

Identifying strengths and weaknesses of computational network inference methods on single-cell RNA-seq data

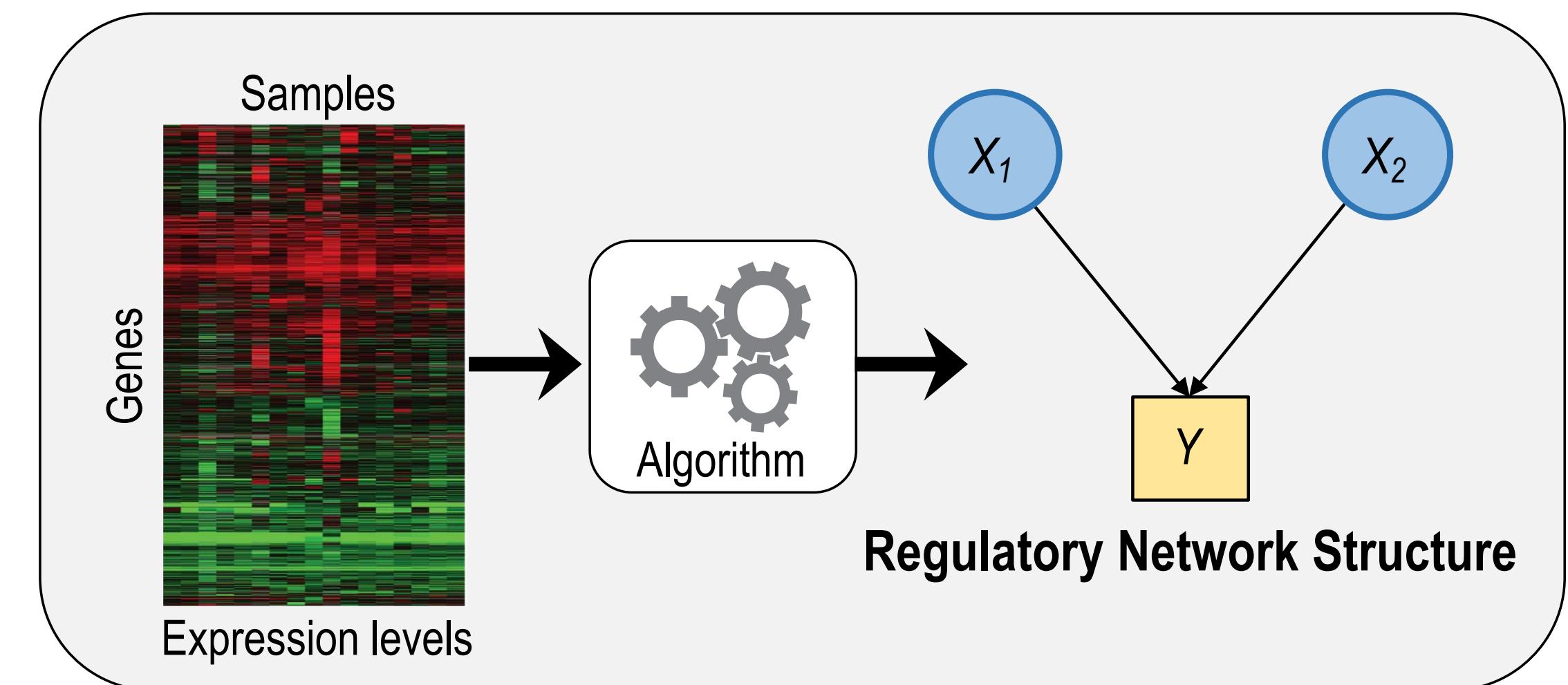
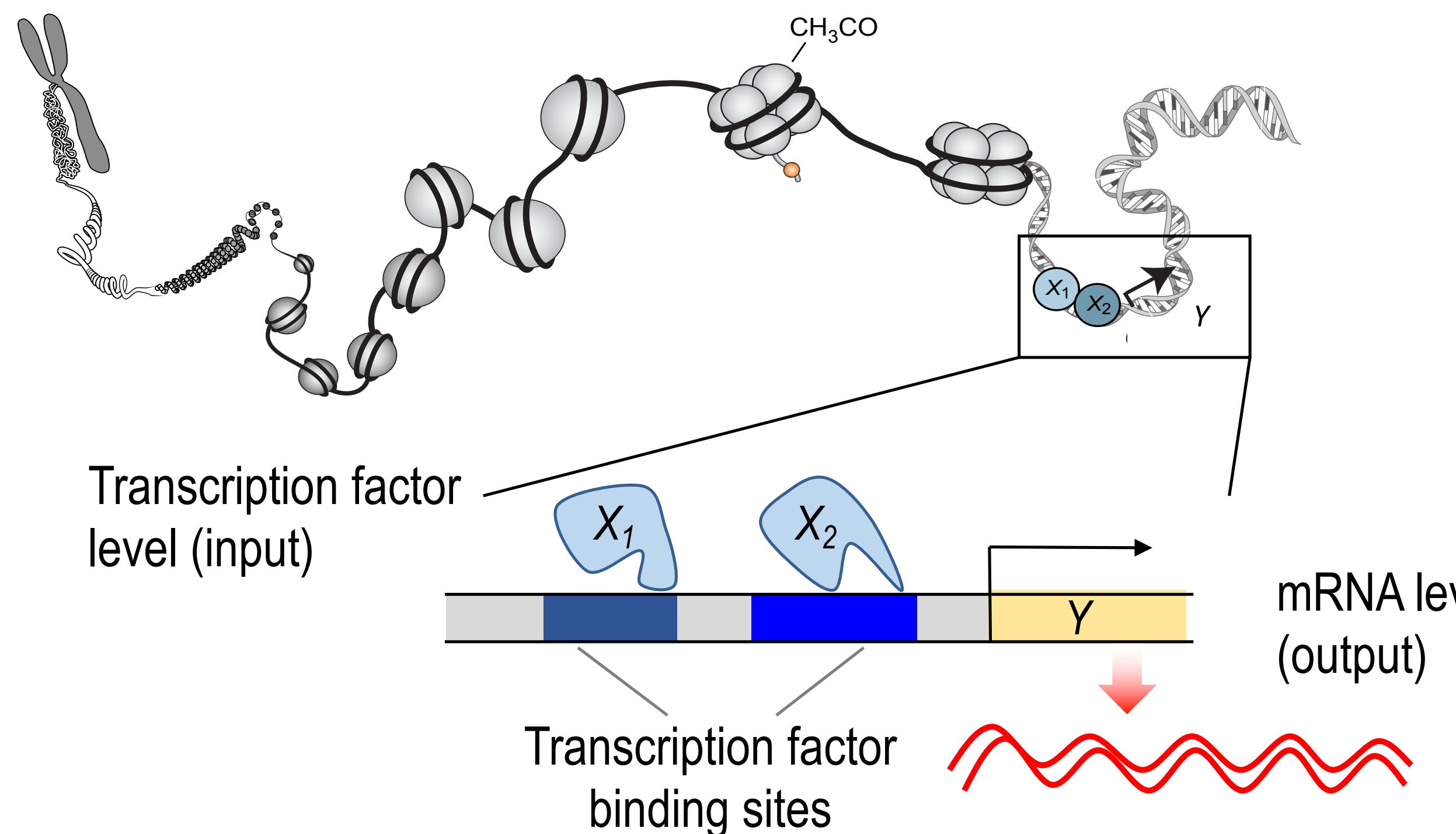
Matthew Stone, Viswesh Periyasamy, Sunnie Grace McCalla, Alireza Fotuhi Siahpirani,
Sushmita Roy



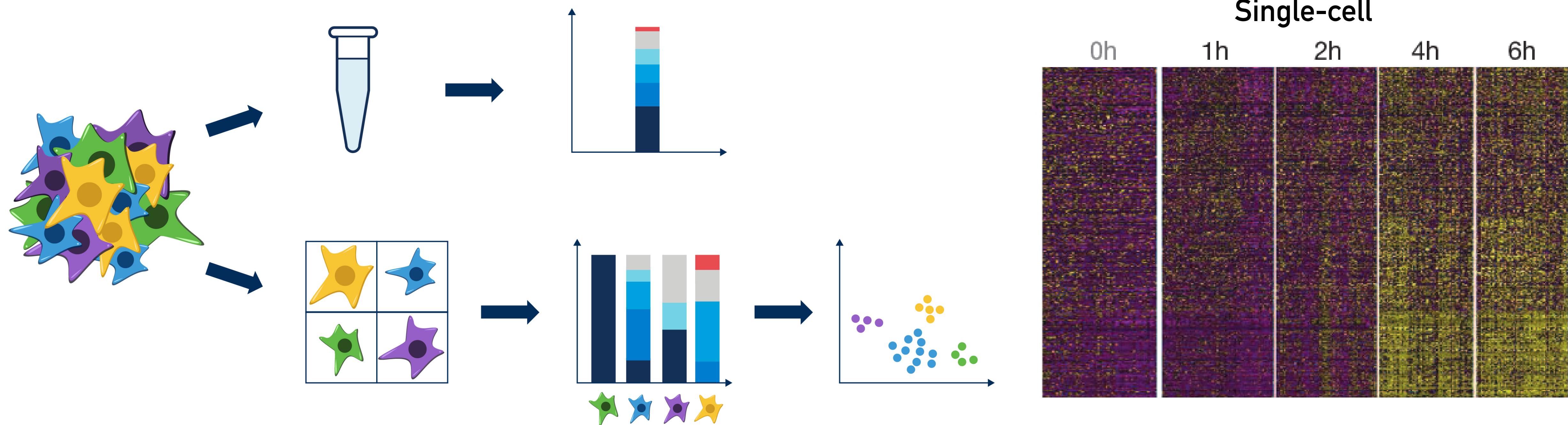
ISMB 2019 — 23 July 2019



Regulatory network inference



Single-cell RNA-sequencing (scRNA-seq)



Open questions in network inference

- How do network inference algorithms scale to large scRNA-seq datasets?
- How accurate are the networks inferred by each method?
- How similar are the networks inferred by different methods?
- Does imputation of sparse expression data improve network accuracy?

Network inference algorithms evaluated

Algorithm	Citation	Methodology
HurdleNormal	McDavid <i>et al.</i> , arXiv 1610.05857	
MERLIN**	Roy <i>et al.</i> , PLoS Comp Biol 2013	
SCENIC	Aibar <i>et al.</i> , Nature Methods 2017	Graphical models and dependency networks
SCHiRM	Intosalmi <i>et al.</i> , bioRxiv 2018	
SILGGM	Zhang <i>et al.</i> , PLoS Comp Biol 2018	
kNN-DREMI	van Dijk <i>et al.</i> , Cell 2018	
PIDC	Chan <i>et al.</i> , Cell Systems 2017	Information theoretic
Scribe*	Qiu <i>et al.</i> , bioRxiv 2018	
BTR	Lim <i>et al.</i> , BMC Bioinformatics 2016	Boolean networks
SCODE*	Matsumoto <i>et al.</i> , Bioinformatics	Ordinary differential equations
LEAP*	Specht and Li, Bioinformatics 2017	
Pearson	n/a	Correlation

