Construction and Use of Biological Network Models in Risk Assessment

Szostak Justyna

Nice, ENOW 2017

19 Mars 2017

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

www.pmiscience.com

Risk Assessment

Risk assessment, in the context of public health,
is the process of quantifying the
probability of a harmful effect to individuals
or populations from human activities.

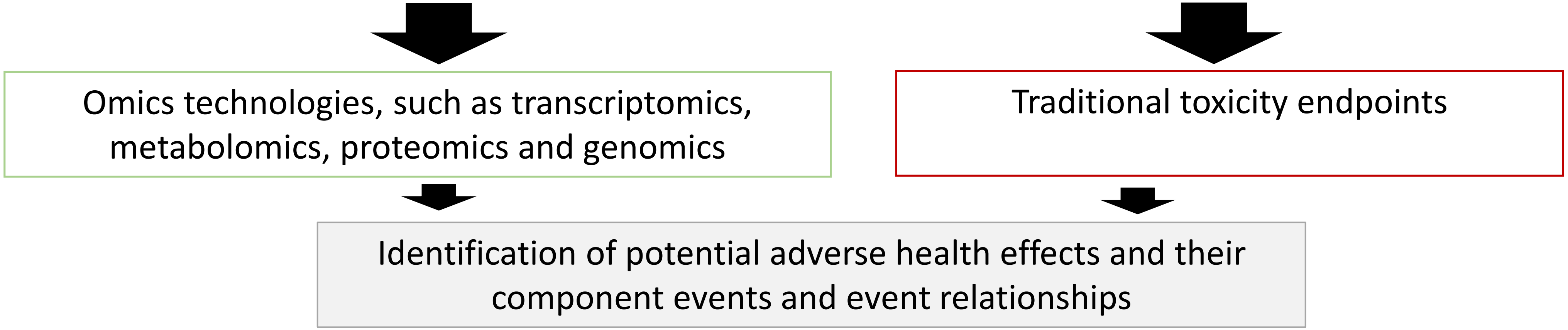
Developing such an approach requires a
mechanistic understanding of the process
by which xenobiotic substances perturb biological systems and lead to
adverse outcome

21st Century Toxicology

The National Research Council, commissioned by the US EPA, developed a vision for 21st-century toxicity testing characterized by a shift in focus away from traditional toxicity testing and toward the exploration of human signaling pathways whose perturbation by biologically active substances or their metabolites causes adverse health effects

Robert Kavlock, director of the EPA the U.S. Environmental Protection Agency (EPA) says:

> “The main difference between Tox21 and prior molecular research in toxicology is one of scale”
>
> “Scientists have generally focused on hypothesis-driven investigations, such as how a chemical interacts with a specific cell target assumed to play a role in toxicity, he explains. Tox21, on the other hand, relies on unbiased screening methods that don’t assume any prior knowledge about what a chemical might do in the cell”

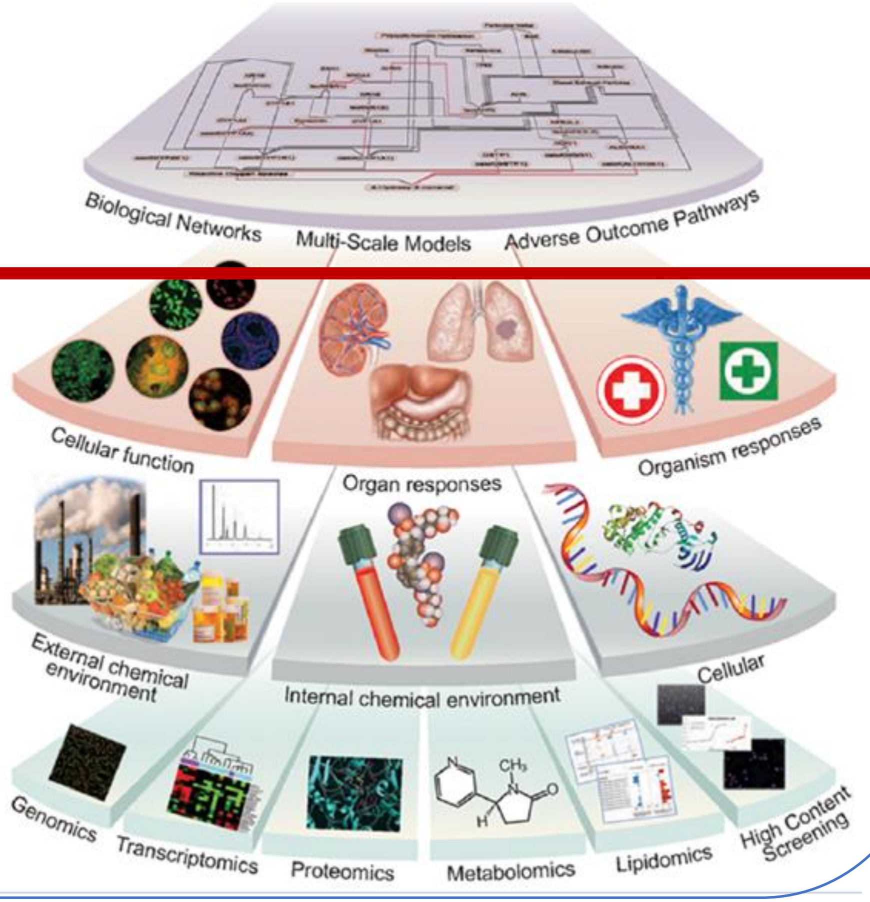


Omics technologies, such as transcriptomics, metabolomics, proteomics and genomics

Traditional toxicity endpoints

Identification of potential adverse health effects and their component events and event relationships

Risk Assessment- System Toxicology at PMI



Computational models

Apical measurements

Molecular measurements

Enabling technologies

Causal Biological Network Models

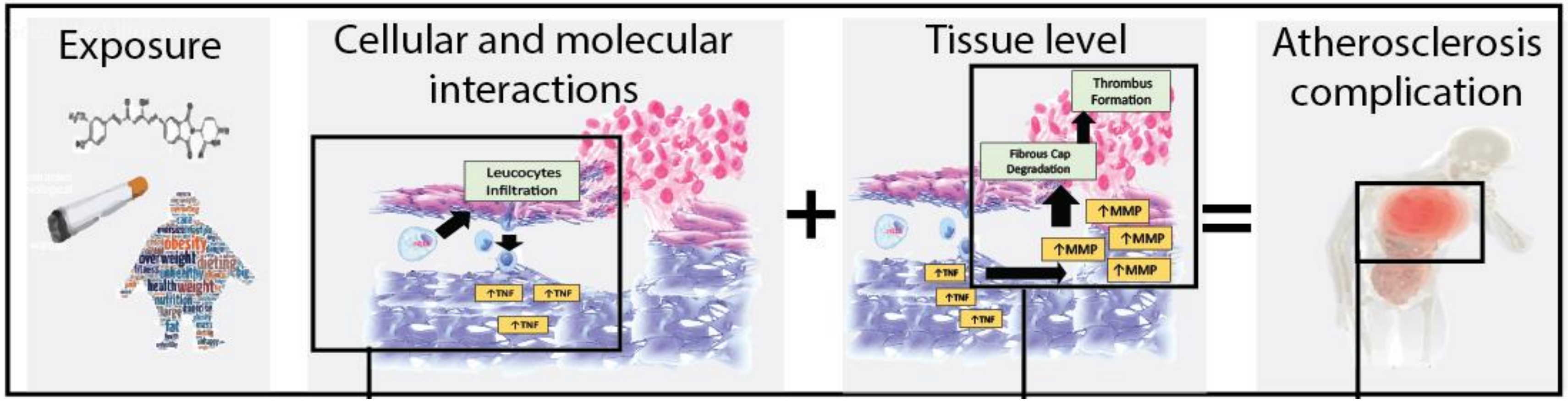
Mechanistic understanding of the process
by which xenobiotic substances perturb biological systems and lead to adverse outcome

