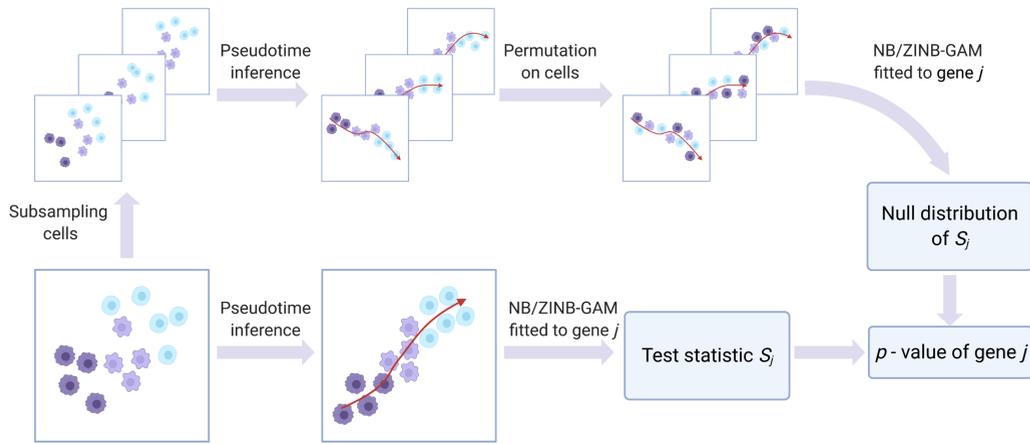




## Abstract

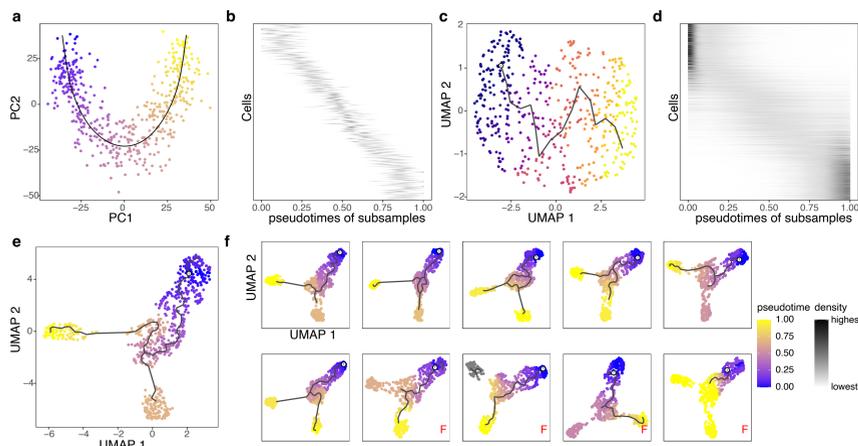
To investigate molecular mechanisms underlying cell state changes, a crucial analysis is to identify differentially expressed (DE) genes along the pseudotime inferred from single-cell RNA-sequencing data. Therefore, we propose PseudotimeDE, a DE method that adapts to various pseudotime inference methods, accounts for pseudotime inference uncertainty, and outputs well-calibrated  $p$ -values.

