

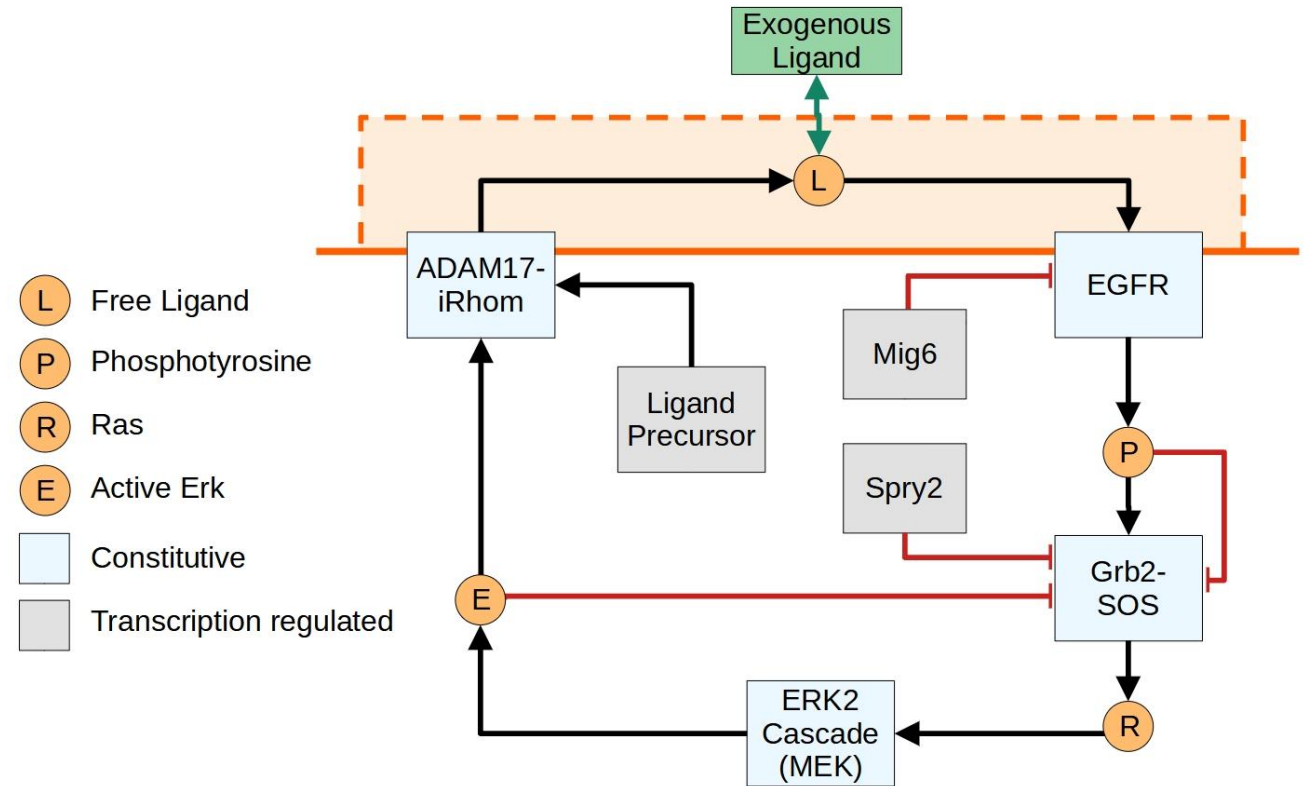
# Modeling the EGF Pathway: Challenges and Pitfalls

Michael Kochen

2022 CSBC West Coast Symposium

# Coarse-grained modeling

- Granularity in line with quantitative data
- Facilitate ease of interpretation
- Expand as warranted
- EGF model ->

