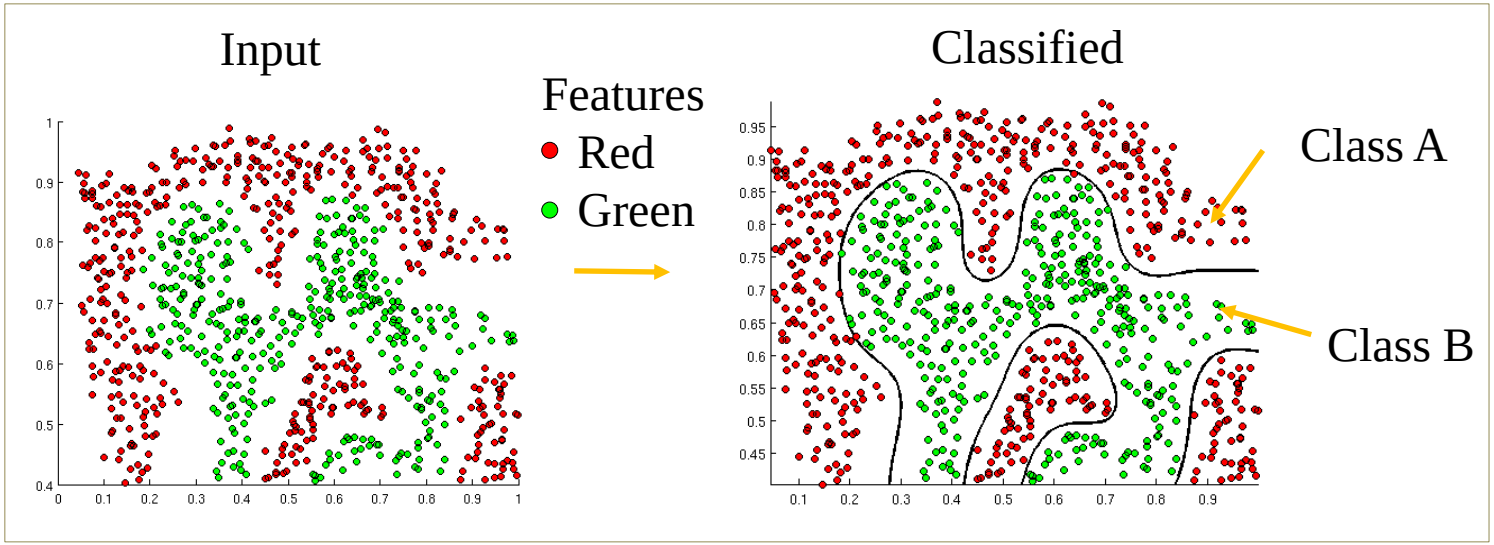
| Disease | Number of unique Drugs |
|---|---|
| 1. Psoriasis | 50 |
| 2. Atopic dermatitis | 25 |
| 3. Rosacea | 12 |
| 4. Acne vulgaris | 58 |
| 5. Alopecia | 3 |
| 6. Melanoma | 26 |
| 7. Eczema | 4 |
| 8. Keratosis | 6 |
| 9. Pruritus | 42 |



# Machine learning classification pipeline




2D classification example using *Support vector machine (SVM)*



# Classification

Step 1

**Features extraction**

- DDI-based features
- DGI-based feature

Step 2

**Data balance - SMOTE**

Annotations
3,767 true ADR
22,216 false ADR

Step 3

**Classification**

- DDIs
- ADR types

**ADR classes**

1. adverse effect
2. effects at molecular level
3. effects related to pharmacokinetics
4. drug interactions without known ADRs
5. others


**5 ML classifiers**

- Bayesian Network
- Decision tree
- Random tree
- Random forest
- K-nearest neighbors

*SMOTE - Synthetic Minority Oversampling Technique*

# Feature selection

**DDI**:

- DDI features + total no of words, drus min no of features before, between after each drug pair
- 24 features (after FR alg)

**DGI**:

- interaction genes + associated genes in CTD
- 20 features (after FR alg)

Feature reduction alg (FR alg): stepwise heuristic alg to identify **top significant features**

# DDI features

| | | Stepwise logistic regression model (p-value) | Mean impurity decrease | | | |
|---|---|---|---|---|---|---|
| | | | DDI classification | | ADR categorization | |
| | | | DDI Features | DDI+DGI Features | DDI Features | DDI+DGI Features |
| **DDI Features** | increase | 6.57e-14 | 0.24 | 0.21 | 0.12 | 0.14 |
| | effect (as negation) | 5.86e-12 | 0.21 | 0.17 | 0.00 | 0.00 |
| | patients | 8.38e-11 | 0.10 | 0.06 | 0.19 | 0.22 |
| | decrease | 3.57e-08 | 0.28 | 0.24 | 0.11 | 0.11 |
| | absorption | 5.90e-07 | 0.13 | 0.11 | 0.12 | 0.12 |
| | decreased | 1.15e-07 | 0.05 | 0.19 | 0.10 | 0.12 |
| | levels | 1.86e-06 | 0.20 | 0.14 | 0.19 | 0.20 |
| | auc | 7.95e-06 | 0.16 | 0.19 | 0.08 | 0.08 |
| | effects | 2.92e-05 | 0.10 | 0.06 | 0.16 | 0.16 |
| | metabolism | 1.65e-05 | 0.18 | 0.11 | 0.15 | 0.15 |
| | administration | 1.31e-05 | 0.19 | 0.20 | 0.15 | 0.15 |
| | enhance | 3.34e-05 | 0.07 | 0.15 | 0.09 | 0.11 |
| | significantly (as negation) | 4.61e-05 | 0.13 | 0.07 | 0.00 | 0.00 |
| | inhibited | 0.0201 | 0.12 | 0.14 | 0.04 | 0.05 |
| | increasing | 0.0046 | 0.12 | 0.13 | 0.02 | 0.30 |
| | antihypertensive | 0.0021 | 0.19 | 0.20 | 0.05 | 0.04 |
| | alter (as negation) | 0.0014 | 0.17 | 0.10 | 0.00 | 0.00 |
| | pressure | 0.0010 | 0.06 | 0.13 | 0.00 | 0.00 |
| | approximately | 0.0009 | 0.15 | 0.13 | 0.24 | 0.25 |
| | potentiate | 0.0005 | 0.18 | 0.16 | 0.00 | 0.00 |
| | resulted | 0.0004 | 0.16 | 0.02 | 0.05 | 0.08 |
| | monitored | 0.0003 | 0.15 | 0.19 | 0.11 | 0.11 |
| | administered | 0.0001 | 0.16 | 0.11 | 0.18 | 0.16 |
| | clearance | 0.0001 | 0.15 | 0.20 | 0.11 | 0.13 |
| **Additional features** | total words between drug pairs | - | 0.20 | 0.43 | 0.31 | 0.33 |
| | total drugs between drug pairs | - | 0.22 | 0.35 | 0.18 | 0.20 |
| | minimum number of features preceding drug pairs | - | 0.26 | 0.23 | - | - |
| | minimum number of features between drug pairs | - | 0.30 | 0.21 | - | - |
| | minimum number of features succeeding drug pairs | - | 0.31 | 0.26 | - | - |

