How to infer accurate gene regulatory networks from perturbations?

A number of techniques exist to infer gene regulatory networks (GRNs) with causal regulatory influences between genes. The most accurate methods use gene expression responses to specific system perturbations, typically gene knock-down by RNAi or overexpression. We show using simulations that the best inference methods can achieve high accuracy if the data is highly informative, but when using data with realistic properties their accuracy is quite low, with Matthew's correlation coefficients around 0.3 (Tjärnberg et al., 2015). My group has developed several new methods to improve accuracy of GRN inference: optimisation of network sparsity (Tjärnberg et al., 2013), usage of functional association data as a prior (Studham et al., 2014), and prediction of inference accuracy based on data properties (Tjärnberg et al., 2015).

Ongoing work is focussing on optimising experimental design to improve data informativeness. We are applying these methods in a project employing siRNAs to perturb the GRN around the oncogene MYC. 40 genes in the MYC pathway were each perturbed in a cancer cell line, and the effect on the other genes was measured at steady state by qPCR. We used a bootstrapped LASSO strategy to infer a preliminary GRN with optimal sparsity and bootstrap cutoff. This GRN contains 157 links, while expected from randomly shuffled data with the same parameters is 0 links. The GRN was validated with 45 two-gene perturbations that were not included when inferring the network.

Studham ME, Tjärnberg A, Nordling TEM, Nelander S, Sonnhammer ELL (2014) Functional Association Networks as Priors for Gene Regulatory Network Inference, Bioinformatics, 30:i130-8

Tjärnberg A, Nordling TE, Studham M, Sonnhammer EL. (2013) Optimal sparsity criteria for network inference. J Comput Biol., 20:398-408

Tjärnberg A, Nordling TE, Studham M, Nelander S, Sonnhammer ELL (2015) Avoiding pitfalls in L1-regularised inference of gene networks. Mol Biosyst. 11:287-96.