# Edinburgh Pathway Editor

**visualization and data management tool for system biology**
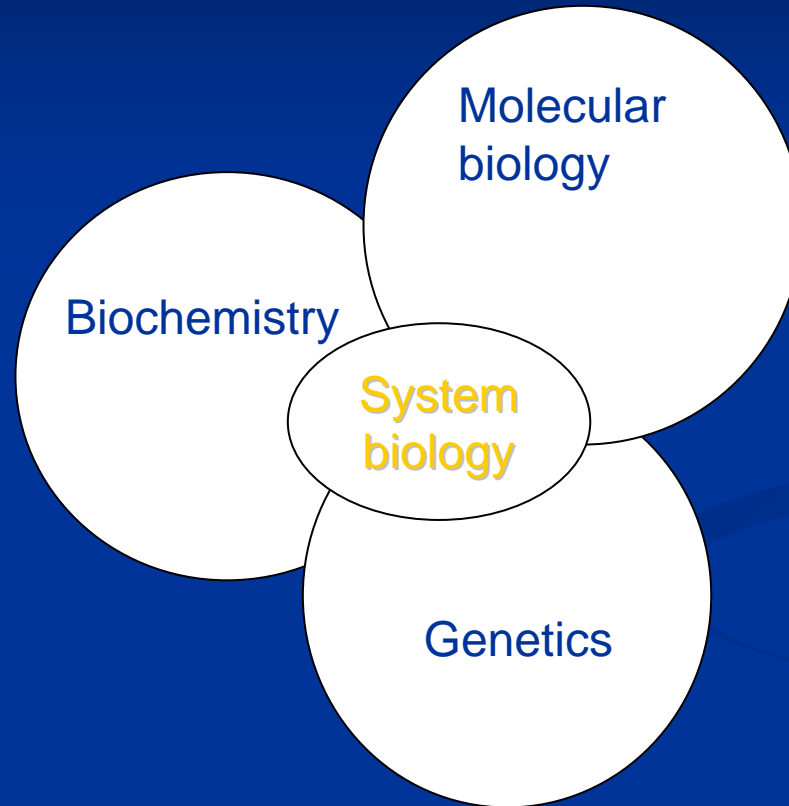
Anatoly Sorokin

Edinburgh Centre for System biology

# Outline

- Introduction to System biology
- Model exchange problem in System biology
- SBML & SBGN – standard way exchange model
- Edinburgh Pathway editor

A. Sorokin

# System biology



Molecular biology

Biochemistry

System biology

Genetics

# System biology

- Put together:
    - Metabolic pathways
    - Gene network
    - Signalling networks
    - Molecular interaction networks

