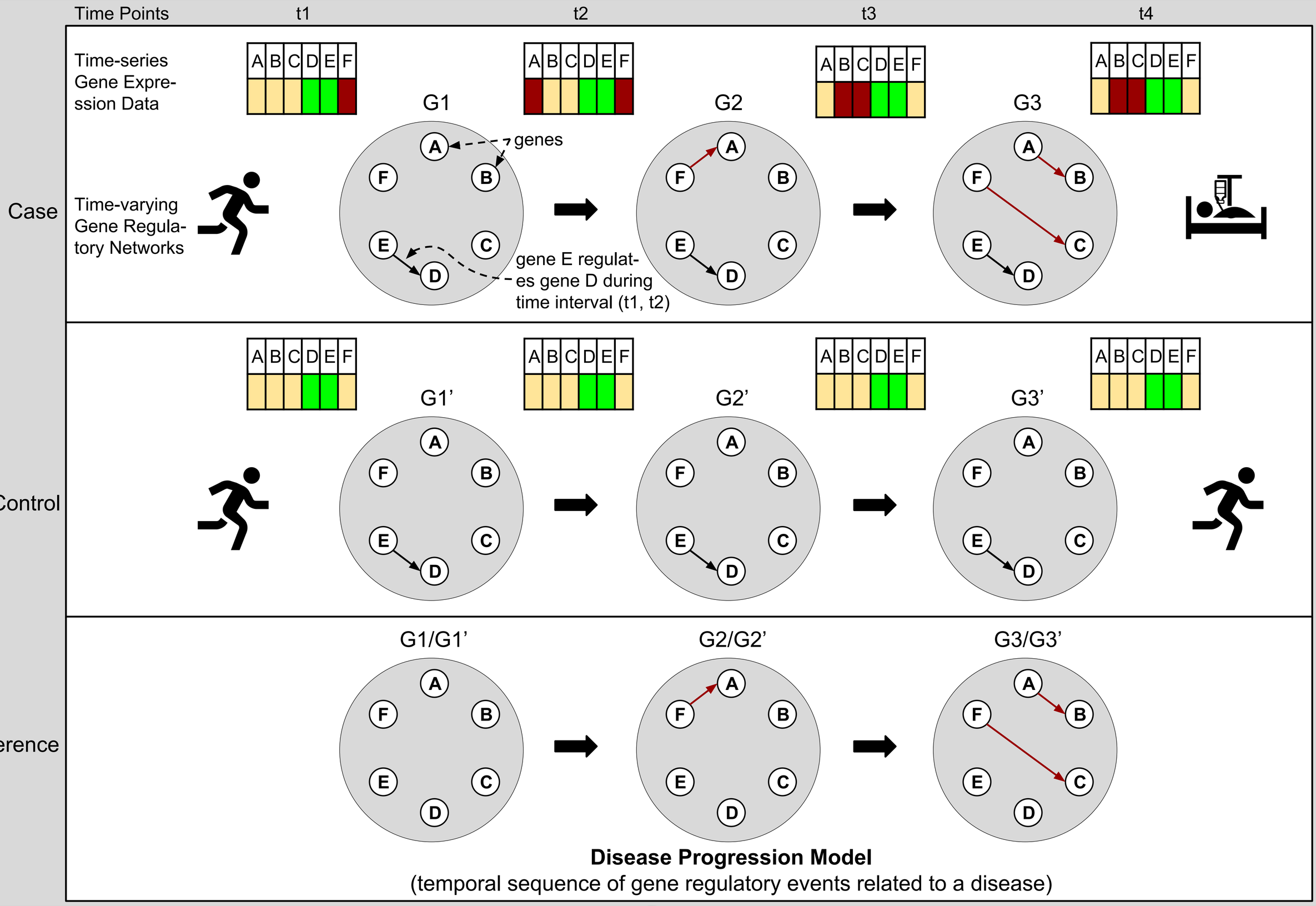
Towards Genome-scale Disease Progression Models

Saptarshi Pyne, Alok Ranjan Kumar and Ashish Anand
Department of Computer Science and Engineering, Indian Institute of Technology Guwahati, Guwahati, India 781039

1. Objective

Downloadable PDF of this Poster:
https://goo.gl/3nLSEz

Discover how the genome-scale *gene regulatory network* changes with time during progression of a given disease.



