MESI-STRAT

Metabolic Signalling Stratification

Modelling methods in the MESI-STRAT project

Dr Ciaran Welsh

Newcastle University

Overview

- Introduction to the MESI-STRAT project from a high level biological perspective
- An introduction to our use of mathematical modelling to better understand biological systems
- A brief formal description of the mathematical framework I’m using
- A brief look at one of my models in active development
- A discussion of the limitations of my current methodologies
- A discussion of some new software that implements a new way of working that also addresses some of these limitations (non-published)

MESI-STRAT: Metabolic Signalling Stratification of Breast Cancer Patients

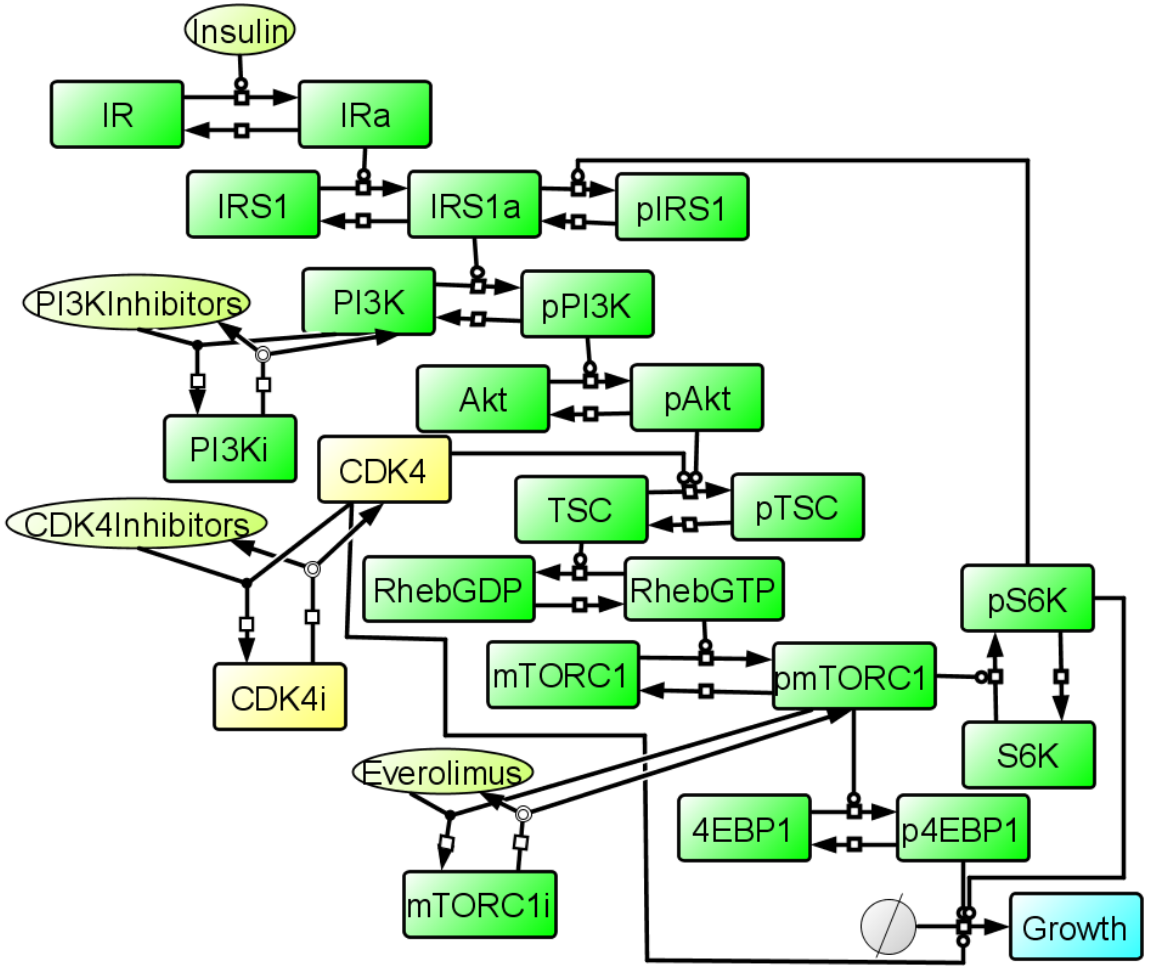
- Breast cancer is a complex disease with mortality rates at approx. 100K per year in the EU alone
- The majority of BC patients (~75-80%) are oestrogen receptor positive (ER+) and are treated with endocrine therapies (ET), which essentially block ER related tumour growth
- ET works, but ~30% of patients eventually relapse with metastatic breast cancer (MBC)
- The mechanisms contributing towards ET resistance broadly include:
  - Loss of ER
  - Mutations in ER
  - Changes in the expression of ER coregulators
- Since the mechanisms leading to ET resistance are still not well understood, many BC patients may be receiving **inadequate or inappropriate treatment**
- The MESI-STRAT project is a large EU funded consortium of researchers collectively aimed at understanding ET resistance from a mechanistic perspective.
  - An emphasis on understanding the complex interactions between metabolism and signalling aspects of the disease