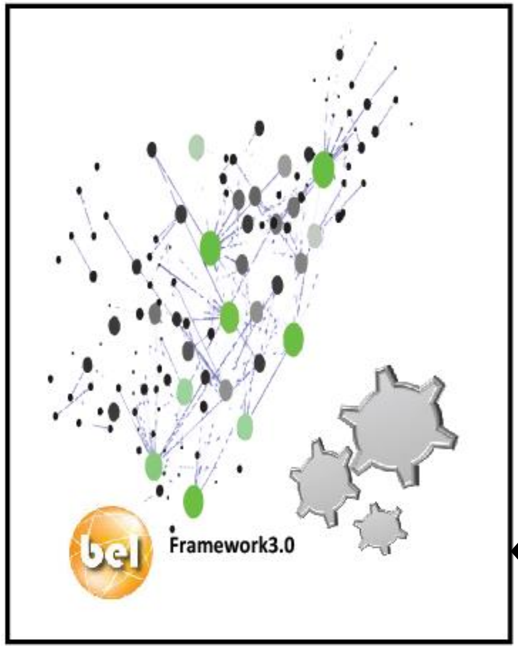
Computational Models Approach

Causal Biological Network Models

Structured Knowledge

Cellular and molecular interactions taking place during disease development or progression are written into computable BEL statements



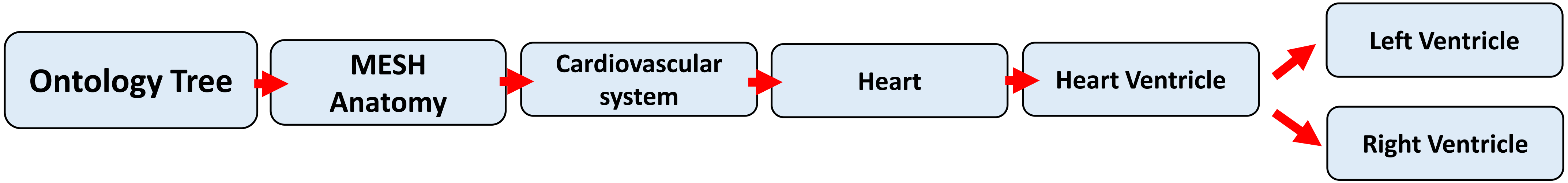


Definition of Ontologies

Ontologies are needed to convert natural language into computable form

U.S. National Library of Medicine

A formal ontology is a controlled vocabulary, a terminology that has well-defined hierarchy, where every term has a relationship to other terms



Every term has a description of what it means:

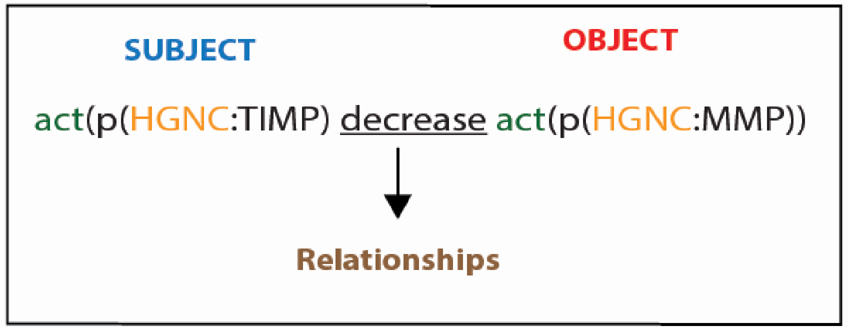
HEART:
A hollow organ located slightly to the left of the middle portion of the chest. It is composed of muscle and it is divided by a septum into two sides: the right side which receives de-oxygenated blood from the body and the left side which sends newly oxygenated blood to the body. Each side is composed of two chambers: the atrium (receiving blood) and ventricle (ejecting blood).

Ontologies control name-spaces by mapping synonyms (and acronyms) to reference terms that represent well-defined entities.

“name-space-control” facilitates the conversion of information into a computer-readable form

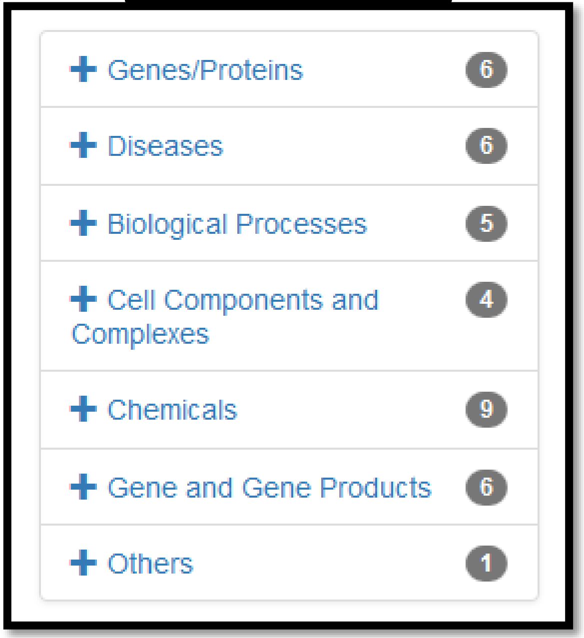
Biological Expression Language (BEL)

BEL STATEMENT

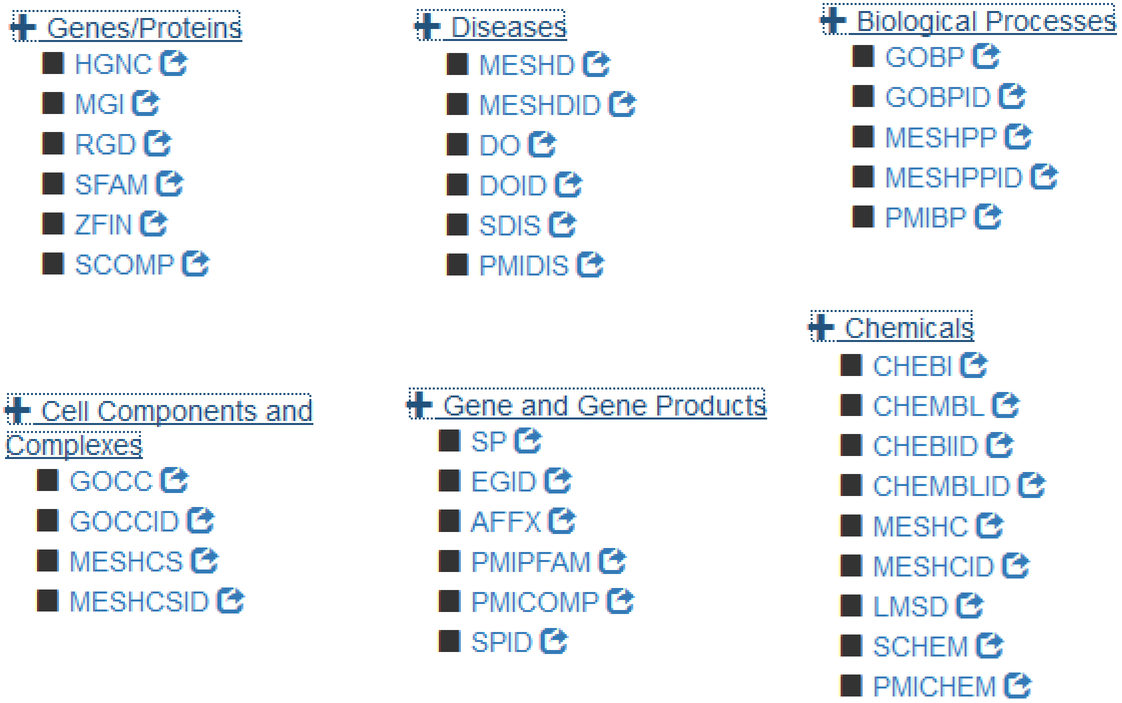


BEL is an example of a computable language that is easy to understand for researchers

