smiRk R package: For time series microRNA-mRNA data

BSU meeting: 4th Feb

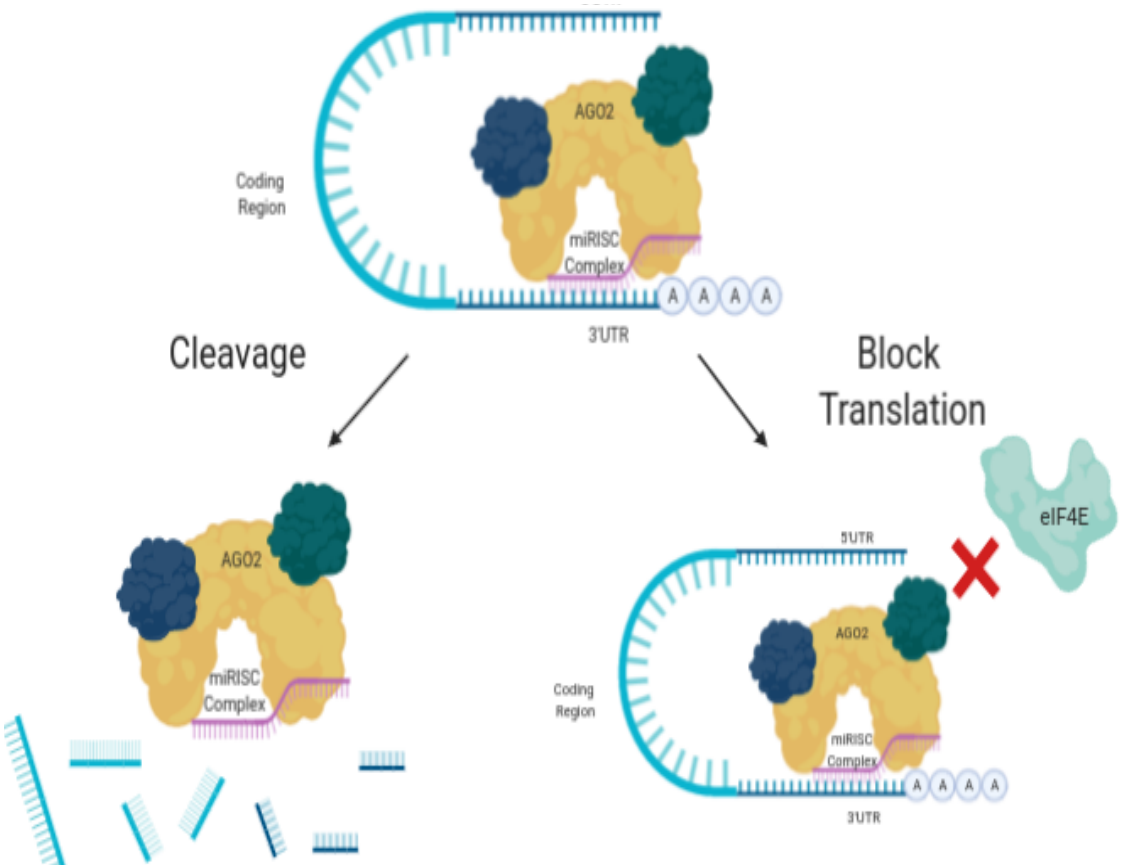
Krutik Patel

Newcastle University

4th February 2020

microRNAs are essential regulators

microRNAs (miRs) are important regulators of mammalian systems.



Why invest in research of miRs?

- miRs are small (16-22nt long) functional single strands of RNA which regulate expression levels of genes [Warnefors, 2014].
- Using a protein complex they can target mRNAs for repression via complementary targetting.

miRs are useful in biological and clinical research

- miRs are highly evolutionarily conserved so it is useful to test their function in animal models. miR related diseases can be studied in animal models e.g. miR140-5p is up regulated during chondrogenesis and down regulated in OA [Miyaki, 2009].
- miRs can be secreted from cells to surrounding fluids e.g. miR21 has been found in urine of CKD/ kidney injury patients. So they can be used as non-invasive biomarkers for conditions [Chen, 2017].
- Research of their use in a clinical setting to treat conditions is of growing interest [Schwarzenbach, 2014].

What are the difficulties of understanding miR-mRNA interactions?

The complex functions of miRs is difficult to comprehend

- A single miR can target multiple mRNAs and a single mRNA can be targetted by multiple miRs.

- This level of complexity is difficult to explore on the molecular level.

