Glucocorticoid receptor interactome

Emyr Bakker[1], Kun Tian[2], James Andrews[2], Constantinos Demonacos[2], Jean-Marc Schwartz[2*] & Marija Krstic-Demonacos[1*]

[1]University of Salford, Salford, UK [2]University of Manchester, Manchester, UK.

*Corresponding Authors: Email: m.krstic-demonacos@salford.ac.uk Email: jean-marc.schwartz@manchester.ac.uk

Introduction

Glucocorticoid hormones are used in the treatment of a variety of diseases, due to their diverse range of effects. Acute lymphoblastic leukaemia (ALL), the most common form of childhood cancer, is one such disease treated by glucocorticoids (GCs). Although there has been much success in the treatment of ALL with GCs, drug resistance remains a problem, despite the accumulated knowledge of the signalling networks underlying the GC actions through the glucocorticoid receptor (GR). Systems biology offers the opportunity to gain a holistic view of protein interaction networks, whilst also allowing for the generation of predictions through *in silico* knockouts which can be validated through conventional laboratory approaches. Systems biology has shown great success in identifying novel signalling pathways. Here we present the preliminary model of the GR interactome, constructed through MATLAB using the add-on CellNetAnalyzer, and visualised through Cytoscape. The model currently consists of 56 nodes representing genes and proteins that interact with the GR, with the glucocorticoid hormone as an input node. During curation tissue-specific reactions were built into the model, leading to nine forms of the model each reflecting a cell type, with a tenth reference model consisting of all interactions. Currently there are 87 interactions in the model overall. These interactions were extracted automatically from the STRING database using the UltraEdit text editor, and then manually curated by literature mining to remove false positives. The curated interactions were imported into CellNetAnalyzer and the dependency matrices were generated, showing 2304 dependencies. These dependencies changed depending on the cell type. Future work will be directed towards expanding and connecting this model to drugs or processes like apoptosis, to detecting potential therapeutic targets, whilst also validating it through analysis of microarray data from ALL patients, paving the way for personalized treatment.

The GEB057 logical model of the glucocorticoid receptor interactome

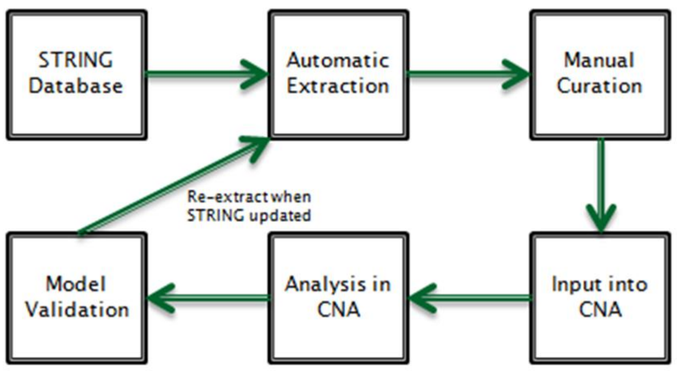

Figure 1: Workflow diagram of model construction

The STRING database lists known and predicted protein-protein interactions. These interactions were downloaded from the protein actions file on STRING, and all interactions involving the glucocorticoid receptor were extracted using the search function on the UltraEdit text editor. Interactions were then filtered by retaining only high confidence scores (greater than or equal to 0.7) and confirming the predictions through extensive literature searching. CellNetAnalyzer, a MATLAB package, was used to simulate the network, whilst also allowing for the possibility of further *in silico* analysis. Visualisation of the network was undertaken using Cytoscape, an open source program.