Dedicated Ontology



- Genes/Proteins 6
- Diseases 6
- Biological Processes 5
- Cell Components and Complexes 4
- Chemicals 9
- Gene and Gene Products 6
- Others 1



Genes/Proteins
- HGNC
- MGI
- RGD
- SFAM
- ZFIN
- SCOMP

Cell Components and Complexes
- GOCC
- GOCCID
- MESHCS
- MESHCSID

Diseases
- MESHD
- MESHDID
- DO
- DOID
- SDIS
- PMIDIS

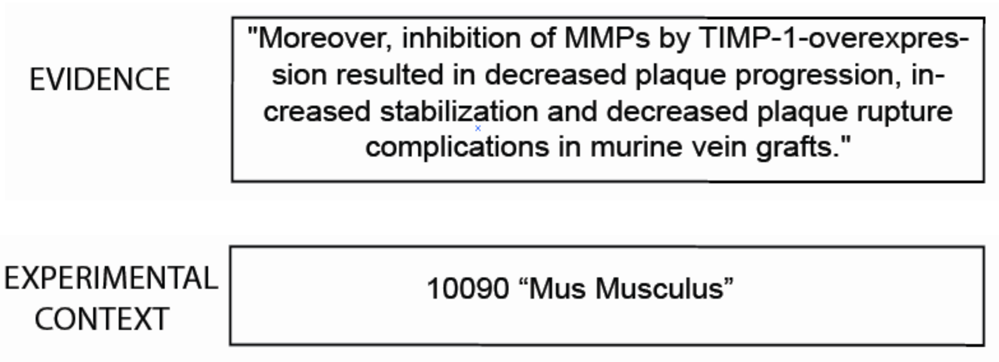
Gene and Gene Products
- SP
- EGID
- AFFX
- PMIPFAM
- PMICOMP
- SPID

Biological Processes
- GOBP
- GOBPID
- MESHPP
- MESHPPID
- PMIBP

Chemicals
- CHEBI
- CHEMBL
- CHEBIID
- CHEMBLID
- MESHC
- MESHCID
- LMSD
- SCHEM
- PMICHEM

Biological Expression Language (BEL)

BELIEF



EVIDENCE

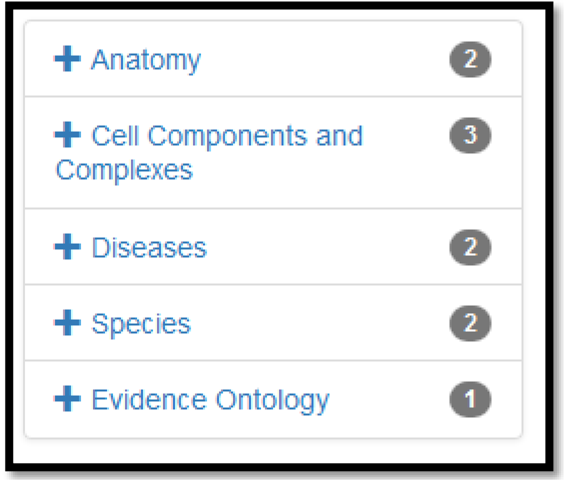
"Moreover, inhibition of MMPs by TIMP-1-overexpression resulted in decreased plaque progression, increased stabilization and decreased plaque rupture complications in murine vein grafts."

EXPERIMENTAL CONTEXT

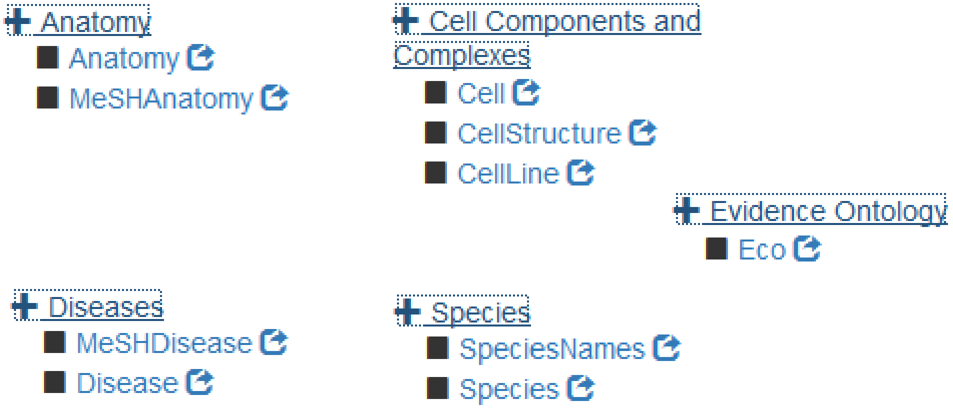
10090 "Mus Musculus"

BEL allows the annotation of each statements

Dedicated Ontology



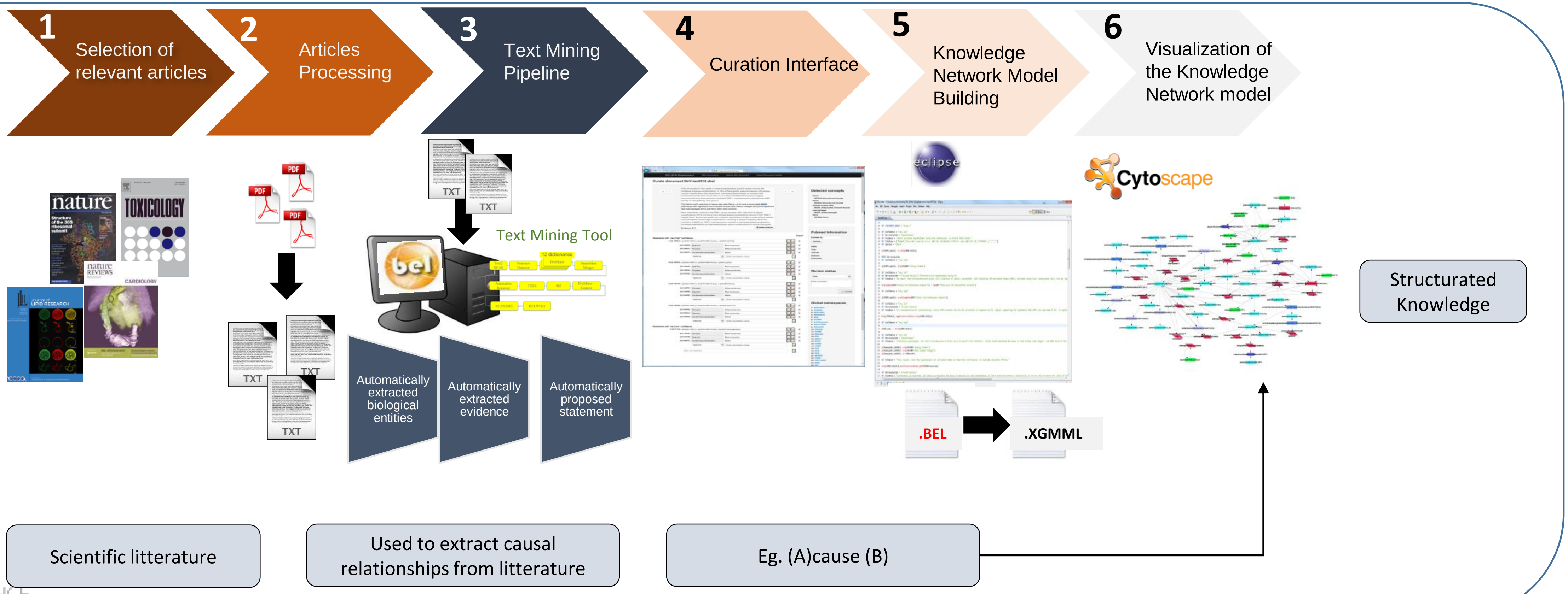
+ Anatomy	2
+ Cell Components and Complexes	3
+ Diseases	2
+ Species	2
+ Evidence Ontology	1



