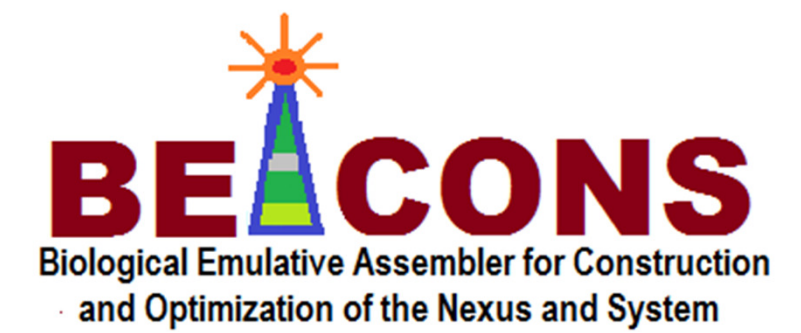


Predicting Transcriptome of *Escherichia coli* from “Marker” Genes

Maurice HT Ling, Chueh Loo Poh
School of Chemical and Biomedical Engineering
Nanyang Technological University, Singapore

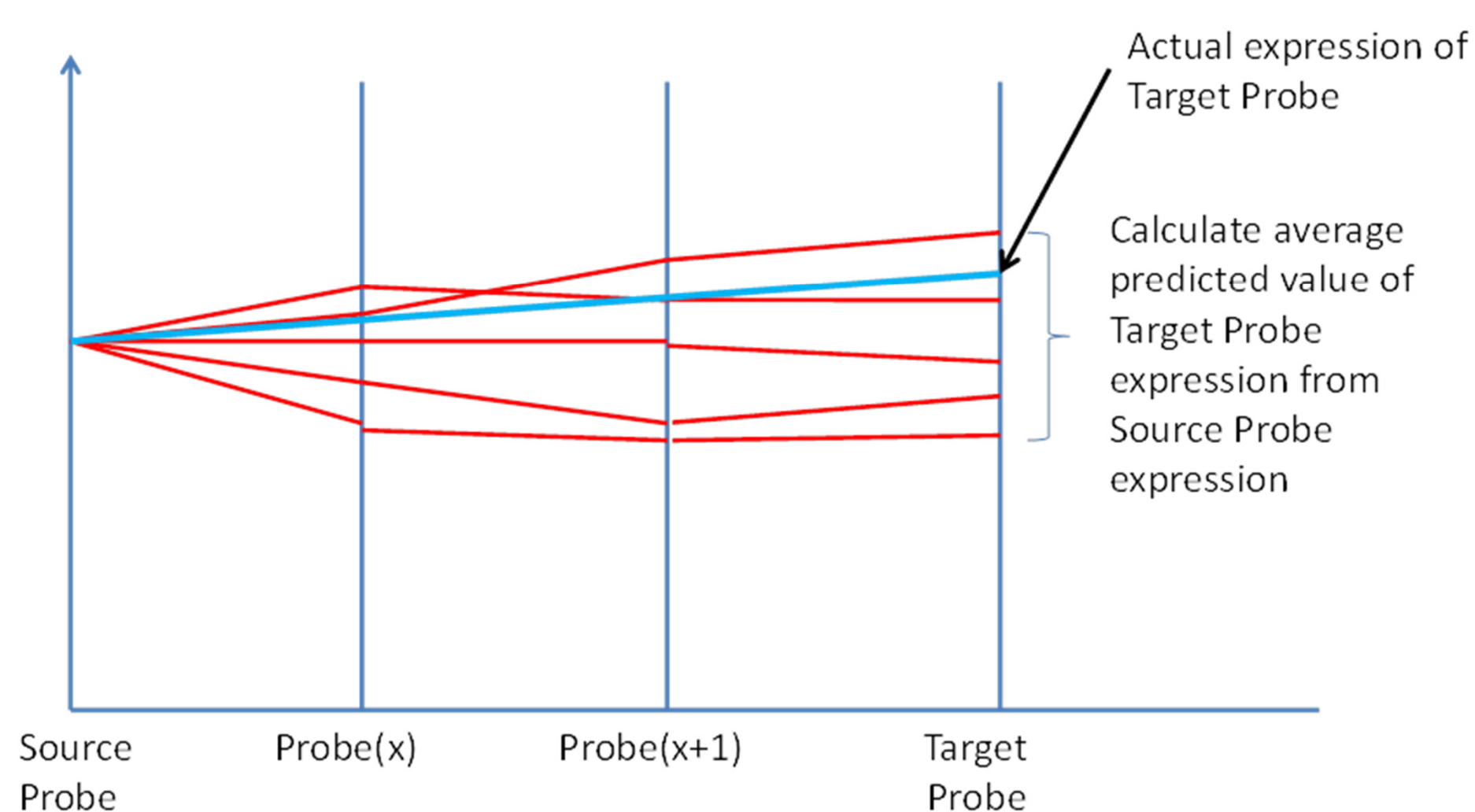


Motivations

- Need to gauge the effects of transgenes to the native system
- Effects of transgenes profiled after the cloning process is completed