# System biology

- Sequencing – gene prediction, functional assignment
- DNA array -- gene activity data
- Mass-spectrometry – biochemical pathway data
- Double Yeast – protein interaction data
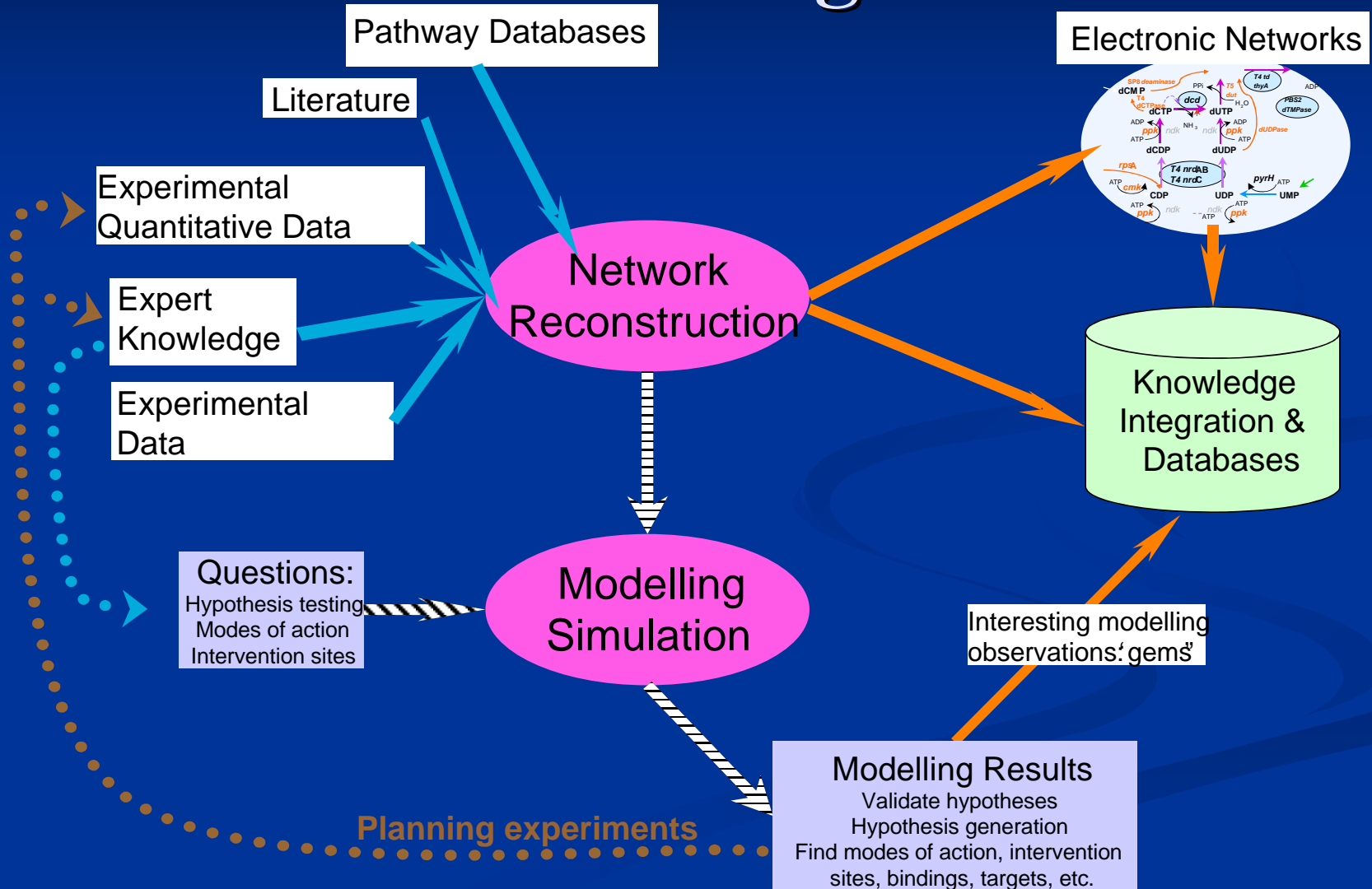- X-ray crystallography – protein structure