With this better understanding, we aim to stratify BC patients to improve clinical decision making and treatment strategies

MESI-STRAT: Importance of cell signalling

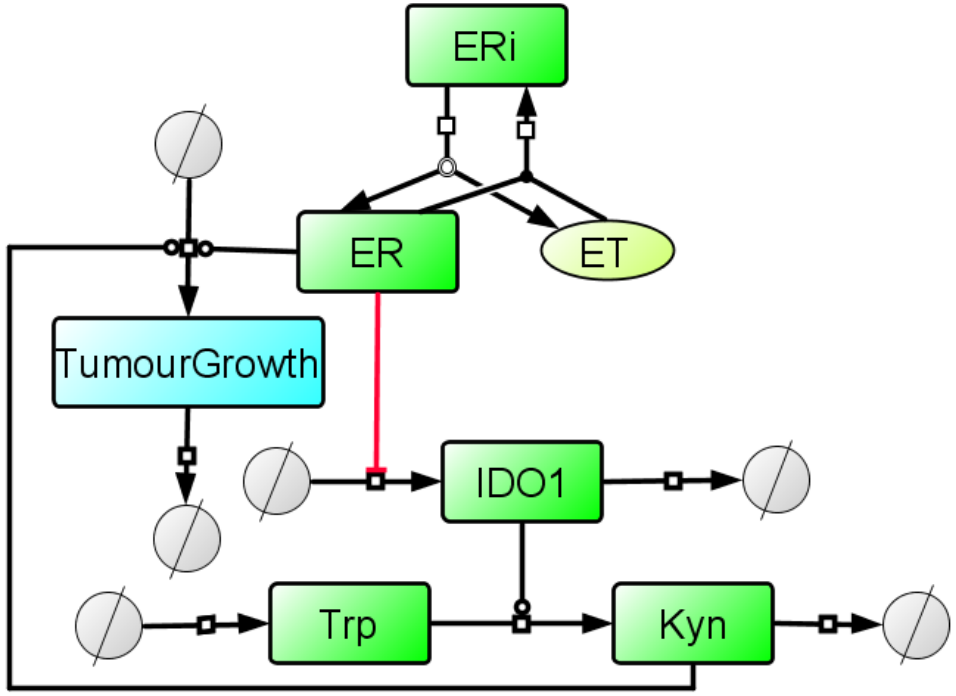
- Given that ER is a signalling molecule, much of the work in understanding ET resistance has focused on cross-talk with other oncogenic signalling pathways
- mTOR and MAPK pathways are two such examples and are thought to be among the drivers of ET resistance
- Activation of mTOR or MAPK leads to enhanced ER activity, both by PTMs and *de novo* synthesis of the ER
- PI3K inhibitors are among the treatment strategies for ER+ BC, aimed at lowering mTOR activity
  - One example of combinatorial therapy is the combination of CDK4 and PI3K inhibitors.
  - CDK4 is a driver of cell cycle but also inhibits TSC2 activity and therefore converges with the PI3K/Akt pathway in mTORC1 activation.
  - Therefore, inhibition of CDK4 leads to sensitisation of BC to PI3K inhibitors.
- An issue in treating ER+ BC is that the contribution of MAPK to ET resistance is not therapeutically explored
  - Not considering MAPK may lead to serious consequences because some drugs (i.e. everolimus, a mTORC1 inhibitor) result in enhanced MAPK activity which may actually exacerbate the disease (perhaps explaining why everolimus has had limited sucess)