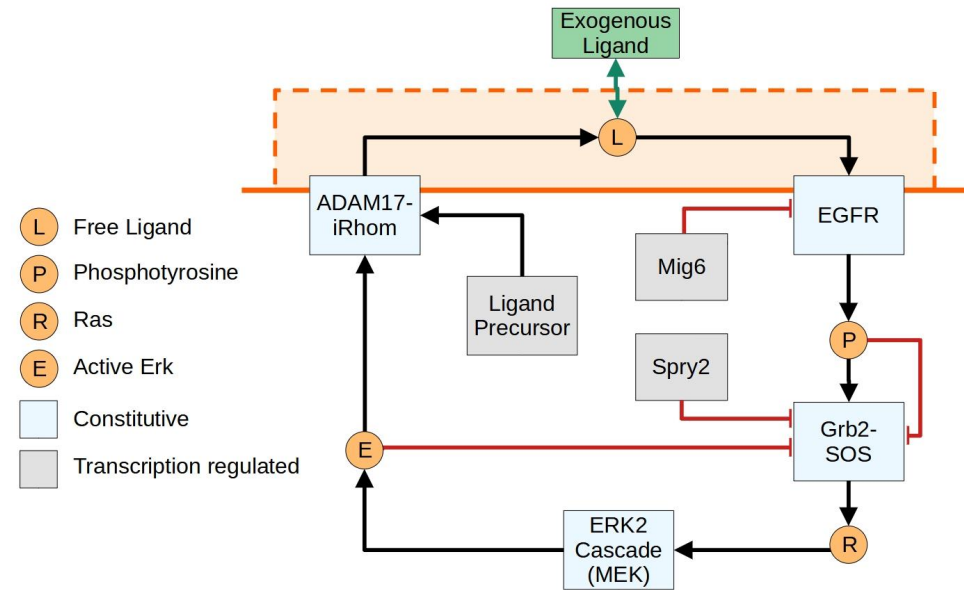
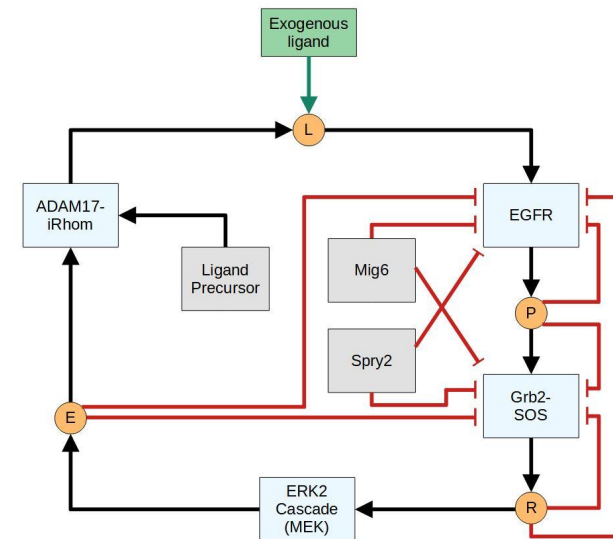
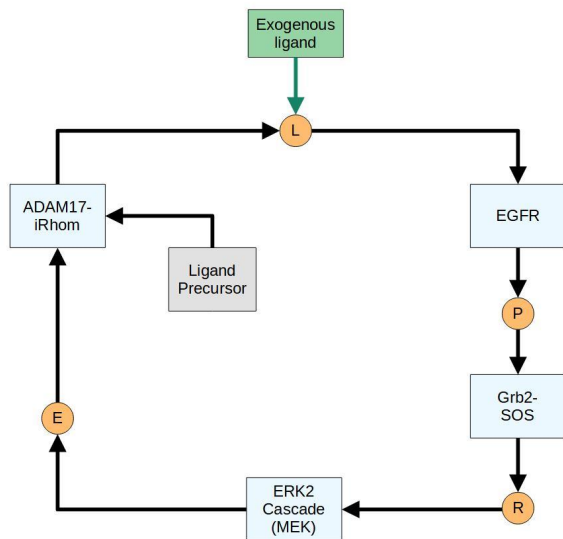


EGFR rate equation

$$\left( \frac{1}{\left( 1 + \frac{Mig6}{k_i} \right)} \right)^w \frac{k_s \left( \frac{L}{k_s^m} \right)^{m_s}}{\left( 1 + \frac{L}{k_s^m} \right)^{m_s}} - \frac{k_p \left( \frac{P}{k_p^m} \right)^{m_p}}{\left( 1 + \frac{P}{k_p^m} \right)^{m_p}}$$