# System biology resorces

- Metabolc and Signal Transduction Pathways Databases
  - KEGG, WIT, EMP, BioCarta
  - Science Signal Transduction Knowledge Environment (STKE), CSNDB, SPAD
  - PathDB, BIND, BRITE, EcoCyc, MetaCyc, TransPath, GeneNetworks
- Interactions Databases
  - Ingenuity
  - DIP, PIM, BRITE , Interact, ProNet, ProChart, Proteome
- Modelling activities
  - SBML, CellML
  - GeneNET
  - E-CELL
  - BioSpice
- Visualisation and editing
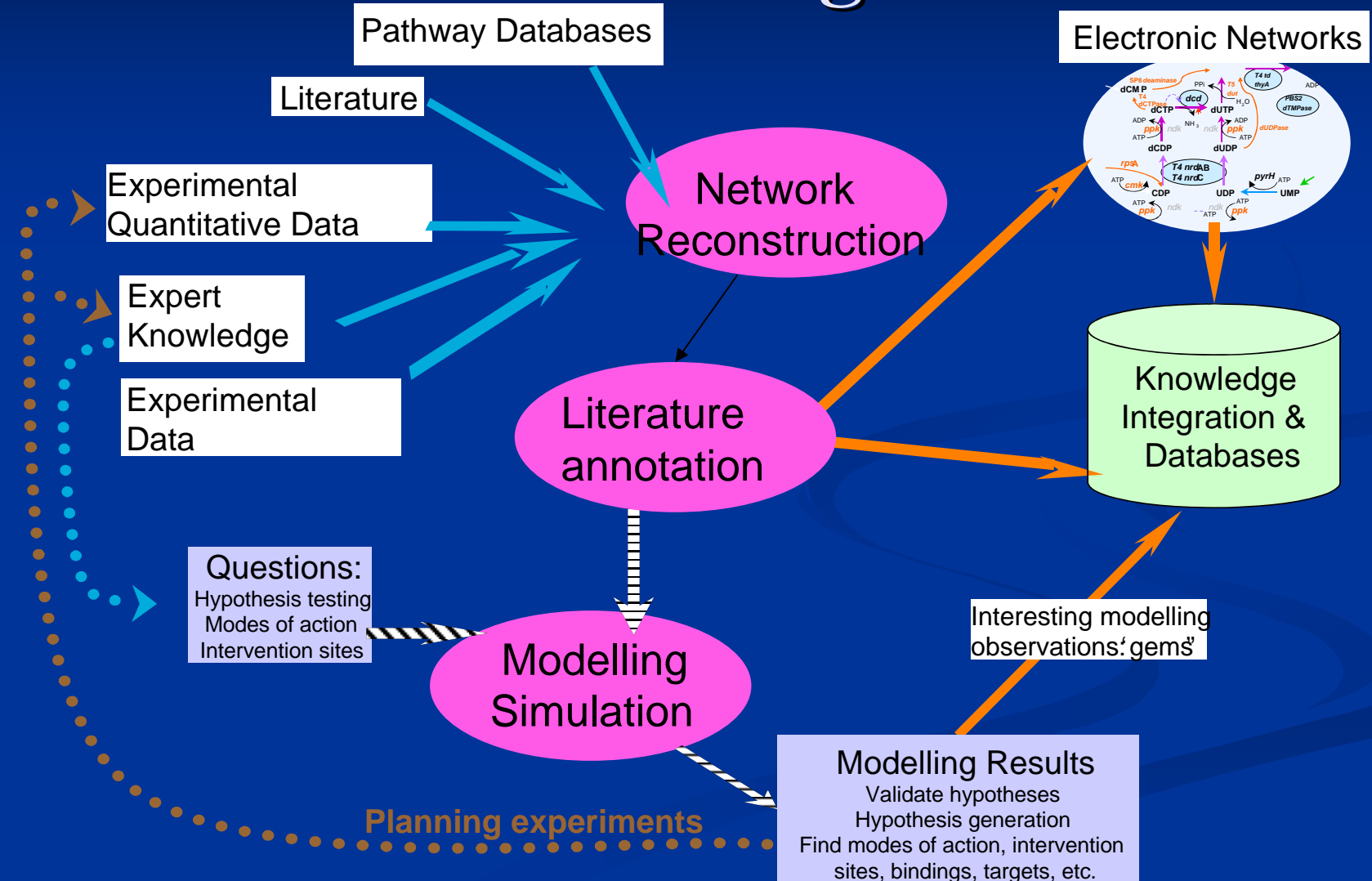  - CellDesigner, Teranode, Pathway Editor

# System biology modelling gap

"Biologist can be divided into two classes: experimentalists who observed things that cannot be explained, and theoreticians who explain things that cannot be observed."

Katzir-Katchalsky

# Network Reconstruction and Modeling

Pathway Databases

Electronic Networks

Literature

Experimental Quantitative Data

Expert Knowledge

Experimental Data

**Network Reconstruction**

**Knowledge Integration & Databases**

Questions:
Hypothesis testing
Modes of action
Intervention sites

**Modelling Simulation**

Interesting modelling observations: "gems"

**Planning experiments**

Modelling Results
Validate hypotheses
Hypothesis generation
Find modes of action, intervention sites, bindings, targets, etc.

A. Sorokin

# Network Reconstruction and Modeling



Pathway Databases

Literature

Experimental Quantitative Data

Expert Knowledge

Experimental Data

Questions:
Hypothesis testing
Modes of action
Intervention sites

Network Reconstruction

Literature annotation

Modelling Simulation

Electronic Networks

Knowledge Integration & Databases

Interesting modelling observations: "gems"

Modelling Results
Validate hypotheses
Hypothesis generation
Find modes of action, intervention
sites, bindings, targets, etc.

**Planning experiments**

A. Sorokin

# Two types of model

- **Heuristic model**
  - Keep expert knowledge
  - Proved by experiment
  - Lack of details

- **Numerical model**
  - Based on heuristic model
  - Ready for simulation
  - Contains arbitrary assumptions