CDK4 inhibition can sensitise cells to PI3K inhibition



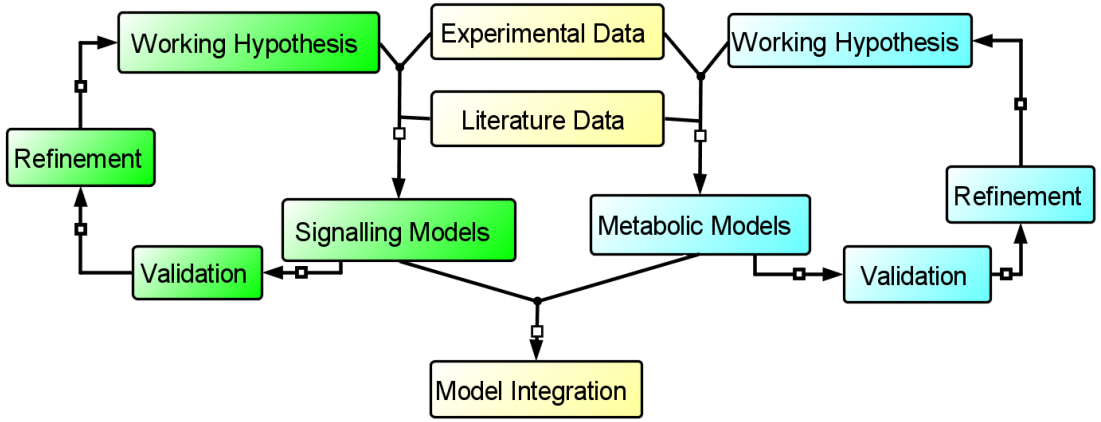
MESI-STRAT: Importance of metabolism

- Tumour metabolism is also an important (but often neglected) player in ET resistance
- Kyurenine (Kyn) is the metabolic product of Tryptophan (Trp), a reaction catalysed by an enzyme called IDO1
- IDO1 is under the regulation of the ER which is blocked by ET
- High Kyn levels have been associated with increased tumour immune evasion and aggression and inhibition of IDO1 have as been positively associated with BC treatment



Hence, one mechanism of ET resistance may involve the unintended enhancement of Kyn along with inhibition of the ER

The MESI-STRAT project: An integrative systems biology project



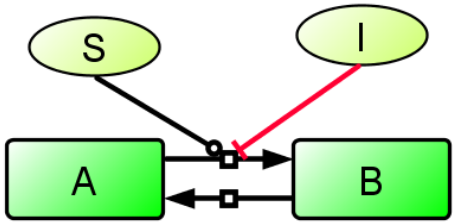
- The regulatory mechanisms that govern cell signalling, metabolism and the interactions between the two are particularly complex
- To help comprehend this complexity, the essential biology is abstracted into mathematical frameworks so that we can simulate biological processes
  - This is the systems modelling approach which sits on the interface between molecular biology, mathematics and computer science
- These models depict our working hypothesis of a biological process and are refined when invalidated, either with our own experiments or by literature findings
- Using these models, we predict the dynamics of key network components in order to understand their behaviour in different biological contexts

Systems modelling methods

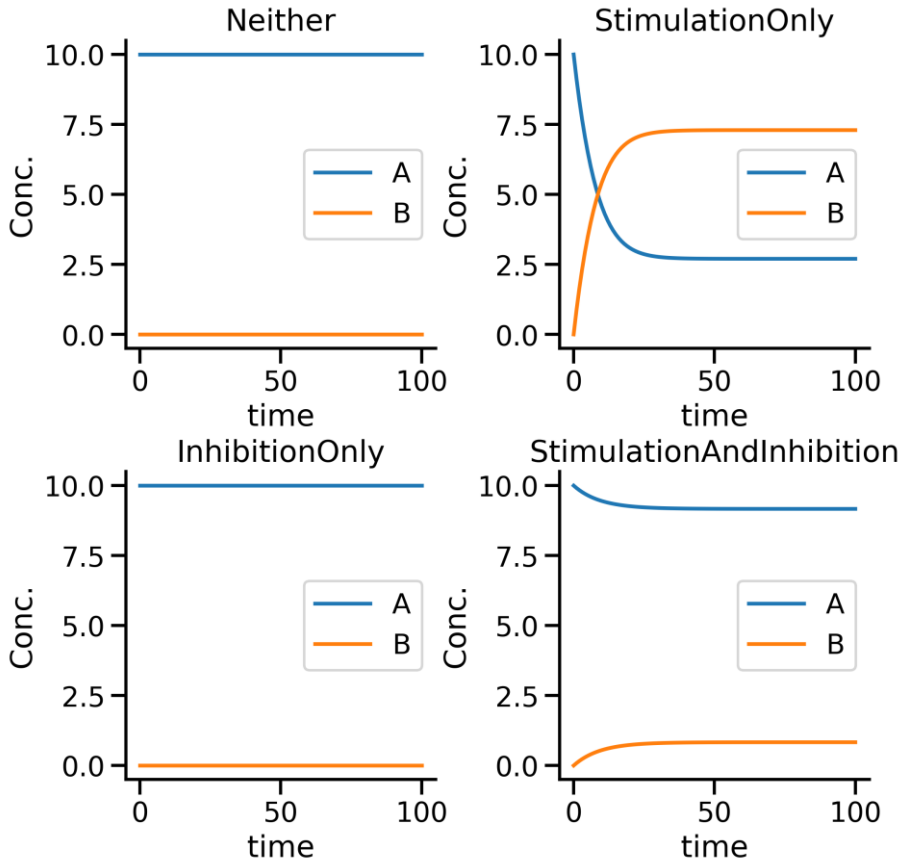
- The goal of a systems modeller in biology is to abstract the essential biology into a mathematical frameworks so that dynamic processes can be simulated
- We aim to recapitulate experimental findings to predict how a system might behave in other biological contexts
- Many simulation frameworks exist including:,
  - ODEs, PDEs, Gillespie, spatial Gillespie, agent based models, cellular automata
  - The choice of model is dependent on your research question
- ODE's are particularly well suited for predicting the dynamics of biological systems because
  - They are quick and easy to built and simulate (i.e. numerical integration)
  - Well developed mathematical methods for parameter optimization and model selection
- Once calibrated, a dynamic model can be perturbed *in silico* to predict the outcome of an intervention