# DGI features

| | | Stepwise logistic regression model (p-value) | Mean impurity decrease | | | |
|---|---|---|---|---|---|---|
| | | | DDI classification | | ADR categorization | |
| | | | DDI Features | DDI+DGI Features | DDI Features | DDI+DGI Features |
| **DGI Features** | acetylation : glutathionylation | - | - | 0.30 | - | 0.22 |
| | chemical synthesis : hydrolysis | - | - | 0.07 | - | 0.21 |
| | expression : hydroxylation | - | - | 0.25 | - | 0.14 |
| | expression : glucuronidation | - | - | 0.17 | - | 0.14 |
| | activity : oxidation | - | - | 0.18 | - | 0.14 |
| | binding : response to substance | - | - | 0.24 | - | 0.12 |
| | hydroxylation : hydroxylation | - | - | 0.20 | - | 0.11 |
| | oxidation : response to substance | - | - | 0.20 | - | 0.09 |
| | activity : chemical synthesis | - | - | 0.21 | - | 0.09 |
| | expression : splicing | - | - | 0.08 | - | 0.09 |
| | expression : stability | - | - | 0.15 | - | 0.07 |
| | acetylation : response to substance | - | - | 0.15 | - | 0.07 |
| | import : transport | - | - | 0.12 | - | 0.06 |
| | glutathionylation : response to substance | - | - | 0.12 | - | 0.06 |
| | degradation : methylation | - | - | 0.11 | - | 0.06 |
| | localization : phosphorylation | - | - | 0.16 | - | 0.03 |
| | binding : methylation | - | - | 0.09 | - | 0.03 |
| | activity : mutagenesis | - | - | 0.06 | - | 0.03 |
| | sulfation : sulfation | - | - | 0.15 | - | 0.02 |
| | oxidation : oxidation | - | - | 0.12 | - | 0.02 |

# Performance of lexicon on drug extraction

Performance of using the chemicals and drugs lexicon on identifying the drugs present in DDI corpus.

| Dataset | | True positive | False positive | False negative | FP1 | Precision | Recall | F-score | P1 | F1 |
|---|---|---|---|---|---|---|---|---|---|---|
| Training (Cross validation) | DrugBank | 11,051 | 2,060 | 932 | 373 | 0.84 | 0.92 | 0.88 | 0.97 | 0.94 |
| | MedLine | 1,372 | 484 | 335 | 6 | 0.74 | 0.80 | 0.77 | 1.00 | 0.89 |
| | Overall | 12,423 | 2,544 | 1,267 | 379 | 0.83 | 0.91 | 0.87 | 0.97 | 0.94 |
| Test | DrugBank | 279 | 61 | 17 | 46 | 0.82 | 0.94 | 0.88 | 0.86 | 0.90 |
| | MedLine | 288 | 191 | 58 | 34 | 0.60 | 0.83 | 0.70 | 0.89 | 0.86 |
| | Overall | 567 | 252 | 75 | 80 | 0.69 | 0.88 | 0.78 | 0.88 | 0.88 |

# ML workflow on DDI/ADR types classification

**DDI** Prediction comparison on **DDI corpus** training data

| | DDI Features | | | DDI and DGI Features | | | DGI Features | | |
|---|---|---|---|---|---|---|---|---|---|
| **Classifier** | **Precision** | **Recall** | **F-score** | **Precision** | **Recall** | **F-score** | **Precision** | **Recall** | **F-score** |
| Bayesian network | 0.93 | 0.69 | 0.79 | 0.93 | 0.69 | 0.79 | 0.54 | 1.00 | 0.71 |
| Decision tree | 0.98 | 0.63 | 0.76 | 0.83 | 0.72 | 0.77 | 0.62 | 0.61 | 0.62 |
| Random tree | 0.76 | 0.77 | 0.76 | 0.79 | 0.77 | 0.78 | 0.69 | 0.71 | 0.70 |
| Random forest | 0.82 | 0.78 | 0.80 | 0.84 | 0.78 | 0.81 | 0.70 | 0.71 | 0.70 |
| K-nearest neighbors | 0.76 | 0.73 | 0.74 | 0.76 | 0.77 | 0.76 | 0.69 | 0.73 | 0.71 |

# ML workflow on DDI/ADR types classification

Performance of classification on **ADR** types using **DDI** features on **DDI corpus** training data.

| Classifier | ADR Type | Precision | Recall | F-score | Average Precision | Average Recall | Macro Average F-score |
|---|---|---|---|---|---|---|---|
| Bayesian network | Adverse effect | 0.73 | 0.76 | 0.74 | 0.71 | 0.67 | 0.69 |
| | Effect at molecular level | 0.79 | 0.52 | 0.62 | | | |
| | Effect related to pharmacokinetics | 0.61 | 0.47 | 0.53 | | | |
| | Drug interaction without known ADR | 0.72 | 0.70 | 0.71 | | | |
| Decision treeRandom tree | Adverse effect | 0.82 | 0.95 | 0.88 | 0.87 | 0.86 | 0.86 |
| | Effect at molecular level | 0.87 | 0.85 | 0.86 | | | |
| | Effect related to pharmacokinetics | 0.82 | 0.77 | 0.79 | | | |
| | Drug interaction without known ADR | 0.92 | 0.88 | 0.90 | | | |
| | Adverse effect | 0.83 | 0.94 | 0.88 | 0.87 | 0.85 | 0.86 |
| | Effect at molecular level | 0.86 | 0.85 | 0.85 | | | |
| | Effect related to pharmacokinetics | 0.81 | 0.77 | 0.79 | | | |
| | Drug interaction without known ADR | 0.93 | 0.85 | 0.89 | | | |
| Random forest | Adverse effect | 0.84 | 0.95 | 0.89 | 0.88 | 0.86 | 0.87 |
| | Effect at molecular level | 0.88 | 0.86 | 0.87 | | | |
| | Effect related to pharmacokinetics | 0.84 | 0.78 | 0.81 | | | |
| | Drug interaction without known ADR | 0.94 | 0.86 | 0.90 | | | |
| K-nearest neighbors | Adverse effect | 0.83 | 0.95 | 0.88 | 0.87 | 0.85 | 0.86 |
| | Effect at molecular level | 0.86 | 0.84 | 0.85 | | | |
| | Effect related to pharmacokinetics | 0.81 | 0.76 | 0.79 | | | |
| | Drug interaction without known ADR | 0.93 | 0.85 | 0.89 | | | |

# ML workflow on DDI/ADR types classification

Performance of classification on **ADR** types using **DDI and DGI** features on **DDI corpus** training data.