# System biology modelling gap

- Models are distributed in large number of papers
- Models are published in informal style
- Difficult to trace all details
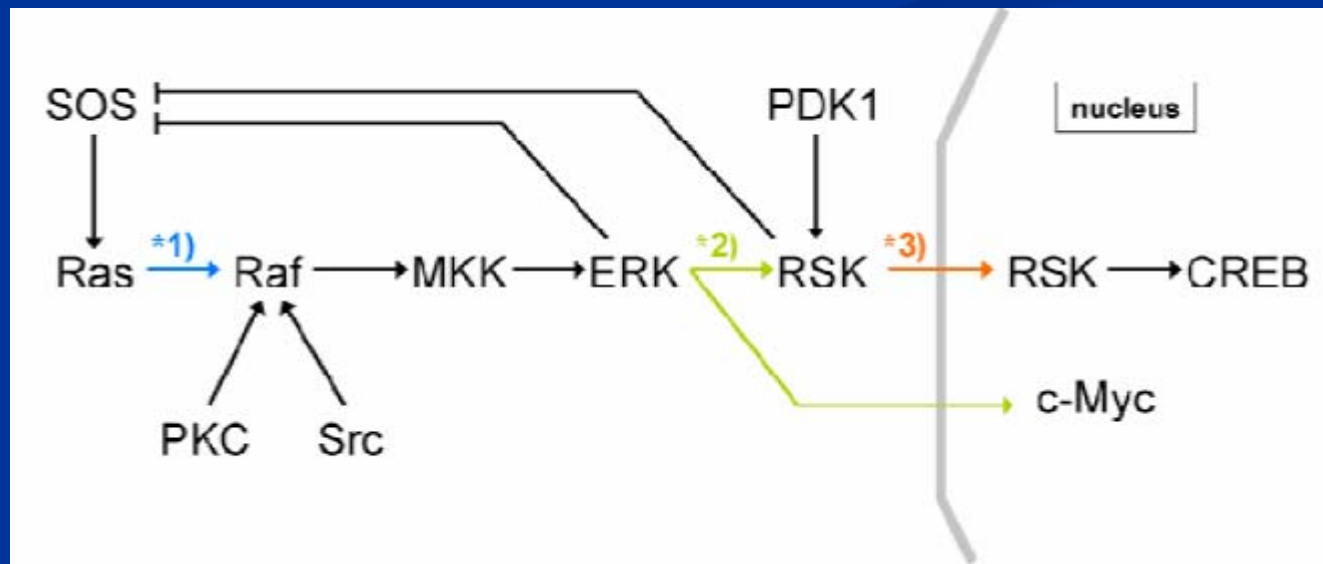
# Systems Biology Markup Language (SBML)

- XML-based language to share model information
- Model of biological process
- Simulation oriented
- Expect detailed description of the process
- No structure to store annotation data

# System Biology Graphical notation (SBGN)

- Standard Graphical notation
- Requirements
  - (1) **Expressiveness**: The notation system should be able to describe every possible relationship among genes and proteins, as well as biological processes.
  - (2) **Semantically unambiguous**: Notation should be unambiguous.
  - (3) **Visually unambiguous**: Each symbol should be clearly identified and cannot be mistaken with other symbols. This feature should be maintained with low-resolution displays, as well as black/white printings.
  - (4) **Extension Capability**: The notation system shall be flexible enough to add new symbols and relationship in a consistent manner. This may include the use of color-coding to enhance expressiveness and readability, but information shall not be lost even with black and white displays.

A. Sorokin

# System Biology Graphical notation (SBGN)

- Drawbacks
  - Created by modellers for modellers needs
  - Expect detailed description of the process
  - Lost of annotation data

A. Sorokin

# Link model types together

- Experimentalist oriented
- Support heuristic models
- Support "Round trip" development
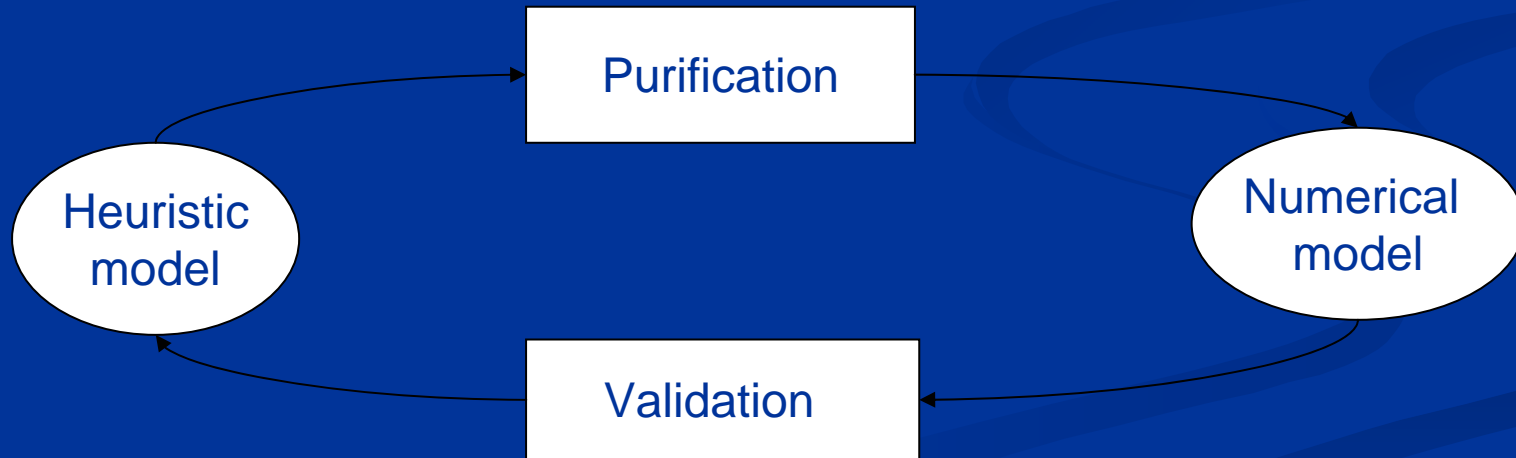
# Requirements