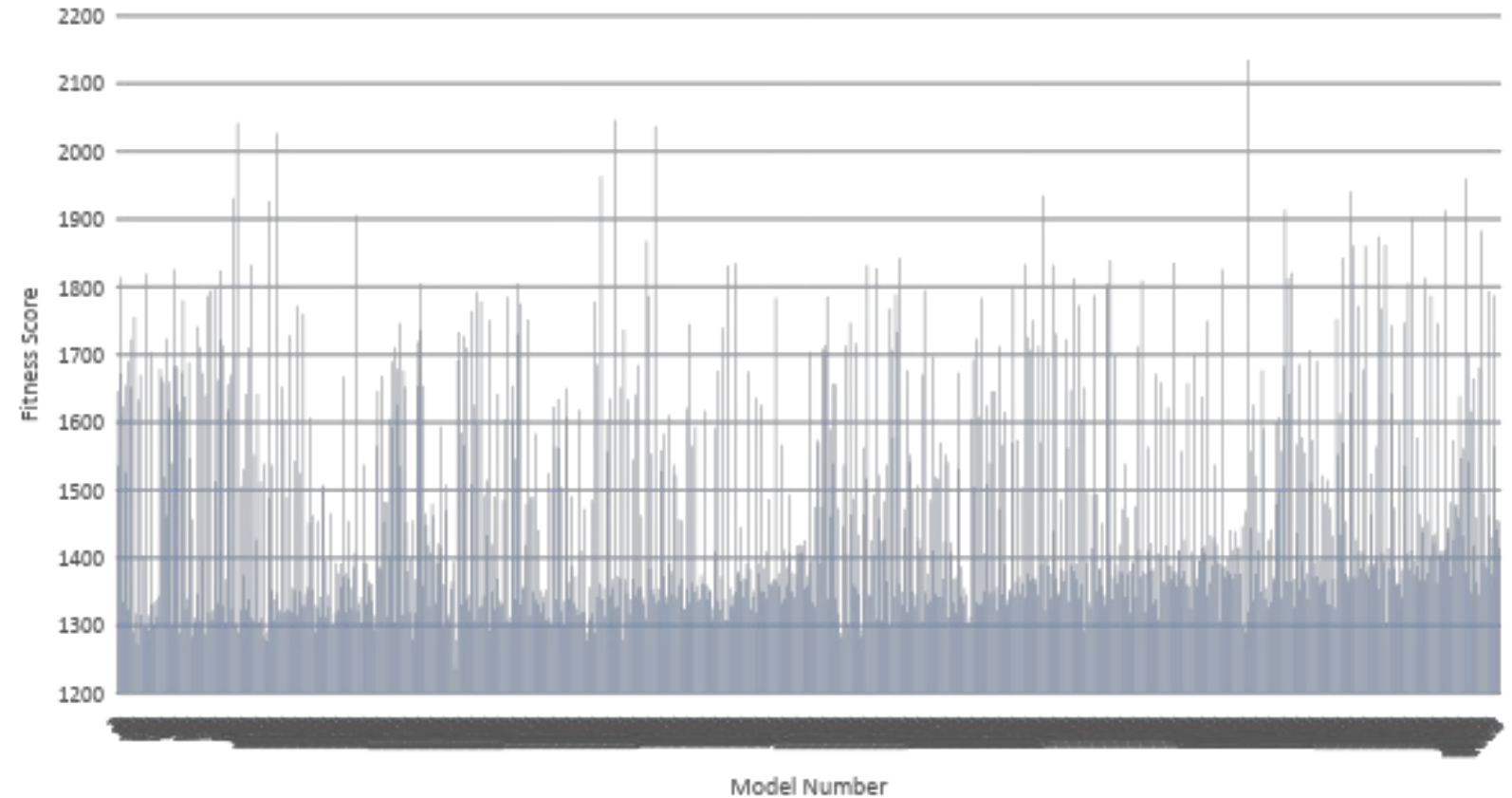
# Ensemble modeling

- True model is always unknown
- Ensemble of plausible models
- Approach
  - Generate variants (1024)
  - Collect best scoring models