A concrete example

$$\vec{x}_{(t)} = f(\vec{x}, \vec{u}, \vec{p})$$



$$\frac{d[A]}{dt} = -\frac{k_{cat} \cdot S_t \cdot A_t}{k_m + A_t + \frac{k_m \cdot I_{(t)}}{k_i} + k2 \cdot B_t}$$



- PyCoTools is a 3rd party API for a modelling toolbox called COPASI for
- PyCoTools enhances productivity by automating repetitive tasks such as
  - Parameter estimation configuration
    - Plus multiple models for model selection
  - Parallel execution of parameter estimations (local or HPC)
  - Identifiability analysis
  - EDA on parameter estimation data

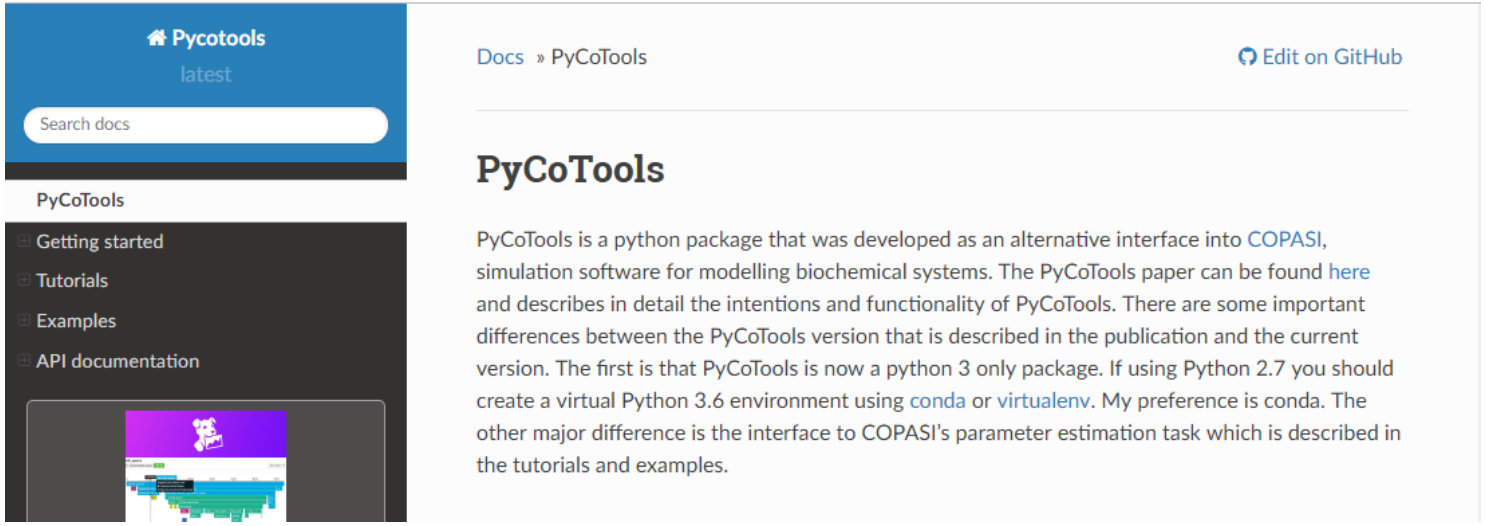
Bioinformatics, 34(21), 2018, 3702–3710
doi: 10.1093/bioinformatics/bty409
Advance Access Publication Date: 22 May 2018
Original Paper