## Methods



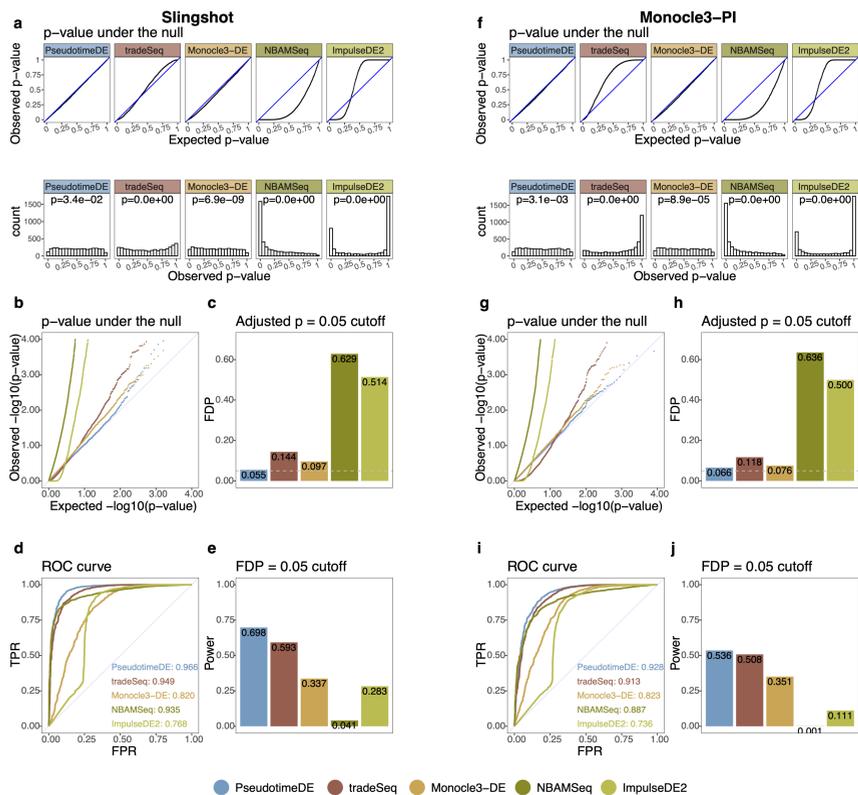
## Results: Uncertainty of Pseudotime

Using subsampling, PseudotimeDE captures the "linear" uncertainty (b & d) and "topology" uncertainty (f) of inferred pseudotime.



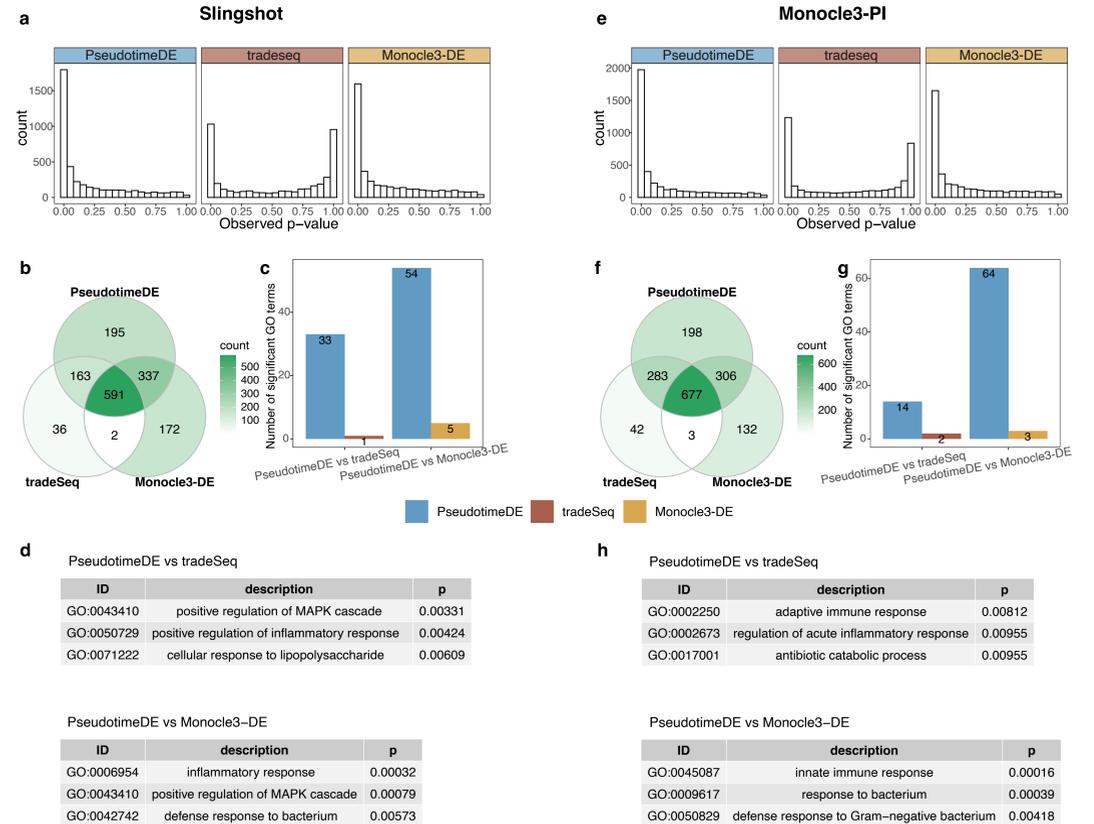
## Results: Comparison to Existing DE Methods by Simulation

We compare PseudotimeDE to existing DE methods on simulated datasets. PseudotimeDE generates well-calibrated  $p$ -values (a,b,f,g), better FDR control (c & h) and higher power (e & j).