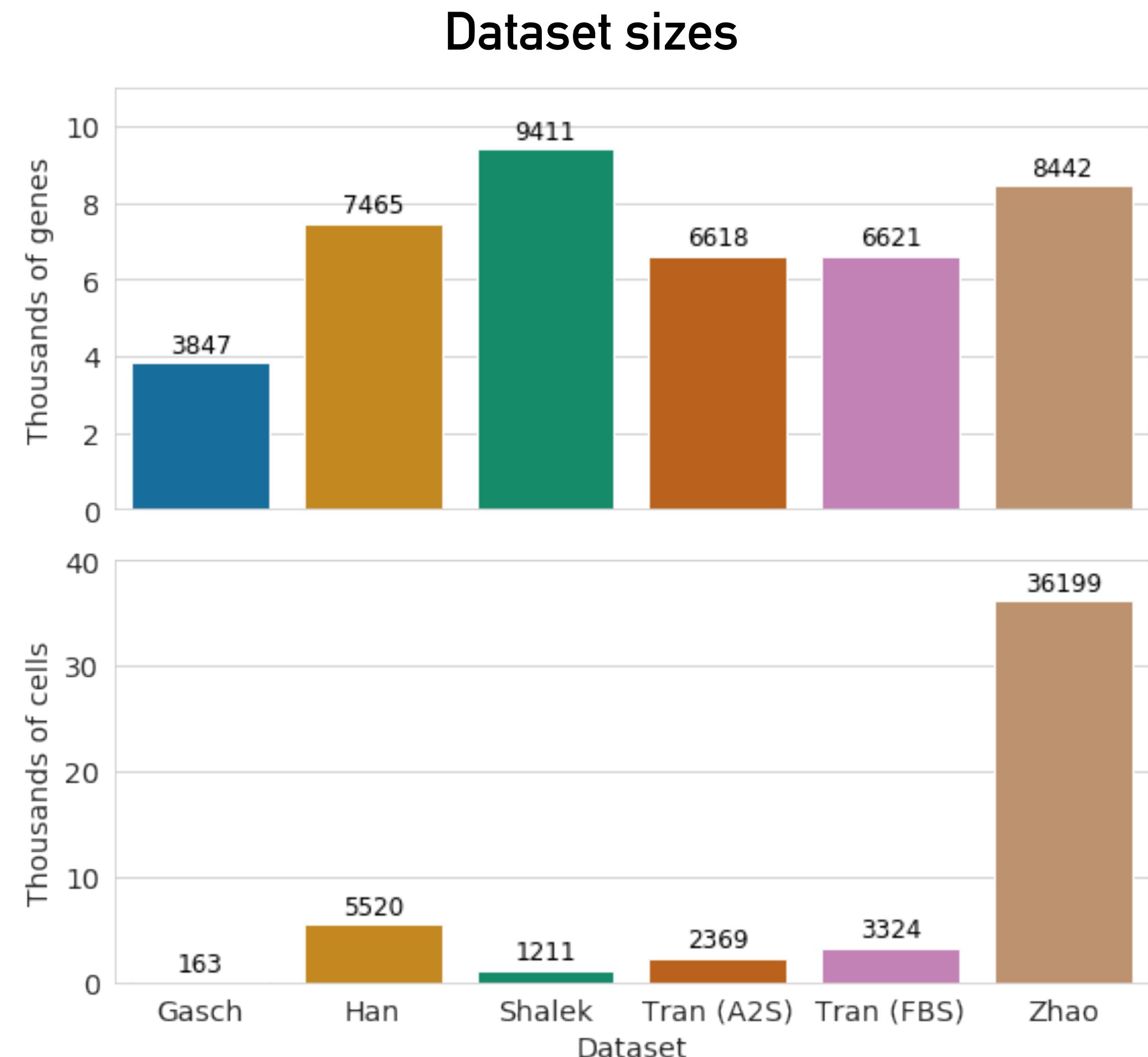
*Incorporate trajectory learning

**Designed for bulk transcriptomic data

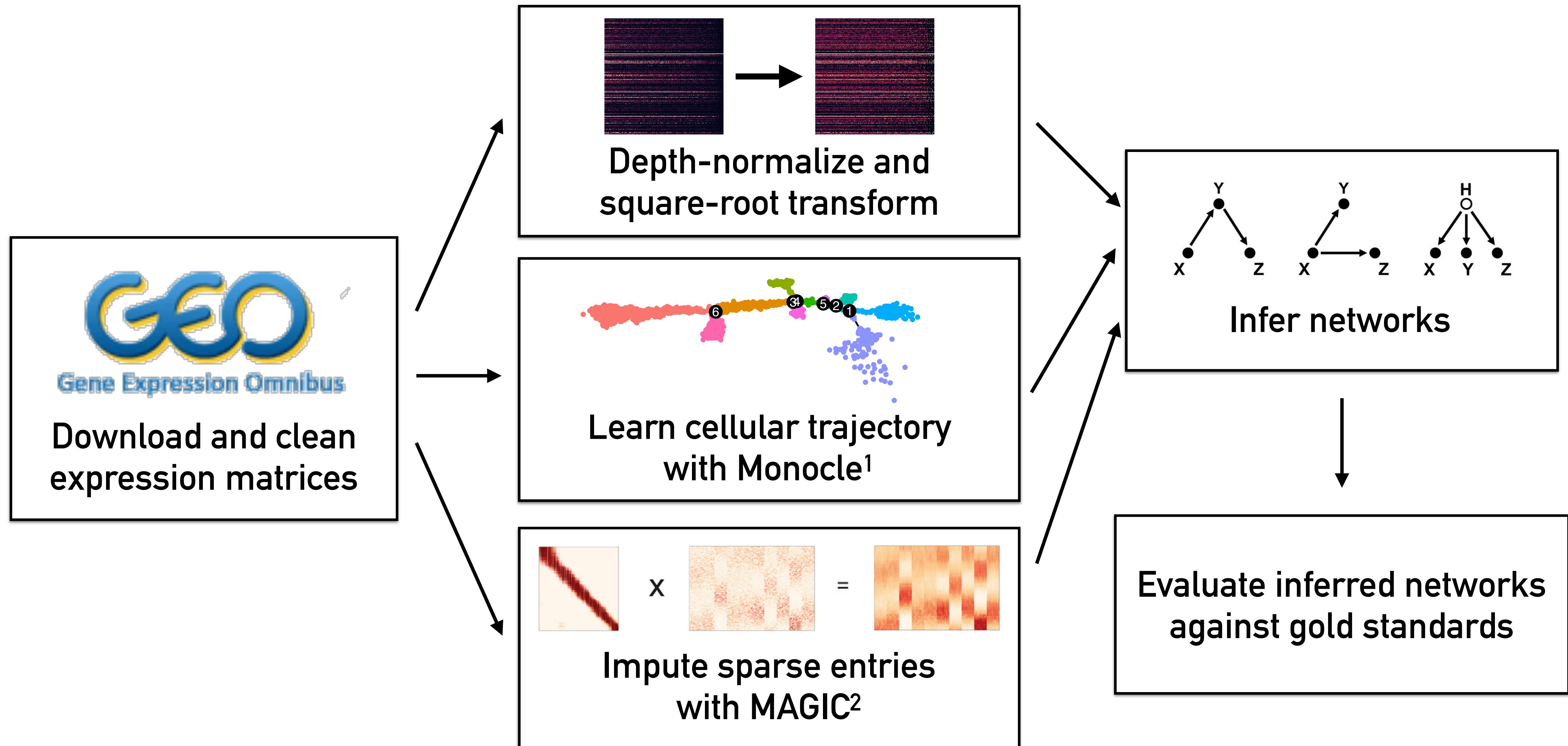
Datasets used for evaluation

Citation	Species	Cell type	GEO ID
Gasch et al., <i>PLoS Biol</i> 2017	Yeast	Yeast	GSE102475
Han et al., <i>Genome Biol</i> 2018	Human	hESC	GSE107552
Shalek et al., <i>Nature</i> 2014	Mouse	DC	GSE48968
Tran et al., <i>Cell Reports</i> 2019*	Mouse	mESC	GSE108222
Tran et al., <i>Cell Reports</i> 2019*	Mouse	mESC	GSE108222
Zhao et al., <i>Cell Stem Cell</i> 2018	Mouse	mESC	GSE114952

*Tran et al. includes two experimental datasets



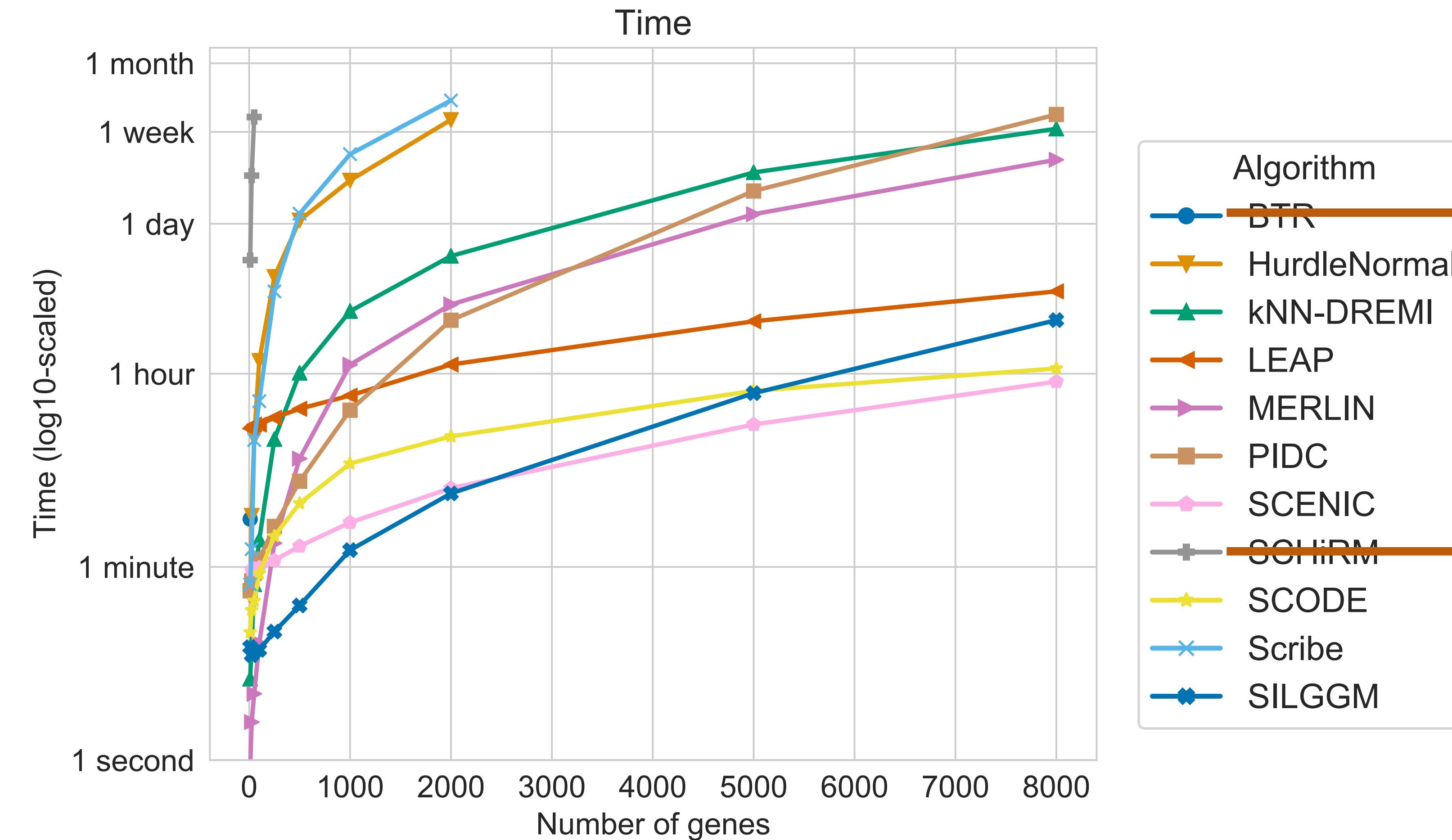
Overview of evaluation pipeline



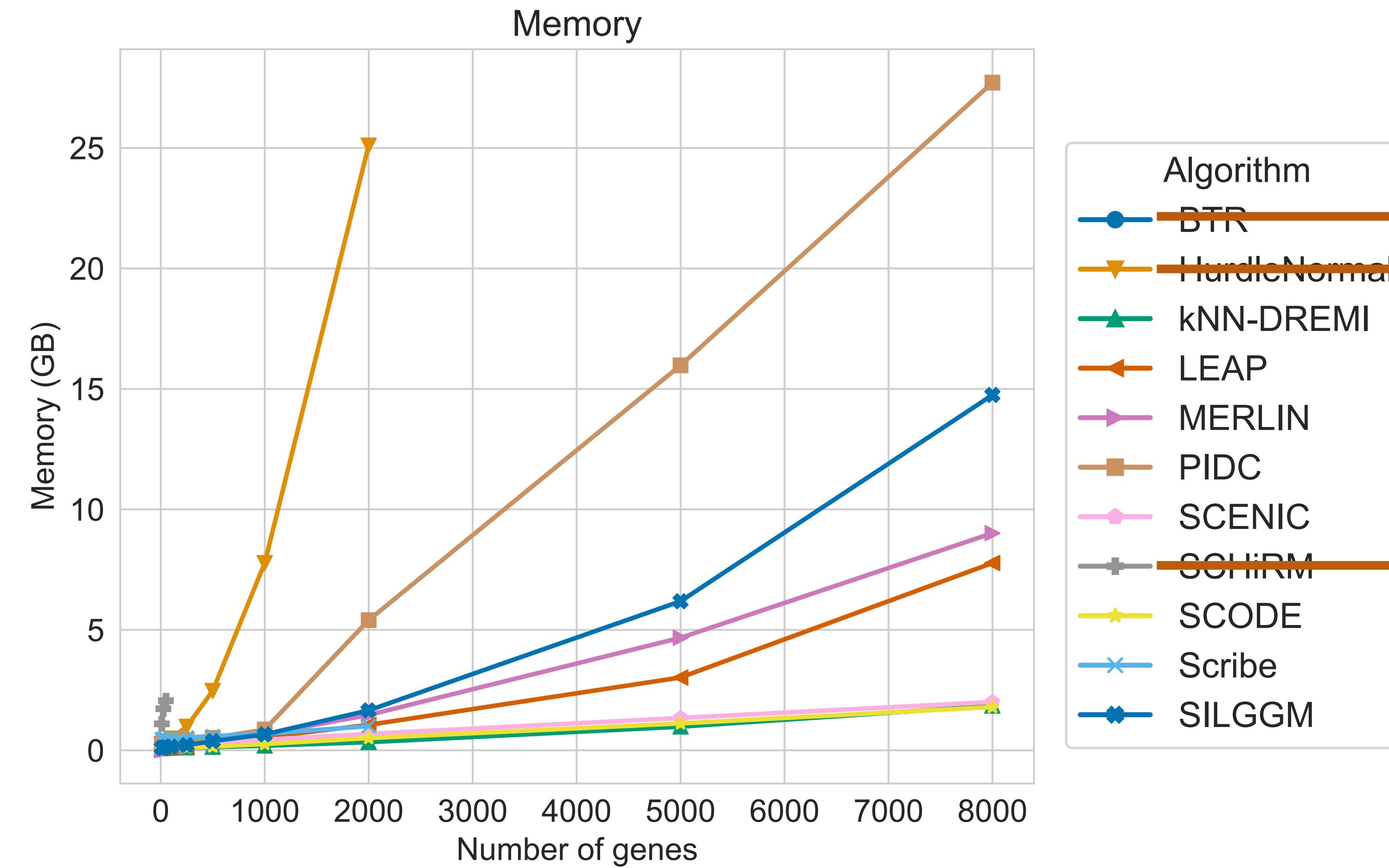
Open questions in network inference

- How do network inference algorithms scale to large scRNA-seq datasets?
- How accurate are the networks inferred by each method?
- How similar are the networks inferred by different methods?
- Does imputation of sparse expression data improve network accuracy?