- Handy to estimate the effects of transgene prior to cloning

Background

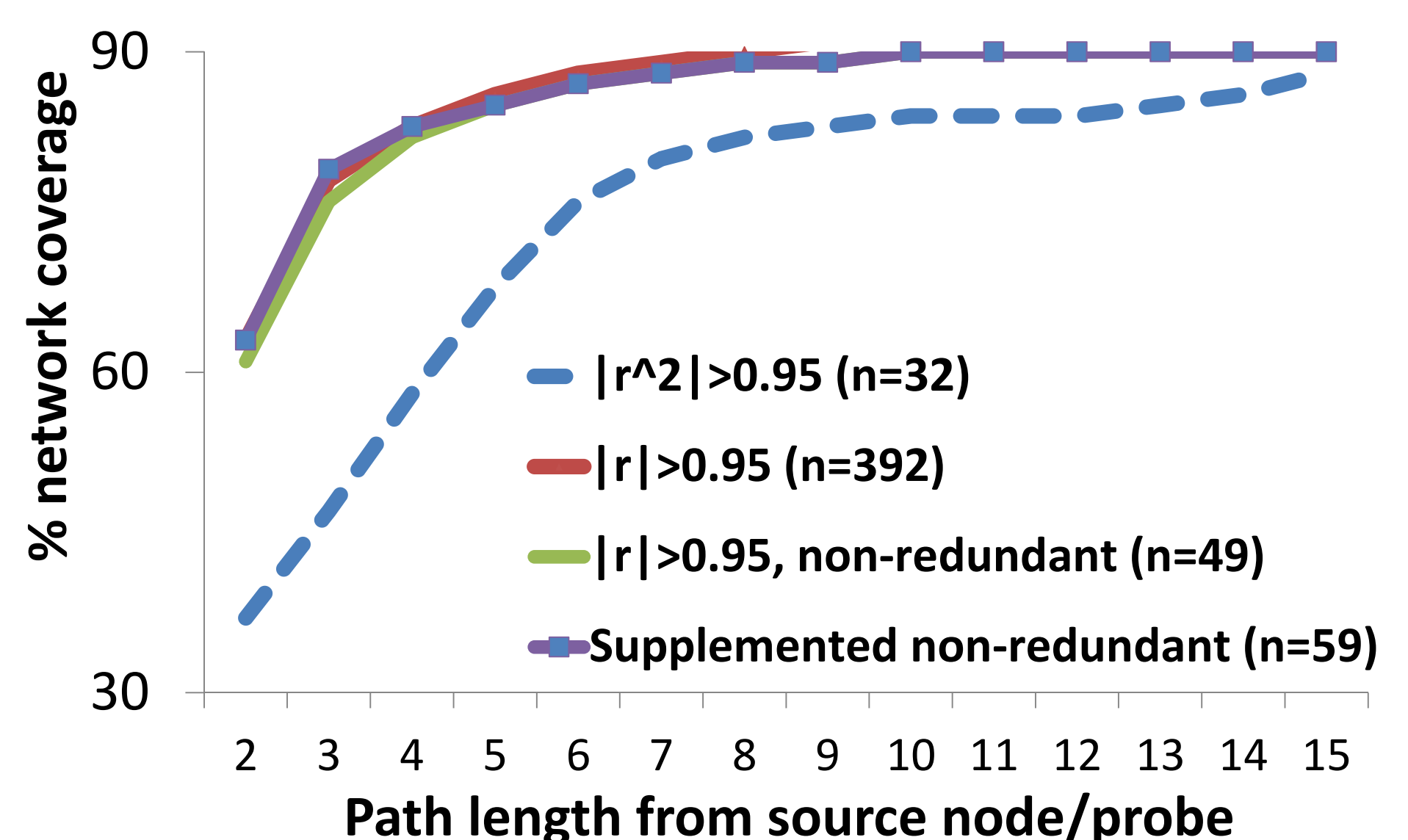
- Gene expression prediction is useful to study the effects of gene over-expression (Selinger et al., 2003)
- Synthetic biologists to predict the effects of transgene from a small number of marker genes (Porcar et al., 2011)
- Gene co-expression network (GCN) to study expression of genes (Obayashi et al., 2013)
- GCN may be used to predict gene expressions.