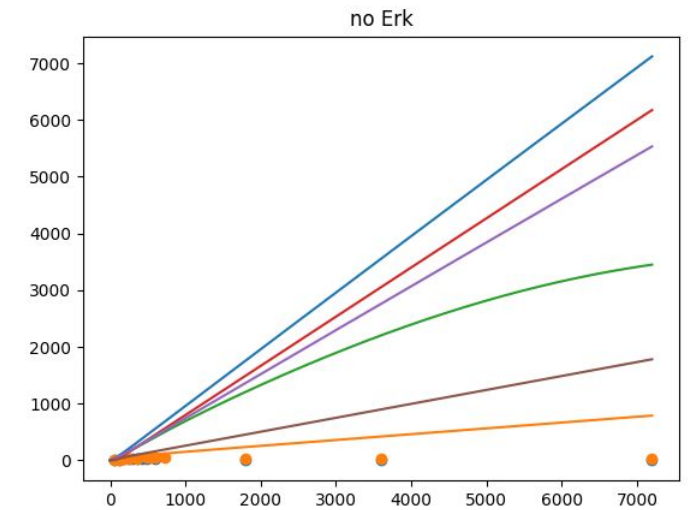
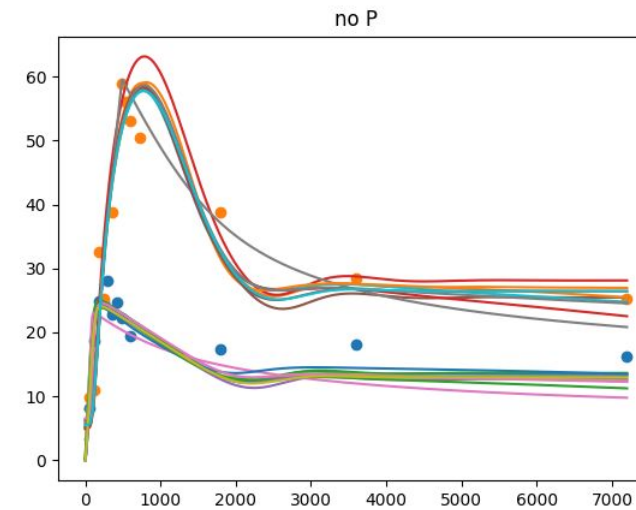
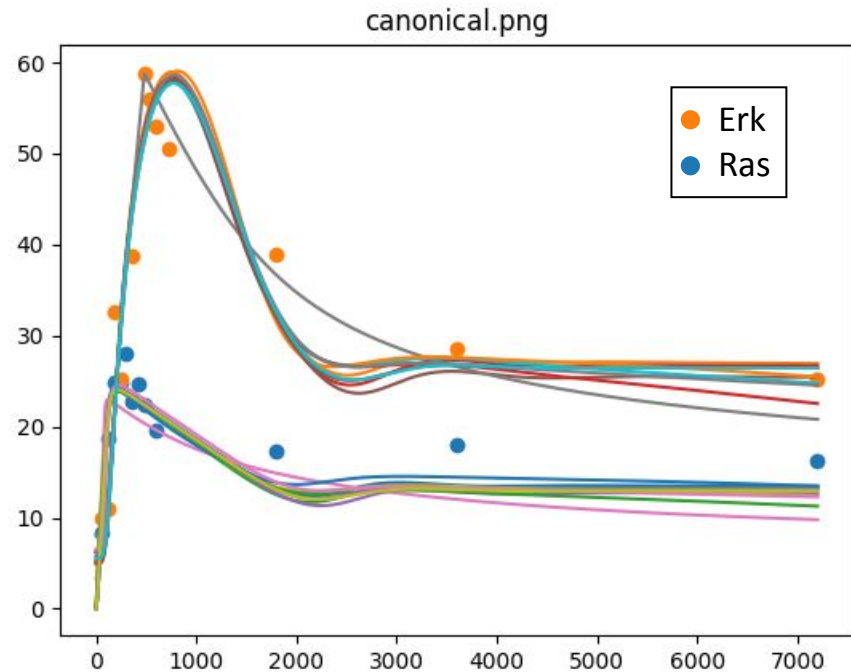