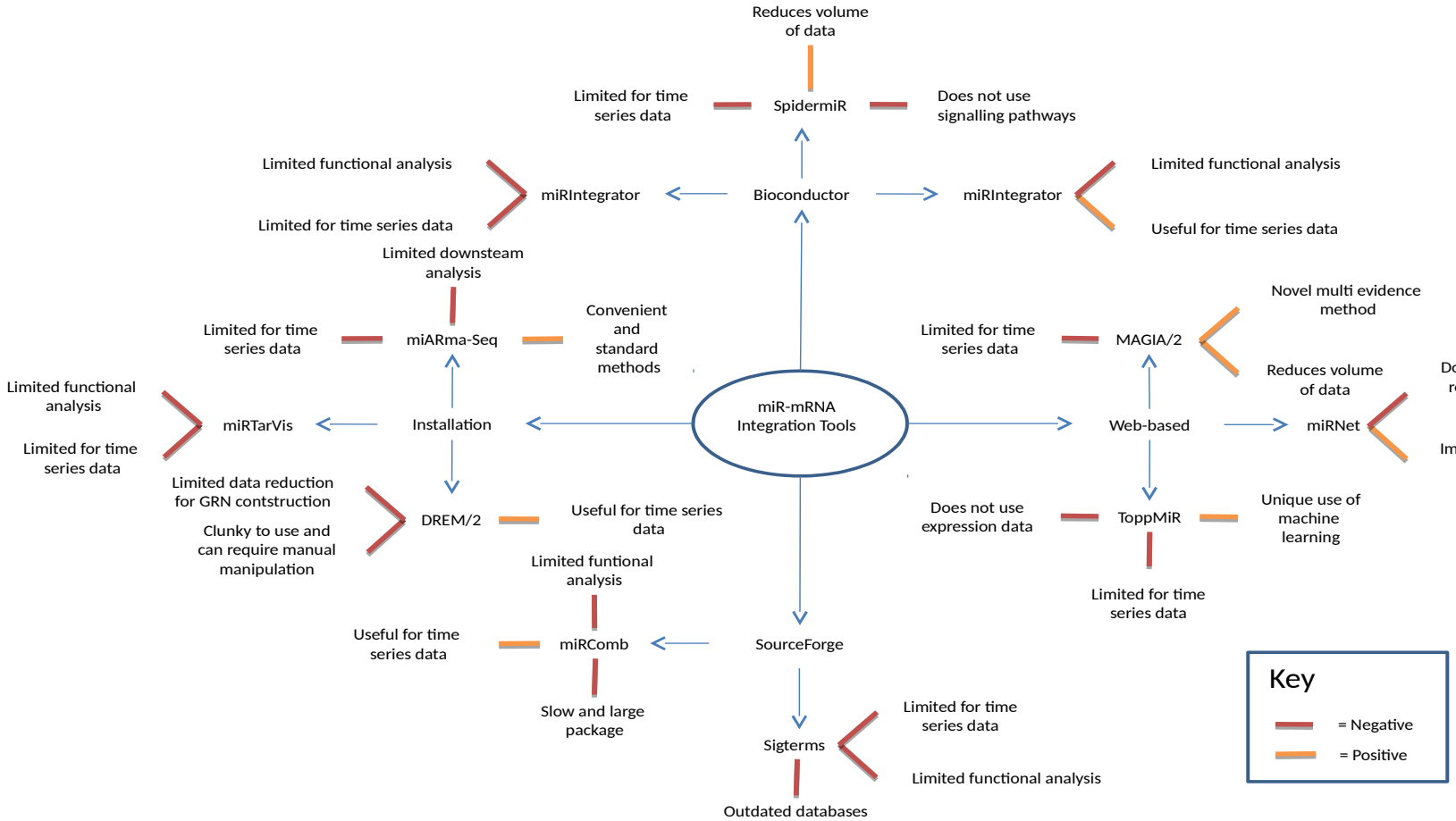
What methods can be employed to gain more knowledge about miR-mRNA interactions?

Big data/ systems methods can help to explore miRs

- Longitudinal Data
  - Generation of mRNA and miR data for multiple time points.
  - Better understand molecular behaviour.
  - Can perform time series differential expression.
- Gene Regulatory Networks
  - A schematic of how the molecular interactions in a system occur.
  - Blueprint for modelling, resource for experimental design and bioinformatic discovery/ analysis.
- Kinetic Modelling
  - Capture the inherent complexity of a biological system using maths.
  - *In silico* predictions of modulations which are not experimentally feasible can be performed.

Can we create GRNs from longitudinal data using current tools?

There are no current tools which can reduce the volume of data from longitudinal miR-mRNA data enough to start the generation of GRNs



So we started making our own!

smiRk (systems microRNA kinetics) R package

Systems microRNA kinetics is the most complete microRNA-mRNA integration tool available.

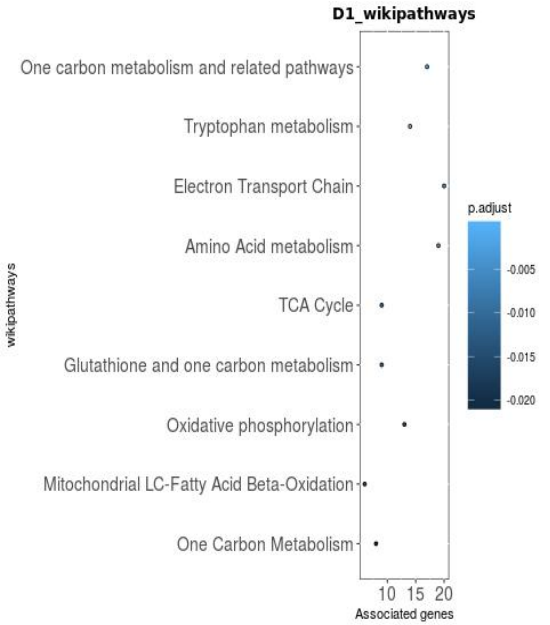
- Input are results from differential expression data (log2fc, FDR, adj.pvals, ect).
- Output is miR-integrated signalling networks for clear hypothesis generation.
- Uses Wikipathways to functionally analyse miR and mRNA data.
- Currently only works with Human and Mouse data.
- Can analyse miR-mRNA data combined or separately.
- Works well for temporal or conditional data.
- Links R with Cytoscape and Pathvisio for open ended analysis.