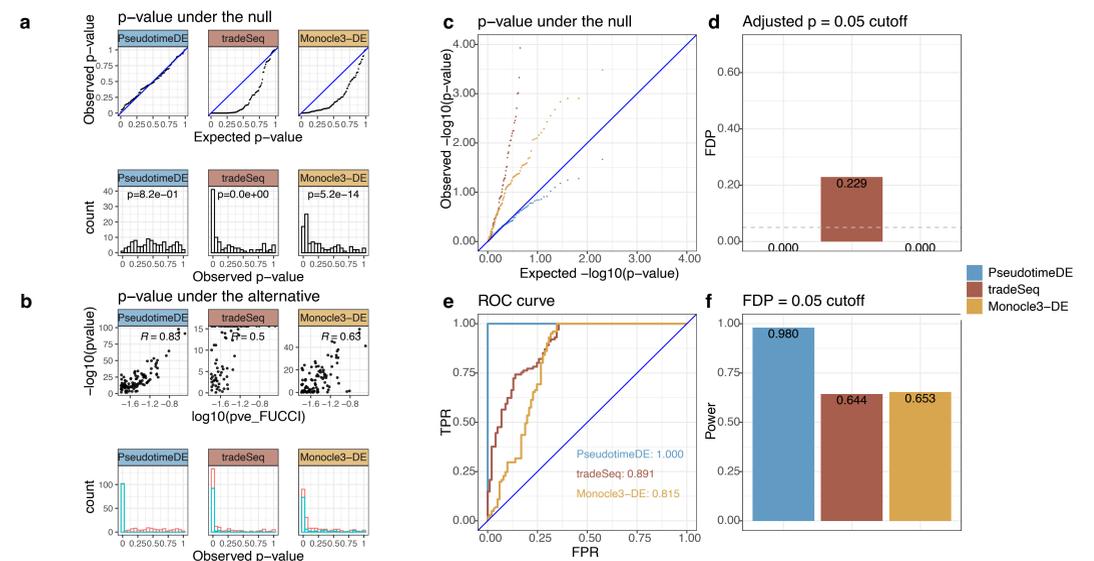
## Results: Comparison to Existing DE Methods on LPS-Dendritic Cell Dataset

Histograms show all genes'  $p$ -values by the three DE methods (a & e). Compared to other DE methods, many more GO terms are enriched in the PseudotimeDE-specific DE genes (c & g). Many of these GO terms are related to LPS, immune process, and defense to bacterium (d & h).



## Results: Comparison to Existing DE Methods on Cell Cycle Phase Dataset

The cell cycle phase dataset contains 50% cell cycle-related genes (DE genes) and their random permutations (non-DE genes).  $p$ -values of PseudotimeDE are both well-calibrated (a) and strongly correlated with true cell-cycle signals (b). PseudotimeDE also yields highest power (f).



## Discussion

We propose a statistical method PseudotimeDE to identify DE genes along inferred cell pseudotime. PseudotimeDE focuses on generating well-calibrated  $p$ -values while taking into account the randomness of inferred pseudotime. PseudotimeDE is flexible to accommodate cell pseudotime inferred in a standard format by any method. Its use of NB-GAM and ZINB-GAM allows it to capture diverse gene expression dynamics and to accommodate undesirable zero inflation in data.

## Software

The R package PseudotimeDE is available at <https://github.com/SONGDONGYUAN1994/PseudotimeDE>.

## References

[1] Dongyuan Song and Jingyi Jessica Li. PseudotimeDE: inference of differential gene expression along cell pseudotime with well-calibrated  $p$ -values from single-cell rna sequencing data. *Genome biology*, 22(1):1–25, 2021.