Runtime as potential regulatory network grows



Memory as potential regulatory network grows

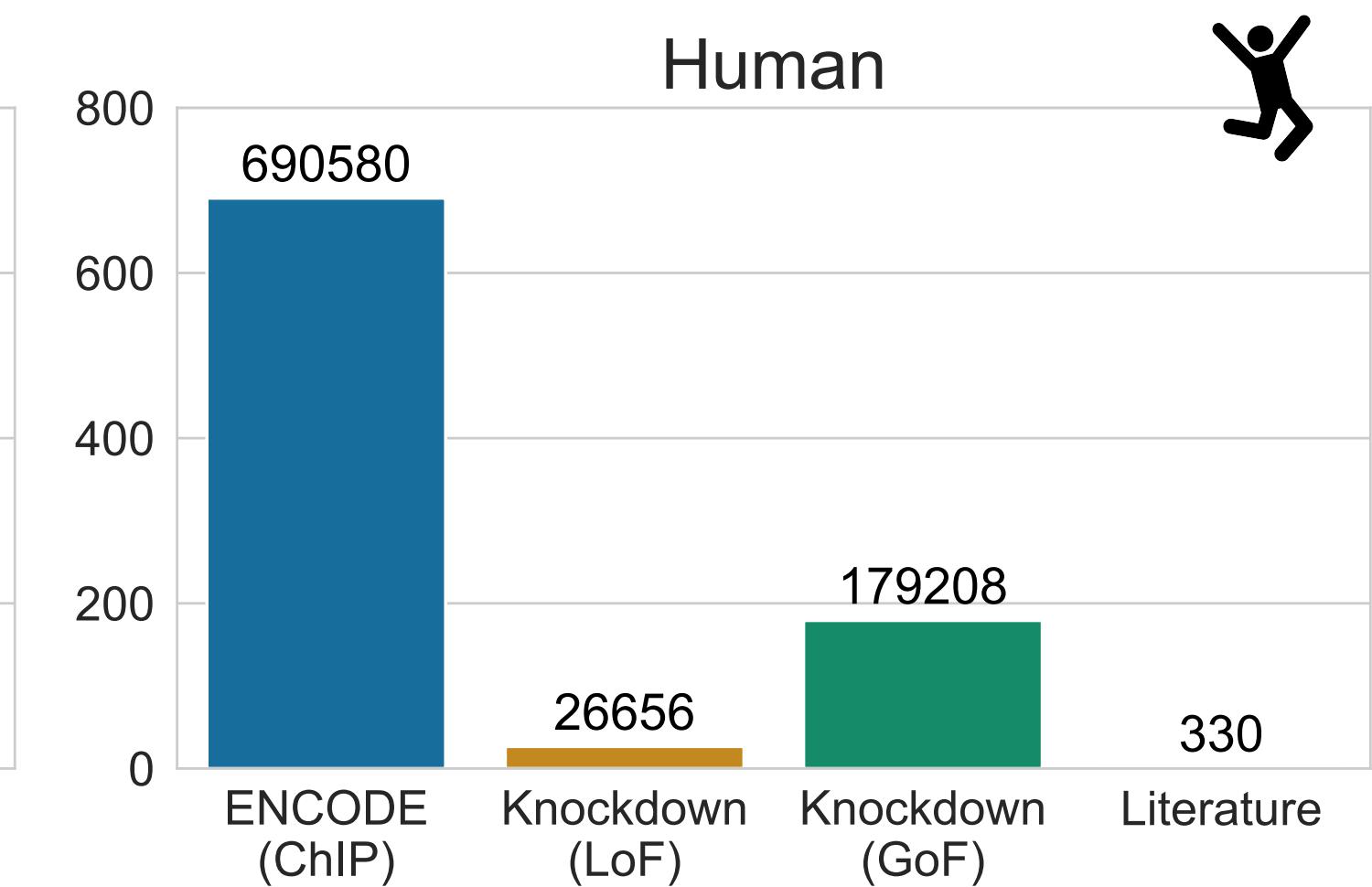
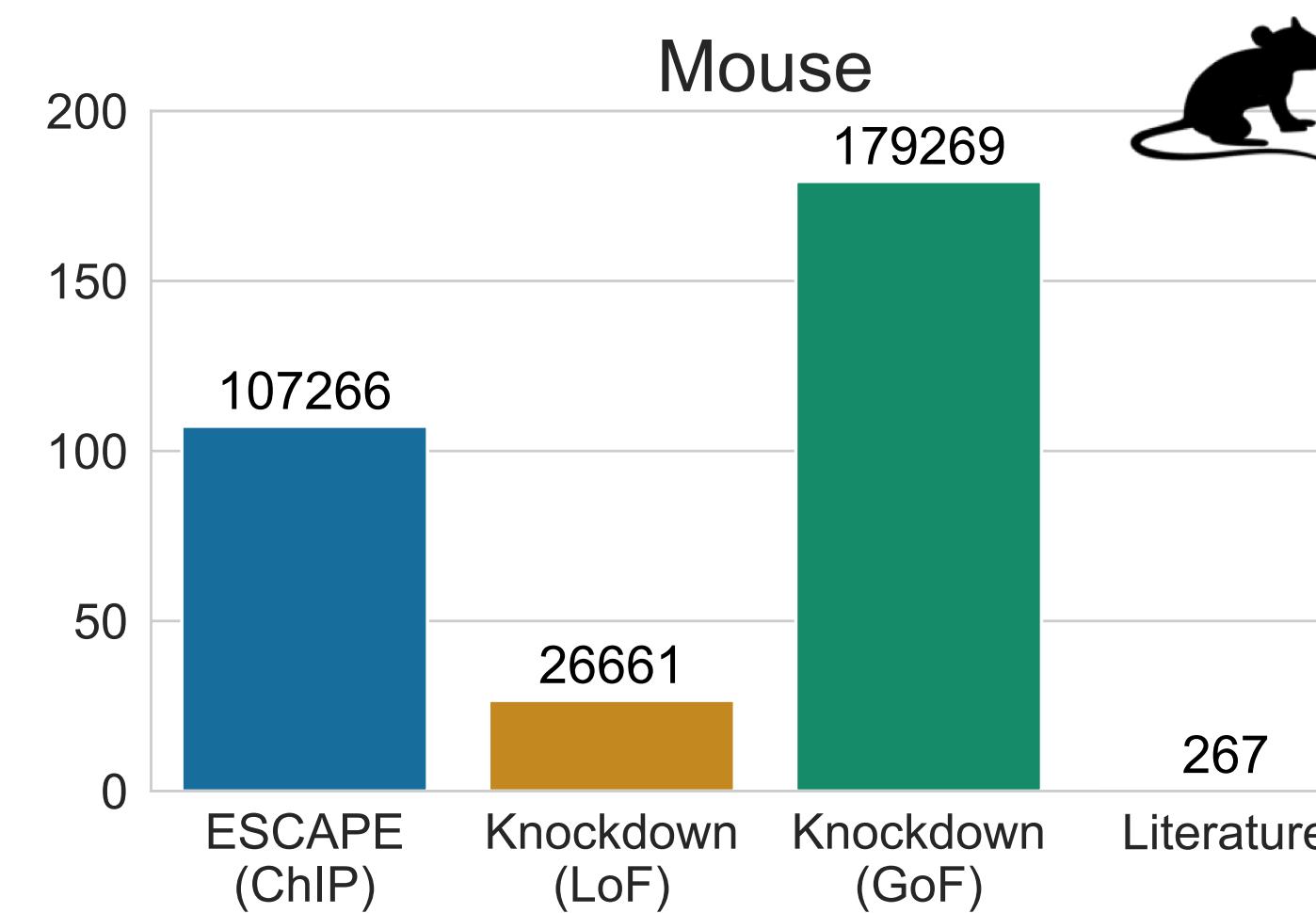
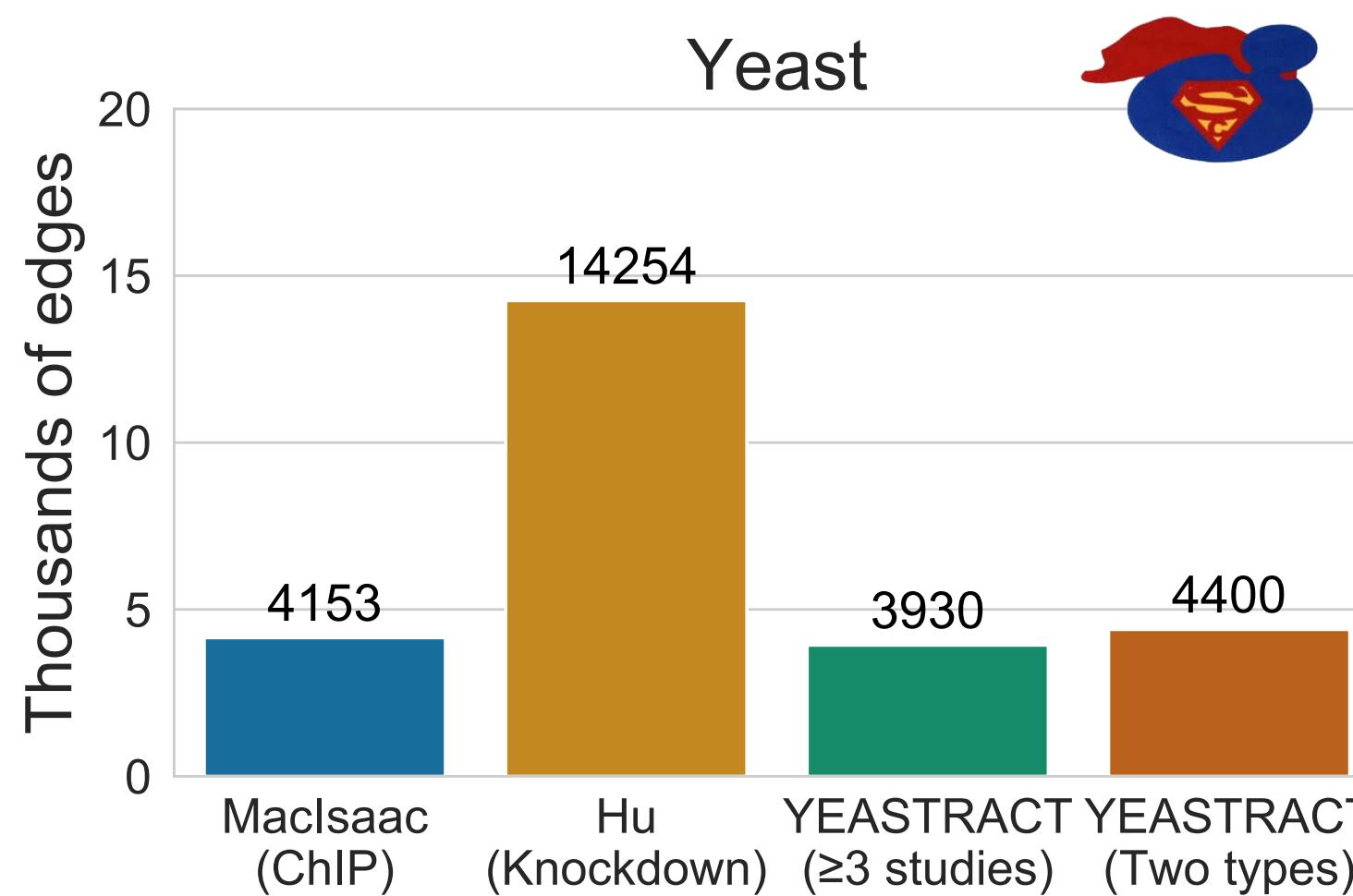
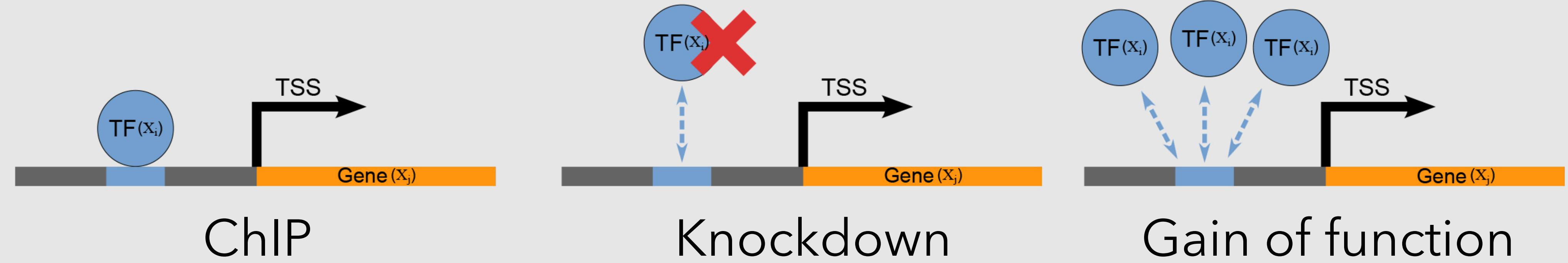


Open questions in network inference

- How do network inference algorithms scale to large scRNA-seq datasets?
- How accurate are the networks inferred by each method?
- How similar are the networks inferred by different methods?
- Does imputation of sparse expression data improve network accuracy?

