Learning cell identities and (post-)transcriptional regulation using single-cell data

Michielsen, L.C.M.

Stellingen behorend bij het proefschrift getiteld

Learning cell identities and (post-) transcriptional regulation using single-cell data

1. Rare or diseased cell types in single-cell RNA sequencing data cannot be detected using the posterior probability as a rejection option (this thesis).
2. Researchers' trust in manual annotations obtained using clustering instead of classification is unjustified since the former has more mistakes and lacks consistency (this thesis).
3. We need sequence-based models trained on single-cell RNA sequencing data to understand transcriptional regulation in heterogeneous tissues (this thesis).
4. Sequence-based models must be trained on individual genomes and perturbation data to increase the variation in the training data (this thesis).
5. Using popular, trending terms, such as foundation models, distracts users and reviewers from a model's actual functionalities and performance and hinders scientific progress.
6. Cells form a continuum, but discrete cell types are needed to facilitate downstream analysis.
7. The human cell atlas will always be incomplete so a correctly working rejection option will remain necessary for classification methods.
8. Single-cell reference atlases are better built per tissue than organism-wide.
9. Men must take the same time of paternity leave as women to enforce equal career opportunities.
10. Generative models such as ChatGPT and Dall-E must show users their carbon footprint.