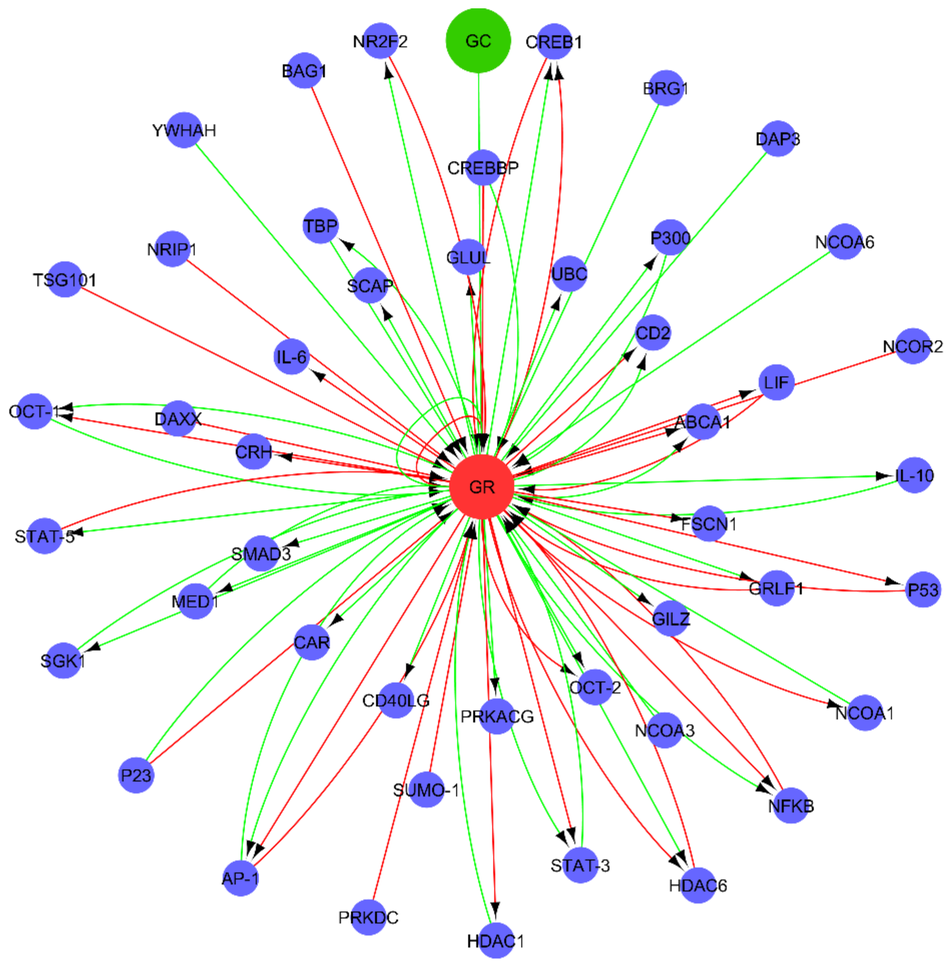

Figure 2: Cytoscape visualisation of the GEB057 model

The GEB057 model above contains 57 nodes (proteins or inputs such as a glucocorticoid) and 87 reactions. Inhibition reactions are shown as red directed arrows, whilst activations are shown as green directed arrows. All interactions are to and from the glucocorticoid receptor. The input node, a glucocorticoid, is highlighted in orange whilst the central protein (glucocorticoid receptor) is highlighted in red.