OXFORD

Systems biology

PyCoTools: a Python toolbox for COPASI

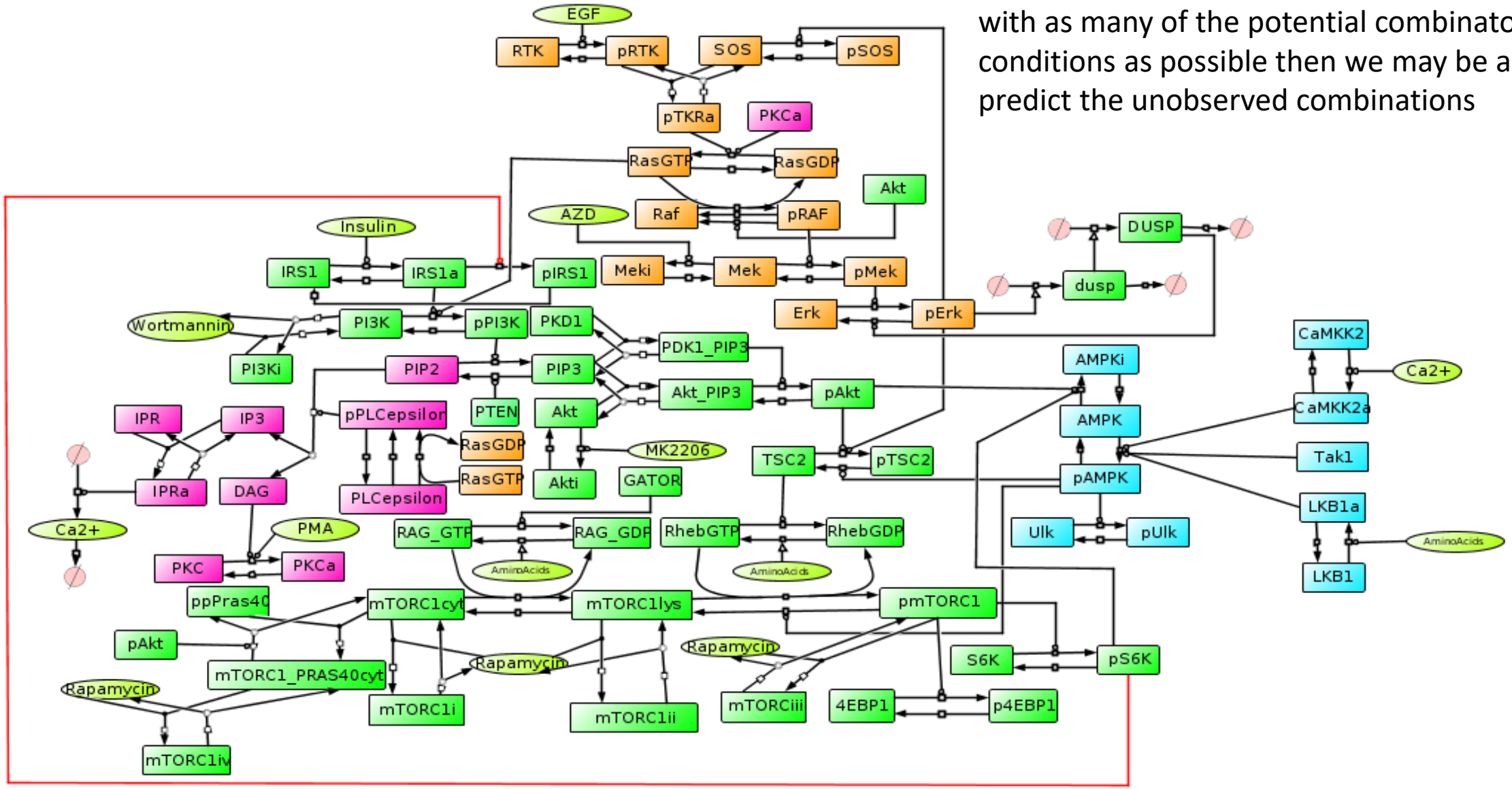
Ciaran M. Welsh[1], Nicola Fullard[3], Carole J. Proctor[2], Alvaro Martinez-Guimera[1], Robert J. Isfort[4], Charles C. Bascom[4], Ryan Tasseff[4], Stefan A. Przyborski[3,*] and Daryl P. Shanley[1,*]



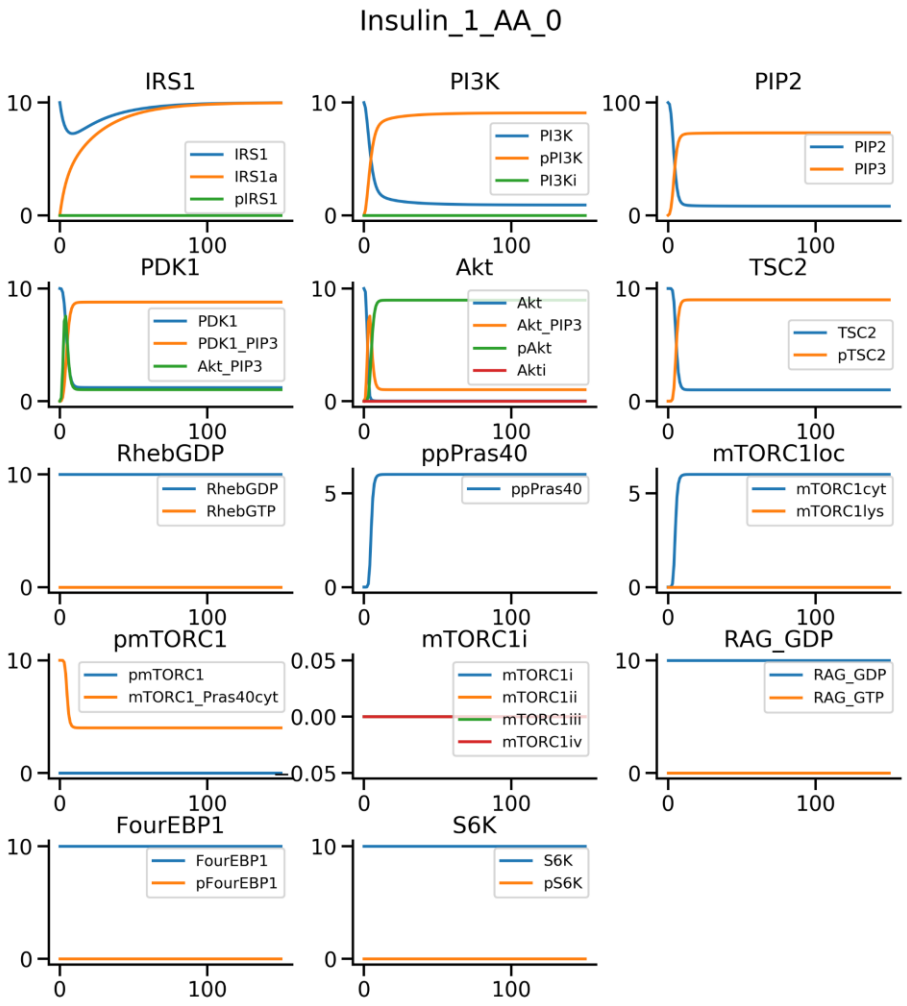
Modelling the interplay between PI3K and MAPK