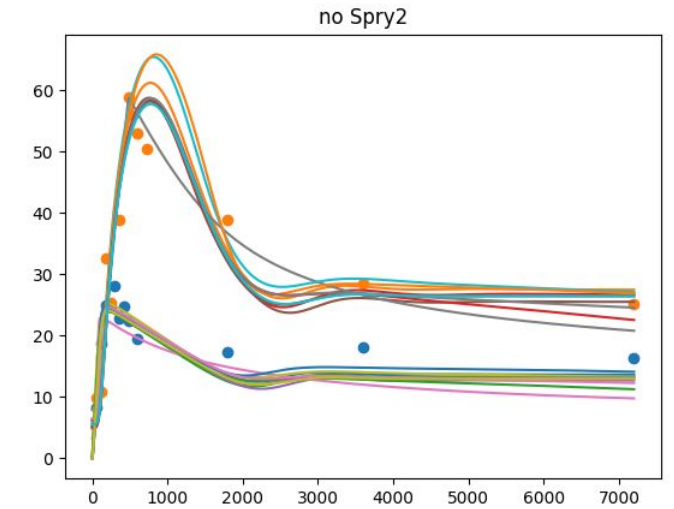
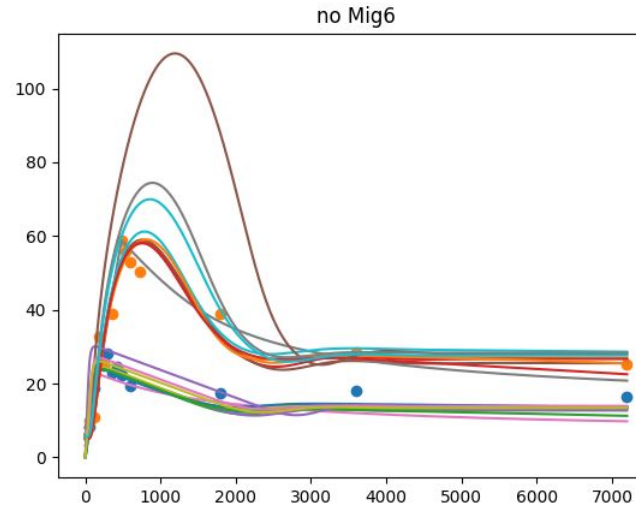
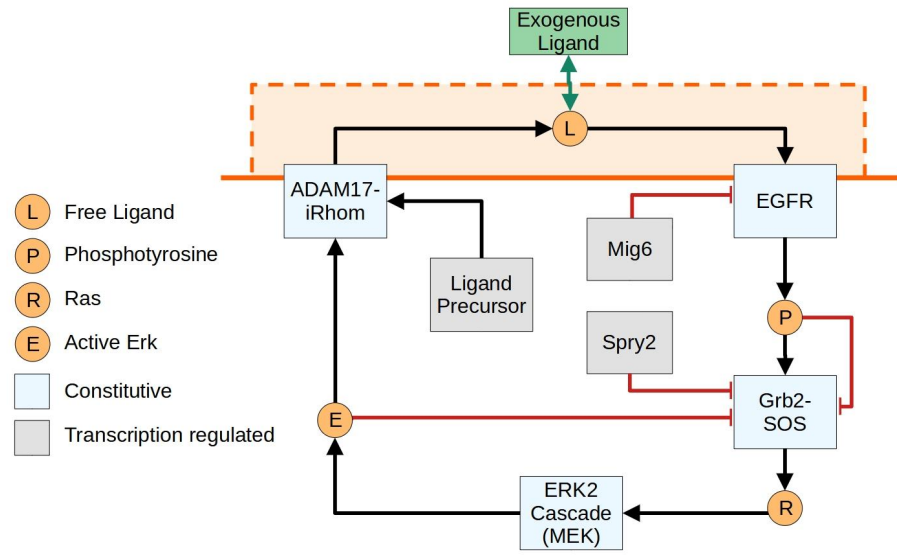