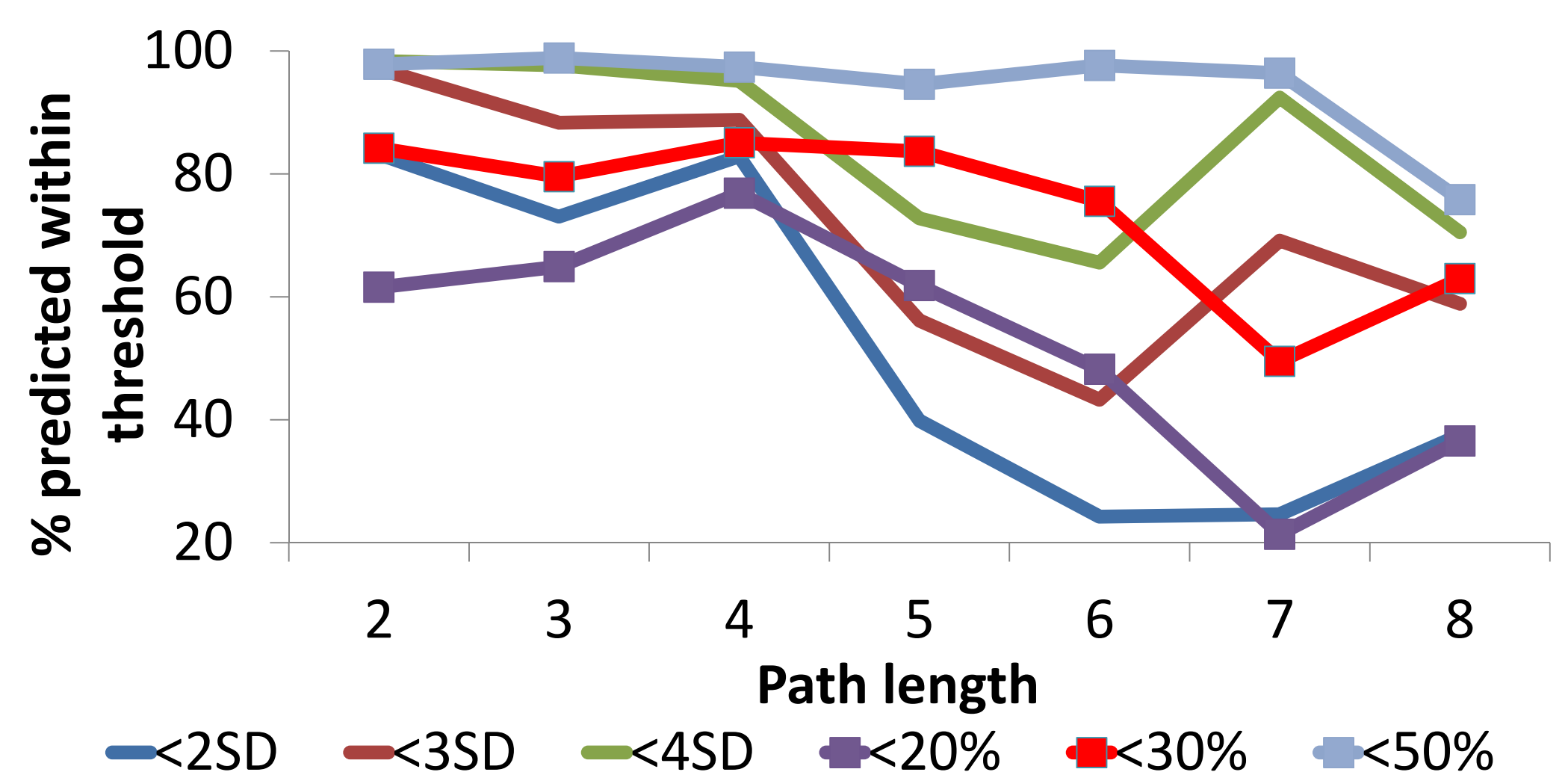
Procedure

1. Collect 605 microarrays from GEO
2. Build correlation network, $|r| > 0.75$
3. Collect microarrays for testing
4. Run predictions and evaluate accuracy

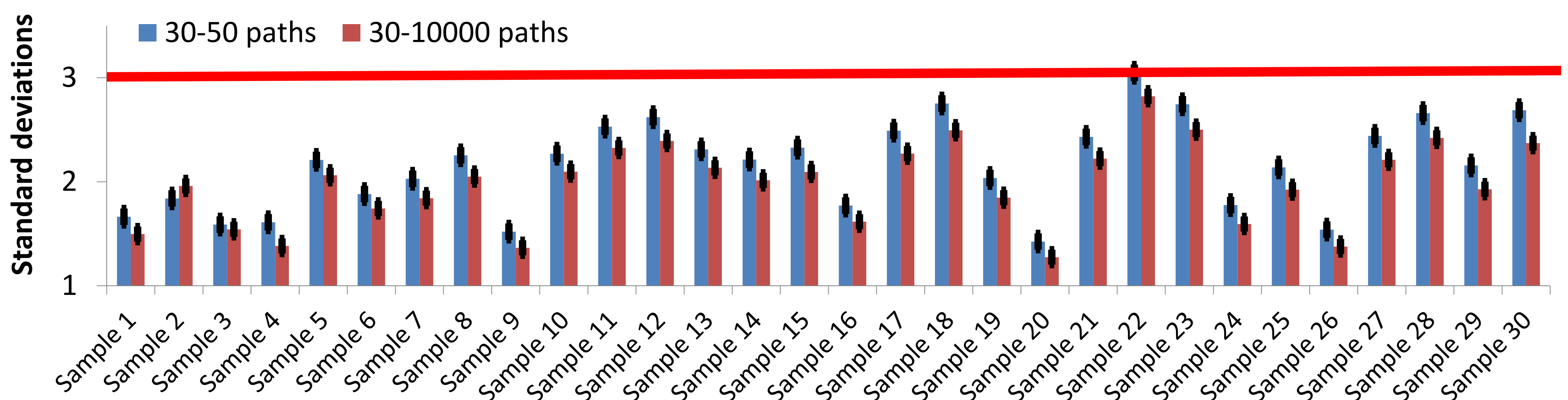
Findings 1: 90% of network reached by 59 “marker” genes/probes



Findings 2: Accurate within 4 degrees



Findings 3: Transcriptome predicted within 3 SD from 59 “marker” genes (4 degrees)



References:

Obayashi T, Kinoshita K: Rank of correlation coefficient as a comparable measure for biological significance of gene coexpression. *DNA Res* 2009, 16(5):249-260.
Porcar M, Danchin A, de Lorenzo V, Dos Santos VA, Krasnogor N, Rasmussen S, Moya A: The ten grand challenges of synthetic life. *Syst Synth Biol* 2011, 5(1-2):1-9.
Selinger DW, Wright MA, Church GM: On the complete determination of biological systems. *Trends in Biotechnology* 2003, 21(6):251-254.

Contacts:

Maurice HT Ling (mauriceling@acm.org)
Chueh Loo Poh (clpoh@ntu.edu.sg)