Modeling, calibration, and power analysis of single-cell massively parallel reporter assays

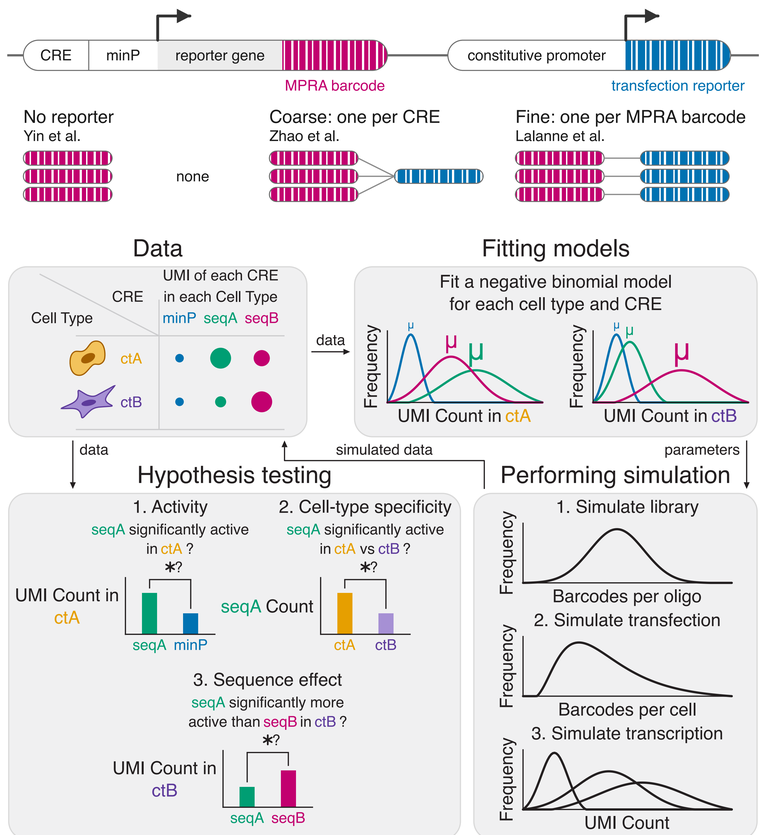
Mackenzie C. Noon[1], Erin N. Gilbertson[1], Grace L. Oualline[1], and Steven K. Reilly[1,*]

[1]Department of Genetics, Yale School of Medicine, New Haven, CT, USA
{mackenzie.noon, erin.gilbertson, grace.oualline, steven.k.reilly}@yale.edu
*Correspondence: steven.k.reilly@yale.edu