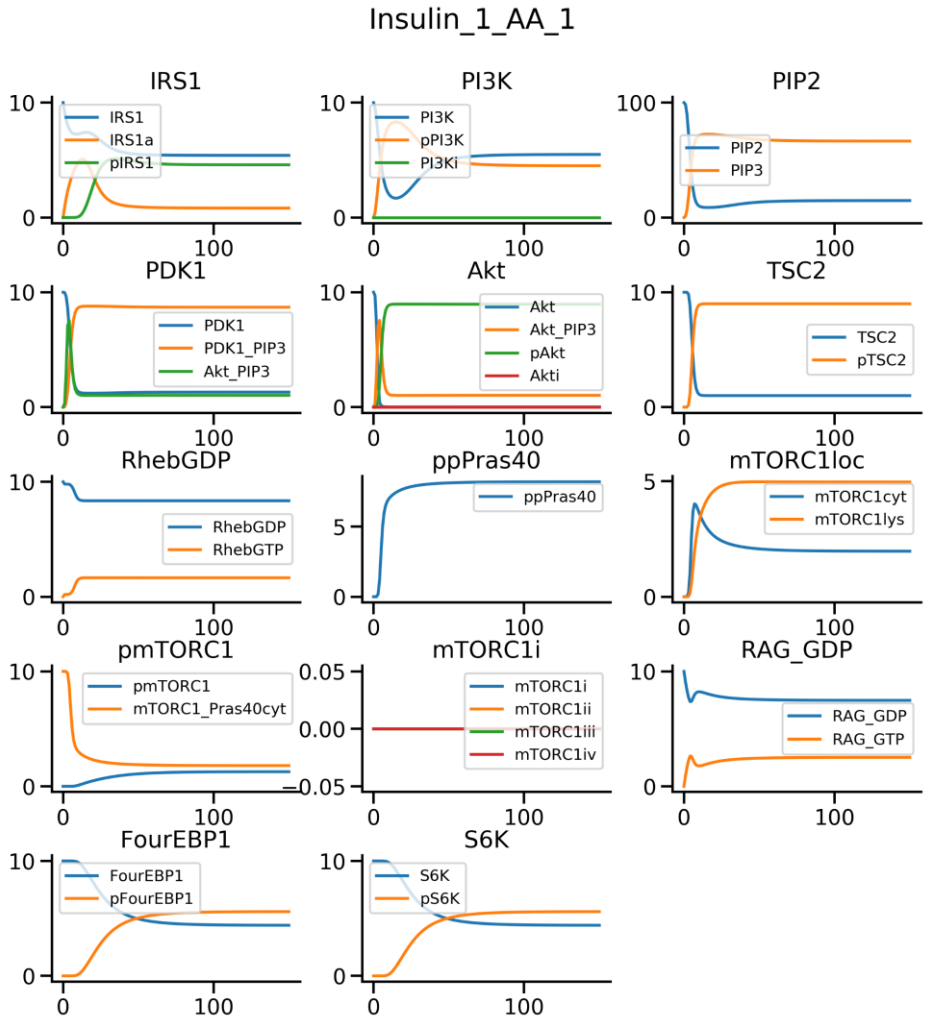
This model has 9 inputs

If I can produce dynamic behaviour consistent with as many of the potential combinatorial conditions as possible then we may be able to predict the unobserved combinations



Simulation of Insulin +/- Amino acids





The model reflects the biology in that insulin only stimulates mTORC1 activation in the presence of ample amino acids