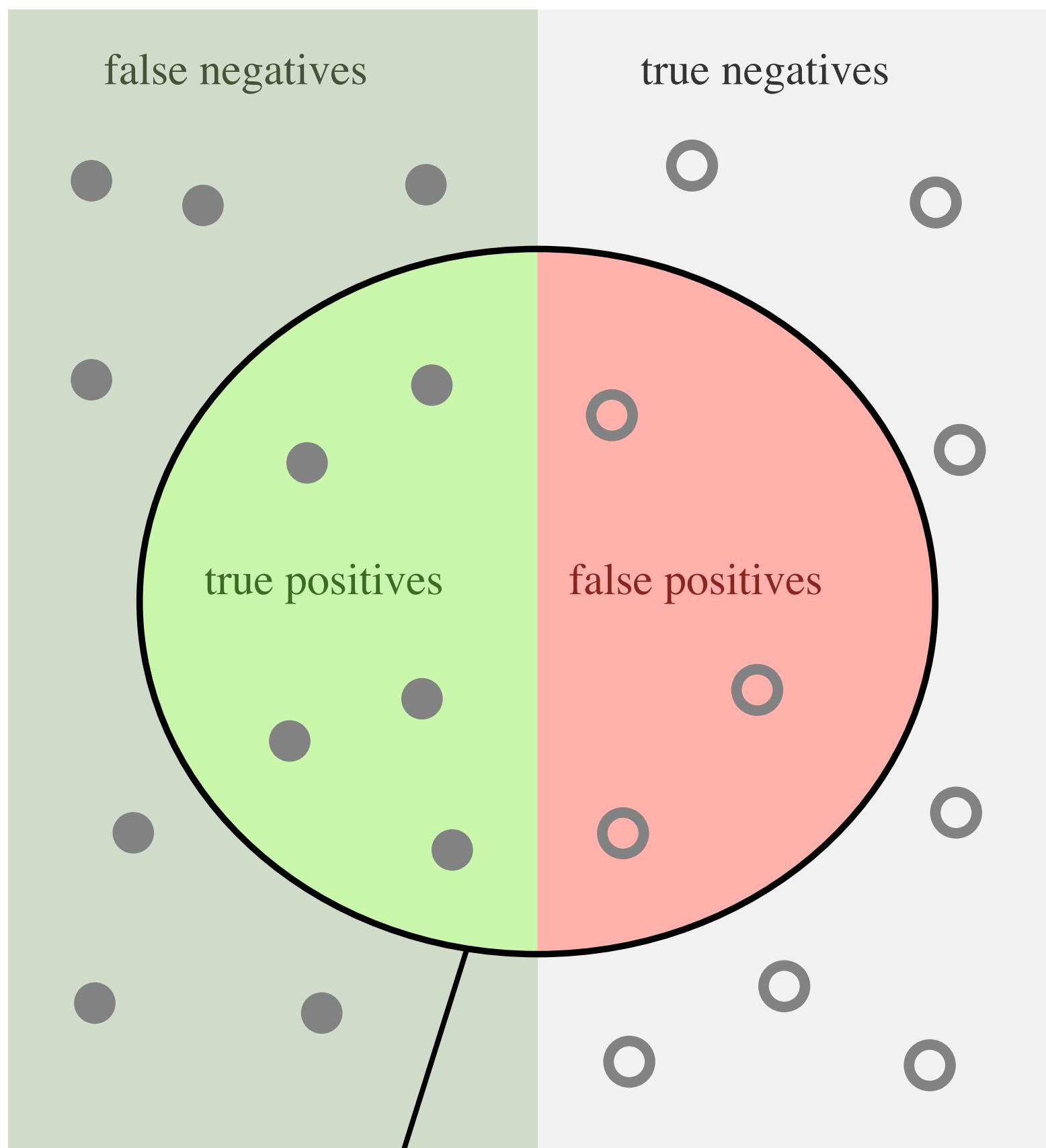
Gold standards

Evidence for regulatory interactions



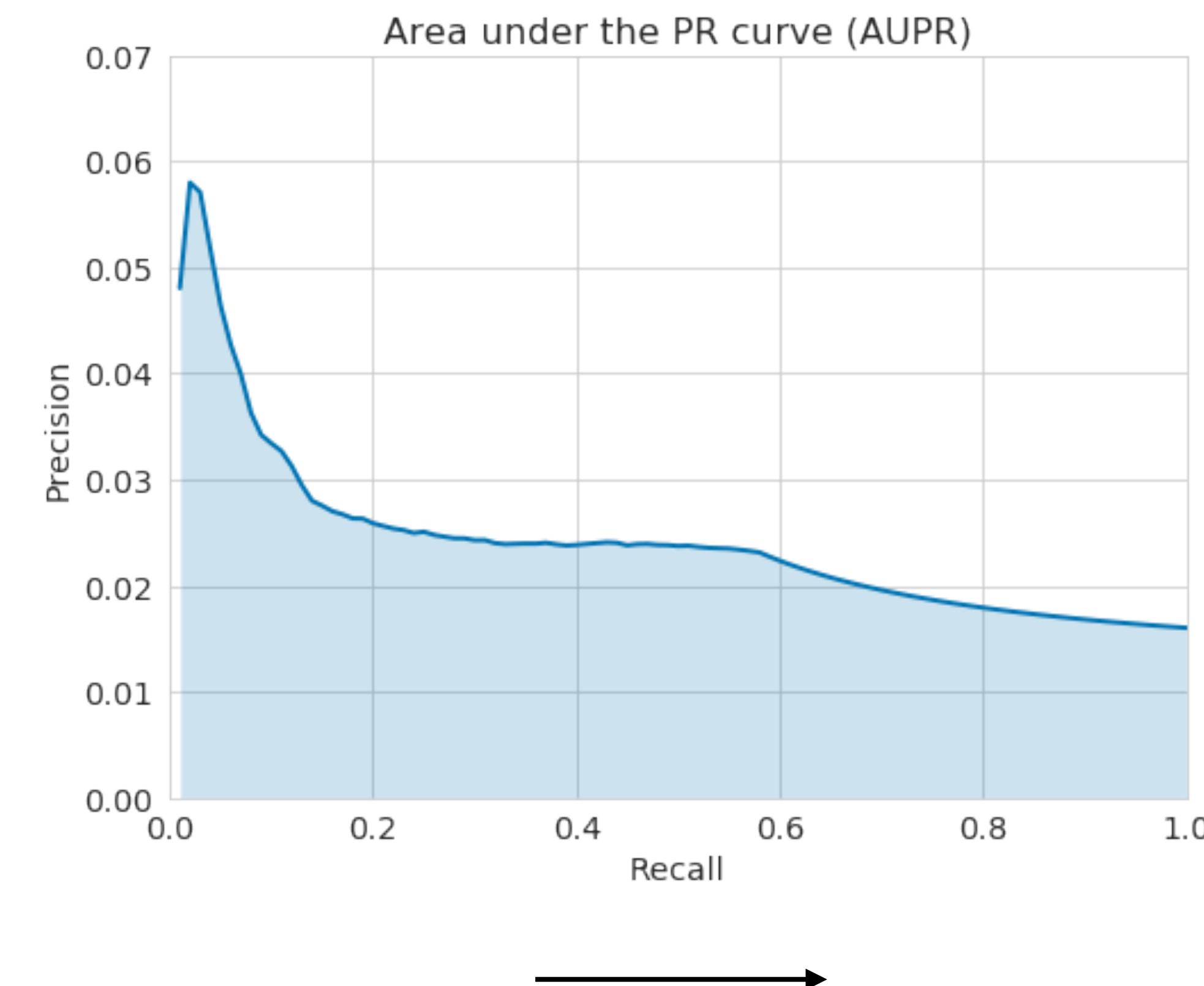
Evaluating network accuracy with precision and recall

Gold standard edges



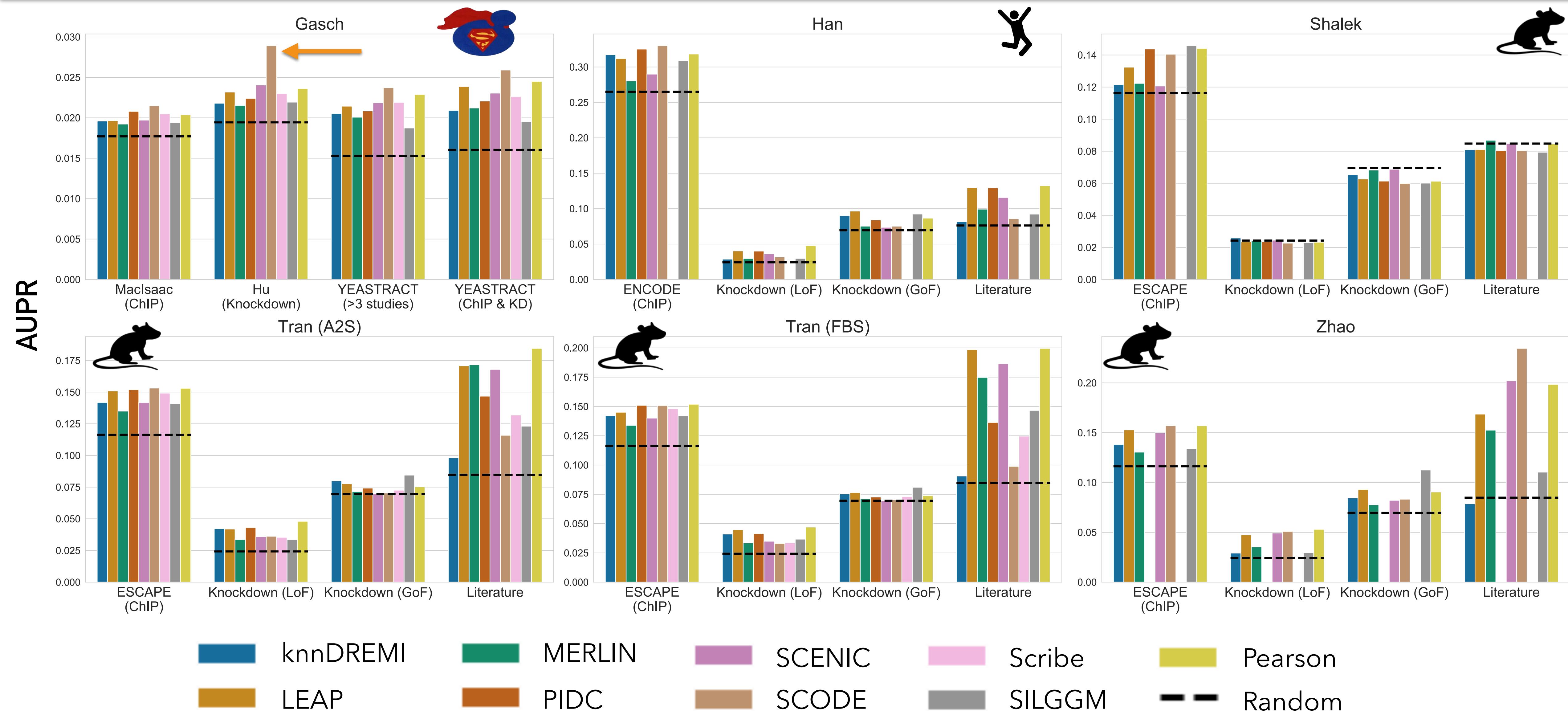
$$\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$$