- + Anatomy
  - Anatomy
  - MeSHAnatomy
- + Cell Components and Complexes
  - Cell
  - CellStructure
  - CellLine
- + Evidence Ontology
  - Eco
- + Diseases
  - MeSHDisease
  - Disease
- + Species
  - SpeciesNames
  - Species

BELIEF Text Mining Tool

Use BELIEF text mining workflow to extract causal mechanistic interaction. BELIEF tool could be used between the different lipid class and proteins, genes and biological processes in order to build a connected knowledge on the causal biological network model.



1 Selection of relevant articles

2 Articles Processing

3 Text Mining Pipeline

4 Curation Interface

5 Knowledge Network Model Building

6 Visualization of the Knowledge Network model

Automatically extracted biological entities

Automatically extracted evidence

Automatically proposed statement

Structurated Knowledge

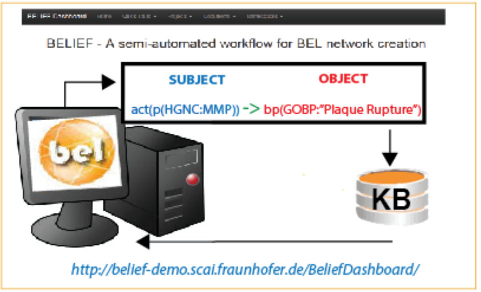
Scientific litterature

Used to extract causal relationships from litterature

Eg. (A)cause (B)

BELIEF Text Mining Tool

BELIEF

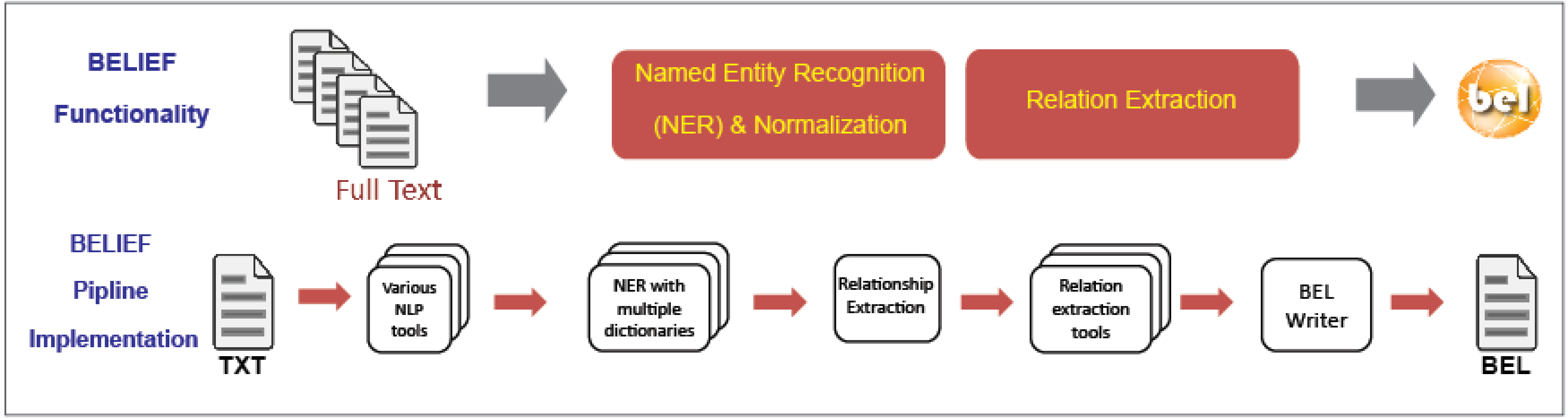


BELIEF - A semi-automated workflow for BEL network creation

http://belief-demo.scai.fraunhofer.de/BeliefDashboard/

BELIEF addresses the biological network model needs by identifying chemical, gene/protein, and biological process and disease terms in scientific articles. Additionally to that BELIEF identifies relationships through a combination of specialized ontologies and linguistics rules.

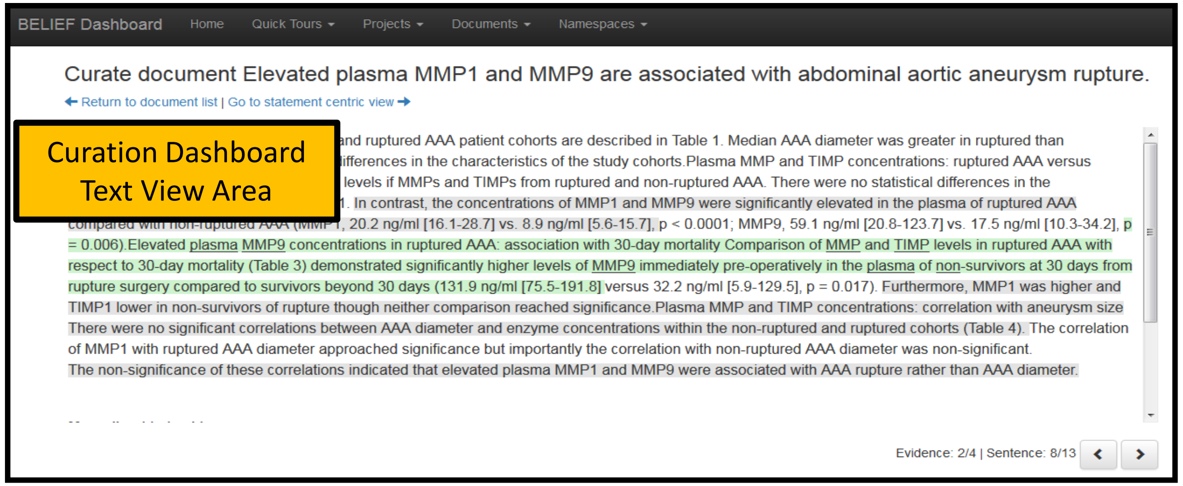
BELIEF Functionalities



Various NLP tools are used for detecting and splitting sentences, identifying words etc. In the next step NER is used to detect relevant entities with given dictionaries, here namespaces. The relationships between these detected entities is captured in the next step and finally a BEL nanopub compliant output is generated.