- Validation of networks vs. controlled vocabularies
    - of Small Molecules, Genes, Proteins, and transcripts to control confusion of multiple names across research areas.

- Networks should allow adding of
    - Literature objects (Medline etc…)
    - Kinetic information to reactions
    - Export in wide variety of formats
    - Any desired data related to model creation

- Build in links to, and load
    - publicly available data sources…

# Edinburgh Pathway Editor

A. Sorokin

# Basic features

- Small number of basic objects to represent main concepts of the biological network
- Metadata-oriented design
- Flexible visual presentation
- Hierarchical data storage
- Customisable direct link to external databases
- Eclipse-based open plug-in design
- XML-based internal presentation
- Export to SBML

# Main objects

- ***Shape*** represents the biological object or subsystem, treated as "a black box"

- ***Port*** interface to a Shape

- *Process* visualises sequences of events e.g. a biochemical reaction or protein interaction

- *Link* is used to represent any pair-wise relation between objects including "identity" or "act on" relations.

# Metadata

- Context define
  - New property types
    - Text
    - Simple data
    - Visual properties (position, size, colour)
    - Collections (text, simple data)
  - List of object properties
  - Inheritance
  - Default value
  - Visualisation properties
  - Links to external databases

# New property

A. Sorokin

# Inheritance, default value

A. Sorokin

# Visual representation

A. Sorokin

# Visual representation

A. Sorokin

# Visual representation

A. Sorokin

# Visual representation



A. Sorokin
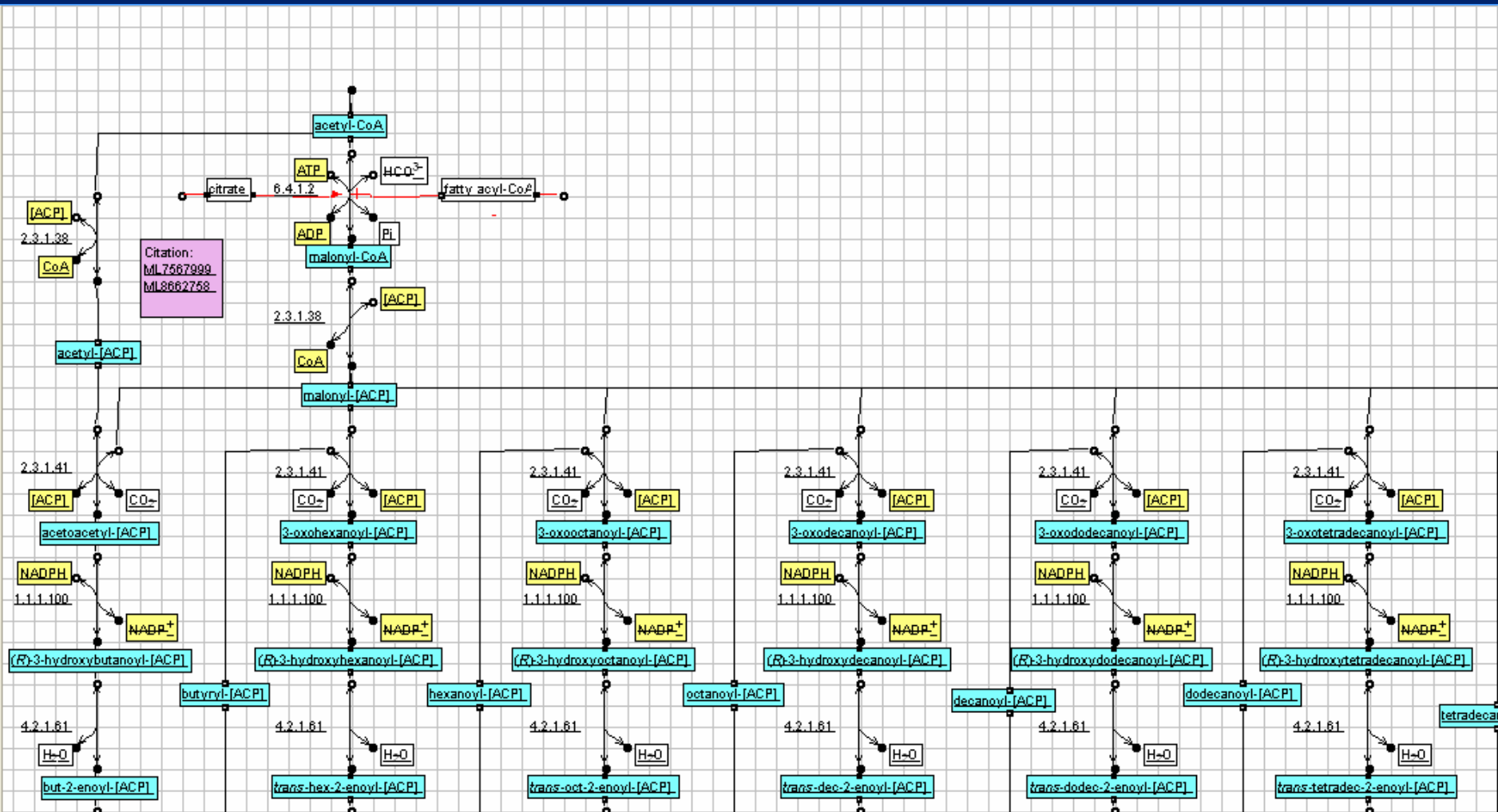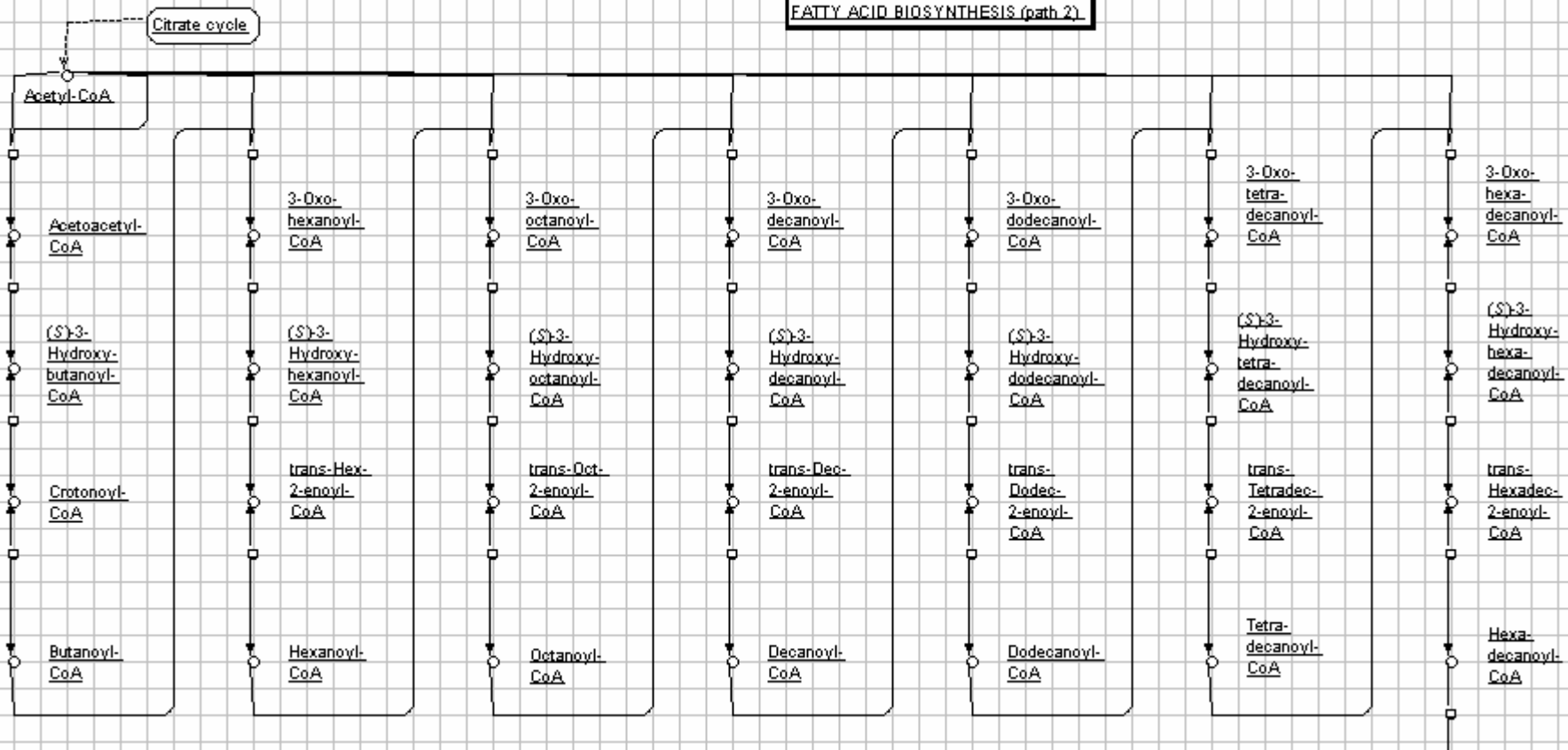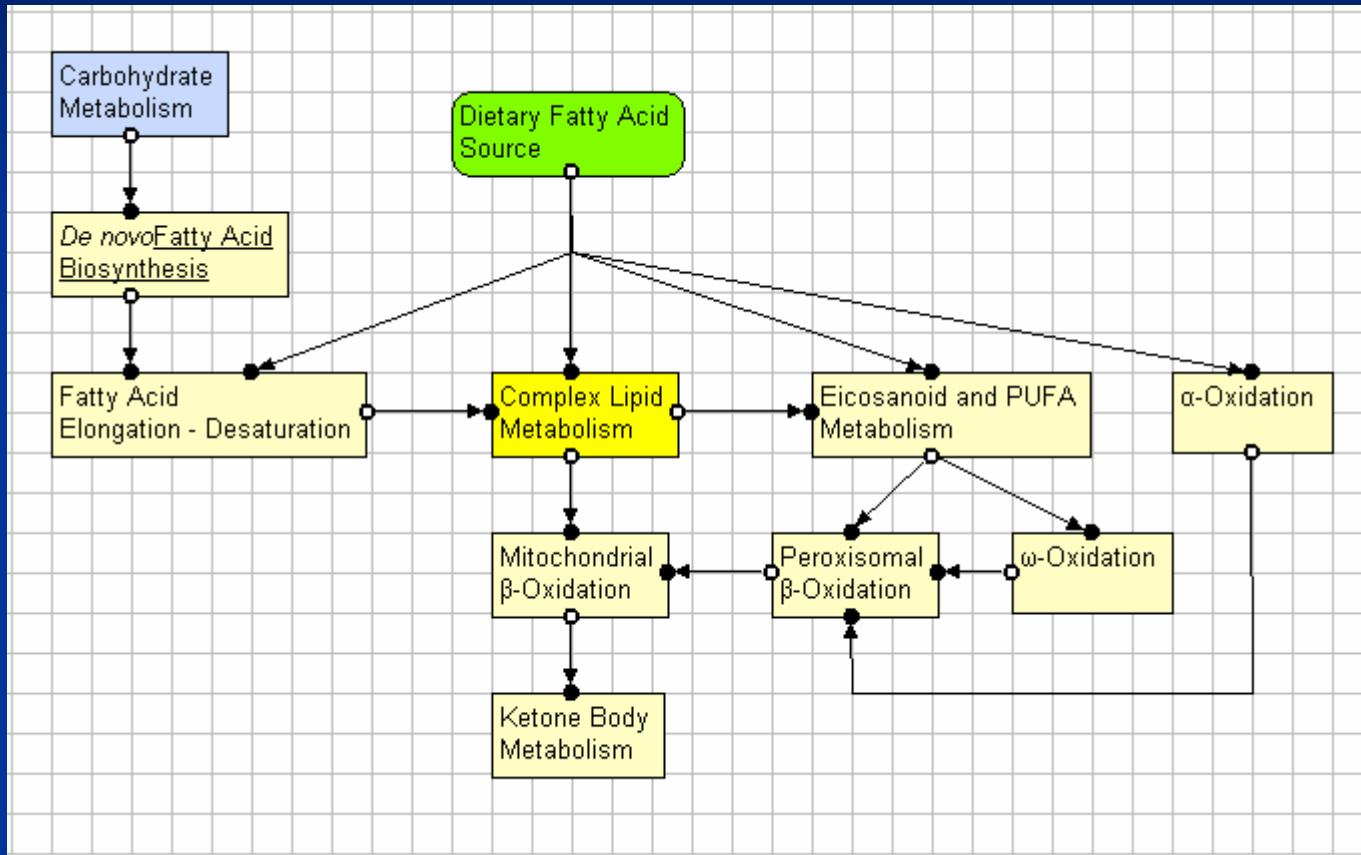
# Visual representation

