Drug Repositioning by Merging Gene Expression Data Analysis and Cheminformatics Target Prediction Approaches

Presented By: Yasaman Kalantar Motamedi

Supervised By: Dr. Andreas Bender

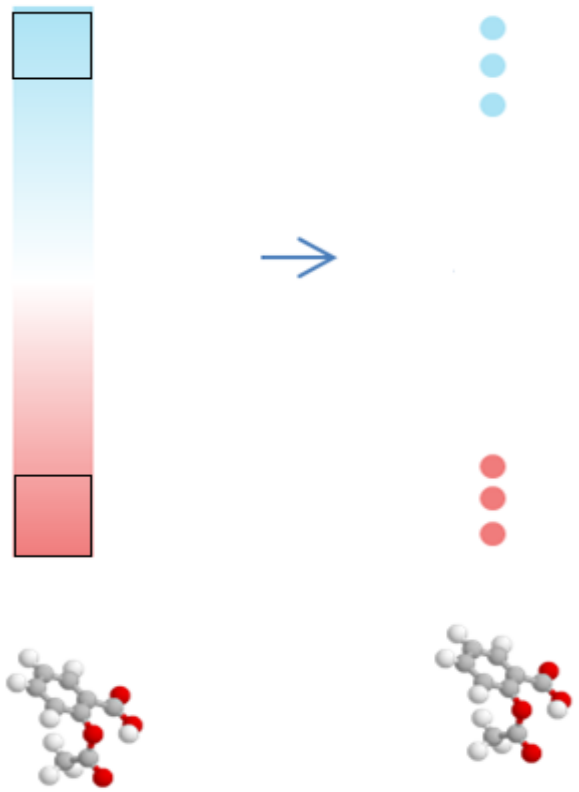
KalantarMotamedi Y,..,Bender A. Drug repositioning based on gene expression data analysis and cheminformatics target prediction identifies small molecules for the targeted differentiation of stem cells into cardiomyocytes, 2014

Introduction

- Predicting compounds for
  - Repurposing drugs: finding new indications for FDA approved drugs
  - Differentiating stem cells to different organs
- Computational approaches:
  - Bioinformatics: Gene expression
  - Cheminformatics: In-silico target prediction

LINCS and Connectivity Map Databases

- CMap
  - 1300 compounds
  - 3 cell lines: MCF7, HL60, PC3
  - Durations and Concentrations
- LINCS
  - 1000 Landmark genes
  - 20,413 Small-molecules
  - ~1,300 FDA-approved drugs
  - 18 cell lines
  - Durations and Concentrations