Some limitations and difficulties often encountered in systems modelling

- A serious limitation of the ODE framework is that a properly informed model requires a lot of good quality data which often isn't available
- The result is that of non-identifiability: a situation where some combinations of the parameters cannot uniquely be determined with the defined optimization problem
- Some techniques to mitigate the non-identifiability problem include bootstrapping or interpolating the data, reducing the complexity of the model and the gold standard - go get more data
- Some other difficulties with ODE modelling:
  - Requirement for normalisation of relative data from different sources before use in parameter estimation
  - Whether to use individual repeats or averages of time points
  - Use of nuisance parameters such as scale and offset parameters
  - What to use for initial conditions?
  - How to deal with non-observed network components?
  - Which optimisation methods to use?
  - How to weight your objective function
  - Tuning algorithm settings to strike a balance between speed and convergence
  - EDA on parameter estimation data for understanding the landscape of the optimization problem (PyCoTools)
  - Effective use of HPC for parameter identification and model selection
  - Whether to use validation data for early stopping criteria (to prevent overfitting)

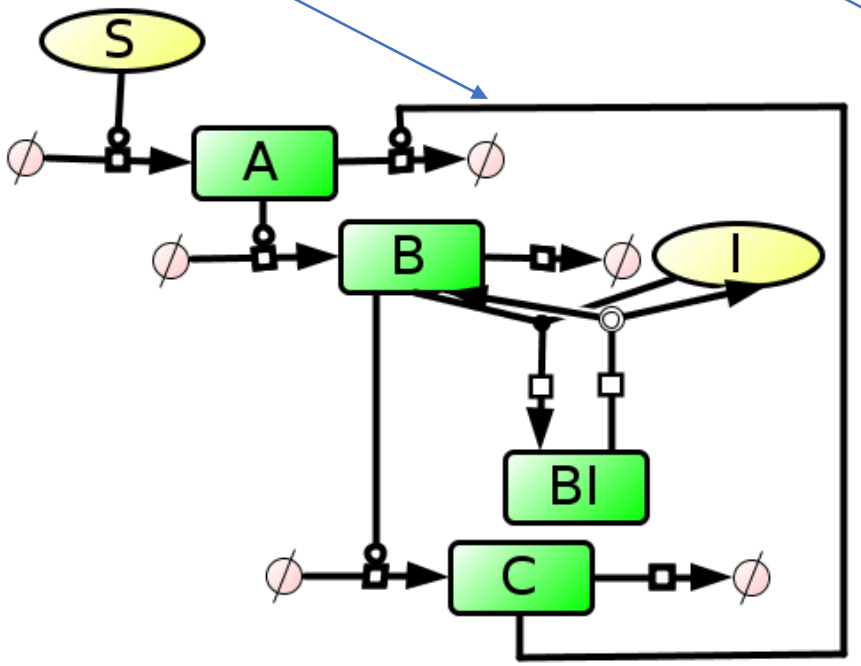
Qualitative Model Fitting: A focus on qualitative model behaviour

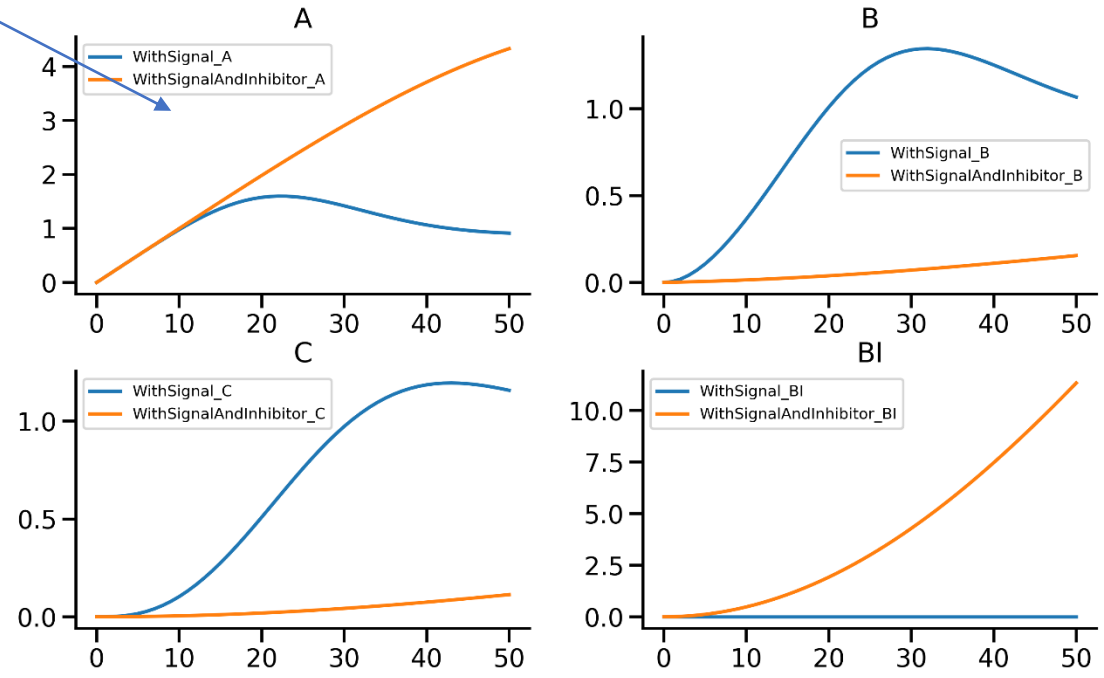
- A normal systems modelling cycle includes a literature review, network design, wet lab experiments, parameter optimisation, identifiability analysis, model refinement and model selection.
- However, by focusing on model behaviour, rather than finding the parameters, we can omit the mathematical rigour of optimization and identifiability analysis in favour of ensuring the behaves well in accordance with data and literature (i.e. the model is well validated)
- This is the focus of a software package I'm currently developing (called QualitativeModelFitting (qmf) on my github)
- Qmf is essentially a unit testing framework for validation of ODE models in systems biology.
- Qualitative observations (i.e. from data or literature) are encoded the syntax of a (yet unnamed) language I invented
- The language is interpreted by a pure Python lexer and parser. TestCase classes and methods are dynamically generated at run time to test the observation
- Hence, an arbitrary number of qualitative observations can be validated simultaneously and highlight important discrepancies with literature findings.
- The package has two font ends:
    1. A manual interface for testing observations and manually tweaking the model
    2. An automatic interface that locates ensembles of parameter sets that are consistent with the observations
        - In active development. Involves designing an objective function that will be
            1. Minimized using standard optimization algorithms (such as hill climbers or evolutionary strategies)
            2. Maximized by a reinforcement learning algorithm (Q-learning based)