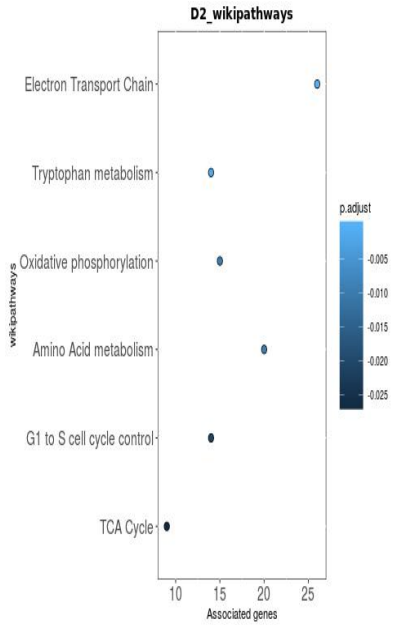
Test data set

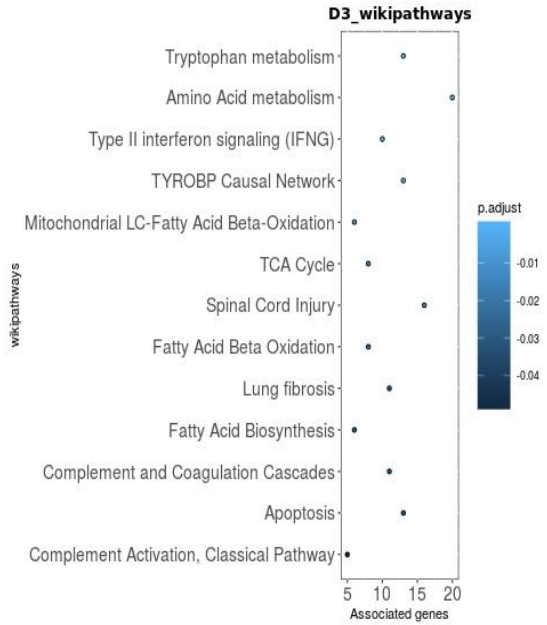
14 day time course of mouse kidney fibrosis, induced by folic acid injection. Limma was used for DE. Time point 0 was compared over all other time points individually using the *makeContrast* function. miR (GSE61328) [Pellegrini, 2016] and mRNA (GSE65267) [Craciun, 2016] data were analysed separately.

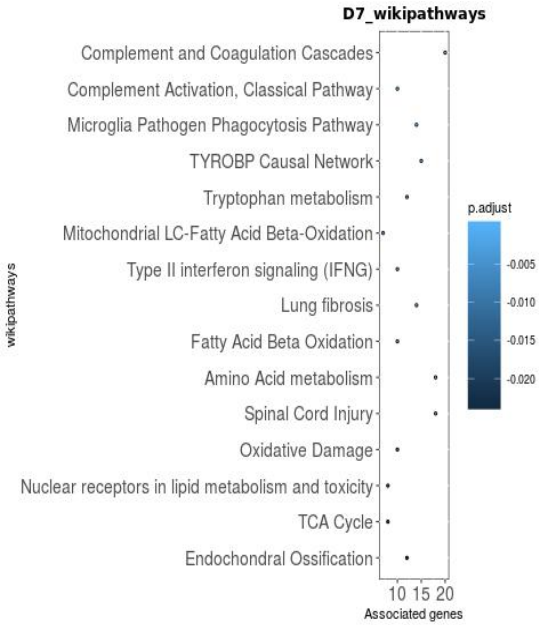
smiRk can use GSEA or fuzzy clustering to find a signalling pathway to investigate miR-mRNA interactions.

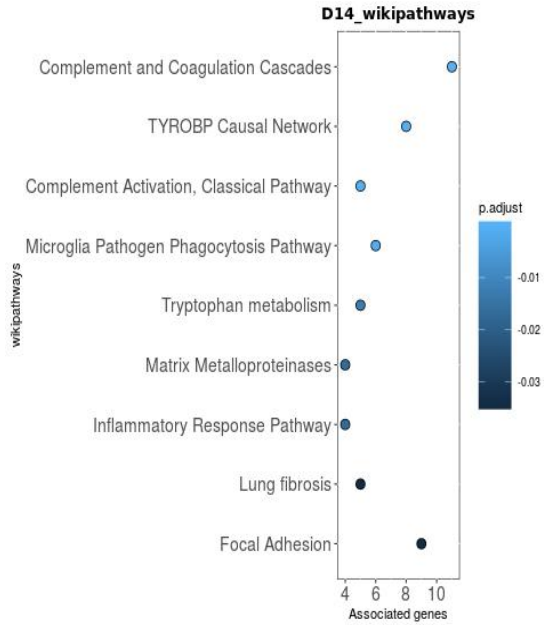
smiRk R package: time series pathway GSEA