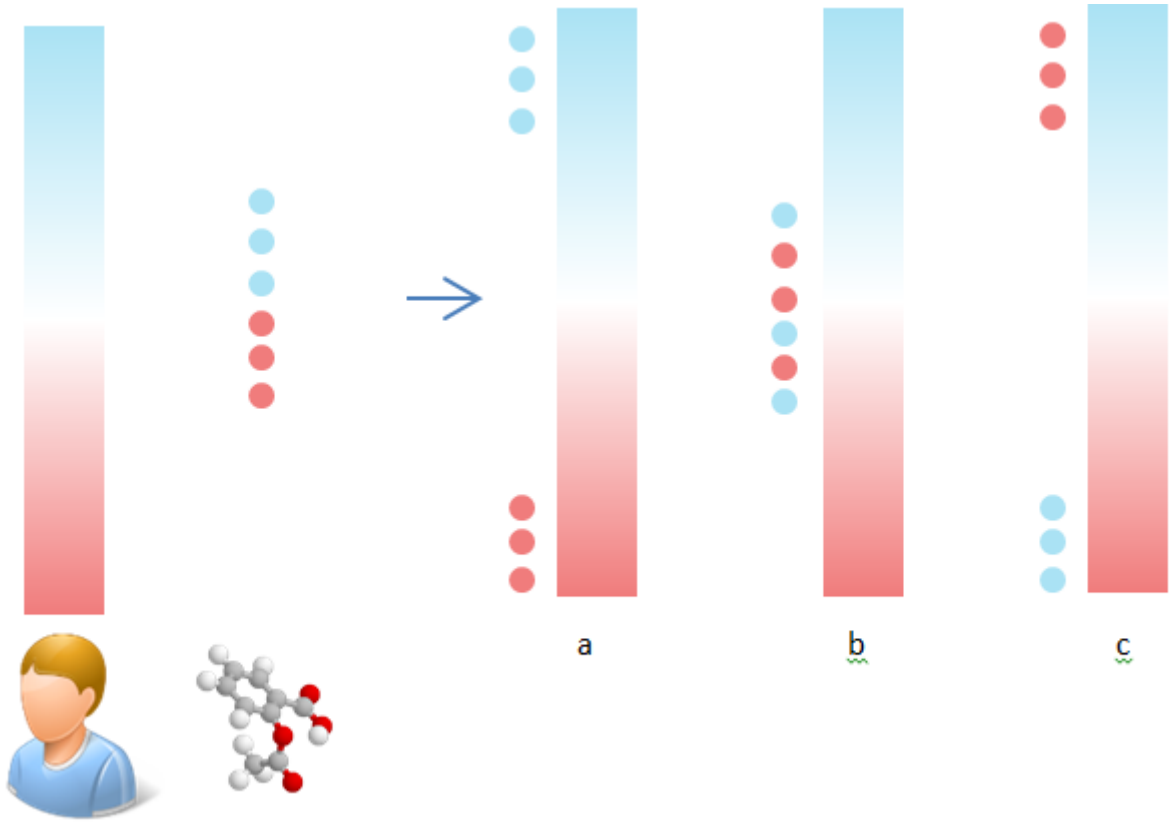


Data Extraction from GEO

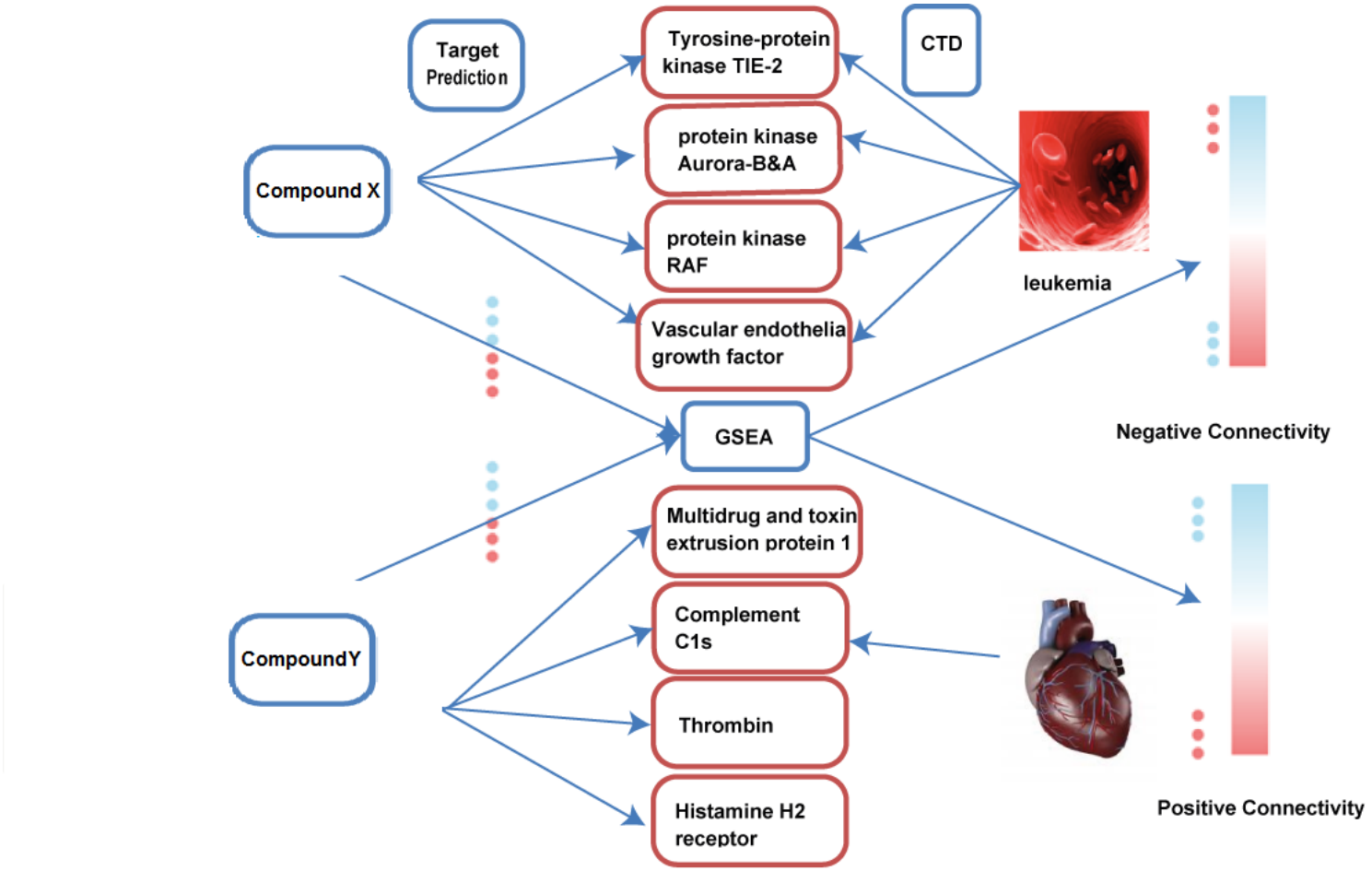
- 661 disease mesh Terms
- 453 disease vs control samples
- 181 GEO databases

- Gene pattern tool
- Unified Format:
  - samples, arrays, organisms were extracted automatically
- Samples Labelling:
  - disease or control based on text mining approaches
- Mesh terms: Disease specification

Gene Set Enrichment Analysis

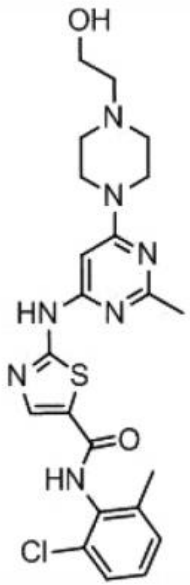


Combination of Gene expression and Cheminformatics approaches



Target Prediction algorithm

Molecule



Chemogenomic data

Ligand1-Target1
Ligand2-Target2
Ligand3-Target3

MODELS

Targets predicted	Bayes Scores
Tyrosine-protein kinase Lyn	79.4
Tyrosine-protein kinase LCK	76
Tyrosine-protein kinase FYN	33
Tyrosine-protein kinase YES	26
Nerve growth factor receptor Trk-A	23
Tyrosine-protein kinase SRC	21
Tyrosine-protein kinase BTK	21
Serine/threonine-protein kinase Aurora-A	20
Tyrosine-protein kinase EGR	19
Serine/threonine-protein kinase Aurora-B	19
Tyrosine-protein kinase BLK	18.3
Tyrosine-protein kinase JAK3	18
Acyl-CoA desaturase	17.4
Tyrosine-protein kinase ABL	14.77
Tyrosine kinase non-receptor protein 2	14.477

Bioactivities ?