$$\text{Recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$



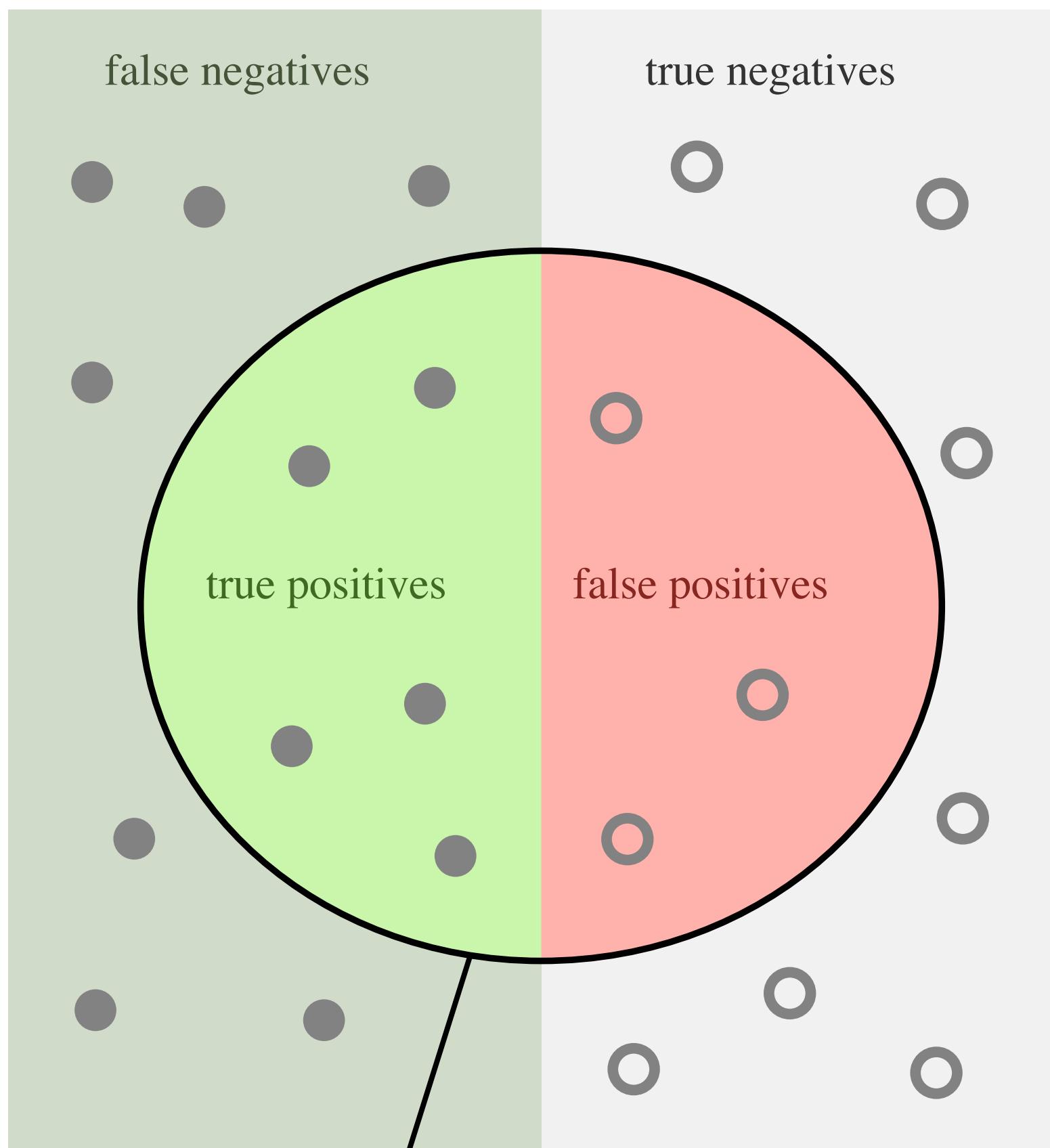
Increasing number of inferred edges

No clear winner emerges based on AUPR



Evaluating network accuracy with precision and recall

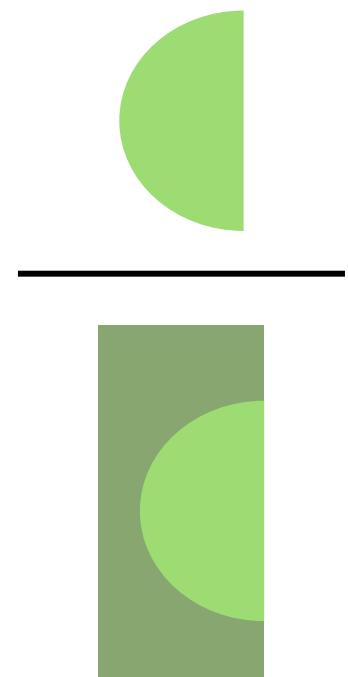
Gold standard edges



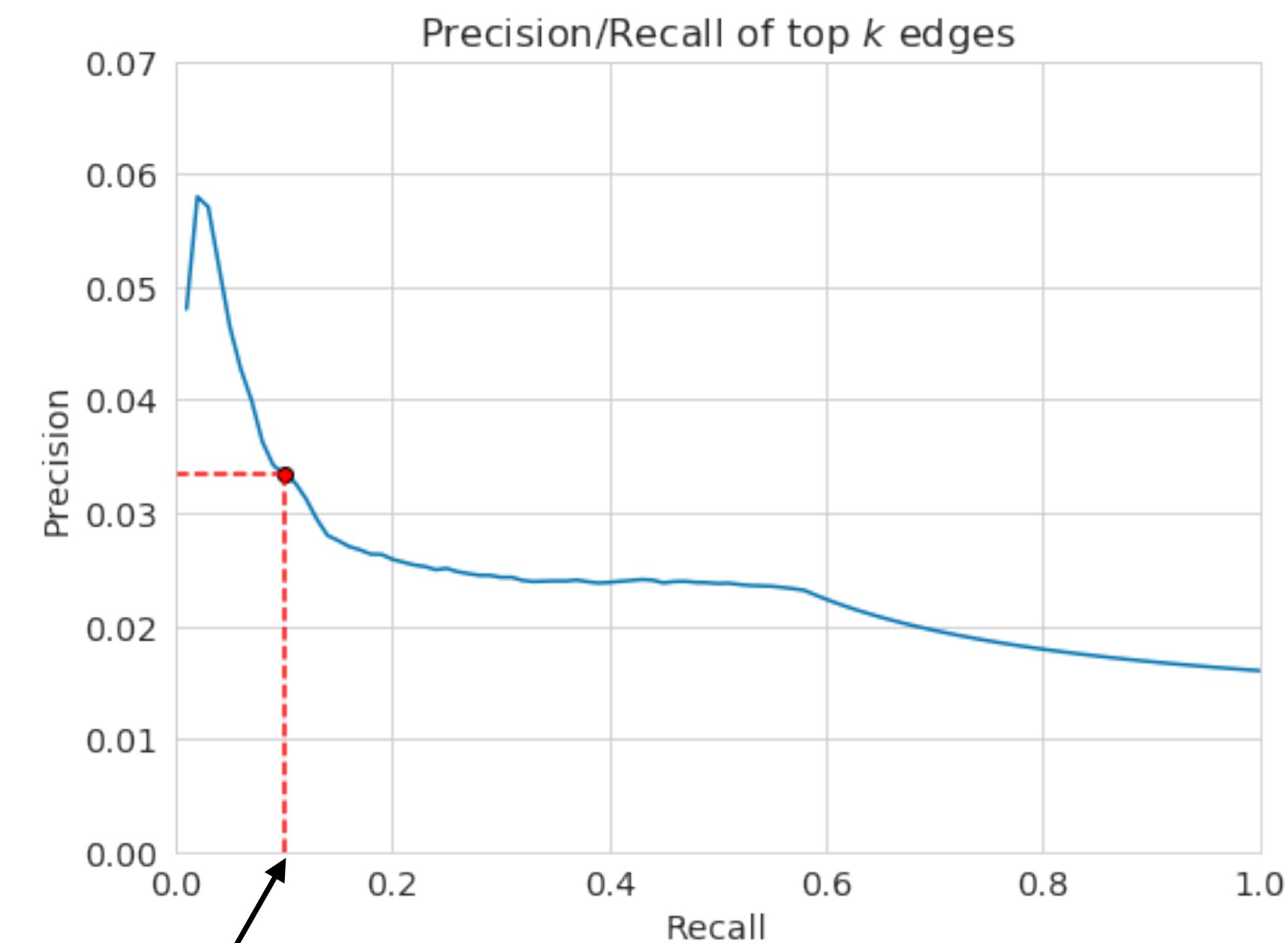
Precision =



Recall =

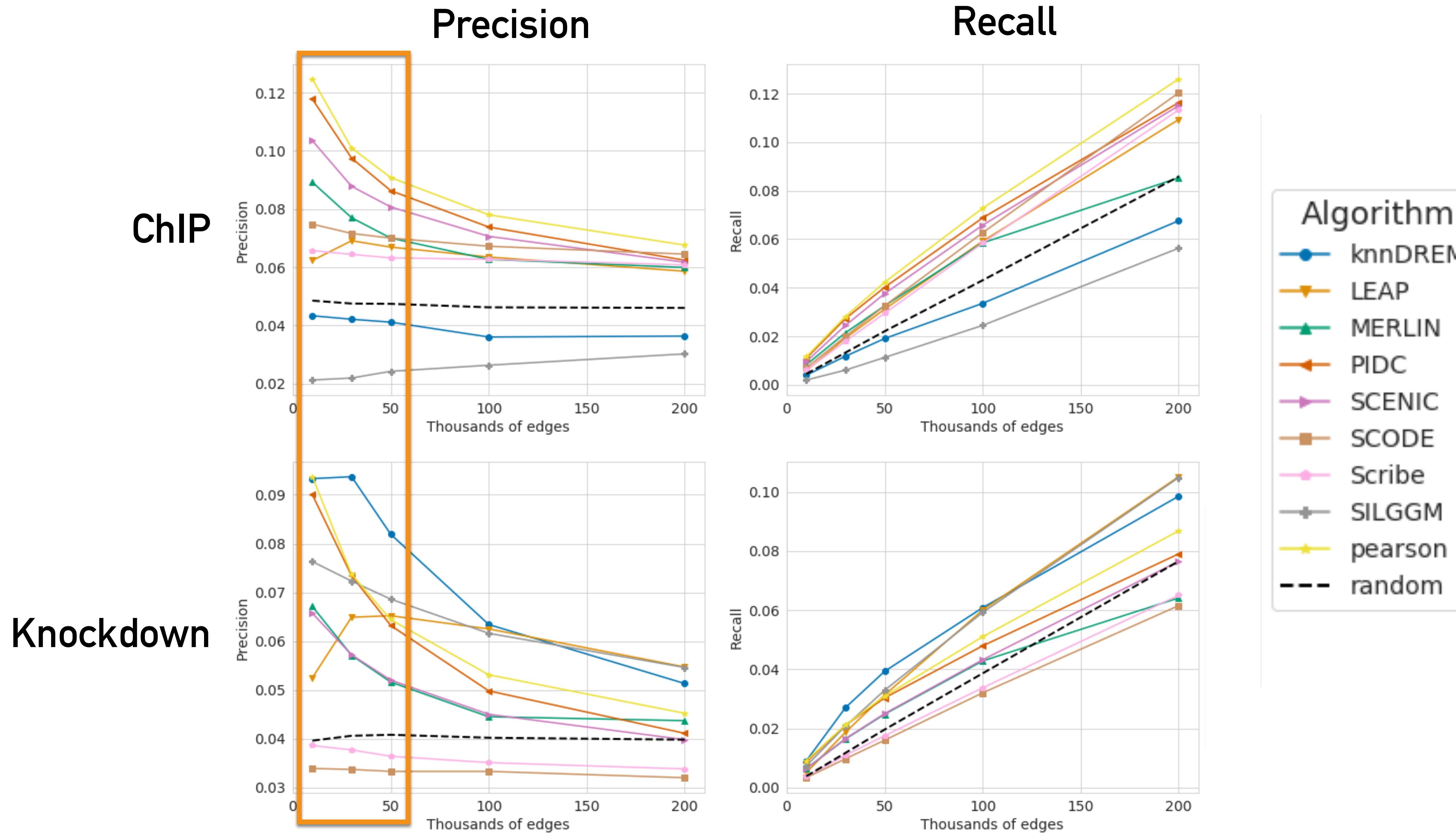


$$F\text{-score} = 2 * \frac{P * R}{P + R}$$