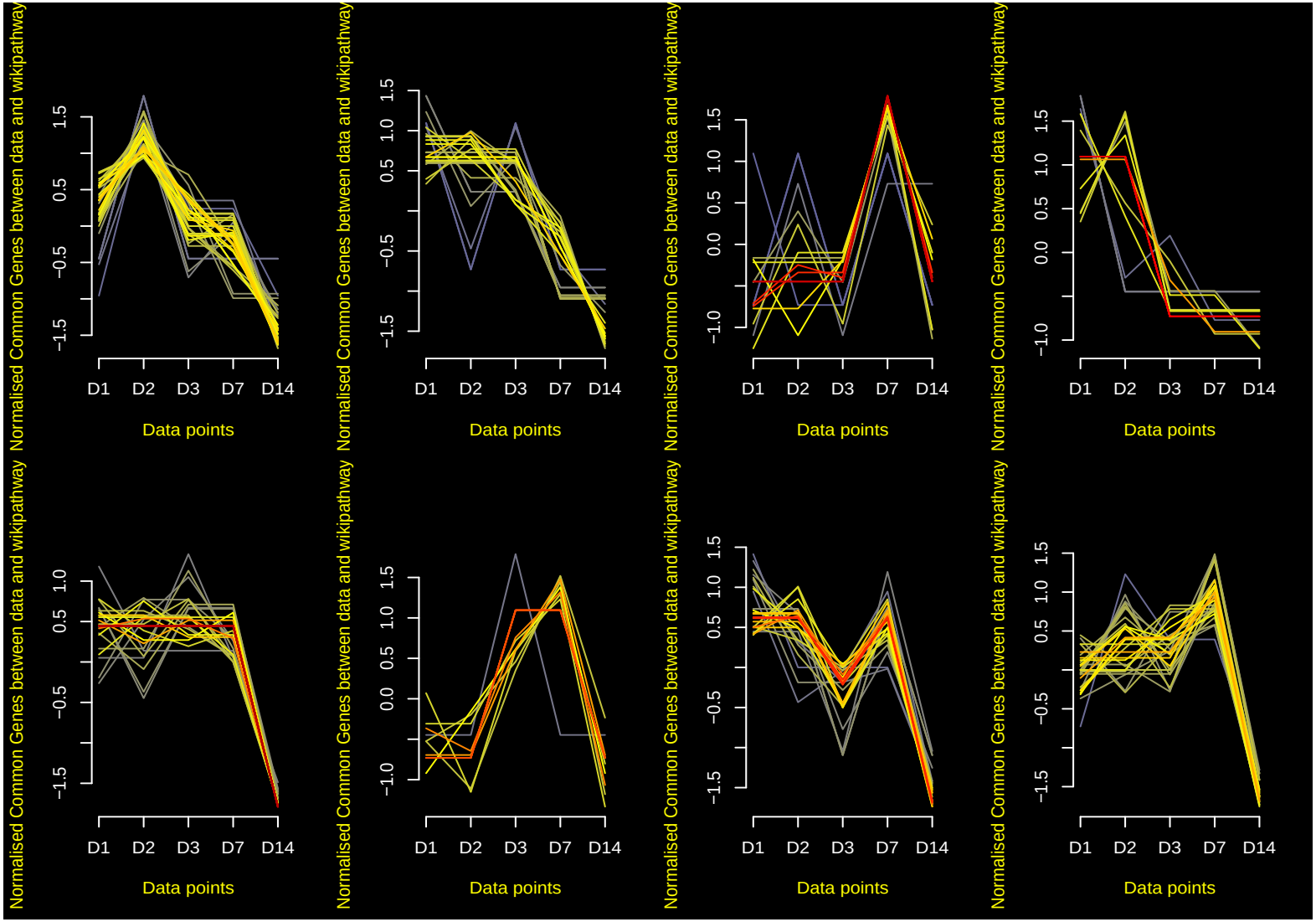








smiRk R package: temporal pathway clustering



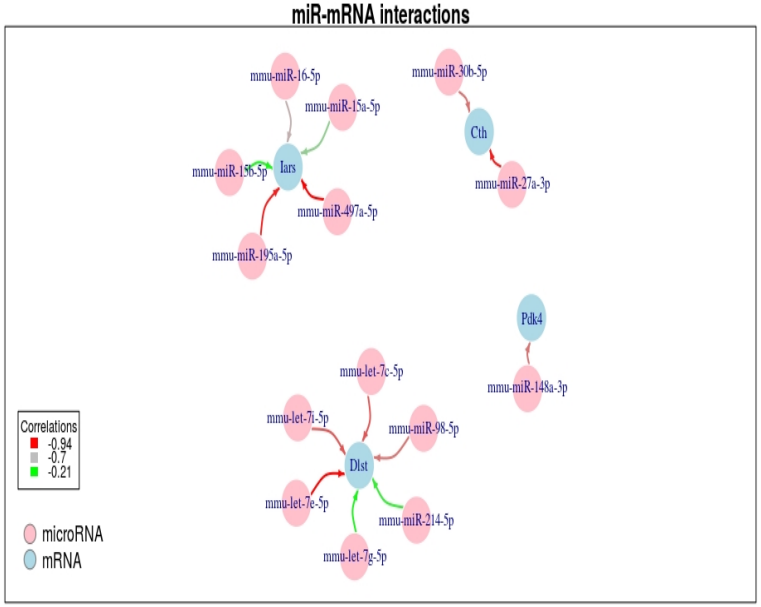
smiRk R package: filtering miR-mRNA interactions

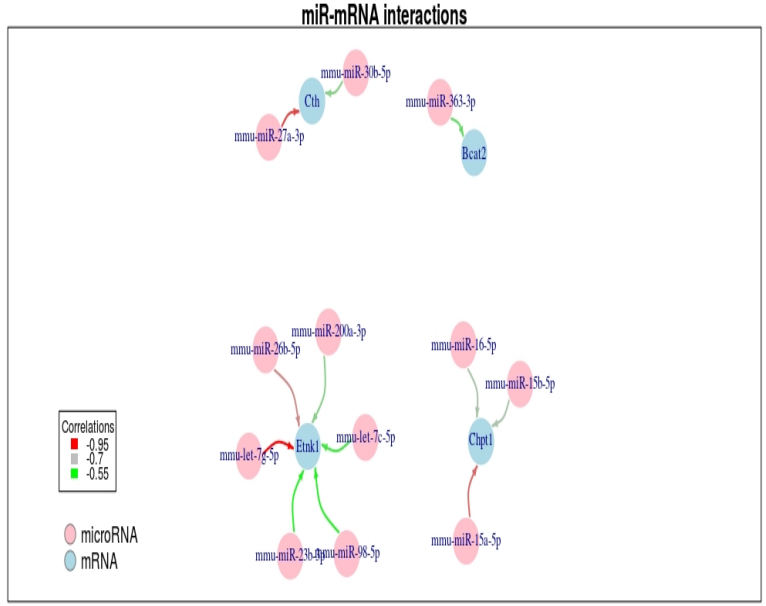
- mRNAs from a selected pathway of interest and all miRs are assumed to have the potential of interacting.