**DDI + DGI** → 90%

| Classifier | ADR Type | Precision | Recall | F-score | Average Precision | Average Recall | Macro Average F-score |
|---|---|---|---|---|---|---|---|
| Bayesian network | Adverse effect | 0.76 | 0.83 | 0.79 | 0.75 | 0.71 | 0.73 |
| | Effect at molecular level | 0.83 | 0.59 | 0.69 | | | |
| | Effect related to pharmacokinetics | 0.67 | 0.47 | 0.56 | | | |
| | Drug interaction without known ADR | 0.74 | 0.72 | 0.73 | | | |
| Decision tree | Adverse effect | 0.85 | 0.96 | 0.90 | 0.89 | 0.88 | 0.89 |
| | Effect at molecular level | 0.94 | 0.87 | 0.90 | | | |
| | Effect related to pharmacokinetics | 0.85 | 0.81 | 0.83 | | | |
| | Drug interaction without known ADR | 0.91 | 0.92 | 0.91 | | | |
| Random tree | Adverse effect | 0.86 | 0.94 | 0.90 | 0.88 | 0.88 | 0.88 |
| | Effect at molecular level | 0.91 | 0.88 | 0.90 | | | |
| | Effect related to pharmacokinetics | 0.83 | 0.79 | 0.81 | | | |
| | Drug interaction without known ADR | 0.91 | 0.90 | 0.91 | | | |
| Random forest | Adverse effect | 0.87 | 0.95 | 0.91 | 0.90 | 0.89 | 0.90 |
| | Effect at molecular level | 0.93 | 0.89 | 0.91 | | | |
| | Effect related to pharmacokinetics | 0.86 | 0.82 | 0.84 | | | |
| | Drug interaction without known ADR | 0.92 | 0.91 | 0.92 | | | |
| K-nearest neighbors | Adverse effect | 0.86 | 0.94 | 0.90 | 0.89 | 0.88 | 0.88 |
| | Effect at molecular level | 0.91 | 0.89 | 0.90 | | | |
| | Effect related to pharmacokinetics | 0.83 | 0.79 | 0.81 | | | |
| | Drug interaction without known ADR | 0.91 | 0.90 | 0.91 | | | |

# Performance comparison with competing methods

Performance **comparison** with the **existing** systems on **DDI corpus** test data

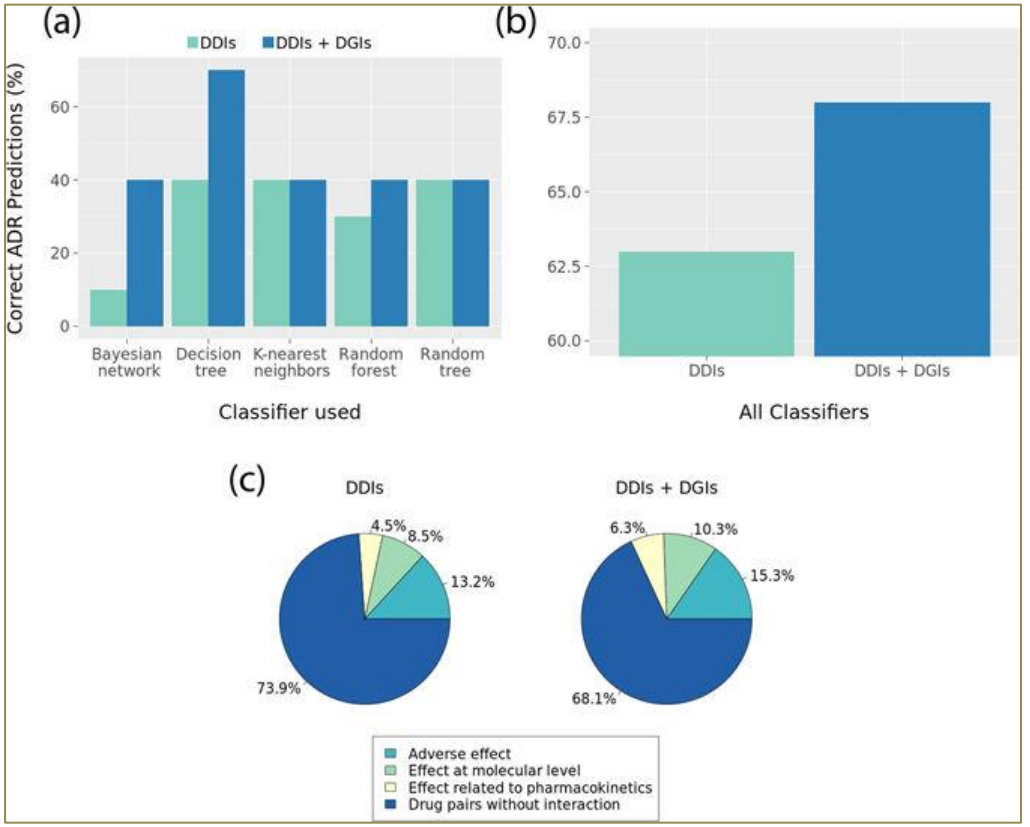
| System | Description | Classifier | DDI classification | | | ADR categorization | | |
|---|---|---|---|---|---|---|---|---|
| | | | P | R | F | P | R | F |
| Our approach | DDI features | Random forest | 0.739 | 0.823 | 0.779 | 0.761 | 0.793 | 0.755 |
| | DDI + DGI features | | 0.875 | 0.790 | 0.831 | 0.839 | 0.761 | 0.798 |
| FBK-irst system | Contextual and shallow linguistic features | Support vector machines | 0.794 | 0.806 | 0.800 | 0.633 | 0.642 | 0.638 |
| WBI system | Ensembles of five different classifiers | Shallow linguistic kernel + a self-developed feature based classifier + Turku event extraction system | 0.801 | 0.722 | 0.759 | 0.642 | 0.579 | 0.609 |
| Uturku system | Deep syntactic features and information from external domain resources | Turku event extraction system | 0.833 | 0.602 | 0.699 | 0.732 | 0.499 | 0.594 |

# ADR predictions for cutaneous diseases

a) Performance of classifiers to predict DDIs and ADR types;

a) Prediction of DDI and ADR types at least by three classifiers;

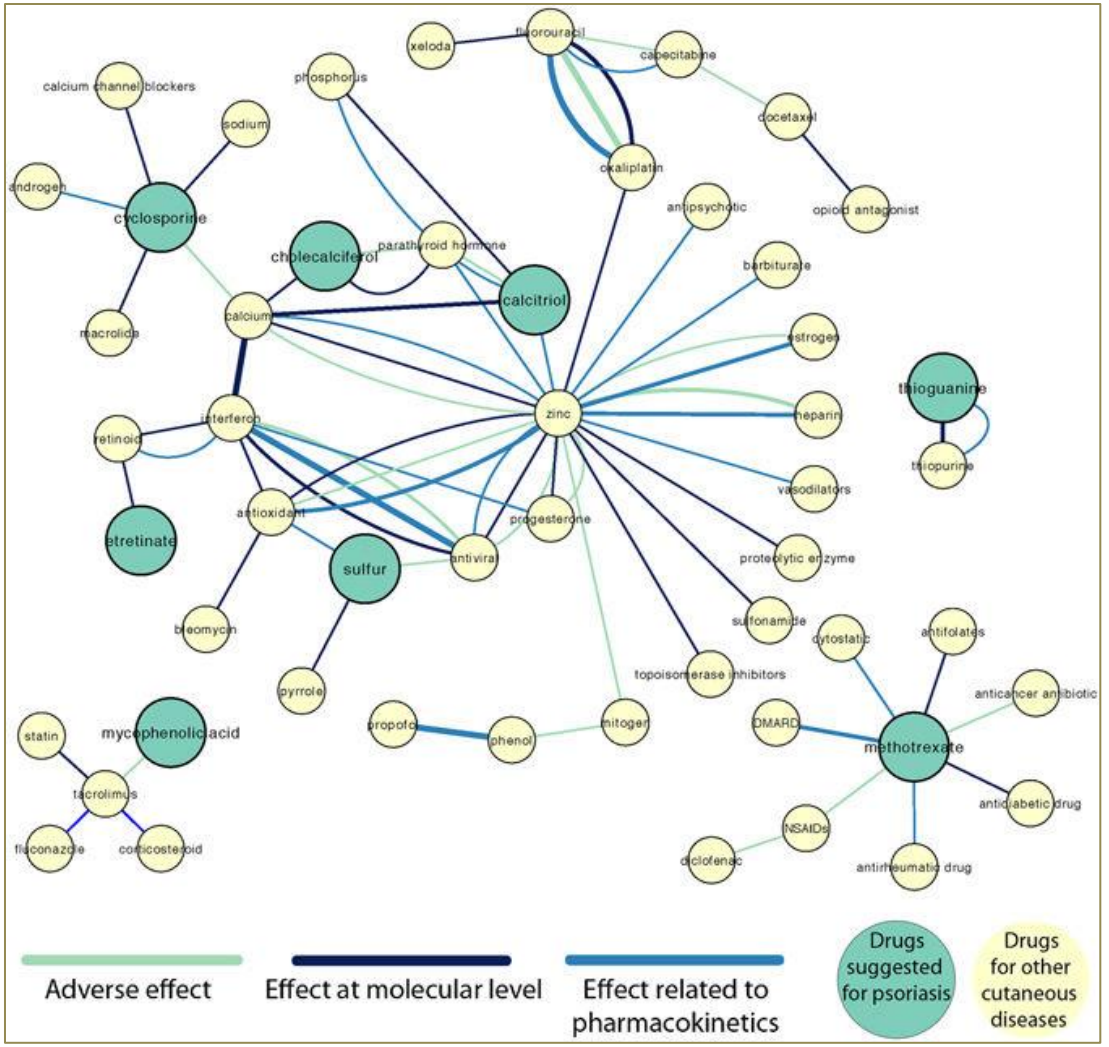
a) Performance of random forest classifier to predict DDIs and ADR types between NDFRT drugs suggested for cutaneous diseases and drugs using DDI features alone and DDI with DGI features.