Biological hypothesis

Training data

- Training compound-target pairs extracted from Chemble v.17
  - Extracted targets with binding affinity less than 10 µM
  - confidence level of 9 or 10
- Leaving us with
  - 385,126 compound-protein pairs
  - 1,643 distinct proteins
  - 226,791 unique compounds

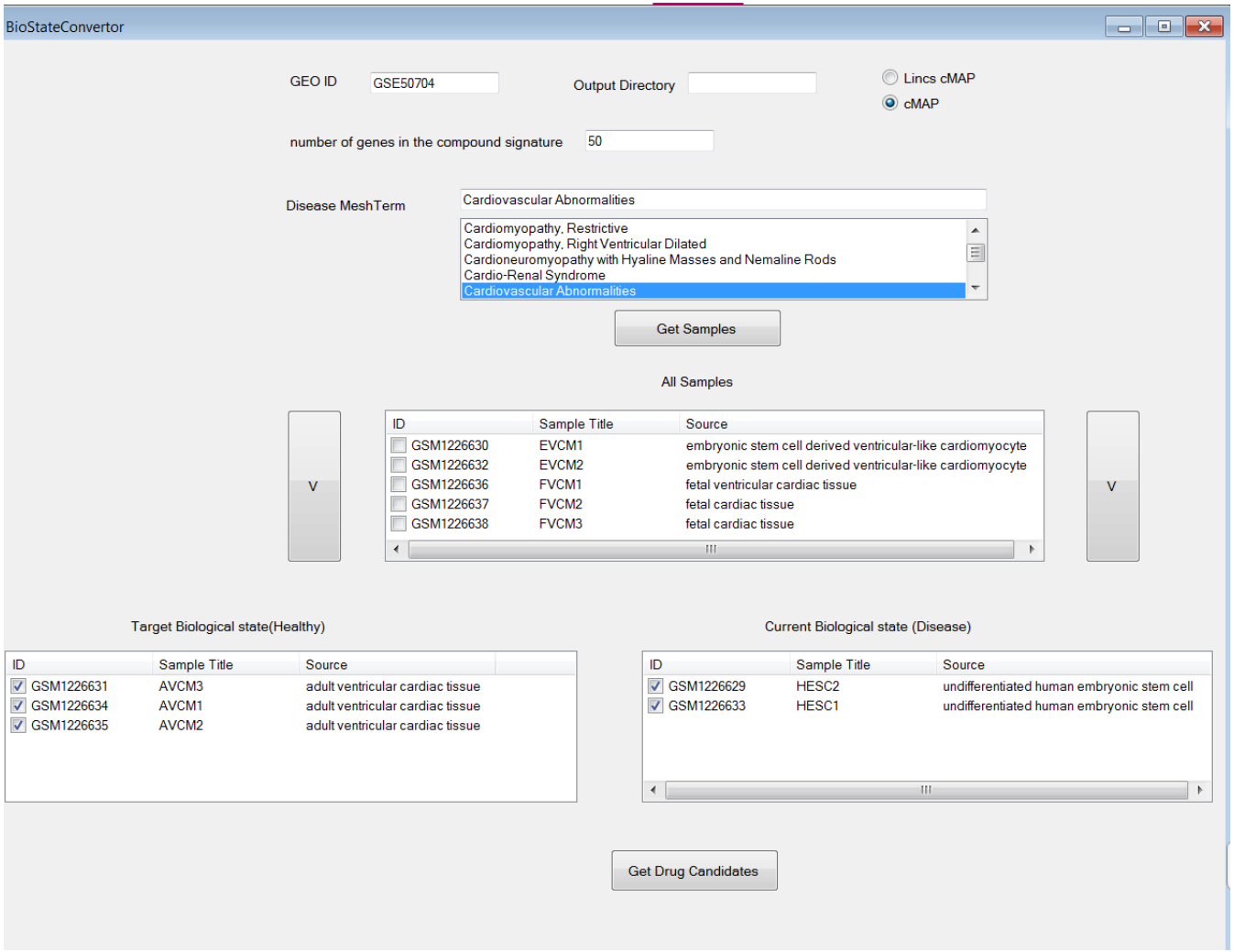
Target Prediction algorithm

- Standardized using ChmAxon Standardizer
- ECFP4 fingerprints were generated using the JChem package of ChemAxon
- Laplacian Modified Naïve Bayes
- Cross validation

  Recall:

  – 74% in top 3 positions

BiostateConvertor: Heart Differentiation application

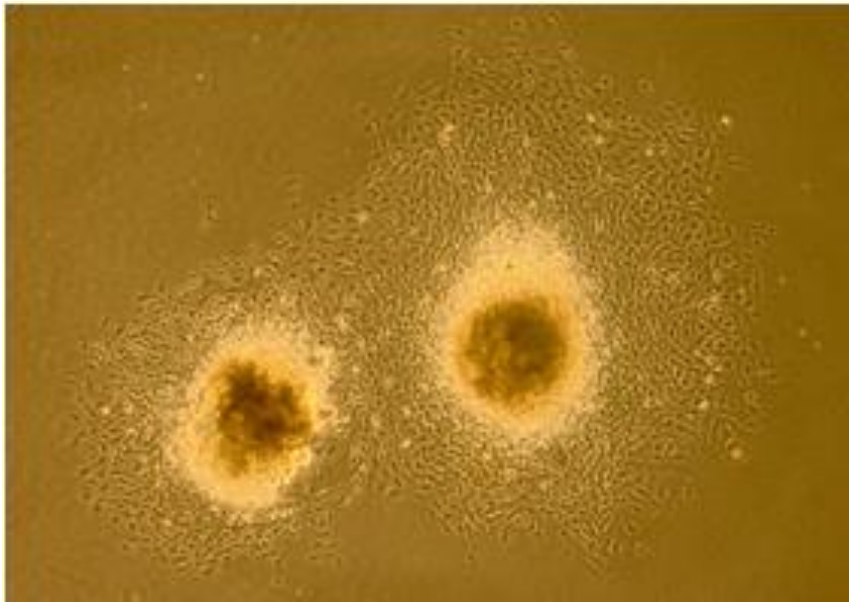


Retrospective validation for differentiation of ESC to cardiomyocytes:

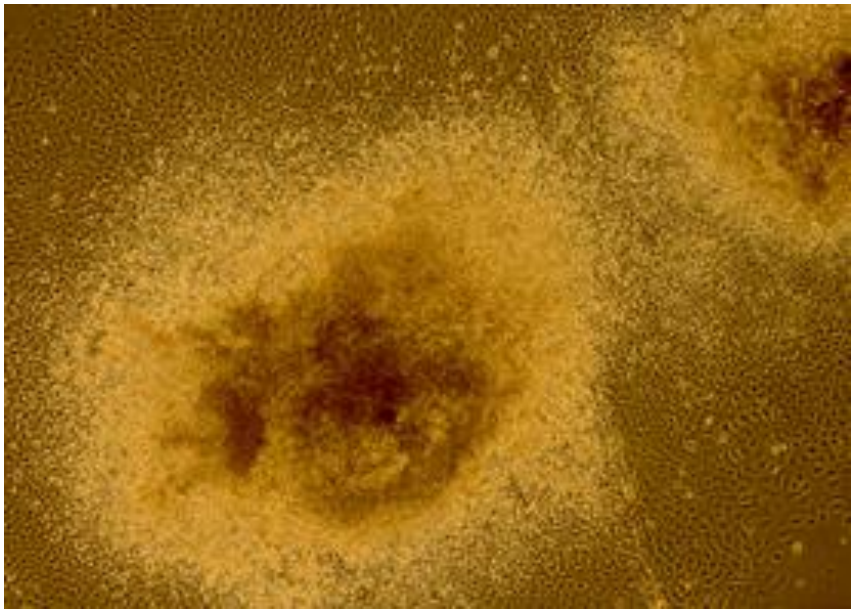
- Meglumine(rank 9):
  - Differentiated BMSC to cardiomyocytes
- Purmycin(rank 3):
  - Differenttiated ESC to pure cardiomyocytes (>99%)
  - Transplanted to the infarcted heart of mouse and improved its function