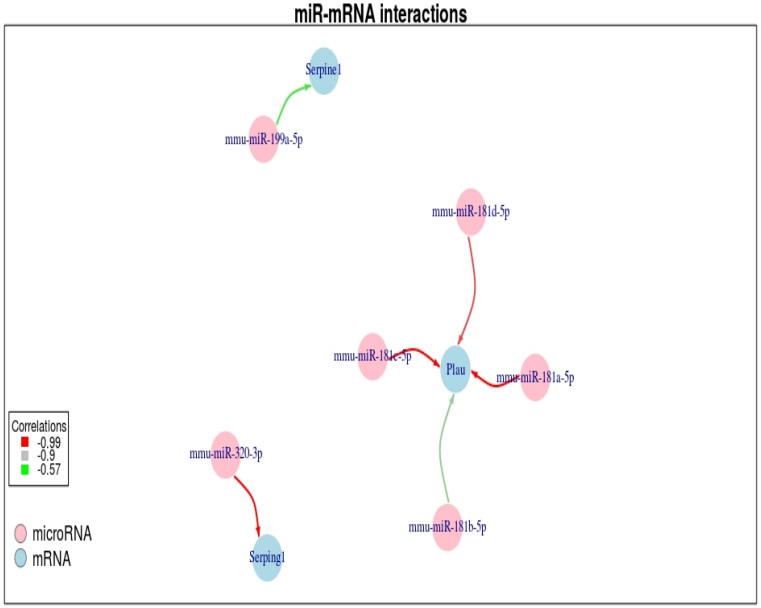
- Databases : TargetScans, miRDB and miRTarBase are used to filter for predicted/ functionally assessed miR-mRNA interactions.

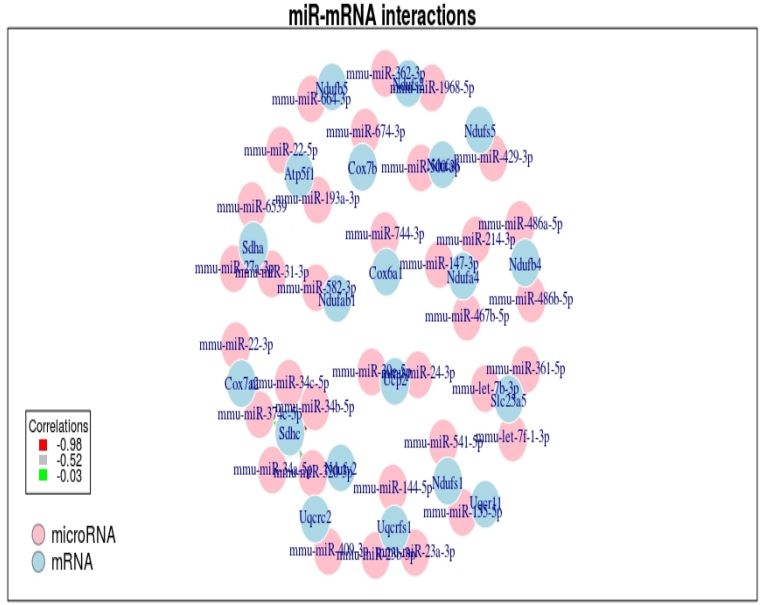
- Average Pearson correlations based on Log2FC / expression values along the time course can be used to filter for negatively correlating interactions.