BELIEF Text Mining Dashboard

BELIEF



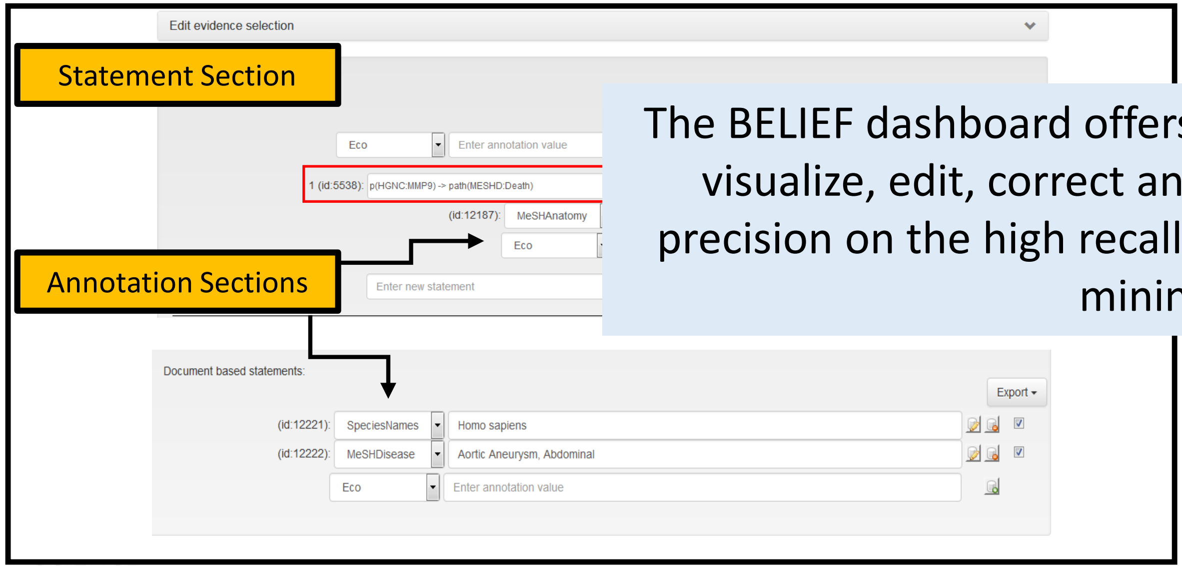
Curate document Elevated plasma MMP1 and MMP9 are associated with abdominal aortic aneurysm rupture.

← Return to document list | Go to statement centric view →

Curation Dashboard Text View Area

…nd ruptured AAA patient cohorts are described in Table 1. Median AAA diameter was greater in ruptured than …ifferences in the characteristics of the study cohorts.Plasma MMP and TIMP concentrations: ruptured AAA versus … levels if MMPs and TIMPs from ruptured and non-ruptured AAA. There were no statistical differences in the …1. In contrast, the concentrations of MMP1 and MMP9 were significantly elevated in the plasma of ruptured AAA compared with non-ruptured AAA (MMP1, 20.2 ng/ml [16.1-28.7] vs. 8.9 ng/ml [5.6-15.7], p < 0.0001; MMP9, 59.1 ng/ml [20.8-123.7] vs. 17.5 ng/ml [10.3-34.2], p = 0.006).Elevated plasma MMP9 concentrations in ruptured AAA: association with 30-day mortality Comparison of MMP and TIMP levels in ruptured AAA with respect to 30-day mortality (Table 3) demonstrated significantly higher levels of MMP9 immediately pre-operatively in the plasma of non-survivors at 30 days from rupture surgery compared to survivors beyond 30 days (131.9 ng/ml [75.5-191.8] versus 32.2 ng/ml [5.9-129.5], p = 0.017). Furthermore, MMP1 was higher and TIMP1 lower in non-survivors of rupture though neither comparison reached significance.Plasma MMP and TIMP concentrations: correlation with aneurysm size There were no significant correlations between AAA diameter and enzyme concentrations within the non-ruptured and ruptured cohorts (Table 4). The correlation of MMP1 with ruptured AAA diameter approached significance but importantly the correlation with non-ruptured AAA diameter was non-significant.
The non-significance of these correlations indicated that elevated plasma MMP1 and MMP9 were associated with AAA rupture rather than AAA diameter.

Evidence: 2/4 | Sentence: 8/13



Statement Section

Annotation Sections

Eco | Enter annotation value

1 (id:5538): p(HGNC:MMP9) -> path(MESHD:Death)

(id:12187): MeSHAnatomy

Eco

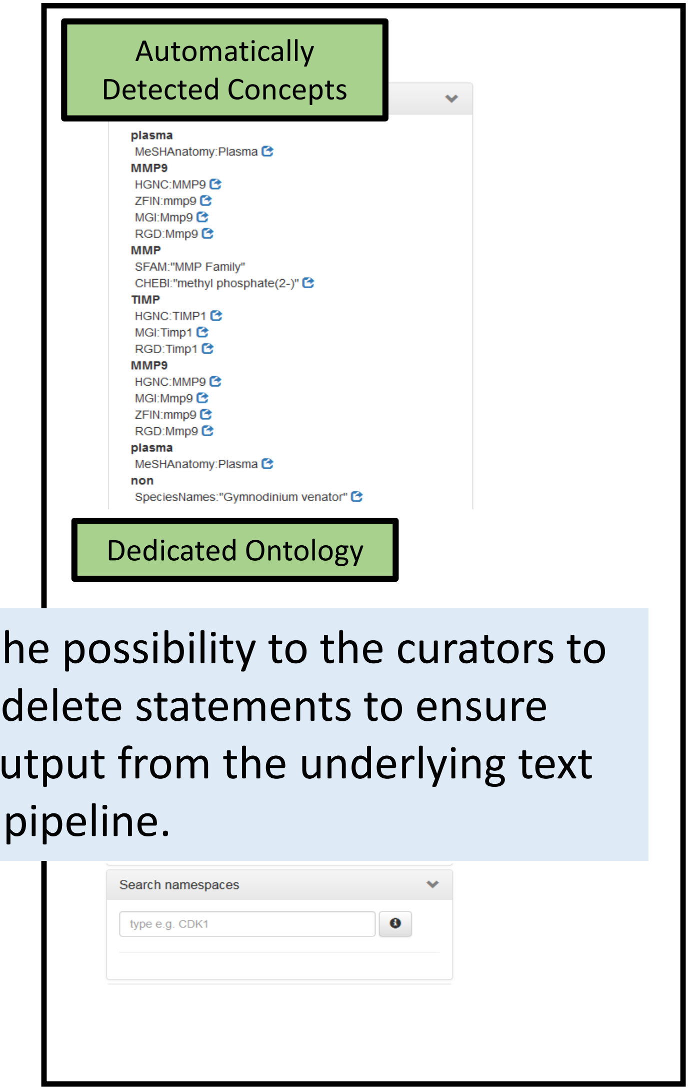
Enter new statement

Document based statements:

(id:12221): SpeciesNames | Homo sapiens ☑

(id:12222): MeSHDisease | Aortic Aneurysm, Abdominal ☑

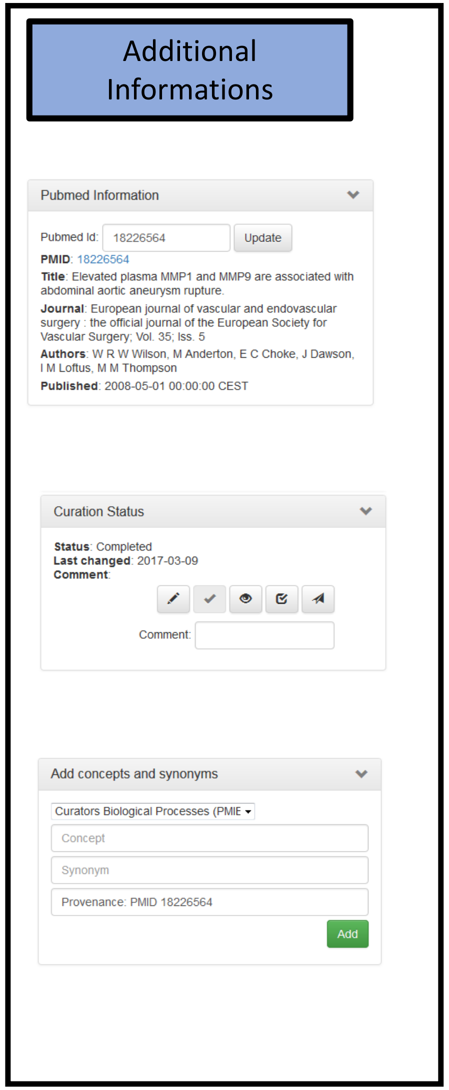
Eco | Enter annotation value



Automatically Detected Concepts

plasma
MeSHAnatomy:Plasma
MMP9
HGNC:MMP9
ZFIN:mmp9
MGI:Mmp9
RGD:Mmp9
MMP
SFAM:"MMP Family"
CHEBI:"methyl phosphate(2-)"
TIMP
HGNC:TIMP1
MGI:Timp1
RGD:Timp1
MMP9
HGNC:MMP9
MGI:Mmp9
ZFIN:mmp9
RGD:Mmp9
plasma
MeSHAnatomy:Plasma
non
SpeciesNames:"Gymnodinium venator"

Dedicated Ontology



Additional Informations

Pubmed Id: 18226564

PMID: 18226564

Title: Elevated plasma MMP1 and MMP9 are associated with abdominal aortic aneurysm rupture.

Journal: European journal of vascular and endovascular surgery : the official journal of the European Society for Vascular Surgery; Vol. 35; Iss. 5

Authors: W R W Wilson, M Anderton, E C Choke, J Dawson, I M Loftus, M M Thompson

Published: 2008-05-01 00:00:00 CEST

Status: Completed
Last changed: 2017-03-09
Comment:

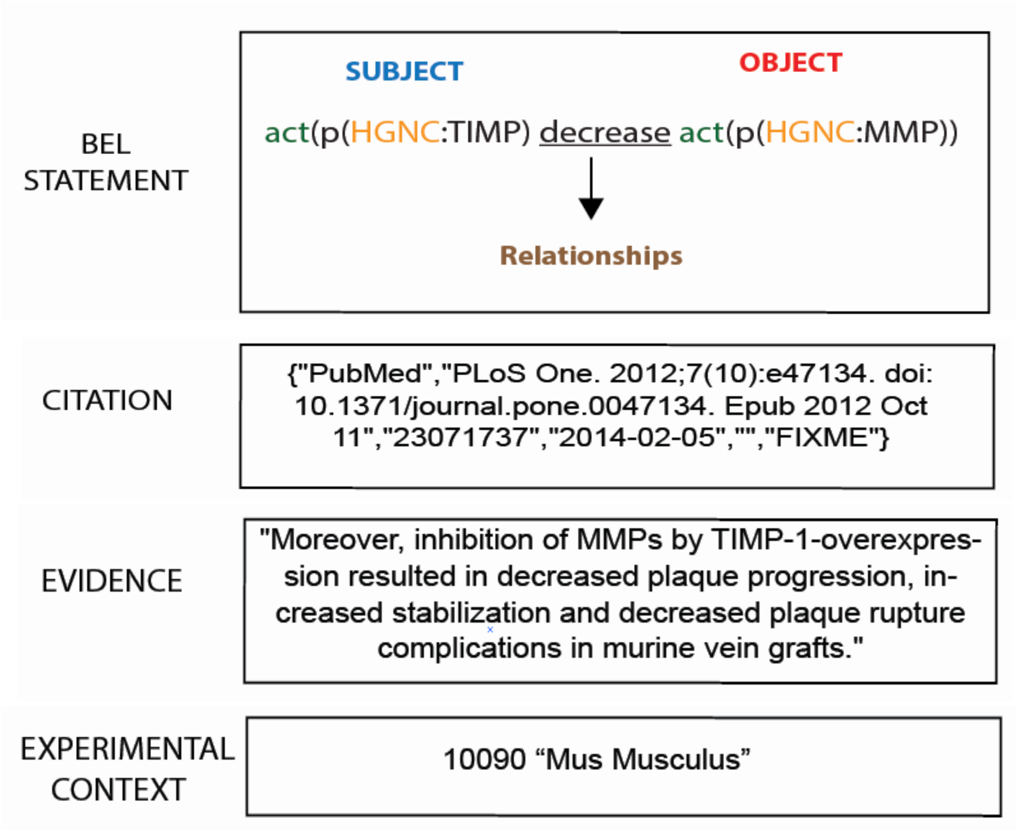
Comment:

Provenance: PMID 18226564

The BELIEF dashboard offers the possibility to the curators to visualize, edit, correct and delete statements to ensure precision on the high recall output from the underlying text mining pipeline.

Biological Expression Language (BEL) Overview

BELIEF



BEL STATEMENT	**SUBJECT** **OBJECT** act(p(HGNC:TIMP) decrease act(p(HGNC:MMP)) **Relationships**
CITATION	{"PubMed","PLoS One. 2012;7(10):e47134. doi: 10.1371/journal.pone.0047134. Epub 2012 Oct 11","23071737","2014-02-05","","FIXME"}
EVIDENCE	"Moreover, inhibition of MMPs by TIMP-1-overexpression resulted in decreased plaque progression, increased stabilization and decreased plaque rupture complications in murine vein grafts."
EXPERIMENTAL CONTEXT	10090 "Mus Musculus"

BEL Nanopub

The three crucial elements of a BEL Nanopub are:

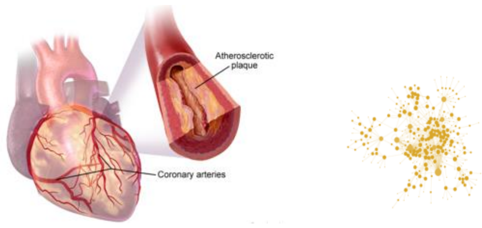
- the BEL statement showing the knowledge statement in a triple and controlled terminology
- the citation information
- evidence sentence

Experiment context is an additional field to simplify the triple assembly into biological network models

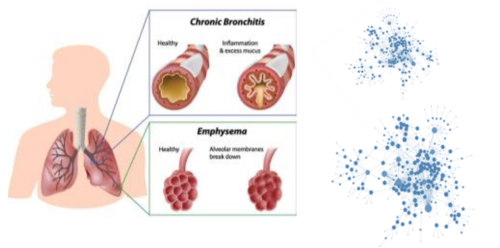
Causal Biological Network Model Database