Yale SCHOOL OF MEDICINE
Genetics

pypi.org/project/scMPRAforge

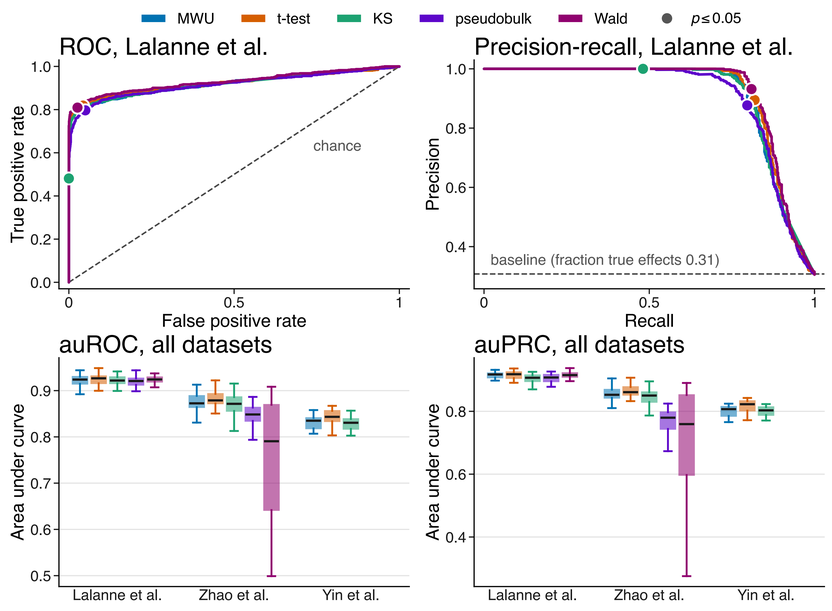
Graphical abstract



Take-aways

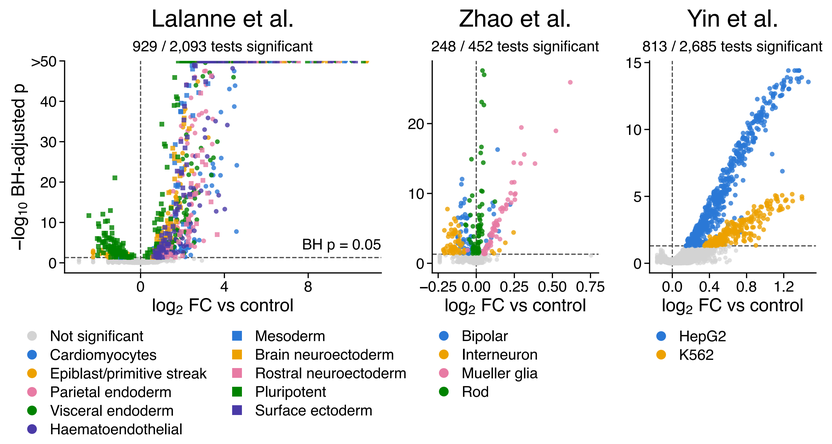
- Massively Parallel Reporter Assays (MPRAs) do not capture cell-resolved information.
- Single-cell MPRA (scMPRA) does, but this new technique lacks standard file formats, analysis pipelines, and principled design guidelines.
- scMPRAforge standardizes file formats and nomenclature, and facilitates hypothesis testing, simulation, and power analysis.
- scMPRAforge recovers additional signal from published data.
- Simulations explain differences in statistical power across published designs.
- Cell number, multiplicity of infection, and the transfection reporter have the largest impact on statistical power.

Standard statistical tests are well powered to detect changes in scMPRA.



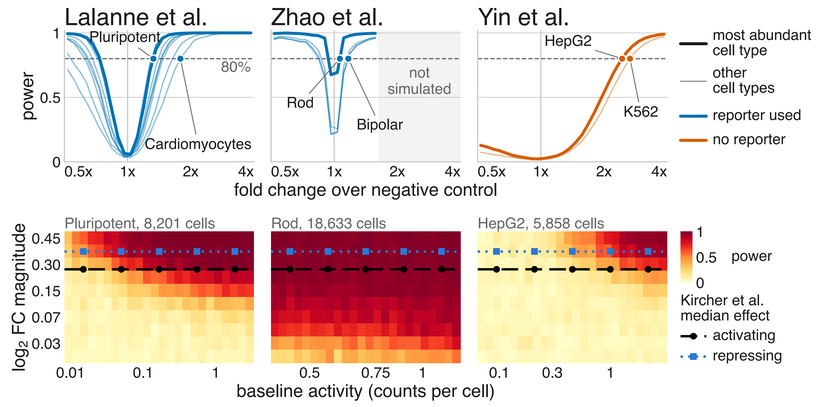
- MWU and the non-pseudobulk t-test perform best.
- Non-pseudobulk t-test is mis-calibrated for some experimental designs, so only MWU is carried forwards.
- Yin et al. has only one replicate and so was not evaluated on Wald or pseudobulk.

scMPRAforge can perform hypothesis testing.



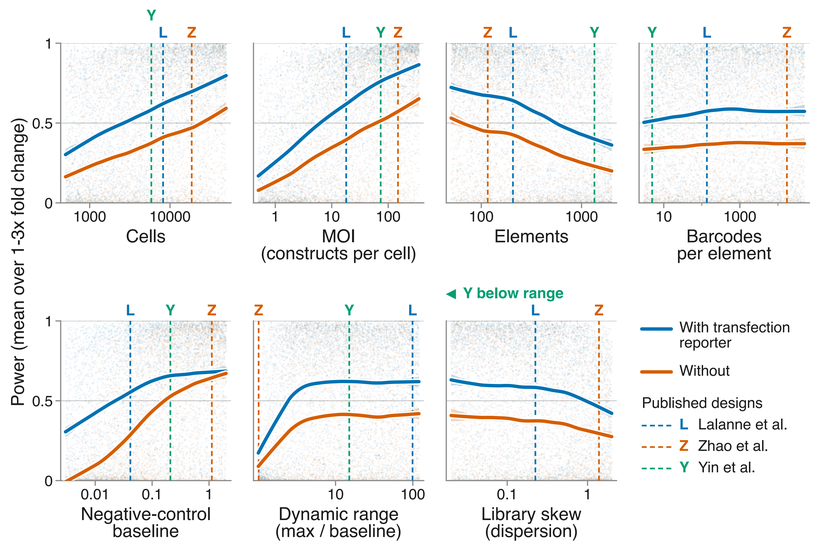
References. Kircher et al. (2019) *Nature Communications* 10:3583. Lalanne et al. (2024) *Nature Methods* 21:983–993. Yin et al. (2025) *Cell Systems* 16:101302. Zhao et al. (2023) *Nature Genetics* 55:346–354.

scMPRAforge facilitates principled power estimation.



- Some existing designs are capable of detecting the median variant effect with 80% power.
- Power increases with baseline activity to a plateau, after which it depends on fold change.
- Simulation not carried beyond power saturation for Zhao et al.

scMPRAforge identifies key axes of variation for the improvement of scMPRA.



- Increasing MOI and cell number uniformly improves power. Increasing number of elements decreases power.
- Increasing dynamic range and negative-control baseline activity also improve power, but only up to an elbow point.