smiRk R package: generate small networks





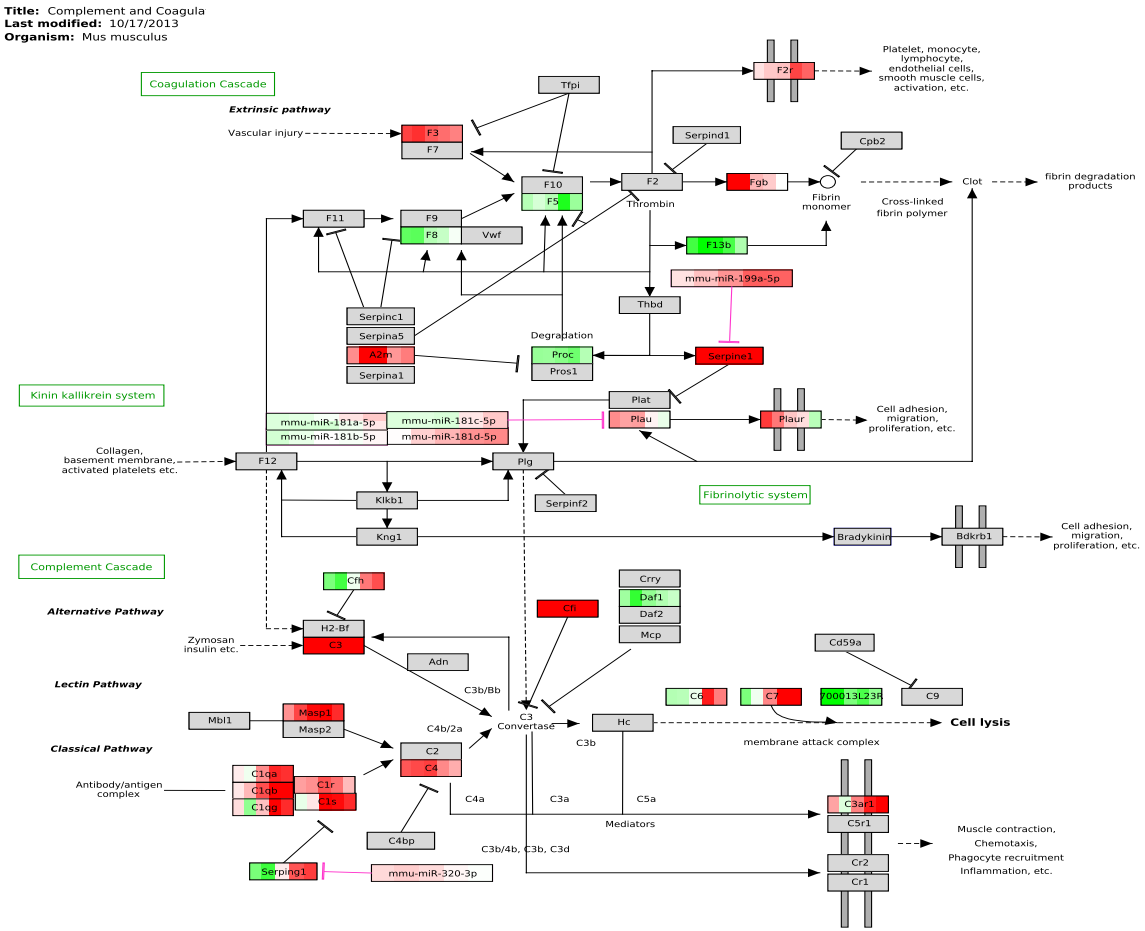




smiRk R package: export to cytoscape

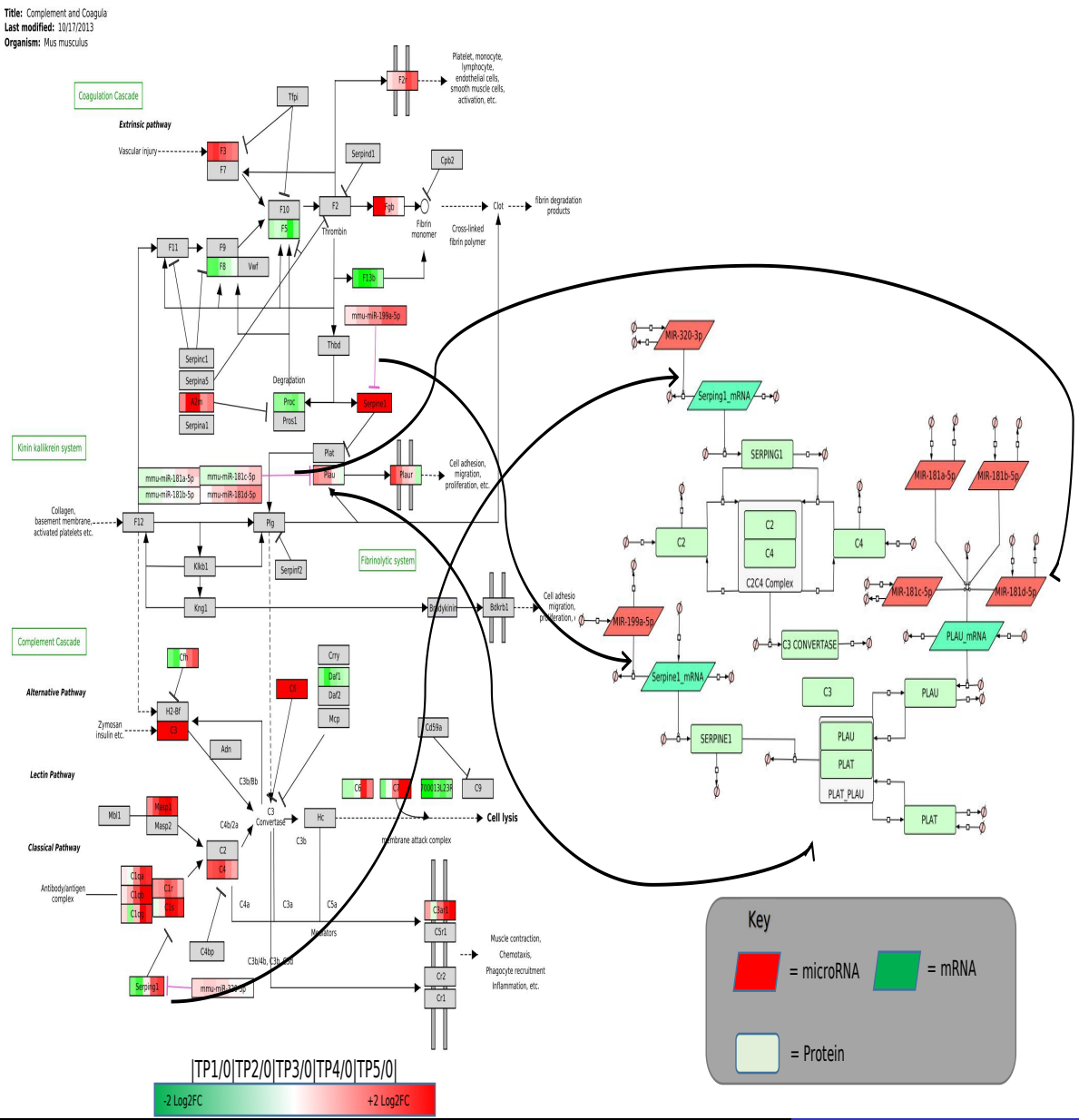
If networks are too large to properly view in R, smiRk can export them to Cytoscape v3.7 or newer [Smoot, 2010].
Users can then continue their exploration in Cytoscape.

smiRk R package: export to pathvisio



smiRk produces two text files for upload onto Pathvisio [van, 2008].
A MAPP file which will allow miRs found on the network to be uploaded, and a dynamics file which fill in kinetic data from DE (log2fc, expression, pvalue) for colour coding.