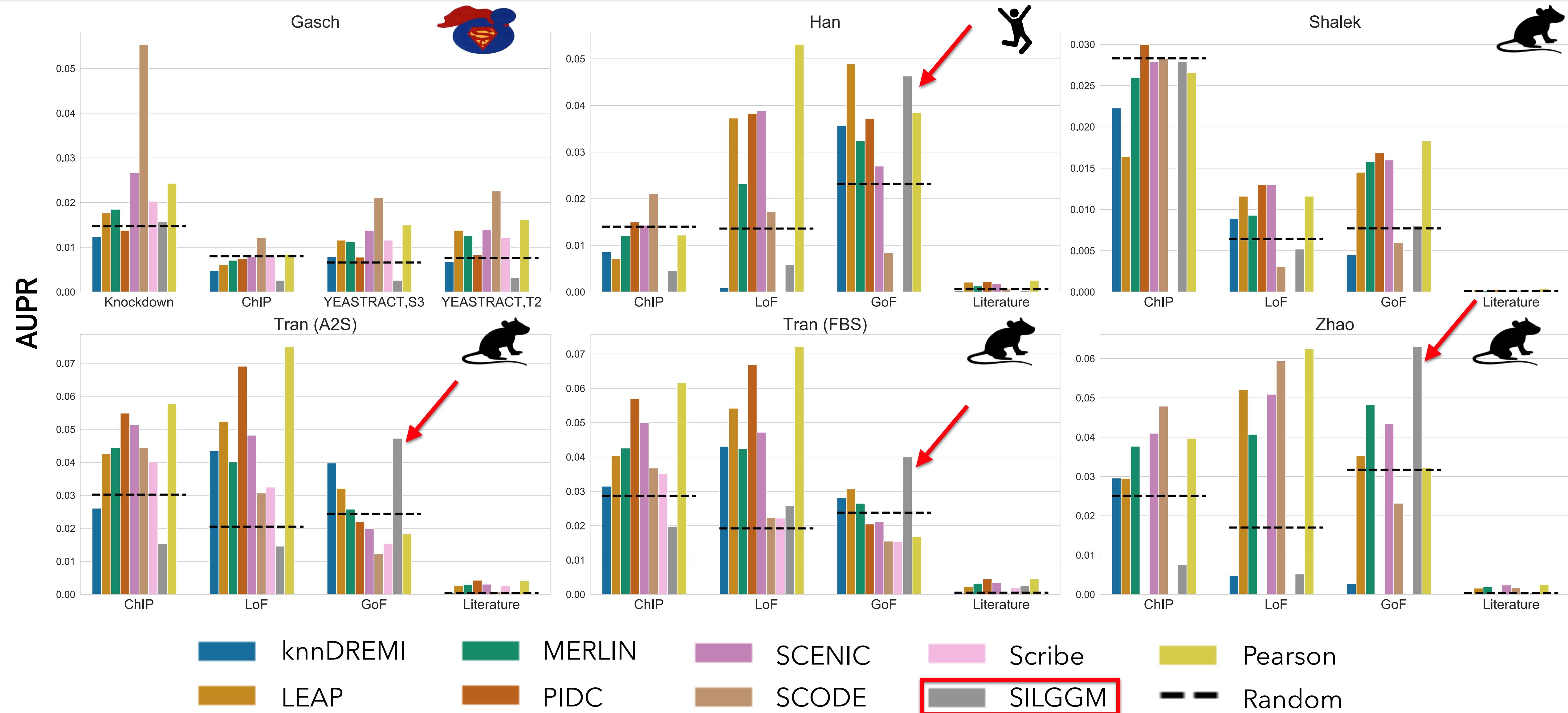


Top k edges

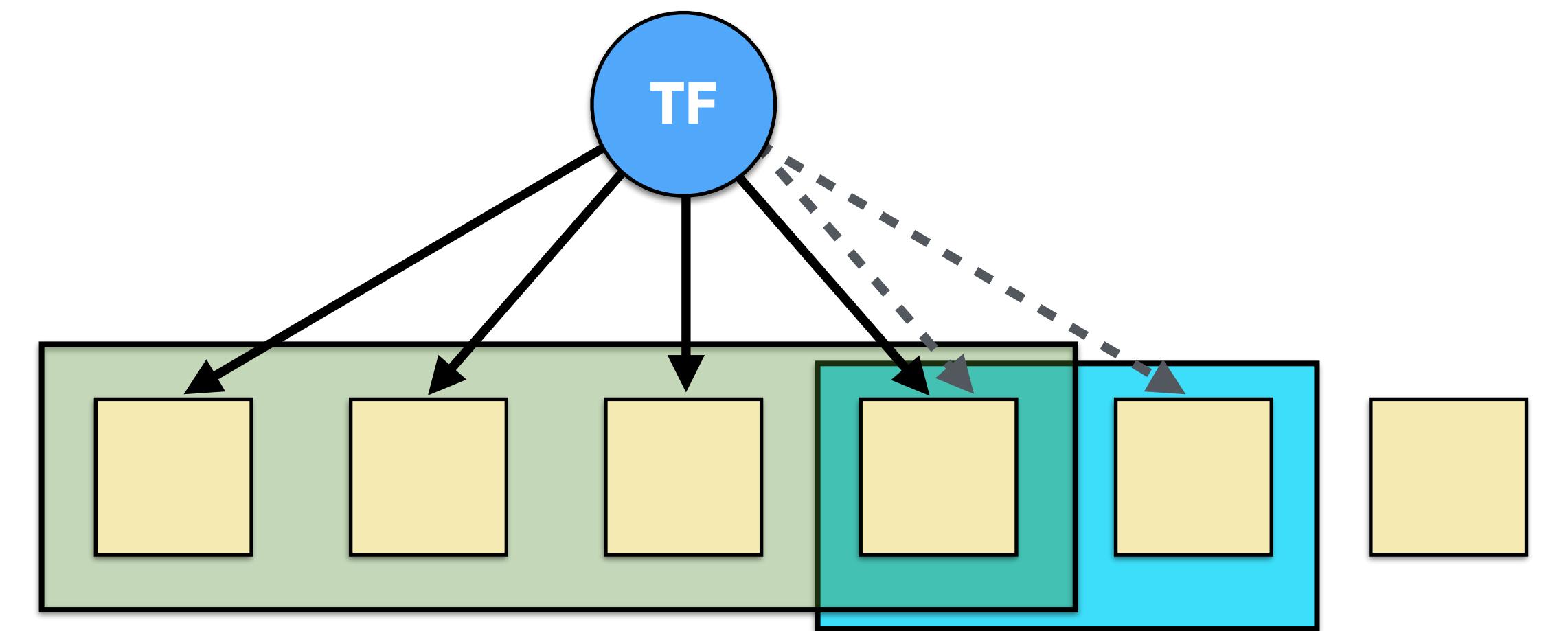
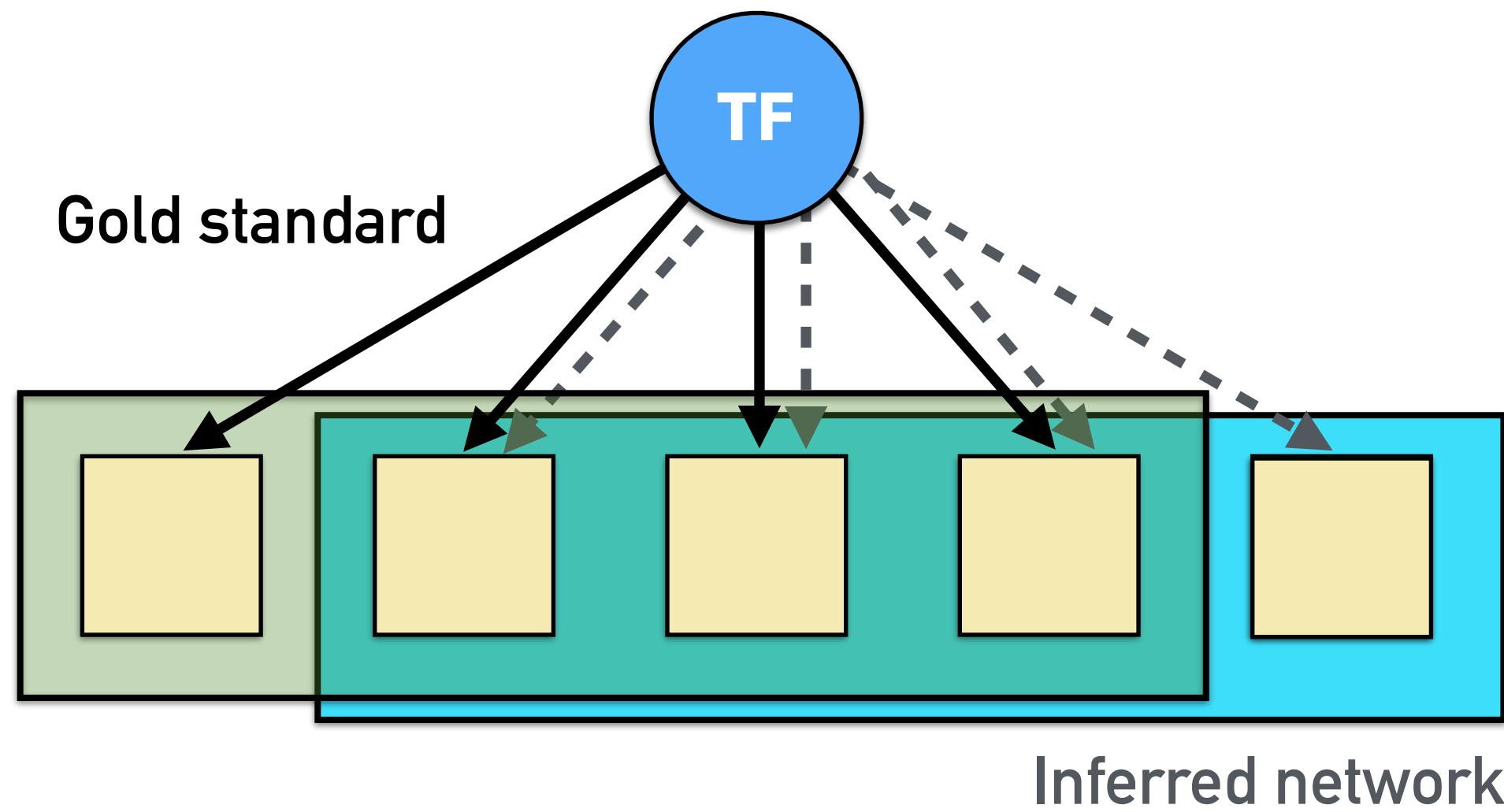
Focusing on top edges discriminates between methods



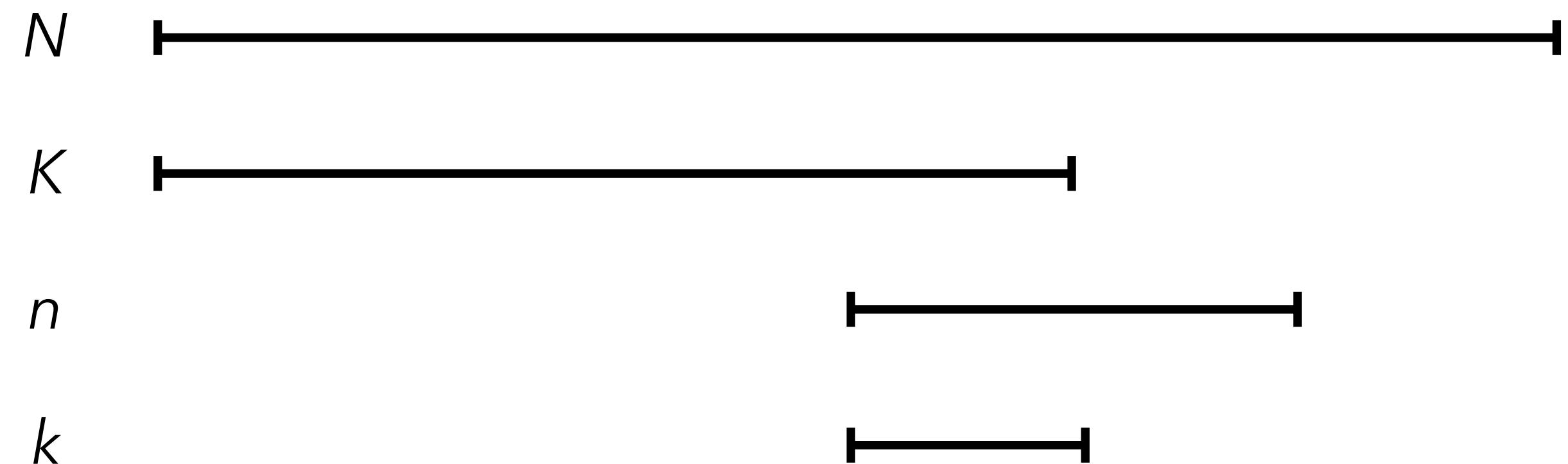
Focusing on top 50k edges distinguishes network accuracy



Predictable transcription factors



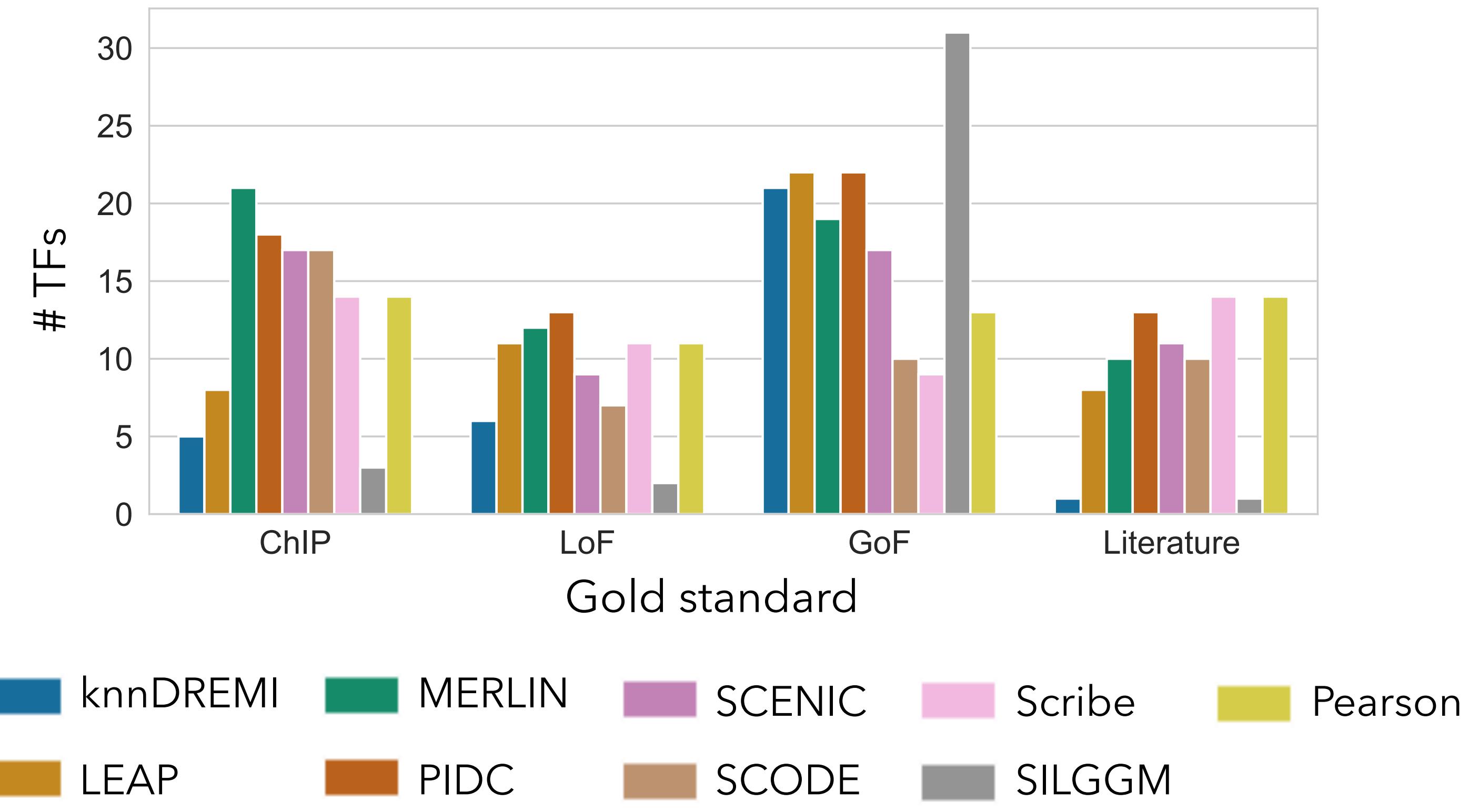
$$p = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}$$