# Case study: ADR predictions related to psoriasis

**DDIs and ADR types predicted for psoriasis**

ADR network for cutaneous diseases showing interaction between NDFRT drugs suggested for cutaneous diseases and drugs. Thickness of the edges correlate with the number of instances to support the ADR predictions.



# Case study: ADR predictions related to psoriasis. Validation.

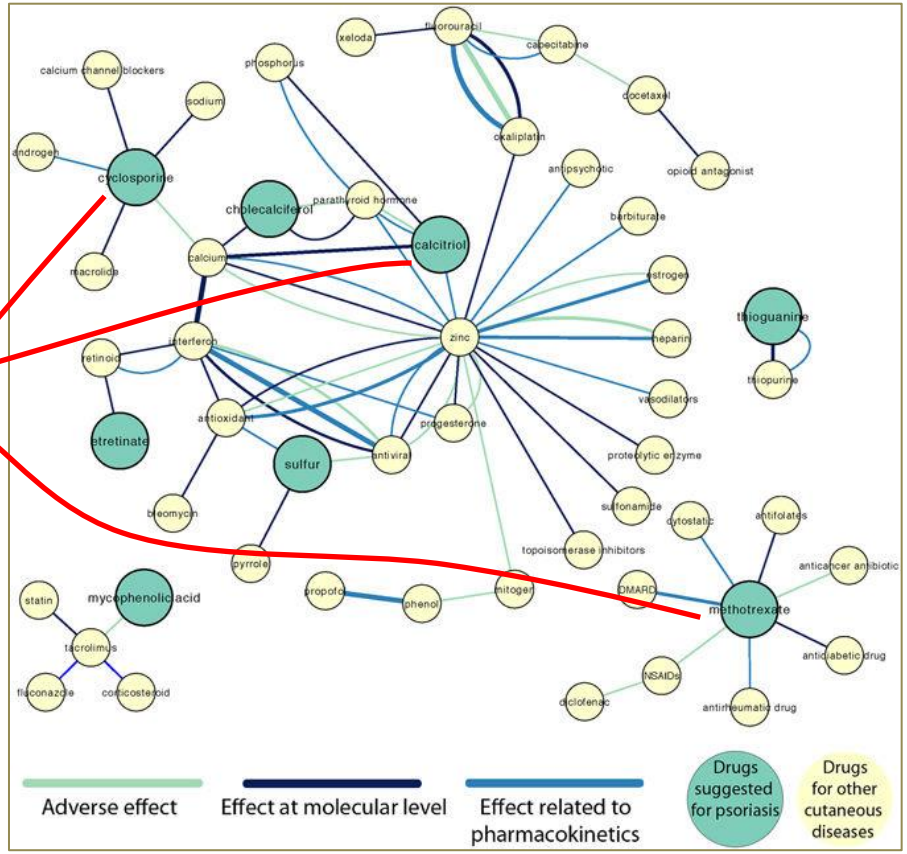
PubMed Sentences with ADR information, predicted by machine learning workflow.

"Simultaneous use of nonsteroidal anti-inflammatory drugs **NSAIDs** probenecid and other drugs has been reported to delay the plasma elimination of **methotrexate** in patients".[32]

"The decreased **parathyroid hormone** levels would then also contribute to a decrease in **calcitriol** synthesis".[33]

"Our findings show that FKBP51 and Cyp40 are positive regulators of androgen receptor that can be selectively targeted by **cyclosporine** A and **FK506** to achieve inhibition of **androgen** induced cell proliferation".[34]

"Albeit its great benefits as immunosuppressant, the use of **Cyclosporine** A has been limited by undesirable nephrotoxic effects, including **sodium** retention, hypertension, hyperkalemia, interstial fibrosis and progressive renal failure in transplant recipients".[28]