Prospective validation of compounds: Morphological changes of EBs

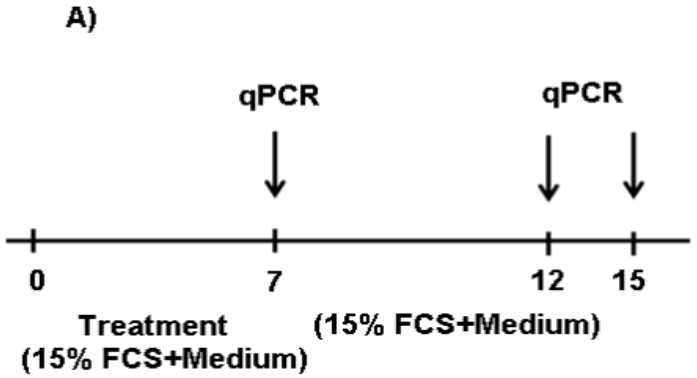
Control + DMSO Day10

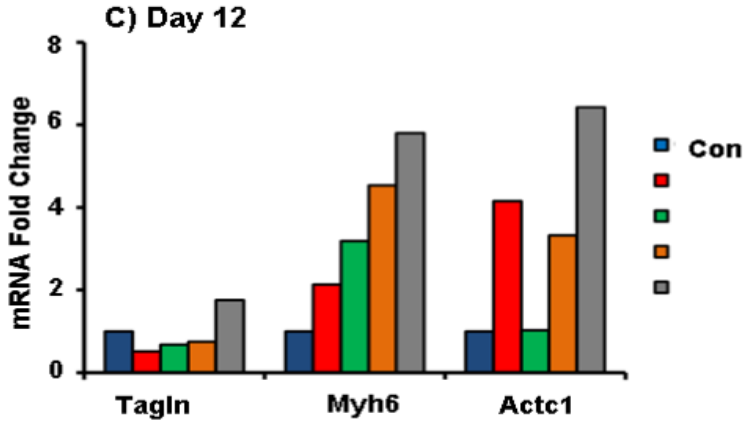


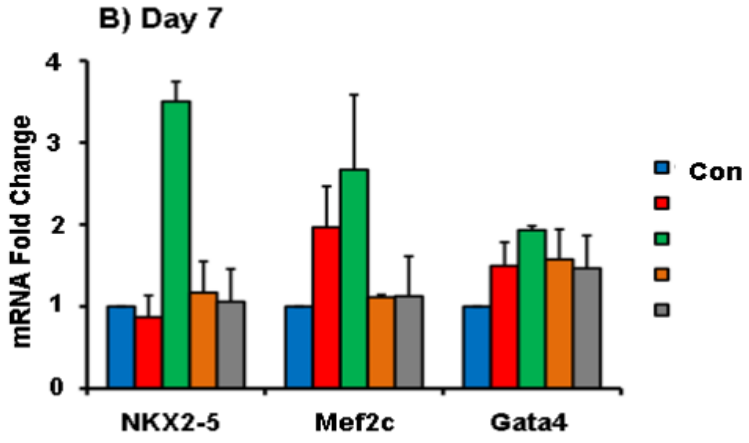
Compound X Day 10

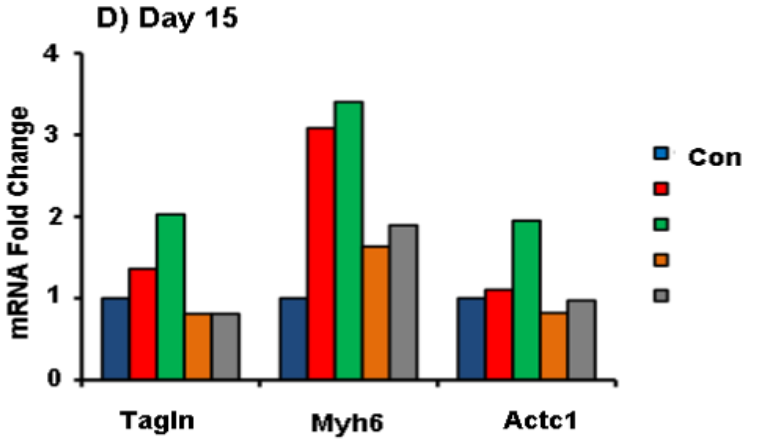


Expression of cardiac precursor and cardiac markers









18/30 of top predicted compounds were supported by literature

- Breast Cancer
  - Tamoxifen: (rank 4 in CMap):
    - Current Therapy
  - Ouabain, Proscillaridin and Digoxin(Top in CMap and Lincs):
    - Inhibited DNA Topoisomerases II and I
    - Inhibited the growth of MCF7 cells

18/30 of top predicted compounds were supported by literature

Breast cancer(continued)

– Niclosamide(rank 3 in LINCS):

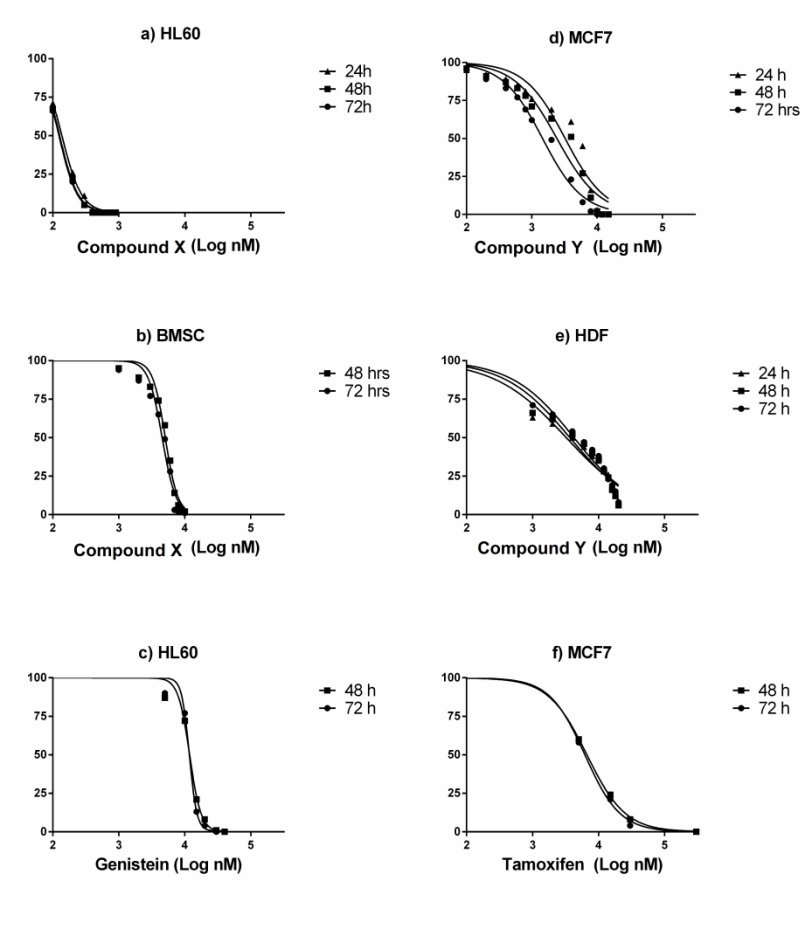
- Induced apoptosis in breast cancer cell lines *in vitro*
- Supressed tumor growth *in vivo*
- Prevented metastases formation

Leukemia

– Thioridazine (rank 11 in CMap):

- Inhibited proliferation of leukemia cells
- Induced apoptosis in them
- Did not affect normal lymphocytes

Prospective validation of cytotoxicity of predicted compounds



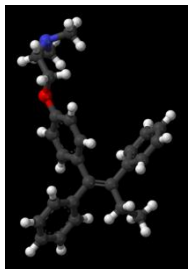
Conclusion

- 18/30 of top predicted compounds were supported by literature
  - For breast cancer and leukemia
- We prospectively validated novel active small molecules
  - For breast cancer and leukemia
  - For differentiating stem cells to cardiomyocytes
- The computational approach is applicable to any other biological state where gene expression data is available

Current computational Directions

- Using Pathway information
- Predicting Synergism and Antagonism for compound combinations

Pathway Based Approach



Tamoxifen

Cell Cycle

Innate Immune System

Pathways in cancer

Glutamatergic synapse

Chemokine signalling pathway

Steroid hormones

Breast Cancer

Future biological directions

- Predicting active small molecules for
- Pancreatic cancer
- Personalized medicine: Sarcoma in children
- inducing transdifferentiation of fibroblasts to oligodandrocytes (finding therapies for MS)
- inducing transdifferentiation of endoderm to beta cells (finding therapies for diabetes)
- increasing success rate of cloning in goats