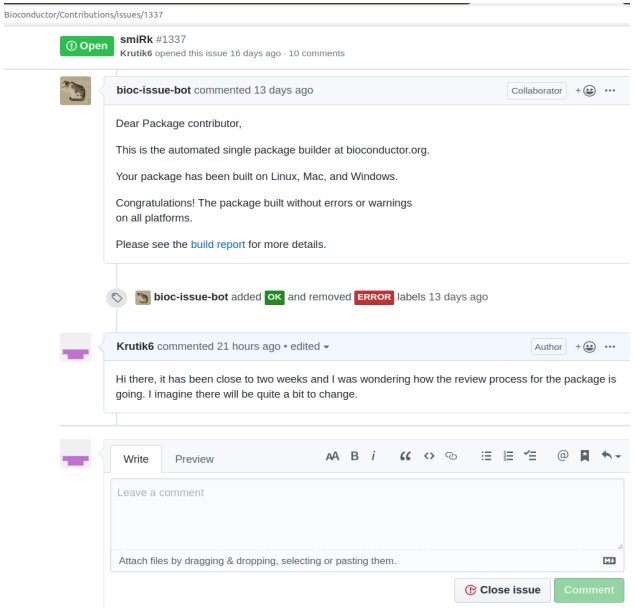
smiRk R package: generation of GRNs from pathvisio



Generation of a GRN from a signalling network is a bottom-up approach as we aim to be minimalist. We have successfully reduced the volume of data while keeping biological signal.

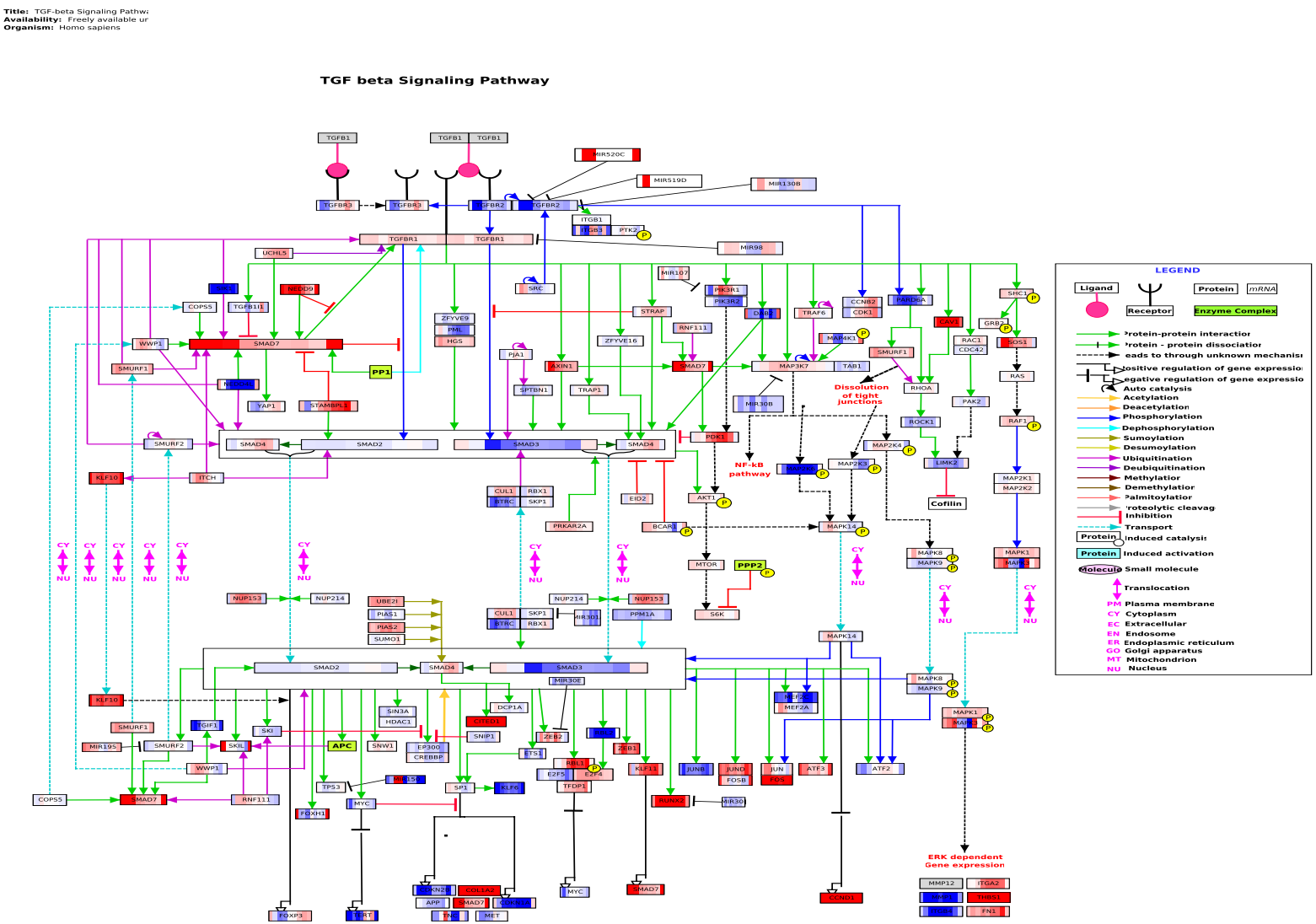
smiRk R package: output

- Package is available on my github repo at the moment. https://github.com/Krutik6/smiRk
- We are passed the first hurdle for a bioconductor upload, it is currently going through intense review by their team.