# Case study: ADR predictions related to psoriasis

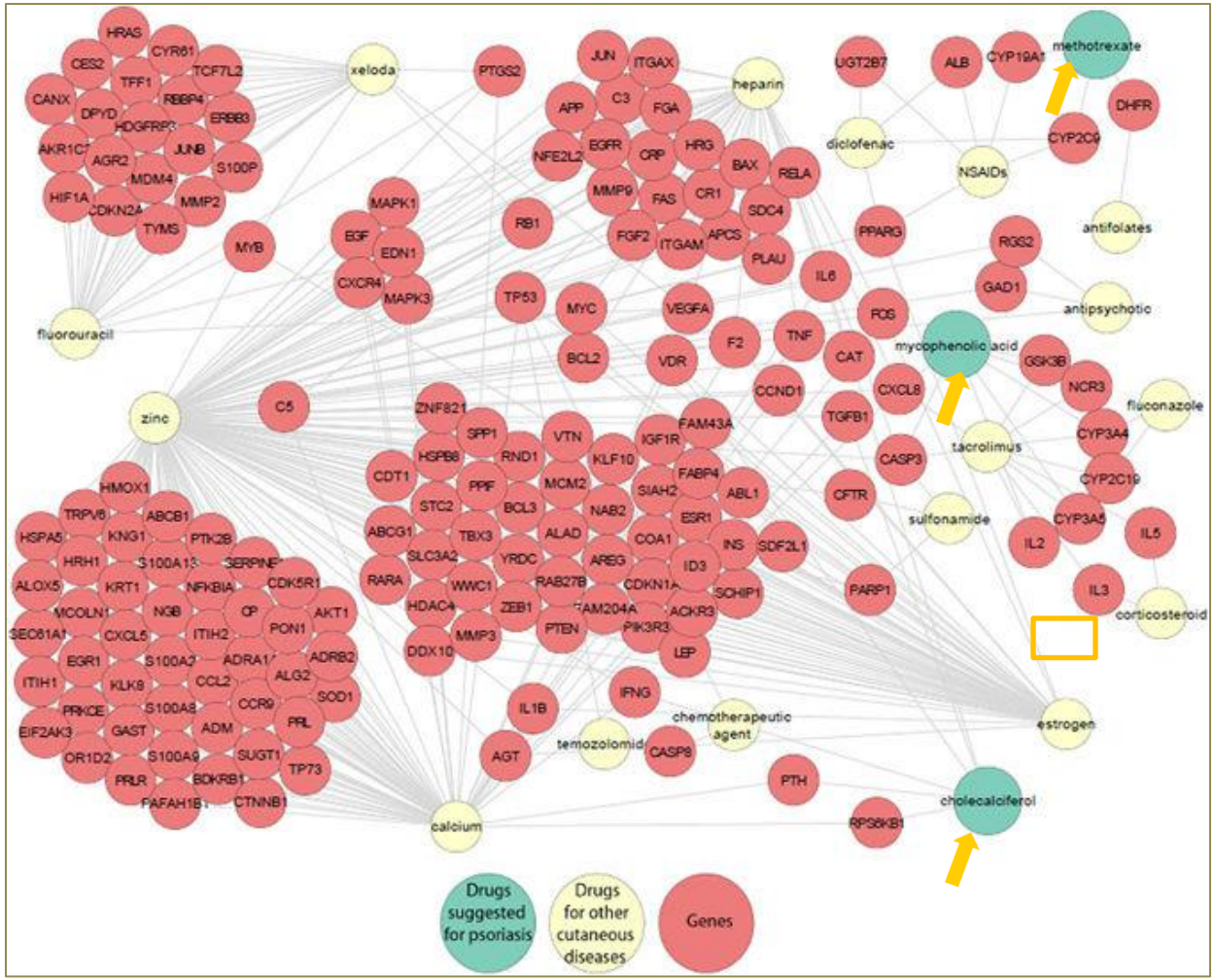
**Genes in ADR prediction**

Gene - DDI network for cutaneous diseases showing interaction between NDFRT drugs suggested for cutaneous diseases/drugs with genes.

177 ADRs predicted for psoriasis medications,

- 31 DDIs are associated with common gene
  - Drug suggested for psoriasis
    - methotrexate
    - cholecalciferol
    - mycophenolic

*not in the same sentence!*



# Case study: ADR predictions related to psoriasis

**Disease comorbidity through DDIs**

**Diseases associated with DDI pairs** from various sources such as

- **NDFRT**
- **DrugBank**
- **UpToDate**
- **CDC**
- **Mayo clinic's** Diseases and Conditions

**Diseases related to ADR prediction**

Crohn's disease
rickets in children
osteomalacia in adults
breast cancer in females
lupus erythematosus
Eczema
hypocalcemia, diabetes, atopic dermatitis, blood pressure, influenza, Raynaud's disease, melanoma and bacterial conjuctivities

# Summary

- Provide an automated approach to predict in advance **medication-related** to **DDIs** and **ADR.**
- Present a workflow that integrates **ML** with **biomedical literature** data to **data-mine** potential **drug-drug interactions for cutaneous diseases.**
- Successfully predict previously known ADRs for drugs prescribed to cutaneous diseases, and are also able to identify promising new ADRs.
- Conducted an intense analysis on DDIs and ADR types related to psoriasis.
- Extend the finding to identify **comorbid diseases related** to **cutaneous** diseases.

# Software for systems biology: from tools to integrated platforms

2011 and before

| | Tools | | Standards | | | Projects |
|---|---|---|---|---|---|---|
| | Software | Resources | Ontologies | File format | Minimum information | |
| **Data and knowledge management** | MAGE-TAB, ISA-TAB, KNIME, caGrid, Taverna, Bio-STEER | BioCatalogue | SBO, OBO, NCBO | MGED (MAGE), PSI, MSI | MIAME, MIAPE, MIBBI, ISO MDR, DCMI | |
| **Data-driven network inference** | R, MATLAB, BANJO | | | | | DREAM Initiative, Sage Bionetworks |
| **Deep curation** | CellDesigner, EPE, Jdesigner, PathVISIO | KEGG, Reactome, Panther pathway database, BioModels.net, WikiPathways | | SBML, SBGN, CellML, BioPAX, PSI-MI | MIRIAM | |
| ***In silico* simulation** | COPASI, SBW, JSim, Neuron, GENESIS, MATLAB, ANSYS, FreeFEM, ePNK, ina, WoPeD, Petri nets, OpenCell, CellDesigner + COPASI, CellDesigner + SOSlib, PhysioDesigner (formerly *insilico*IDE) | | | SED-ML, SBRML, PNML, SBML | MIASE | |
| **Model analysis** | MATLAB, Auto, XPPAut, BUNKI, ManLab, ByoDyn, SenSB, COBRA, MetNetMaker, DBSolve Optimum, Kintecus, NetBuilder, BooleanNet, SimBoolNet | | | | | |
| **Physiological modelling** | JSim, PhysioDesigner (formerly *insilico*IDE), CellDesigner (cellular modelling), FLAME, OpenCell, Virtual Physiology (produced by cLabs), GENESIS, Neuron, Heart Simulator, AnyBody | | | CellML, SBML, NeuroML, MML | | IUPS Physiome Project, Virtual Physiological Human, High-Definition Physiology |
| **Molecular interaction modelling** | AutoDock Vina, GOLD, eHiTS | RCSB PDB, ZINC, PubChem, PDBbind | | | | |

This table summarizes the tools and resources that correspond to each step in a systems biology workflow; please refer to FIG. 1 for an overview of the workflow and to Supplementary information S1 (table) for additional information and Weblinks to these resources.

2017

OpenCell, Flame, Copasi, CellDesigner, NetBuilder, SimBoolNet, PhysioDesigner, etc …

*Ghosh, Samik, et al. "Software for systems biology: from tools to integrated platforms." Nature reviews. Genetics 12.12 (2011): 821.*

# Scope and limitations of computational methods

Models for different purposes -> different modelling technique -> parameters estimation

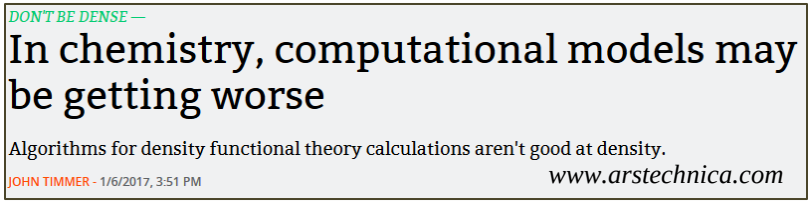
Computational approaches and tools

| Modeling approach | Typical applications | Limitations | Tools |
|---|---|---|---|
| **Individual particle-based stochastic** | Small subcellular signaling processes, apects of bacterial biochemistry | Limited to small systems (in terms of space and chemical complexity) | MCell (32), Smoldyn (315), ChemCell (316), GetBonNie (non- spatial) (49) |
| **Particle number stochastic** | Signaling processes with important stochastic aspects (due to small system size or high sensitivity) | Limited to small systems (in terms of space and chemical complexity), less detail than individual particle simulation | MesoRD (35), SmartCell (33), GetBonNie (non- spatial) |
| **Concentration-based spatial, non-stochastic** | Cellular signaling processes with important spatial aspects | Either high spatial resolution or biochemical complexity, No stochasticity | Virtual Cell (37), Simmune (36) |
| **Concentration-based, non-spatial, non-stochastic** | Cellular signaling processes without spatial aspects | Assumption of global biochemical homogeneity in the simulated system | Copasi (46), E-cell (44), Cellware (45), Systems Biology Workbench (47), GetBonNie |