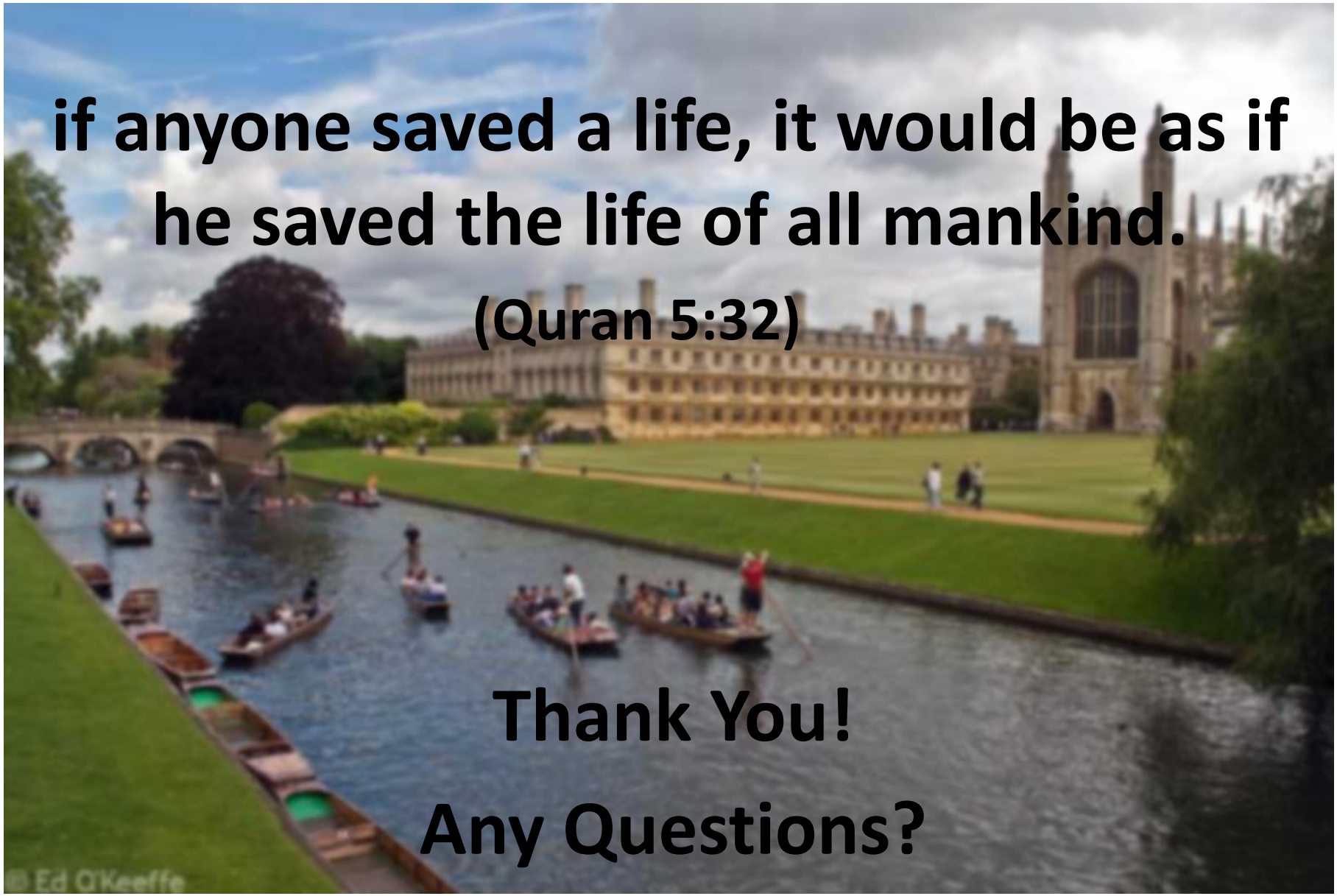
Acknowledgement

- I am grateful to
- Dr. Andreas Bender
  - for funding me through ERC grant
  - For supervising the computational approaches
- Unilever Centre in Cambridge University
  - For funding me for one term
- Dr. MH Nasr-Esfahani
  - For supervising the experimental work

European Research Council
Established by the European Commission

Unilever Cambridge
Centre For Molecular Science Informatics

ROYAN INSTITUTE
Reproductive Biomedicine & Stem Cell


if anyone saved a life, it would be as if he saved the life of all mankind.
(Quran 5:32)
Thank You!
Any Questions?
© Ed O'Keeffe