Predictable transcription factors

	knnDREMI	LEAP	MERLIN	PIDC	SCENIC	SCODE	Scribe	SILGGM	Pearson
Chd1	1.77	1.06	1.55	1.37	1.89	1.09	0.42	1.56	
E2f4	0.62	1.96	1.25	1.43	1.33	1.34	1.17	0.73	1.46
Eed	1.90	0.00	0.45	0.97	0.20	0.50	0.94	1.48	0.67
Esrrb	0.85	1.02	1.21	1.18	1.29	1.17	1.27	1.27	1.14
Ezh2	4.48	0.14	0.61	0.19	0.41	0.41	0.99	1.77	0.50
Gli1	0.00	1.03	3.53	1.09	0.00	0.00	0.00	1.63	1.10
Jarid2	3.01	0.96	0.42	1.23	0.61	0.36	0.67	0.87	1.08
Klf2	4.70	0.71	2.37	3.91	3.32	1.18	1.68	5.81	5.16
Klf4	0.87	1.13	1.18	1.64	1.37	1.49	1.29	0.94	1.69
Klf5	1.24	0.00	2.98	7.46	3.10	0.00	3.82	1.78	9.92
Max	0.83	1.26	1.14	1.08	1.25	1.09	1.03	0.00	
Myc	2.22	1.54	1.31	1.43	1.53	1.31	0.54	1.00	
Mycn	0.30	0.39	1.20	1.06	1.25	1.56	1.35	0.41	1.15
Nanog	0.76	1.21	1.30	1.18	1.19	1.16	1.12	1.30	1.15
Nr0b1	0.64	1.37	1.72	1.73	2.09	1.38	1.67	1.34	1.63
Oct4	0.75	1.00	1.28	1.22	1.38	1.32	1.13	1.01	1.13
Phc1	2.70	0.94	0.79	1.18	0.63	0.68	0.79	0.26	1.08
Prdm14		1.38	1.96	1.36	2.23	2.01	1.00	1.86	
Rest	1.02		0.86	1.21	1.11	0.62	0.69	1.07	1.08
Rnf2	3.13	0.00	0.57	0.17	0.66	0.51	0.78	1.35	0.16
Sall1	0.29	0.00	1.27	1.88	2.03	1.44	2.61	0.99	1.63
Sall4	0.62	1.63	1.33	1.36	1.36	1.59	1.70	0.90	1.31
Smad1	0.00	1.20	1.21	0.53	0.75	1.67	0.87	0.64	0.00
Sox2	0.70	0.96	1.47	1.24	1.43	1.27	1.73	1.33	1.19
Stat3	0.00	0.89	0.67	0.92	1.31	1.33	1.00	1.19	
Suz12	2.37	0.77	0.56	0.44	0.48	0.47	0.84	1.84	0.70
Tbx3	0.65	0.53	1.72	2.05	1.16	0.60	2.46	0.89	0.00
Tcf3	0.87	1.22	1.10	0.92	1.14	1.23	1.14	1.09	1.36
Trim28	0.85	1.07	1.02	1.05	1.07	1.23	1.18	1.10	1.07
Yy1	3.12	1.93	1.02	1.67	1.24	1.24	0.00	2.42	
Zfp281	0.00	1.26	1.59	2.22	1.14	1.62	1.41	0.28	3.41
Zfp42	0.50	0.94	1.40	1.28	1.41	1.26	1.33	1.13	1.18
Zfx	2.16	2.16	1.48	1.66	1.32	1.74	1.39	0.51	
Zic3	0.55	1.77	2.21	1.81	2.65	2.41	2.28	0.83	1.75

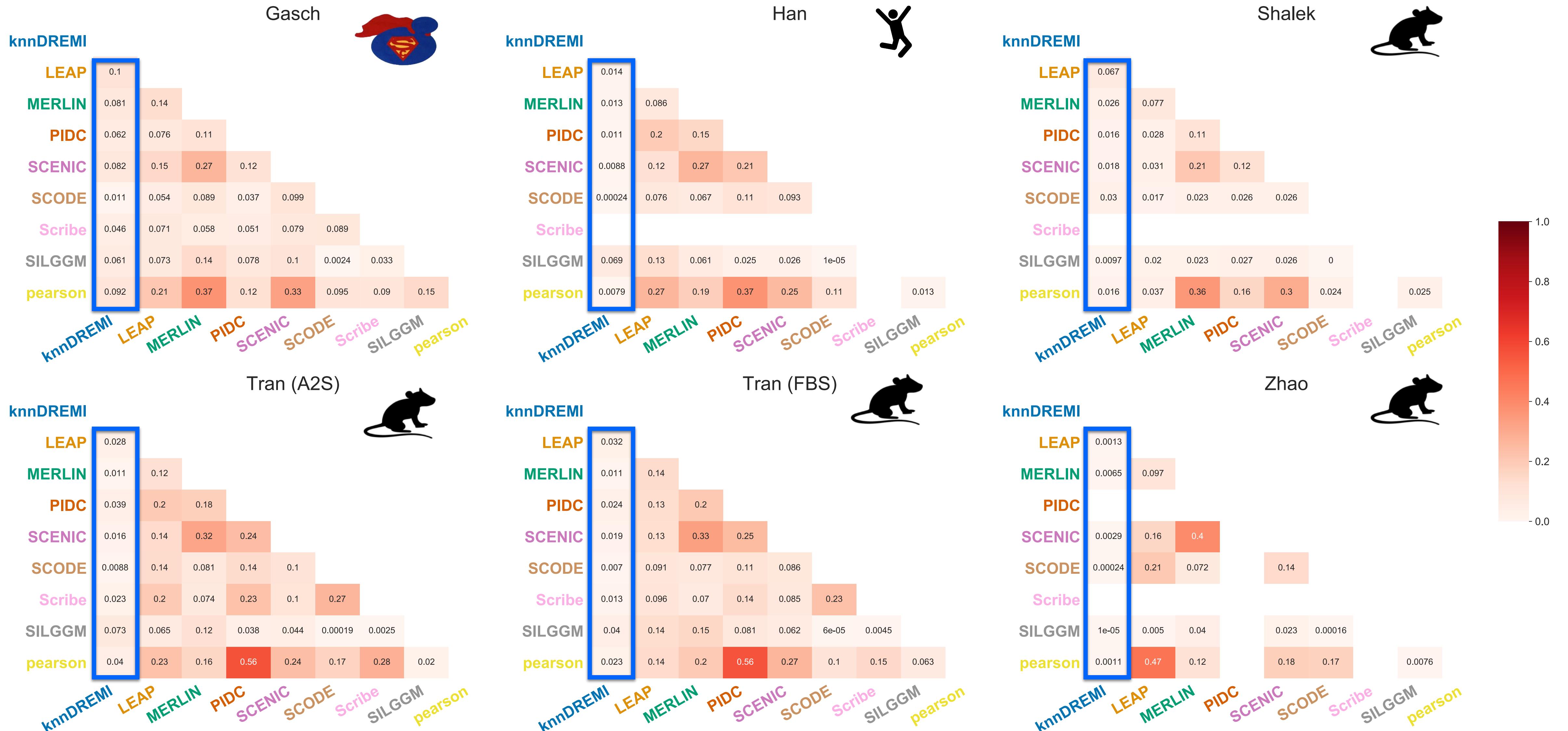
Predictable TFs in Tran (A2S)



Open questions in network inference

- How do network inference algorithms scale to large scRNA-seq datasets?
- How accurate are the networks inferred by each method?
- How similar are the networks inferred by different methods?
- Does imputation of sparse expression data improve network accuracy?

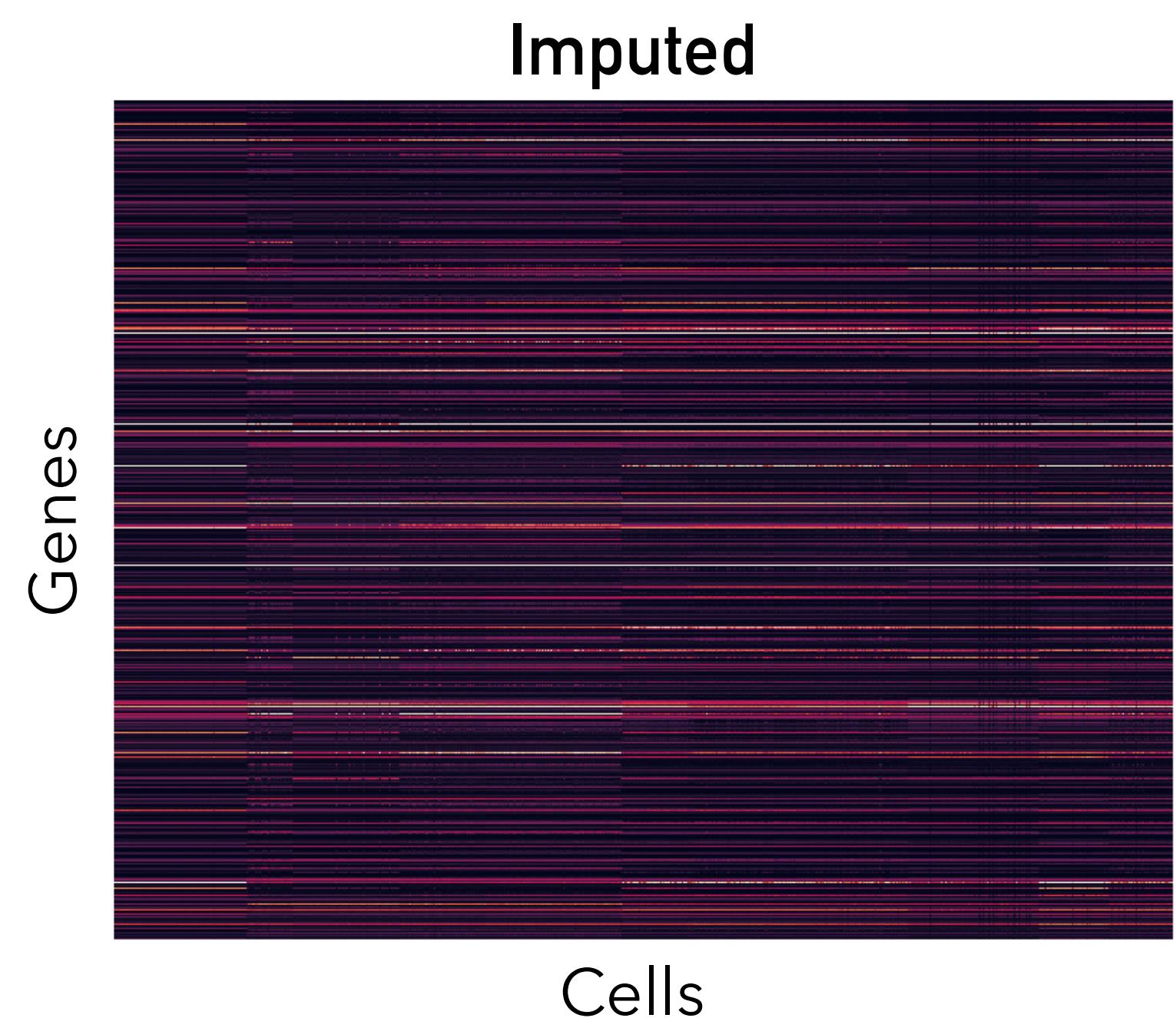
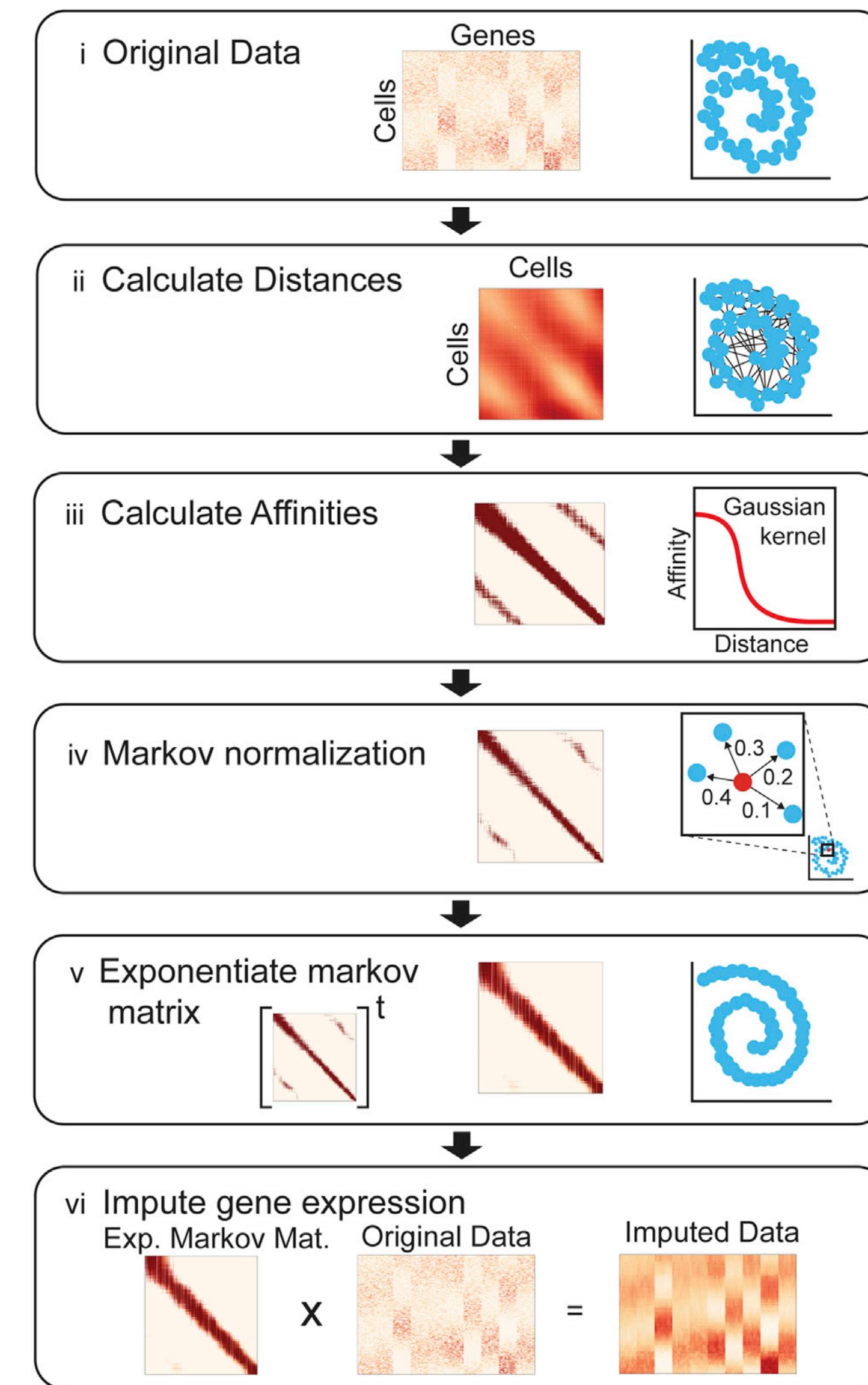
PIDC, SCENIC, MERLIN, correlation infer similar networks