*Germain, Ronald N., et al. "Systems biology in immunology: a computational modeling perspective." Annual review of immunology 29 (2011): 527-585.*

# Limitations

- Models **cannot replace** laboratory experiments
  - Build in virtual world based on laboratory experiments
- Models **cannot prove** mechanisms
  - Are developed based on observed results
  - Can disprove mechanisms / hypotheses

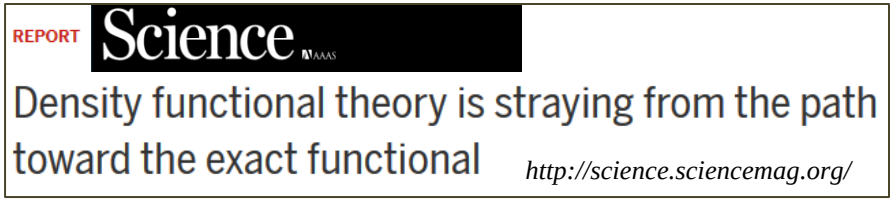

DON'T BE DENSE —

In chemistry, computational models may be getting worse

Algorithms for density functional theory calculations aren't good at density.

JOHN TIMMER - 1/6/2017, 3:51 PM

*www.arstechnica.com*


REPORT

Science AAAS

Density functional theory is straying from the path toward the exact functional

*http://science.sciencemag.org/*

*"Conclusion : that the latest trend of developing functionals using unconstrained forms leads to unphysical electron densities despite the excellent energy-related performance of these methods."*

# Conclusion

- Models have value because they allow their users to peer at deadbolt mechanisms from a different vantage point, sometimes even from inside.

- Computational modelling is transitioning into mainstream science in much the same way that statistics did many years ago.

- Computational models are becoming nearly obligatory, especially when a study argues for a new mechanism or functional relationship.

- More interdisciplinary work.

# Thank you!

*“Essentially, all models are wrong, but some are useful.” (George E.P. Box)*


Acknowledgement: Anna Henzi

# Appendix. Precision, Recall, and F1 Scores
# Machine Learning / Model Evaluation

Positive (P): Observation is positive (for example: is an apple)

Negative (N): Observation is not positive (for example: is not an apple).

True Positive (TP): Observation is positive, and is predicted to be positive.

False Negative (FN): Observation is positive, but is predicted negative.

True Negative (TN): Observation is negative, and is predicted to be negative.

False Positive (FP): Observation is negative, but is predicted positive.

$$Precision = \frac{TP}{TP + FP} = \frac{\text{positive predicted correctly}}{\text{all positive predictions}}$$

$$Recall = TPR = \frac{TP}{TP + FN} = \frac{TP}{P} = \frac{\text{predicted to be positive}}{\text{all positive observations}}$$

$$F1 = 2\frac{Precision * Recall}{Precision + Recall}$$

# Appendix. Cutaneous diseases and their comorbid diseases identified through ADRs

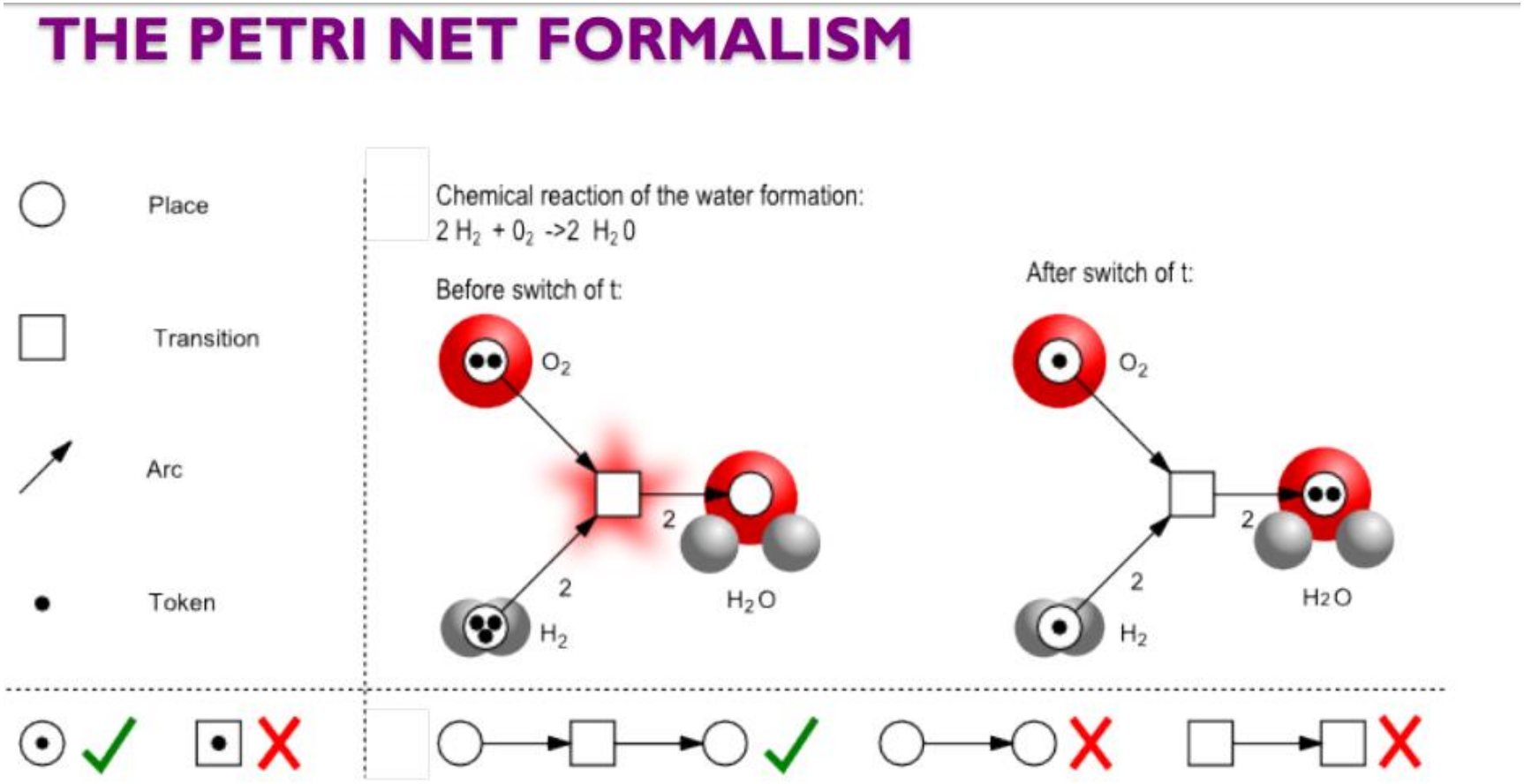
| Drug1 | Drug2 | Disease for Drug1 | Disease for Drug2 |
|---|---|---|---|
| Cyclosporine | Calcium | Psoriasis | Bone related diseases |
| Methotrexate | Anticancer antibiotic | Psoriasis | Cancer |
| Thioguanine | Thiopurine | Psoriasis | Acute lymphoblastic leukemia\|autoimmune disorders (Crohn's disease, rheumatoid arthritis) |
| Calcitriol | Calcium | Psoriasis | Bone related diseases |
| Calcitriol | Phosphorus | Psoriasis | Bone related diseases - Rickets in children, Osteomalacia in adults |
| Methotrexate | DMARD | Psoriasis | Rheumatoid arthritis, Lupus erythematosus, Psoriasis |
| Calcitriol | Zinc | Psoriasis | Eczema |
| Sulfur | Antioxidant | Acne vulgaris, Psoriasis, Rosacea | Cancer |
| Mycophenolic acid | Tacrolimus | Psoriasis | Atopic dermatitis |
| Cyclosporine | Androgen | Psoriasis | Breast cancer in females |
| Cholecalciferol | Parathyroid hormone | Psoriasis | To control hypocalcemia in patients in hypoparathyroidism |
| Calcitriol | Parathyroid hormone | Psoriasis | To control hypocalcemia in patients in hypoparathyroidism |
| Cyclosporine | Sodium | Psoriasis | Blood pressure and blood volume |
| Cholecalciferol | Calcium | Psoriasis | Bone related diseases |
| Methotrexate | Antidiabetic drug | Psoriasis | Diabetes |
| Sulfur | Antiviral | Acne vulgaris, Psoriasis, Rosacea | Viral diseases - Influenza (flu) |
| Methotrexate | Antirheumatic drug | Psoriasis | Rheumatoid arthritis |
| Cyclosporine | Calcium channel blockers | Psoriasis | High blood pressure, Chest pain, Raynaud's disease |
| Methotrexate | Antifolates | Psoriasis | Cancer |
| Methotrexate | NSAIDs | Psoriasis | Fever, pain, inflammation |
| Cyclosporine | Macrolide | Psoriasis | Bacterial conjuctivities |
| Etretinate | Retinoid | Psoriasis | Melanoma |
| Diclofenac | NSAIDs | Keratosis | Fever, pain, inflammation |
| Fluorouracil | Xeloda | Keratosis | Colorectal neoplasms |
| Tacrolimus | Corticosteroid | Atopic dermatitis | Rheumatoid arthritis, Lupus, Asthma, Allergies, Addison's disease |
| Tacrolimus | Fluconazole | Atopic dermatitis | Cryptococcal meningitis, AIDS-related opportunistic infections, Fungemia, Vulvovaginal candidiasis, Histoplasmosis, Chronic mucocutaneous candidiasis, Histoplasmosis, Coccidioidomycosis, Blastomycosis |
| Temozolomide | Chemotherapeutic agent | Melanoma | Cancer |
| Zinc | Heparin | Eczema | Thromboembolism, Thrombophlebitis, Pulmonary embolism, unstable Angina, Myocardial infarction, Cerebral infarction, postoperative complications, Coronary thrombosis |
| Zinc | Progesterone | Eczema | Endometrial hyperplasia, Uterine hemorrhage, female infertility, Amenorrhea |
| Zinc | Calcium | Eczema | Bone related diseases |
| Zinc | Estrogen | Eczema | Menorrhagia, breast neoplasms, premature menopause, primary ovarian insufficiency, Hypogonadism, Prostatic neoplasms, hot flashes |
| Zinc | Antipsychotic | Eczema | Schizoprenia |
| Zinc | Sulfonamide | Eczema | Acne vulgaris, Acne rosacea, Seborrheic dermatitis |