# Fitness scores: Best of 10 for each model

- 1024 models
- 10 runs per model
- ~ 7h per model
- BlackBoxOptim + Roadrunner
- EMSL (PNNL) Tahoma cluster

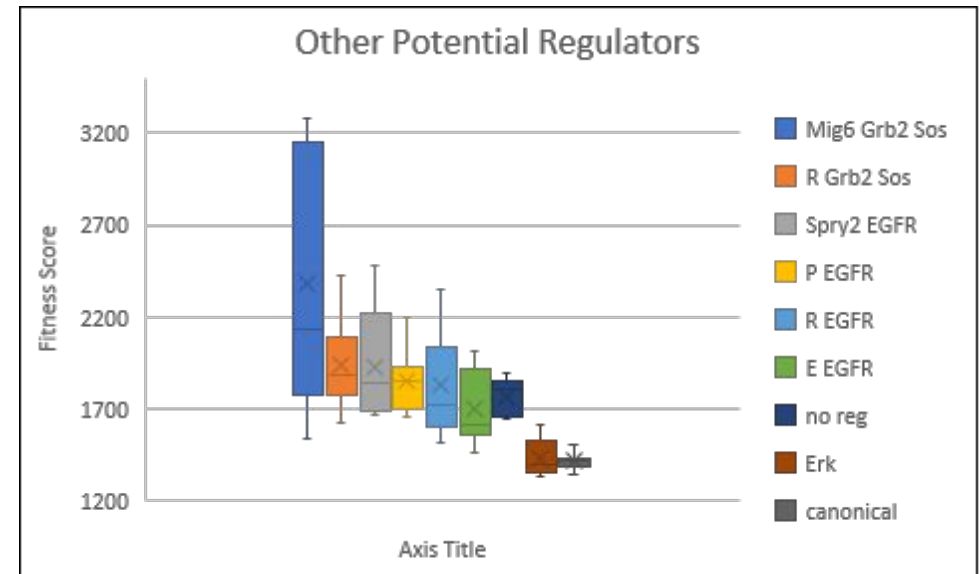
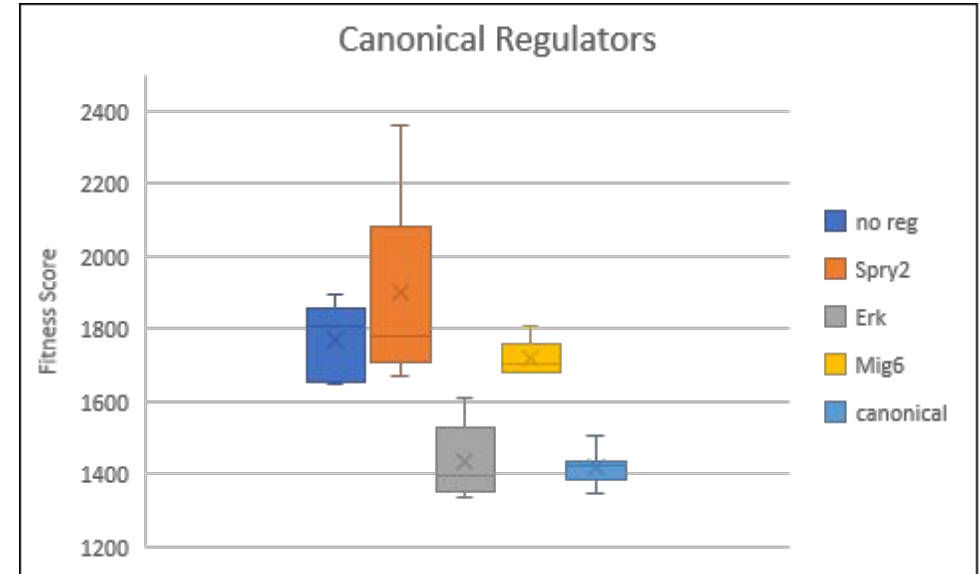
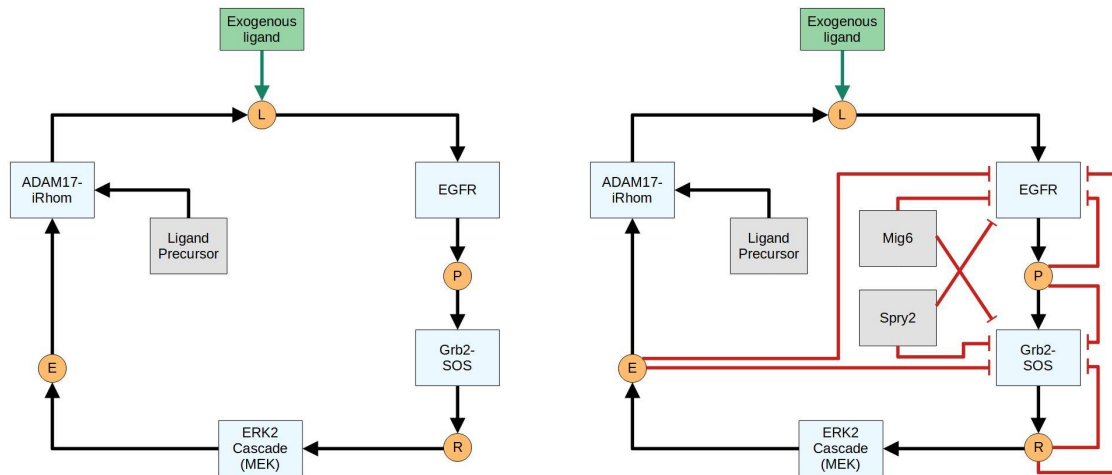