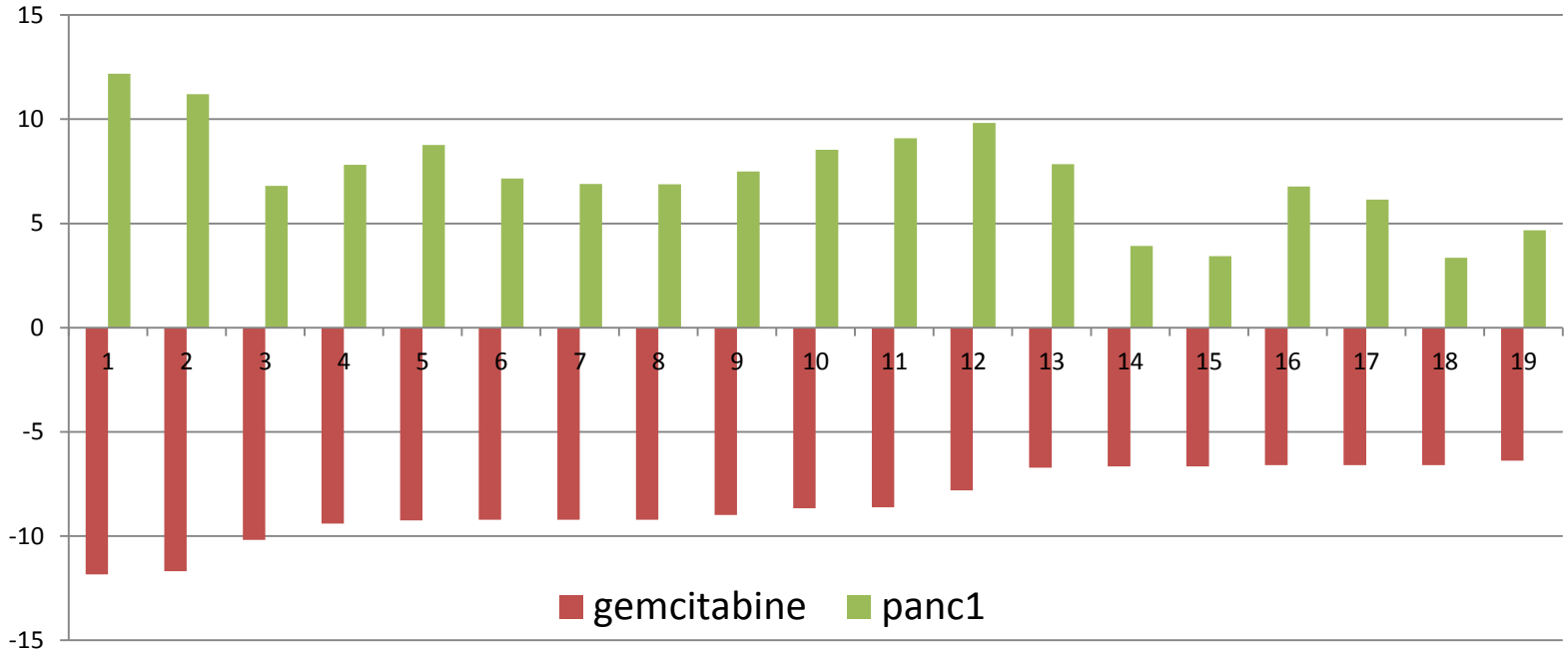
Pathway Based Approach

- BioSystems Pathways
  - 2010 human pathways
- Calculate Enrichment Score for each pathway given drug or disease signature
  - Count number of shared genes between each pathway and gene set of interest
  - Generate 20000 random gene sets as the background
  - Calculate zscore: $z = \frac{x - \mu}{\sigma}$
- Gene Signature of compounds -> pathway Signature
- Gene Signature of disease -> Pathway Signature

Pathway Based Approach

- Find compounds that can enrich the most important pathways in the disease in a reverse way
- Find top x enriched pathways in the disease
- Generate rank order list of compounds using:
  - Correlation
  - Euclidean