Cell Type Specificity

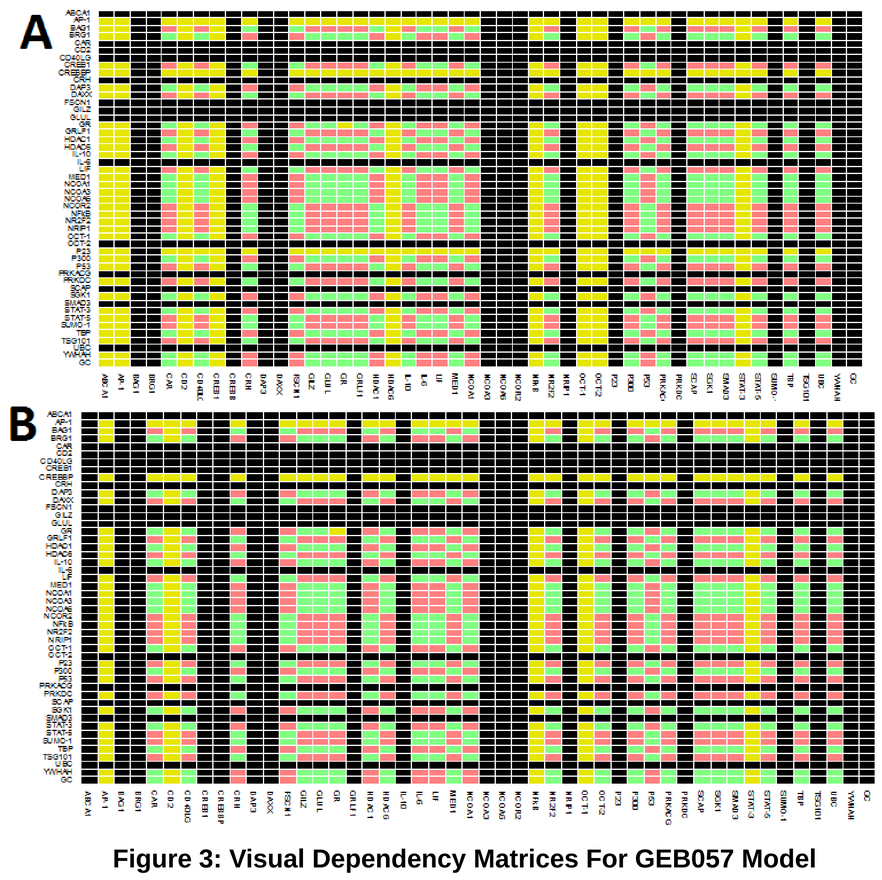

Figure 3: Visual Dependency Matrices For GEB057 Model

Due to the diverse range of effects of the glucocorticoid receptor, cell type-specificity is hugely important in order for the model to be reliable and robust. The model currently consists of nine cell types with ten forms of the model overall: Reference (containing reactions from all tissue types); Bone; EDL Muscle; Embryonic Development; Liver; Lymphoid; Neuronal; Placental; Skeletal Muscle and SOL Muscle. Figure 3A shows the dependency matrix for the Reference form of the model, whilst Figure 3B shows the dependency matrix for Neuronal tissue. In both cases, black represents no effect of Gene A (Vertical) against Gene B (Horizontal), yellow represents an ambivalent relationship (both activating and inhibiting influences), light red represents a weak inhibition, light green represents a weak activation, dark red represents strong inhibition and dark green represents strong activation.

Cell/Dependency Type	No Effect	Ambivalent Factor	Weak Inhibitor	Weak Activator	Strong Inhibitor	Strong Activator
Reference	1182	379	367	376	0	0
Bone	1408	207	345	344	0	0
EDL Muscle	1408	207	345	344	0	0
Embryonic Development	1440	227	315	322	0	0
Liver	1376	209	360	359	0	0
Lymphoid	1376	239	345	344	0	0
Neuronal	1347	214	367	376	0	0
Placental	1380	212	360	352	0	0
Skeletal Muscle	1440	235	315	314	0	0
SOL Muscle	1408	207	345	344	0	0

Table 1: Number of each dependency for each form of the model

Conclusions & Future Work

- We have developed a methodology for the generation of reliable large-scale models and introduced a novel mechanism for cell type-specificity in the model.
- Analysis of the dependency matrices reveals cell type-specific changes in the relationship between genes. It is expected that following the generation of the second layer (interactions between the nodes in the primary layer) further differences in the dependency matrices will be unveiled, possibly generating cell type-specific predictions of changes in regulation of gene expression.
- Future work includes generating outputs by linking the model to cellular processes such as apoptosis, metabolism or inflammation through the Gene Ontology database.
- Logical steady state analysis (LSSA) will then be undertaken to follow the signalling events downstream of the input node (GC → GR), to see the results of which output(s) will be activated.
- *In silico* gene knockouts will then be performed to see changes in the dependency matrix and in LSSA results. This will provide predictions which can then be validated using text mining and wet laboratory approaches.