Causal Biological Network Models

Cardiovascular diseases



Lung Disease

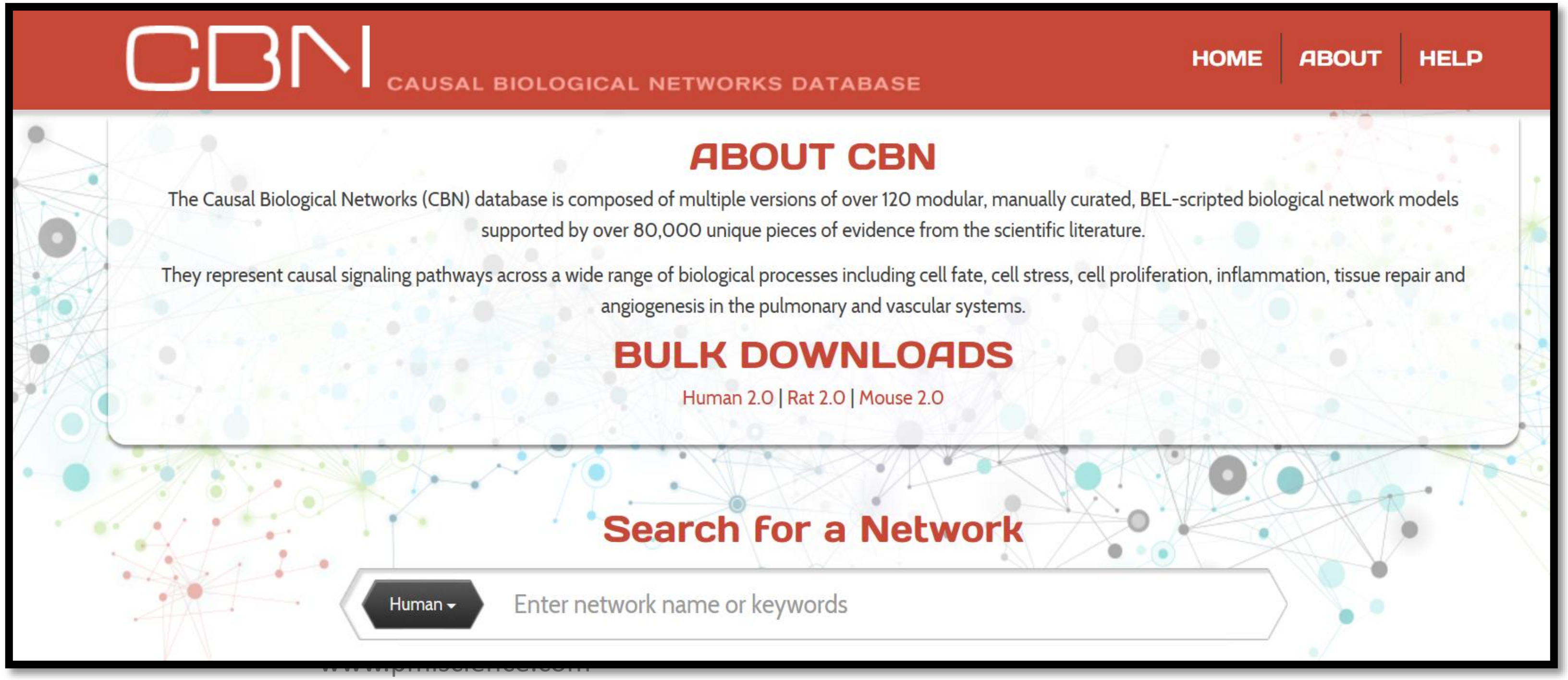


Verification of Causal Biological Networks Using a Web-based Collaborative Competition

sbv IMPROVER
SYSTEMS BIOLOGY VERIFICATION

Verification of Causal Biological Networks Using a Web-based Collaborative Competition

sbv IMPROVER
SYSTEMS BIOLOGY VERIFICATION



ABOUT CBN

The Causal Biological Networks (CBN) database is composed of multiple versions of over 120 modular, manually curated, BEL-scripted biological network models supported by over 80,000 unique pieces of evidence from the scientific literature.

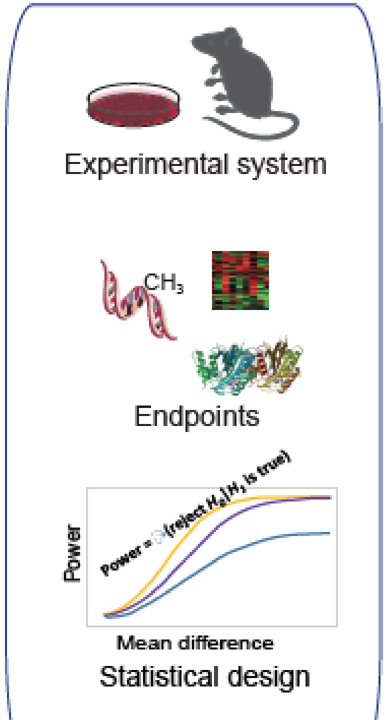
They represent causal signaling pathways across a wide range of biological processes including cell fate, cell stress, cell proliferation, inflammation, tissue repair and angiogenesis in the pulmonary and vascular systems.