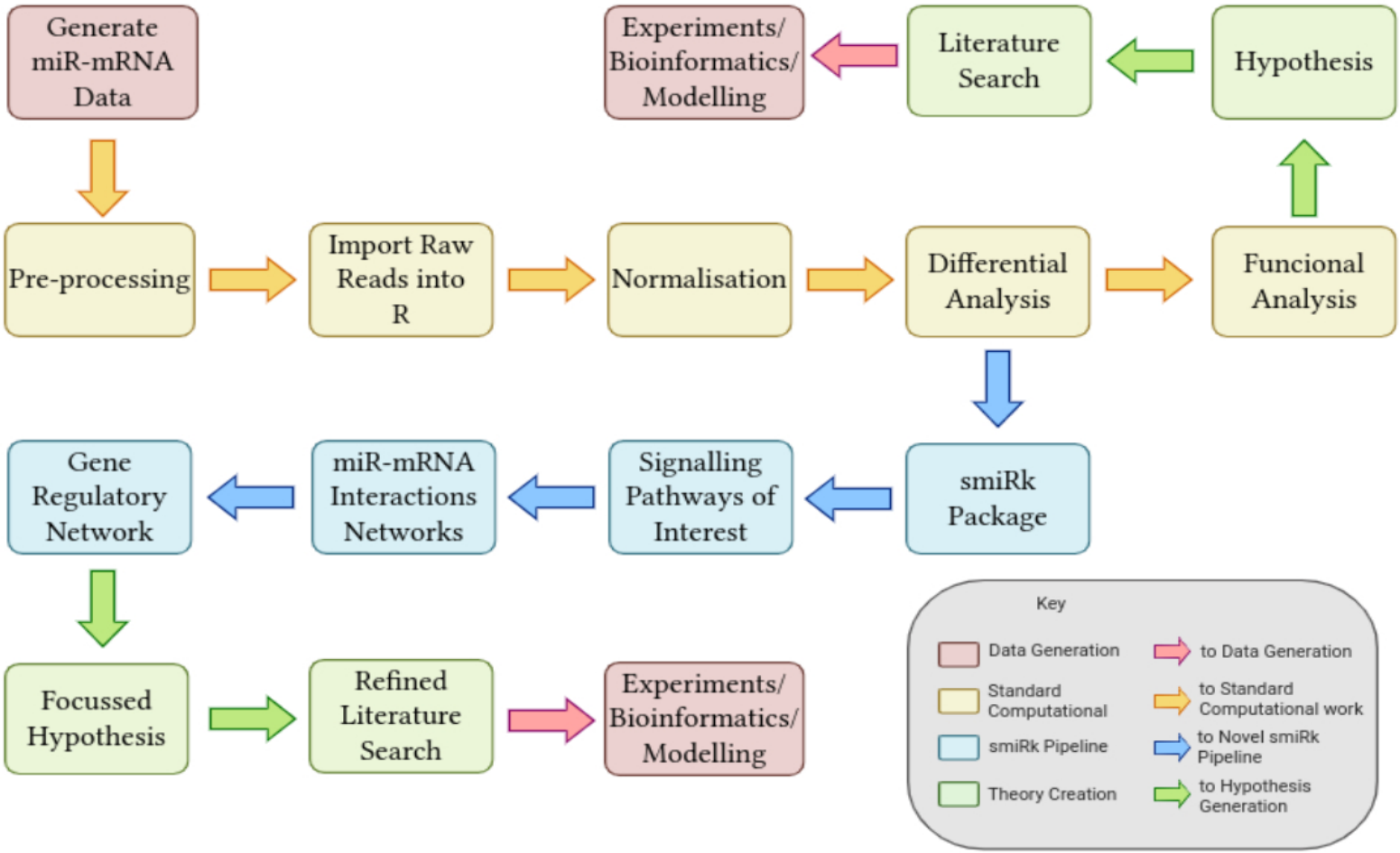
- After the package is on bioconductor we will push for a publication on *bioinformatics*

smiRk R package: Has been tested on multiple data sets



Breast cancer data sets from GSE78169 [Baran, 2016].

Ultimate Goal: I want smiRk to be a staple part of miR-mRNA analysis



Created in BioRender.com

Thanks

Daryl Shanley, and other members of our group : James, Sharmilla, Louise, Kathryn, Peter, Ciaran.

David Young and his lab including Matt and Jamie.

BSU.

References


Maria Warnefors (2014) Conserved microRNA editing in mammalian evolution, development and disease *Genome biology* **15**(6), R83.


Shigeru Miyaki (2009) MicroRNA-140 is expressed in differentiated human articular chondrocytes and modulates interleukin-1 responses *Arthritis & Rheumatism* **60**(9), 2723–2730.


Congcong Chen (2017) Urinary miR-21 as a potential biomarker of hypertensive kidney injury and fibrosis *Scientific reports* **7**(1), 1–9.


Florin Craciun (2016) RNA sequencing identifies novel translational biomarkers of kidney fibrosis *Journal of the American Society of Nephrology* **27**(6), 1702–1713.


Kathryn Pellegrini (2016) Application of small RNA sequencing to identify microRNAs in acute kidney injury and fibrosis *Toxicology and applied pharmacology* **312**, 42–52.


Michael Smoot (2010) Cytoscape 2.8: new features for data integration and network visualization *Bioinformatics* **27**(3), 431–432.


Iersel van (2008) Presenting and exploring biological pathways with PathVisio *BioMed Central* **9**(1), 399.


Alexander Pico (2008) WikiPathways: pathway editing for the people *PLoS biology* **6**(7), e184.


Heidi Schwarzenbach (2014) WikiPathways: pathway editing for the people *Nature Publishing Group* **11**(3), 145.


Jeanette Baran (2016) An integrative transcriptomics approach identifies miR-503 as a candidate master regulator of the estrogen response in MCF-7 breast cancer cells *Cold Spring Harbor Lab* **22**(10), 1592–1603.