# Scope and limitations of computational methods

| Investigation/ prediction of | Computational methods | Scope, limitations |
|---|---|---|
| Structure, function and mechanisms of metabolic enzymes | Homology modelling, quantum mechanics, molecular dynamics simulations, and so on | • Analysis of ligand binding events and enzyme mechanisms at a high level of detail and accuracy<br>• Particularly useful for the investigation of unstable reaction intermediates with very short lifetimes |
| Sites of metabolism | Knowledge-based systems, data mining, machine learning, QSAR models, reactivity models, ligand docking, molecular interaction fields, shape-based methods, and so on | • Able to predict the likely sites of metabolism with adequate accuracy<br>• In general, at least one site of metabolism is correctly identified among the three highest-ranked atom positions of a molecule in 70–90% of all cases[152] within the domain of chemical applicability |
| Metabolites (chemical structure) | Knowledge-based systems, data mining | • Dominated by knowledge-based systems<br>• Can produce large numbers of metabolites<br>• Main challenge: finding ways of ranking metabolites accurately |
| Metabolic rates | Quantum mechanics, molecular dynamics simulations (QSAR models) | • Prediction generally not possible<br>• QSAR-like approaches may work, but only within an extremely narrow chemical space |
| Interactions of drugs with targets related to drug metabolism | QSAR models | • Prediction of ligand affinity and inhibitory activity, in cases in which adequate training data are available<br>• Prediction of mechanism-based inhibitors remains highly challenging |
| | Free-energy calculations | • Accurate prediction of binding affinities without need for extensive training data<br>• Computationally expensive and labour-intensive |
| Bioactivity and toxicological effects | Various ligand- and structure-based approaches | • Target prediction methods have become widely available but high false-positive rates (that is, accurate ranking of targets) remain a limiting factor<br>• Prediction of bioactivities for metabolites hampered by lack of training data<br>• Rule-based approaches are able to detect most toxicophores, but prediction of time-dependent inhibitors remains challenging |

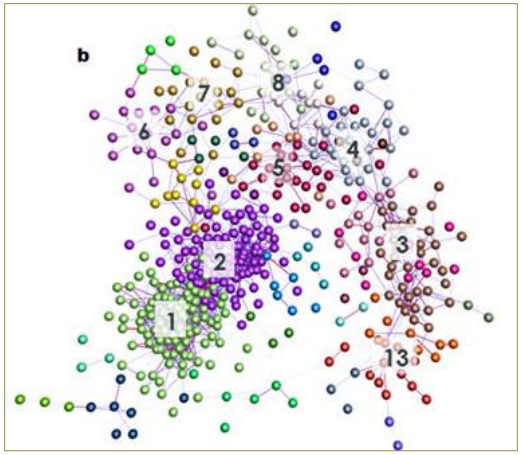
*Kirchmair, Johannes, et al. "Predicting drug metabolism: experiment and/or computation?." Nature reviews. Drug discovery 14.6 (2015): 387.*

# Appendix. Petri Net example




*http://www-dssz.informatik.tu-cottbus.de/publications/tutorial_milano_2013/1b-milano-tutorial2013-PN4SysBio_Blaetke.pdf*