BULK DOWNLOADS

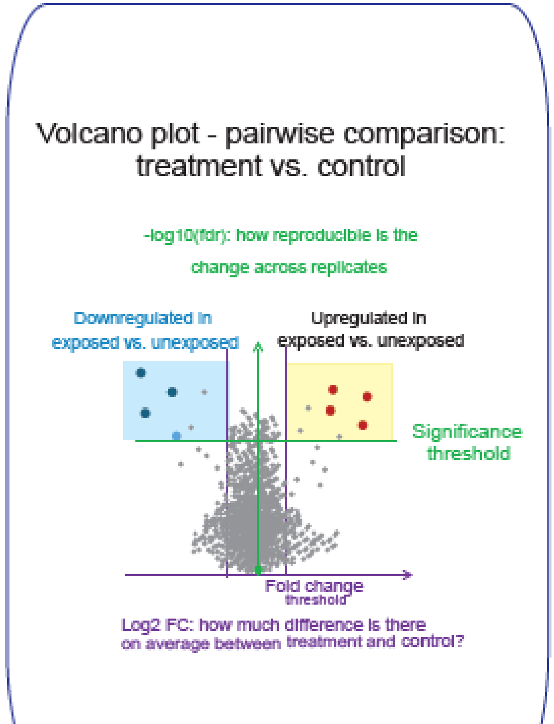
Human 2.0 | Rat 2.0 | Mouse 2.0

Search for a Network

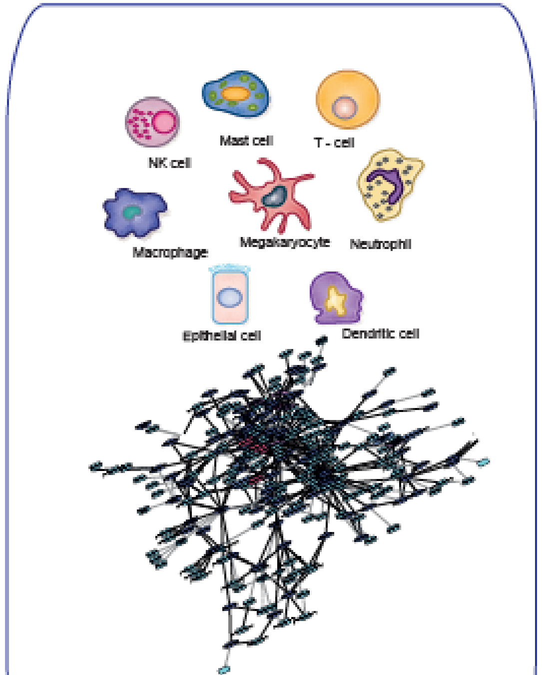
System Toxicology Assessment Approach



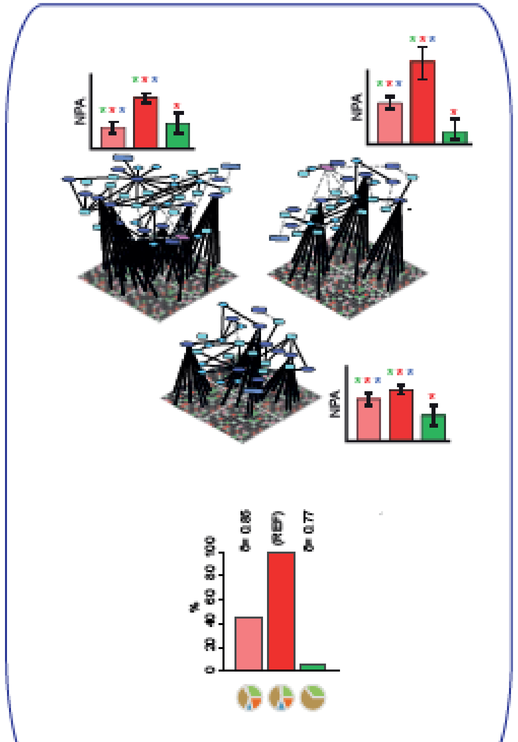
Experimental design



Systems response profile



Select tissue-specific networks from the network suite or build new networks



Compute NPA and BIF

The workflow starts with a careful selection of the experimental system and robust statistical design based on the choice of endpoints to be measured.

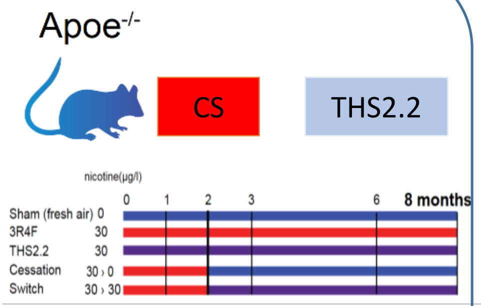
Transcriptomics data can be analysed to obtain the systems response profiles triggered by the exposure.

The transcriptomics data is analysed in the context of casual biological network models for mechanistic interpretation.

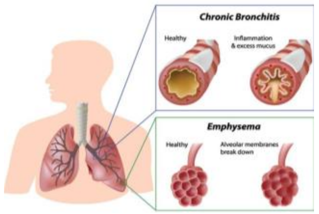
Sophisticated algorithms are used to compute the network perturbation amplitude (NPA) for each network and the aggregated overall biological impact factor (BIF).

System Toxicology Assessment Results

THS2.2 - 8-month systems toxicology inhalation / cessation study in Apoe$^{-/-}$ mice

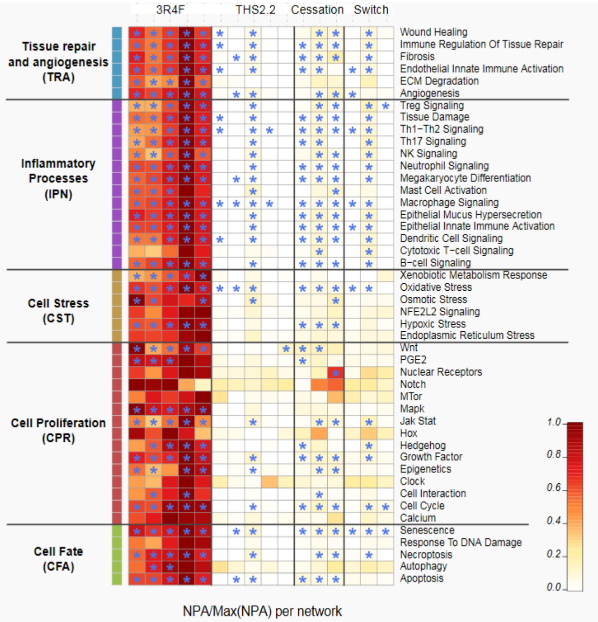


Lung Disease



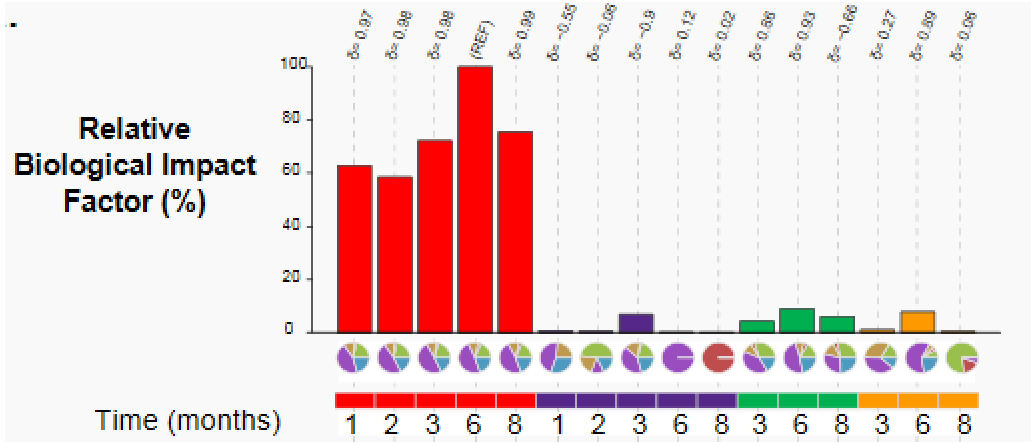
At designated time points (months 1, 2, 3, 6, and 8 after commencing exposure), the animals were examined for multiple parameters to **comprehensively assess** the development and progression of COPD

Network Models

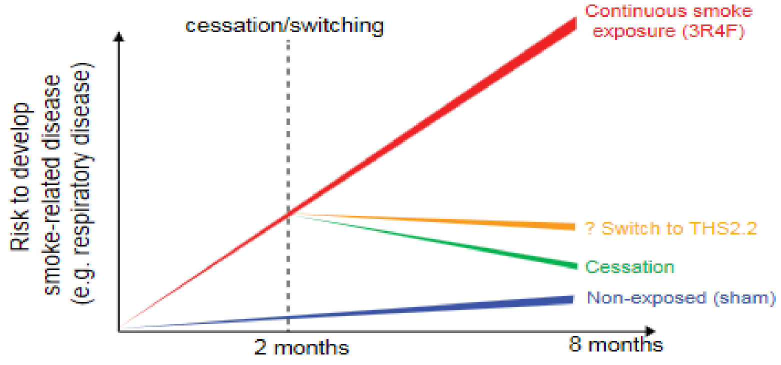


Network perturbation amplitude (NPA) analysis from the lung for treatment versus sham.

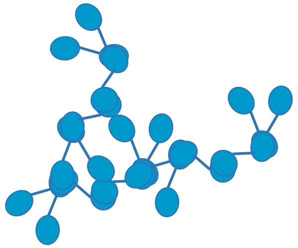
Relative BIF (RBIF) for treatment versus sham



Switching Study



Conclusion



Our approach in using systems toxicological assessment requires computational and knowledge-driven analysis of data from various experimental techniques

Knowledge is extracted from scientific articles and converted into a human and computer-readable format: BEL

BELIEF

BELIEF is supporting the automated extraction of knowledge as well as the manual curation and outputs in BEL

BEL networks can be verified by the crowd using the sbvIMPROVER Network Verification Challenge (NVC)

Verification of Causal Biological Networks Using a Web-based Collaborative Competition

Reviewed and verified networks are shared in the Causal Biological Networks Database (CBN)

"Systems Toxicology allows to create detailed understanding of the mechanisms by which biological systems respond to toxicants and uses this understanding to assess the risk of chemicals, drugs and consumer products"