Qualitative Model Fitting: A toy model

```
MODEL2 = """
model SimpleFeedback()
   compartment Cell = 1;
   var A in Cell;
   var B in Cell;
   var C in Cell;
   const S;
   const I;

   A = 0;
   B = 0;
   C = 0;
   S = 0;
   I = 0;
   BI = 0;

   k1 = 0.1;
   k2 = 0.1;
   k3 = 0.1;
   k4 = 0.1;
   k5 = 10;
   k6 = 0.1;
   k7 = 0.1;
   k8 = 0.1;

   R1: => A          ; Cell * k1*S;
   R2: A =>          ; Cell * k2*A*C;
   R3: => B          ; Cell * k3*A;
   R4: B =>          ; Cell * k4*B;
   R5: B + I => BI    ; Cell * k5*B*I;
   R6: BI => B + I    ; Cell * k6*BI;
   R7: => C          ; Cell * k7*B;
   R8: C =>          ; Cell * k8*C;
end
"""
```

This string defines a set of ODEs representing this network and produces these predictions for the two conditions





Qualitative Model Fitting: specifying model conditions and some observations

Conditions definition

Condition names (arbitrary)

```
inputs = dict(
    WithSignal=dict(
        inputs=dict(
            S=1
        ),
        obs=[
            'A@t=20 > A@t=10',
            'A@t=30 < A@t=20'
        ]
    ),
    WithSignalAndInhibitor=dict(
        inputs=dict(
            S=1, I=1
        ),

        obs=[
            'A@t=20 > A@t=10',
            'A@t=30 < A@t=20'
        ]
    )
)
```

Observations

Interpretation of observations :
With stimulation by S, A should be increasing between t=10 and t=20 and decrease again at t=30

With stimulation by A and inhibition by I, the same conditions should hold.

Qualitative Model Fitting: A simple front end

Input:

Antimony string

Import the manual API

```
from qualitative_model_fitting import manual_interface
results = manual_interface(MODEL2, self.input, 0, 50, 51)
print(results.to_df())
```

Returns a dict like results object

Start, stop and step integration parameters

Inputs and observations

Output:

```
"""
                          observation  truth
WithSignal             0  A@t=20 > A@t=10   True
                       1  A@t=30 < A@t=20   True
WithSignalAndInhibitor 0  A@t=20 > A@t=10   True
                       1  A@t=30 < A@t=20  False
"""
```

The ‘to_df()` method
Turns the results into a standard pandas dataframe.

Key point: we now know that the last condition doesn’t hold without the need for looking at the simulations. For simultaneous model validation of hundreds of observations, this is particularly valuable

To summarise

- The MESI-STRAT project is dealing with some very complicated biology
- My role in this project is to piece together the relevant biological interactions in a mechanistic model, using in-house and literature sourced data
- I have developed a new unit testing framework for validating model predictions that addresses a bottleneck in the development of mechanistic models
- Therefore, this software has the potential to potently enhance productivity within the systems biology community

Thank you for listening

Ciaran Welsh