# Appendix. Gene Ontology enrichment in clusters preferentially induced by TNFα or TSLP signalling.

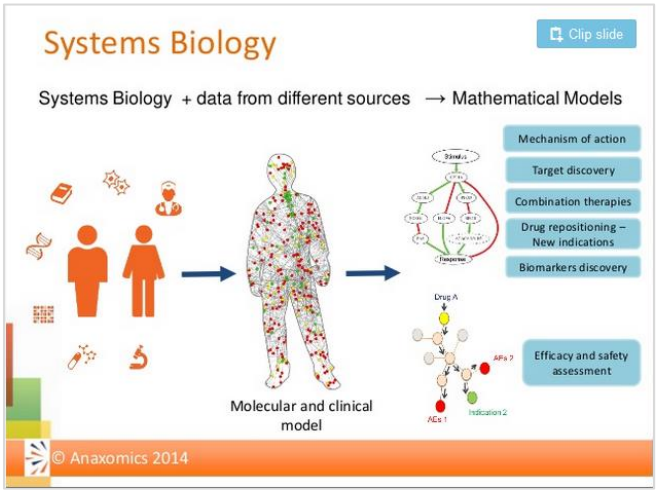



| Cluster | Preferentially regulated by (time, cytokine, two way ANOVA) | gene number | GO (FDR B&H)/gene list for low gene number clusters |
|---|---|---|---|
| 01 | TNFα (p < 0.0001, p = 0.021) | 95 | immune response (p = 0.0051), leukocyte activation (p = 0.0051), proteasome activator complex (p = 0.009) |
| 02 | TNFα (p < 0.0001, p = 0.011) | 84 | Pathways: cell cycle (p = 0.008), HIV infection (p = 0.012), proteasome (p = 0.036), cross-presentation of soluble exogenous antigens (endosomes) (p = 0.036), |
| 09 | TNFα (p < 0.0001, p = 0.052) | 12 | regulation of RNA splicing (p = 0.015) |
| 17 | TNFα (p < 0.0001, p = 0.018) | 6 | CLIP2, IL1R2, OAF, RAB38, TCF7, TMEM184C |
| 18 | TNFα (p = 0.0002, p = 0.002) | 6 | C17orf62, C19orf54, CPNE1, FTSJD2, HECW1, STK25 |
| 03 | TSLP (p < 0.0001, p = 0.005) | 36 | no annotation |
| 04 | TSLP (p = 0.006, p = 0.001) | 25 | no annotation |
| 05 | TSLP (p < 0.0001, p = 0.019) | 25 | JUN kinase binding (p = 0.027) |
| 06 | TSLP (p < 0.0001, p = 0.001) | 18 | peroxisome proliferator activated receptor binding (p = 0.017) |
| 07 | TSLP (p < 0.0001, p = 0.004) | 18 | no annotation |
| 08 | TSLP (p < 0.0001, p = 0.007) | 16 | nucleotide transferase activity (p = 0.026) |
| 10 | TSLP (p < 0.0001, p = ns) | 10 | nucleotide metabolism (p = 0.042) |
| 11 | TSLP (p < 0.0001, p = 0.022) | 10 | mRNA splicing (p = 0.025) |
| 12 | TSLP (p = 0.0007, p = 0.021) | 10 | Golgi aparatus (p = 0.025) |
| 13 | TSLP (p < 0.0001, p = 0.038) | 9 | transferrin receptor activity ((p = 0.001) |
| 14 | TSLP (p = 0.0014, p = 0.003) | 8 | ATP5L, EFHA1, ID2, INIP, RECQL, RPS4X, TMSB4X, UBL5 |
| 15 | TSLP (p < 0.0001, p = 0.054) | 8 | CAMK1D, ELL3, LAP3, MLLT4, MPC1, NET1, NFE2L3, STOM |
| 16 | TSLP (p < 0.0001, p = ns) | 7 | ARAP1, CNDP2, GSDMD, N4BP2L2, NINJ2, PARP10, VPS13B |

# From Systems Biology to Systems Medicine

- Simulations of models can reveal hidden patterns and/or counterintuitive mechanisms in complex systems.
- Systems medicine -> functioning of drug
- Functioning of drug -> cellular- and tissue-level networks, linked to an individual patient's genome, metabolome, and proteome.
- **necessary to depart** from gene-, protein-, and pathway-centric approaches

Systems Biology → Systems Medicine



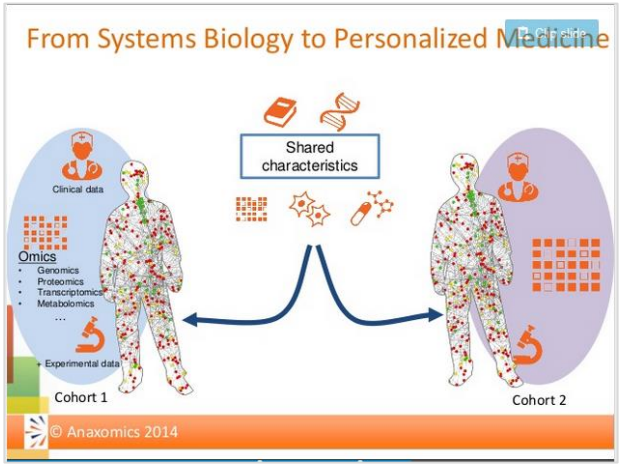



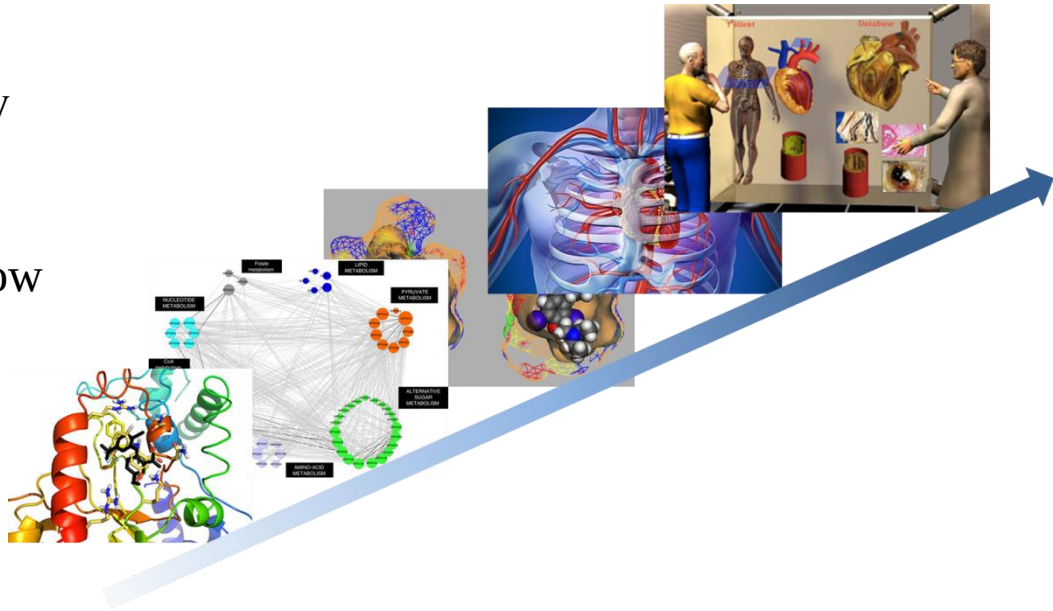

*Image from Teresa Sardon , Anaxomics, System biology approaches to personalized medicine*

*Wolkenhauer, Olaf, et al. "The road from systems biology to systems medicine." Pediatric research 73.4-2 (2013): 502-507.*

# What are computational models?

Computational modeling is the use of computers to **simulate** and study the behavior of complex systems using **mathematics**, **physics** and **computer science**.

**Multiscale modeling**: A key feature of today's computational models is that they are able to study a biological system at **multiple levels**, including molecular processes, cell to cell interactions, and how those interactions result in changes at the tissue and organ level .



*https://www.nibib.nih.gov/science-education/science-topics/computational-modeling*