A. Sorokin
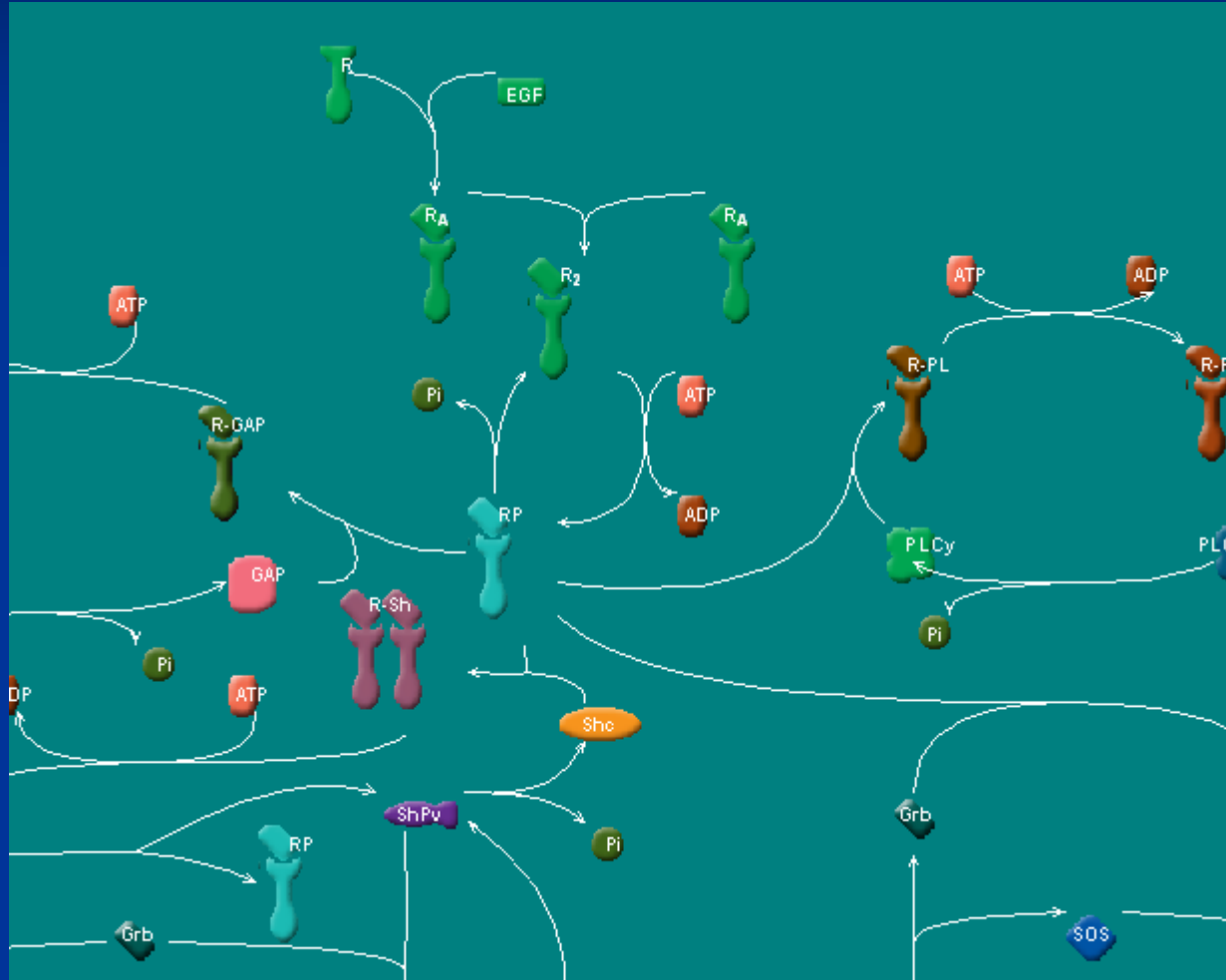
# Development team

- Igor Goryanin, The University of Edinburgh
- Kirill Paliy, EMP Project Inc.
- Alex Selkov, EMP Project Inc.
- Oleg Demin, Moscow State University

www.bioinformatics.ed.ac.uk/epe/