anti-correlation of pathway signatures

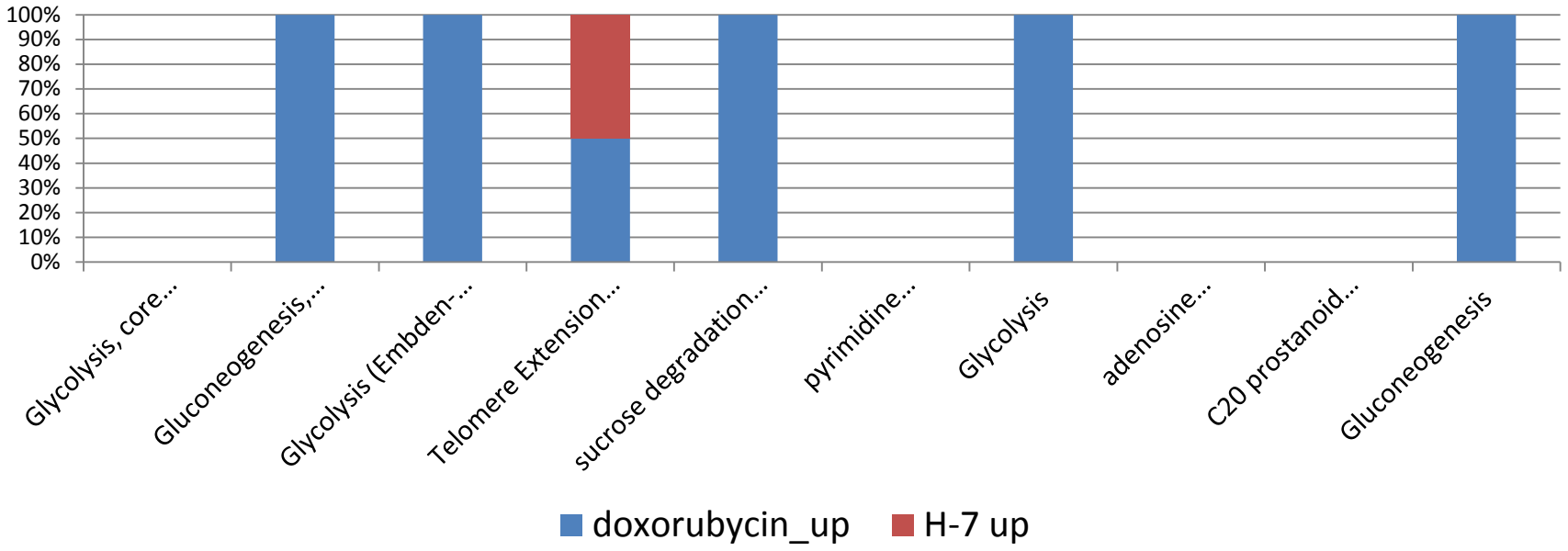


Cell line	Score	rank
panc1	-0.81	18

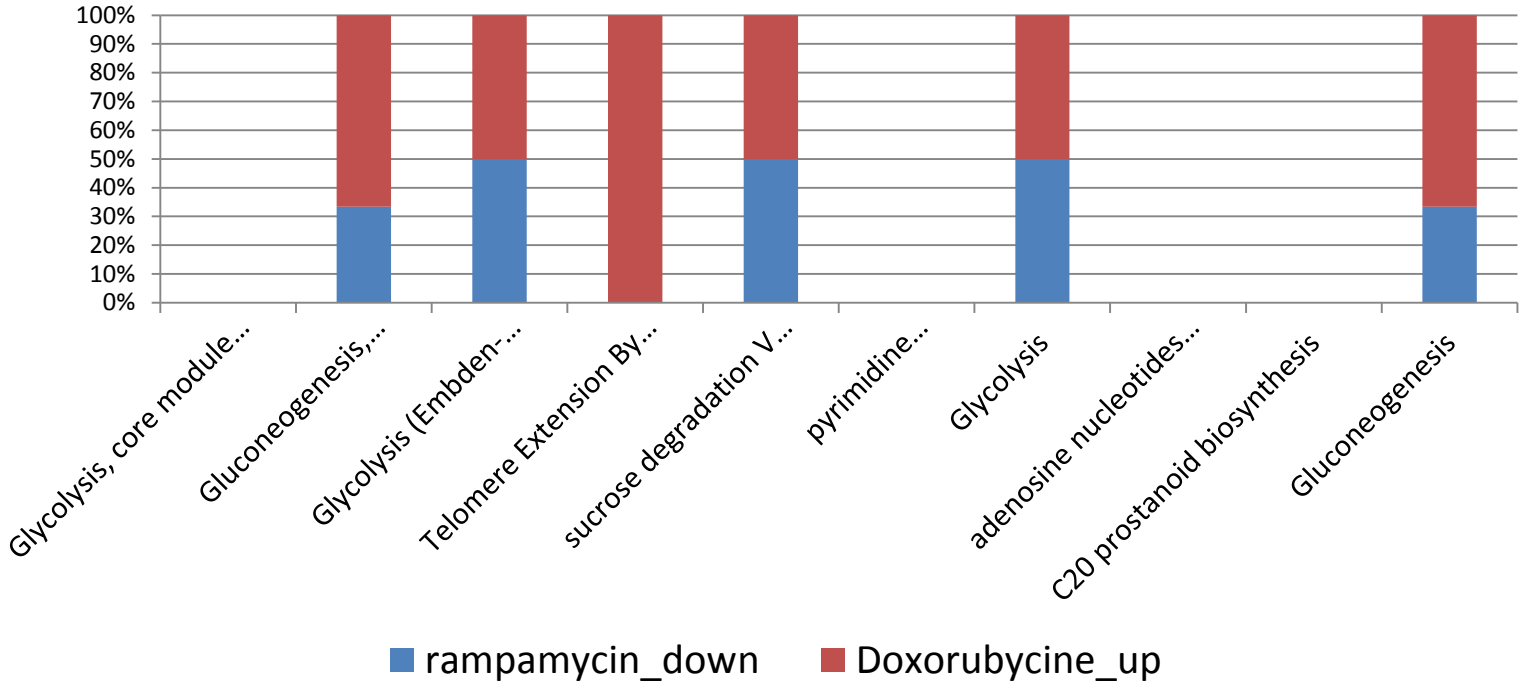
Gemcitabine-Panc1 Pathways

1-Mitotic Prometaphase	11-FOXM1 transcription factor network
2-Resolution of Sister Chromatid Cohesion	12-Aurora B signaling
3-M Phase	13-Gastric cancer network 1
4-Cell Cycle Mitotic	14-Phosphorylation of the APC/C
5-Cell Cycle	15-APC/C:Cdc20 mediated degradation of Cyclin B
6-Separation of Sister Chromatids	16-Activation of NIMA Kinases NEK9 NEK6 NEK7
7-Mitotic Anaphase	17-Phosphorylation of Emi1
8-Mitotic Metaphase and Anaphase	18-Golgi Cisternae Pericentriolar Stack Reorganization
9-Kinesins	19-Regulation of PLK1 Activity at G2/M Transition
10-PLK1 signaling events	

Synergy Prediction



Antagonism prediction



References

- 1) Koutsoukas, A., Lowe, R., KalantarMotamedi, Y., Mussa, H. Y. & John, B. O. In Silico Target Predictions: Defining a Benchmarking Data Set and Comparison of Performance of the Multiclass Naïve Bayes and Parzen-Rosenblatt. *J. Cheminform.* **53**, 1957–1966 (2013).
- 2) Rong, Z., Wei, Y., Yan-hong, L. & Feng-zhi, W. Meglumine cyclic adenylate induces differentiation of bone marrow mesenchymal stem cells into cardiomyocytes in vitro. *Chinese J. Pathophysiol.* **10**, 2040–2044 (2011).
- 3) Kolossov, E. *et al.* Engraftment of engineered ES cell-derived cardiomyocytes but not BM cells restores contractile function to the infarcted myocardium. *J. Exp. Med.* **203**, 2315–2327 (2006)
- 4) Winnicka, K., Bielawski, K., Bielawska, A. & Surazyński, A. Antiproliferative activity of derivatives of ouabain, digoxin and proscillaridin A in human MCF-7 and MDA-MB-231 breast cancer cells. *Biol. Pharm. Bull.* **31**, 1131–1140 (2008).
- 5) Ye, T. *et al.* The anthelmintic drug niclosamide induces apoptosis, impairs metastasis and reduces immunosuppressive cells in breast cancer model. *PLoS One* **9**, e85887 (2014).
- 6) Zhelev, Z. *et al.* Phenothiazines suppress proliferation and induce apoptosis in cultured leukemic cells without any influence on the viability of normal lymphocytes. Phenothiazines and leukemia. *Cancer Chemother. Pharmacol.* **53**, 267–275 (2004).