# Effects from removing regulatory edges



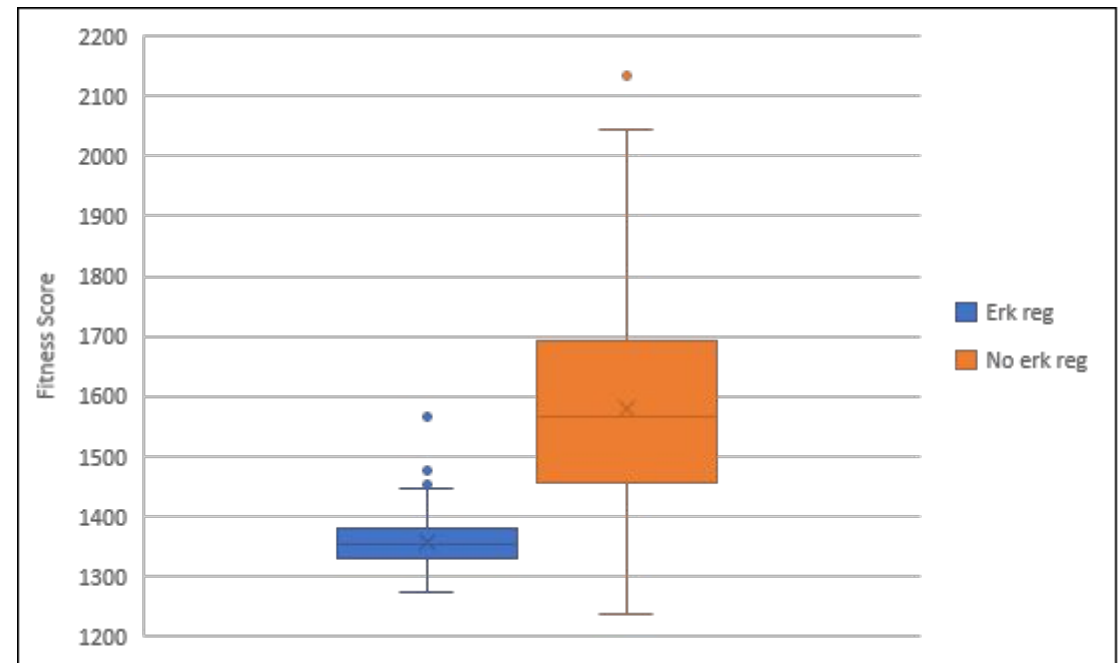
# Single regulator models

Does any single regulatory interaction make a difference?



# Erk regulation of Grb2-Sos: present vs absent

- 1024 models: best of 10
- Two groups
- Erk models have the clear advantage
- Still possible to design a model without Erk regulation of Grb2-Sos that outperforms



# Conclusions and Implications

- Erk inhibition on Grb2-Sos is clearly important.
- Little effect from other rewirings (with a few exceptions).
- Implications for more complicated models.
  - Lower data to complexity ratio.
  - Potentially many model variants may fit.
  - Many implausible models may also fit.
- How much confidence should you really have in your model?
- What kind of data and how much does it take to correctly constrain model topology?