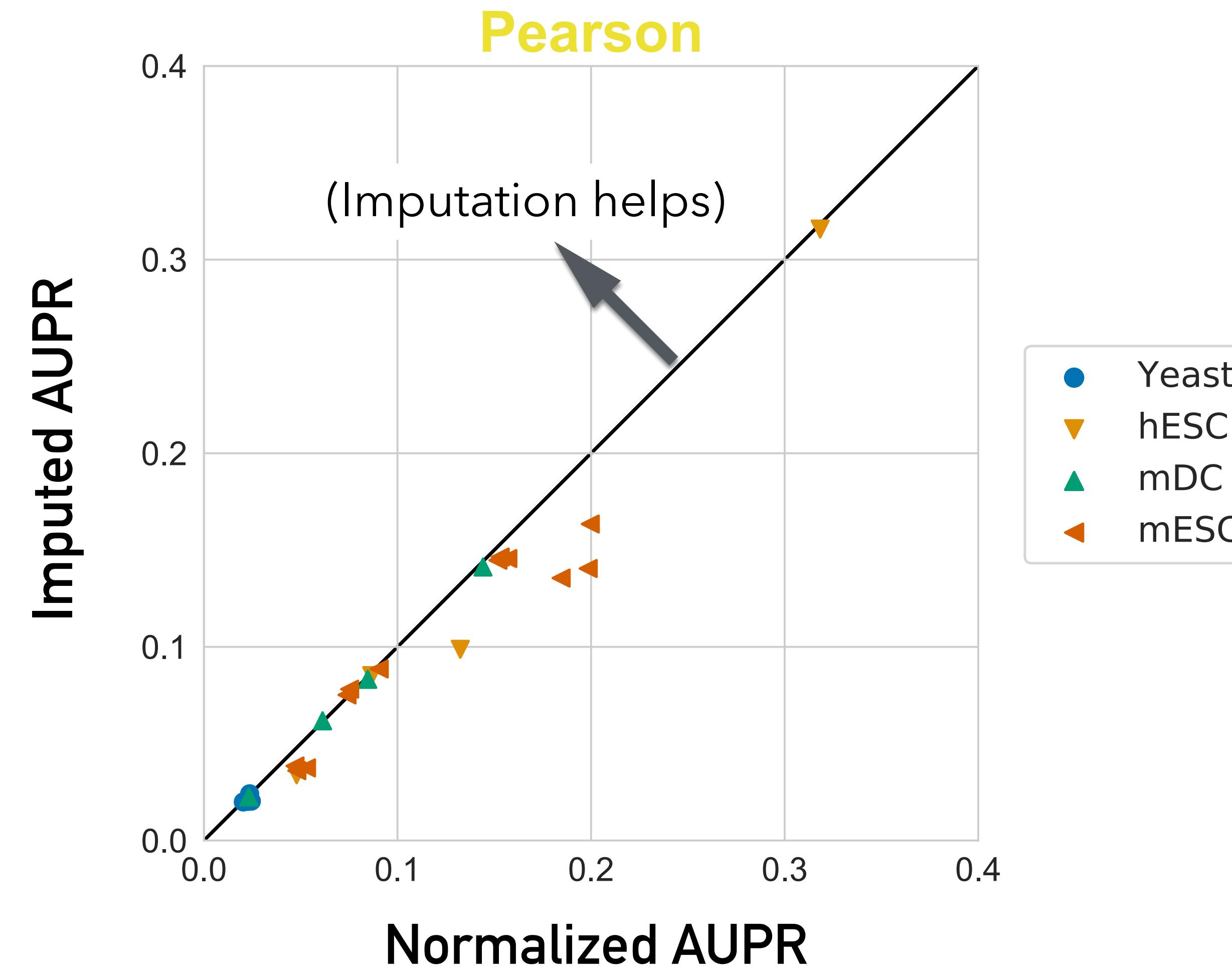
Open questions in network inference

- How do network inference algorithms scale to large scRNA-seq datasets?
- How accurate are the networks inferred by each method?
- How similar are the networks inferred by different methods?
- Does imputation of sparse expression data improve network accuracy?

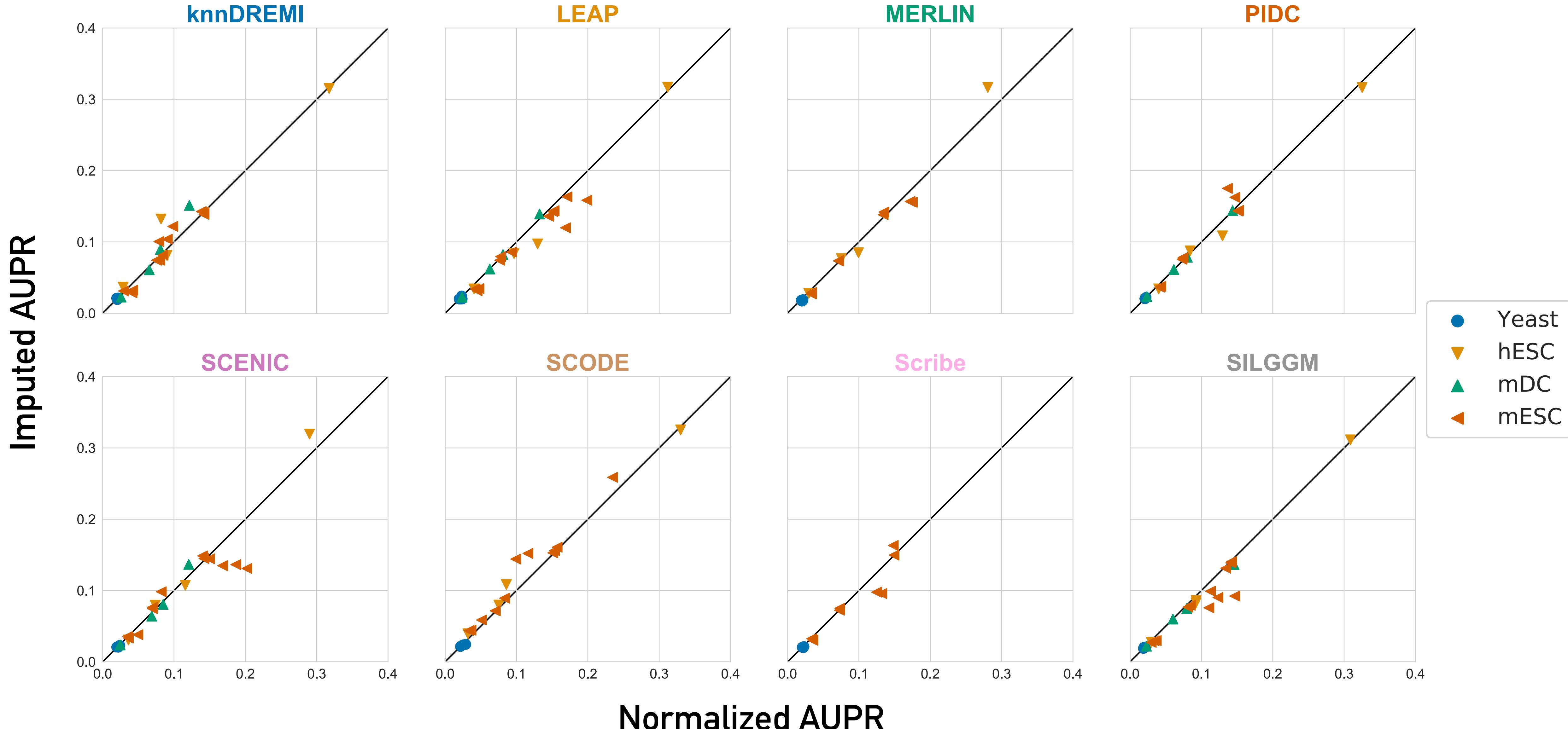
Imputing expression counts with MAGIC



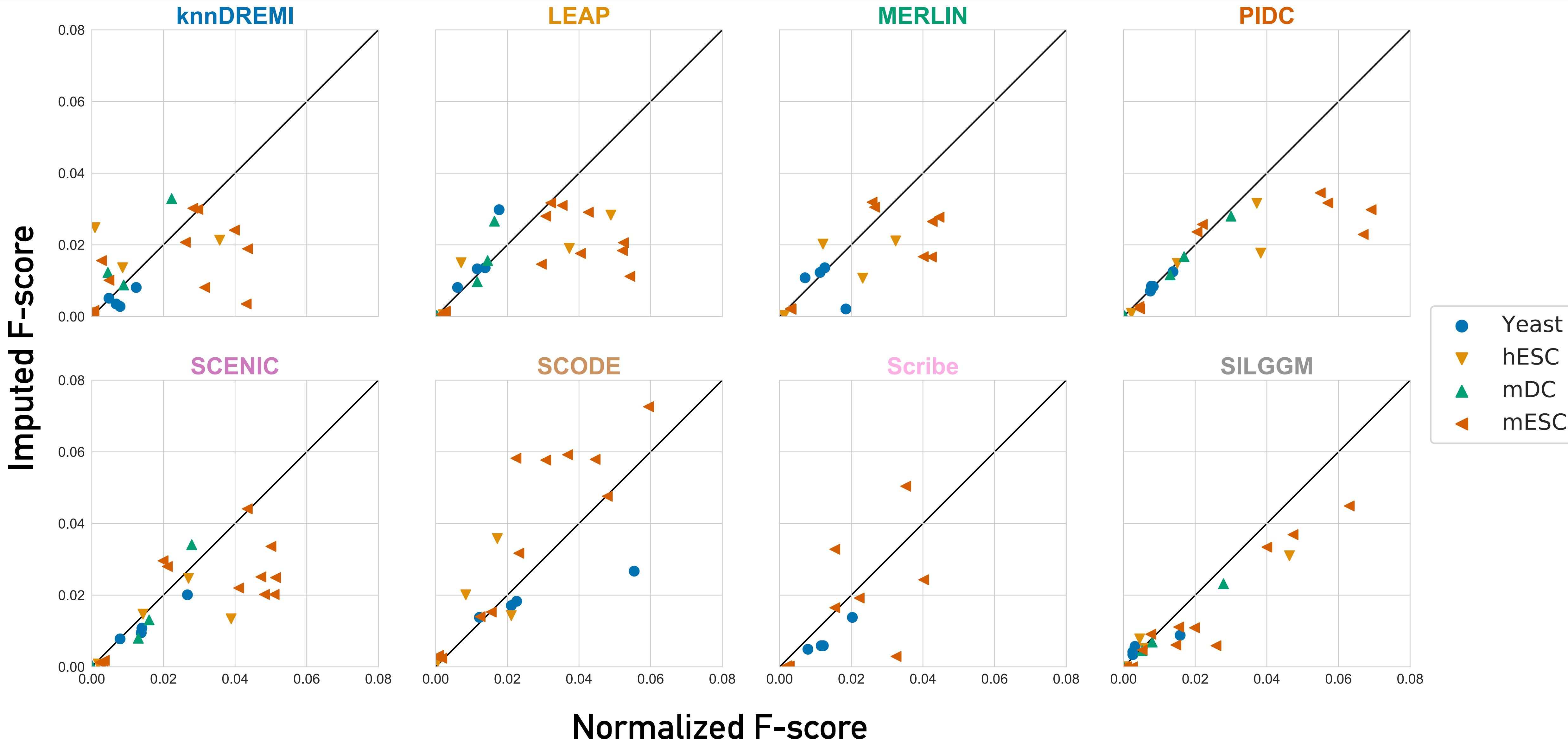
Comparing AUPR of imputed vs un-imputed data



Imputation generally does not improve AUPR



Imputation generally does not improve F-score in top edges



Conclusions

- AUPR doesn't distinguish a clear "best" inference method
- Examining top scoring edges identifies algorithms which outperform others in certain contexts
- In mammalian datasets, correlation and PIDC demonstrate highest accuracy towards ChIP and knockdown gold standards
- Imputation generally is not beneficial for network inference
- Demonstrates need for improved single-cell gold standards

Acknowledgments



Roy Lab



Viswesh
Periyasamy



Sunnie Grace
McCalla



Alireza Fotuhi
Siahpirani

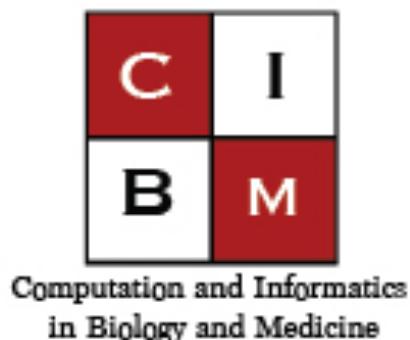


Sushmita
Roy

Brittany Baur
Sara Knaack
Da-Inn Lee

Junha Shin
Shilu Zhang

Funding and Support



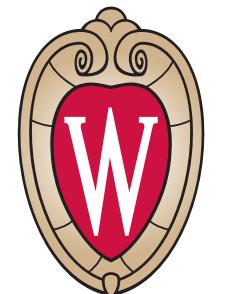
NLM Training Grant
5T15LM007359



ISMB/ECCB
Travel Fellowship



UW Data Science Initiative



Amanda Hua Stone



And the researchers that have shared
their data and code with the community!