Supplementary Resources

Resources to accompany the Crick single-cell workshop at the Festival of Genomics in London, 28th January 2026

Python

- Download Anaconda
- Download VSCODE
  - Python Extension Pack
- Install uv
- Configure Ruff

Crick Training

- Crick Technical Training)

Community Resources

- scverse Discourse
- scverse events
- Single Cell Best Practices Book
- scverse packages list

Tools

Pre Processing & Pipelines

- Download Cell Ranger
- nf-core/scrna-seq
- nf-core/scdownstream

QC

- Scanpy
- DoubletDetetection

Single Cell Atlases/Databases

- Human Cell Atlas Data Portal
- CZI CellxGene
- HuBMAP Data Portal

Converting between R, Python, and Loupe Objects

- sceasy
- loupe browser
- loupeR

Dimension reduction

Summarising complex data with PCA

- PCA summarises data by finding lines of best fit through high-dimensional space
- The first line of best fit (or principal component) explains as much variance as possible
- Subsequent principal components are all orthogonal and explain remaining available variance
- We can use a scree plot to visualise the proportions of variance explained by each component

Other alogorithms for variable gene identification

- If you wanted to apply VST (`flavor='seurat_v3_paper'`, requires `pip install scikit-misc`), you would need to apply this to the raw counts
- `sc.pp.highly_variable_genes(adata, batch_key="sample", flavor='seurat_v3_paper', n_top_genes=2000, layer='counts')`

Considerations for atlas-based annotation

Benefits	Risks & Limitations
Fast and scalable annotation	Misclassification if reference is incomplete
No need for manual marker selection	Sensitive to batch effects
Reproducible across datasets	May miss novel or rare cell types
Integrates biological knowledge	Depends heavily on reference quality

Download pre-built CellTypist models

- Cell Typist has many models that you can download. Here we pick the Immune Mouse Gut which is relevant to our dataset.
- See the Cell Typist documentation for more details about building your own references

```
celltypist.models.models_description()[0:5]
```

```
Detailed model information can be found at `https://www.celltypist.org/models`
```

	model	description
0	Immune_All_Low.pkl	immune sub-populations combined from 20 tissue...
1	Immune_All_High.pkl	immune populations combined from 20 tissues of...
2	Adult_COVID19_PBMC.pkl	peripheral blood mononuclear cell types from C...
3	Adult_CynomolgusMacaque_Hippocampus.pkl	cell types from the hippocampus of adult cynom...
4	Adult_Human_MTG.pkl	cell types and subtypes (10x-based) from the a...

```
# Remove the `model=` argument to download all models
celltypist.models.download_models(model='Adult_Mouse_Gut.pkl')
```

```
Storing models in /Users/campbej/.celltypist/data/models
Total models to download: 1
Skipping [1/1]: Adult_Mouse_Gut.pkl (file exists)
```

Downstream analysis tools

Integration

- BBKNN and Ingest
- Scanorama
- harmony
- scVI-tools

Downstream Analysis

- pseudobulk via decoupler
- DGE via pyDESeq2
- DGE via glmGamPoi
- trajectory analysis
- velocity analysis