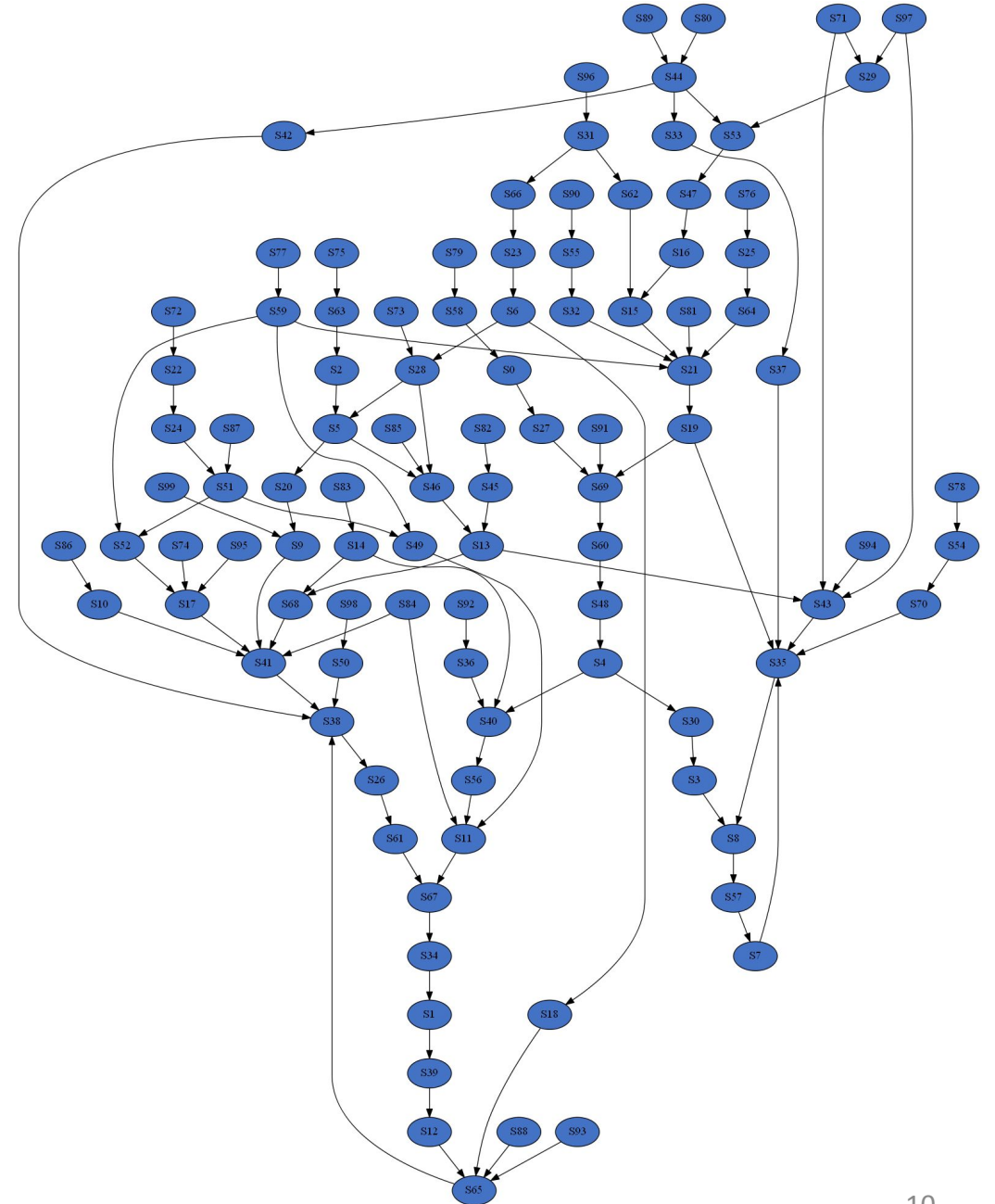
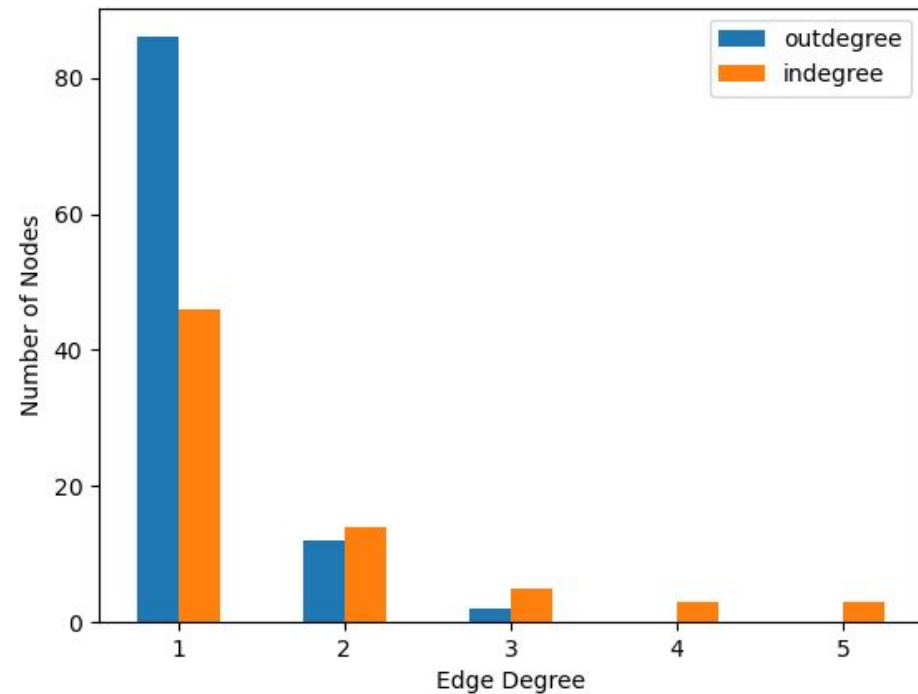
# SBbadger

- Python package for the generation of synthetic reaction networks
  - `pip install sbbadger`
  - <https://github.com/sys-bio/SBbadger>
  - <https://arxiv.org/abs/2202.13004>
- Definable Properties
  - Degree distributions
  - Reaction types
  - Size
  - Reaction kinetics
    - Mass action
    - Lin-log
    - Generalized Michaelis-Menten
    - Modular (Liebermeister)

# in/out degree distributions

Power law:  $\frac{k^{-\alpha}}{\text{zetaeta}(a)}$

- Outdegree:  $\alpha = 3$
- Indegree:  $\alpha = 2$



# Going forward

- Generate mechanistic benchmark models with properties typically seen in the literature and use them to evaluate and develop model inference algorithms.
- Example: various implementations of Sparse Identification of Nonlinear Dynamics (**SINDy**).
- Gauge the required amount, type, and placement of data for accurate model reconstruction.

# Acknowledgements

- Herbert Sauro
- Steve Wiley
- Song Feng
- NIH - National Cancer Institute grant U01CA242992
- Environmental Molecular Sciences Laboratory (EMSL) at PNNL