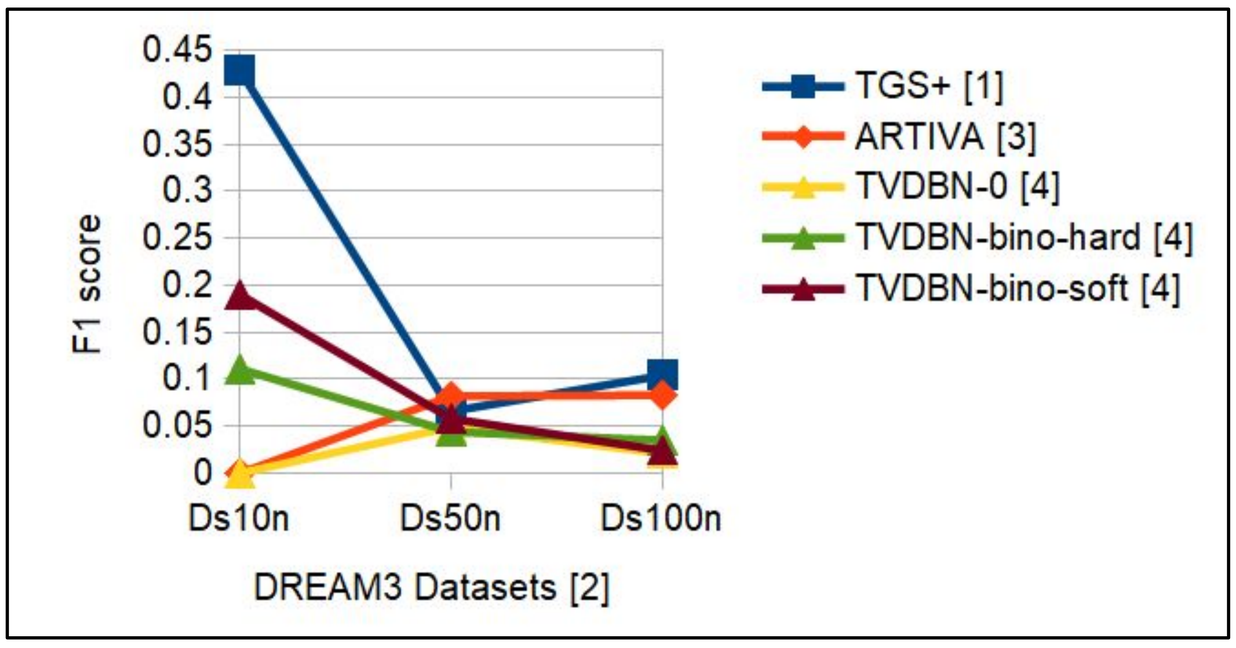
2. Challenges, Methods & Results

Challenges

- Estimation of a large number of parameters which also necessitates high computation time and memory
- Integration of auxiliary data, e.g., TF binding site information, to enhance accuracy

Methods and Results

- A novel algorithm, namely '*an algorithm for reconstructing Time-varying Gene regulatory networks with Shortlisted candidate regulators (TGS) plus* (**TGS+**), is proposed [1]. TGS+ provides the state-of-the-art computational speed and accuracy (F1 score) w.r.t. three in-silico DREAM3 [2] benchmark datasets having 10, 50 and 100 genes.
- TGS+ has two steps. In step 1, for each gene, a shortlist of its candidate regulators is inferred through an information theoretic pipeline. In step 2, the shortlisted candidates are evaluated within a Bayesian network framework to identify the true regulators among them.



#Genes\Algo	**TGS+**	ARTIVA	TVDBN-0	TVDBN-bino-hard	TVDBN-bino-soft
10	**6 sec**	10 min	2 min	2 min	2 min
50	**22 sec**	5 hrs	12 min	10 min	8 min
100	**1 min**	32 hrs	52 min	3 hrs	17 min

3. Conclusions & Future Directions

- A novel algorithm, TGS+, is proposed to reconstruct *time-varying gene regulatory networks* from *time-series gene expression datasets*. It provides the state-of-the-art computational speed and accuracy w.r.t. three benchmark in-silico DREAM3 datasets.
  TGS+ also reconstructs a biologically significant D. melanogaster (Dm) aka 'fruit fly' life cycle model in sub-30 mins when applied on a Dm developmental life cycle dataset with 588 genes and 66 time points.
- However, main memory (RAM) consumption of TGS+ increases exponentially with the number of genes. Reducing its memory requirement remains an important challenge.
- Moreover, TGS+ can be extended to incorporate auxiliary regulatory evidences, e.g., TF binding site information, for enhancing accuracy.

4. References

1. Pyne et al. '*Rapid Reconstruction of Time-varying Gene Regulatory Networks*'. TCBB, 2018.https://ieeexplore.ieee.org/document/8423706/
2. DREAM3 - In Silico Network Challenge. https://www.synapse.org/#!Synapse:syn2853594/wiki/71567
3. Lèbre et al. '*Statistical inference of the time-varying structure of gene-regulation networks*'. BMC Syst Biol, vol 4, no 1, p130, Sep 2010.
4. Dondelinger et al. '*Non-homogeneous dynamic Bayesian networks with Bayesian regularization for inferring gene regulatory networks with gradually time-varying structure*'. Mach Learn, vol 90, no 2, p191–230, 2013.

5. Contacts

AA: anand.ashish@iitg.ernet.in, SP: saptarshipyne01@gmail.com