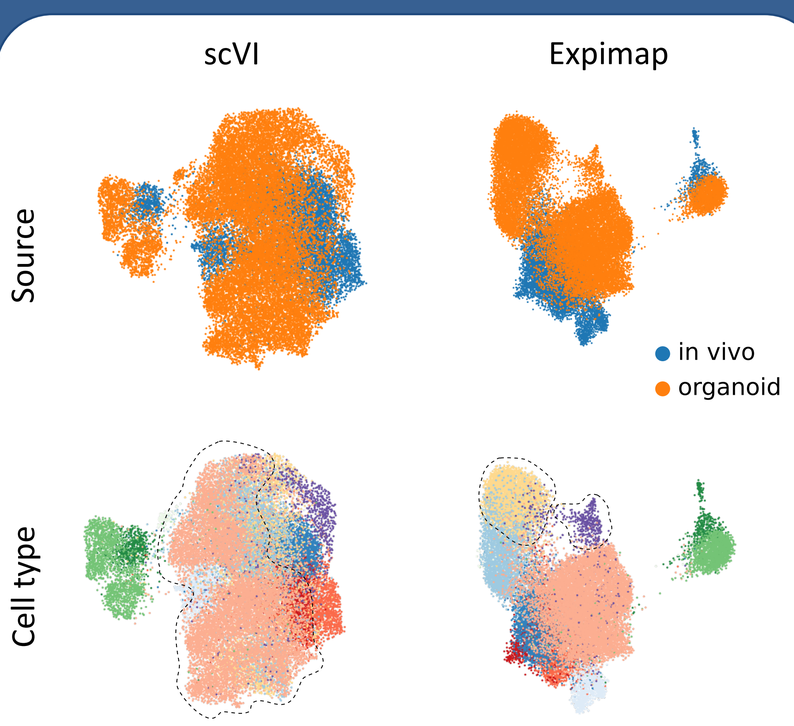
Learning interpretable embeddings for *in vivo/in vitro* comparison of endometrial epithelium

Marie Moullet, Luz Garcia Alonso, Mo Lotfollahi, Roser Vento Tormo
Wellcome Sanger Institute

Background

- The endometrium is the mucosal lining of the uterus. It plays a key role in female reproductive health and disease.
- Endometrial organoids derived from glandular tissue reproduce key features of endometrial biology *in vitro.*[1,2]
- Deep learning models can be used to learn low dimensional representations of single cell omic data. Expimap[3] learns interpretable embeddings by mapping the gene expression matrix to known gene programs (GP).

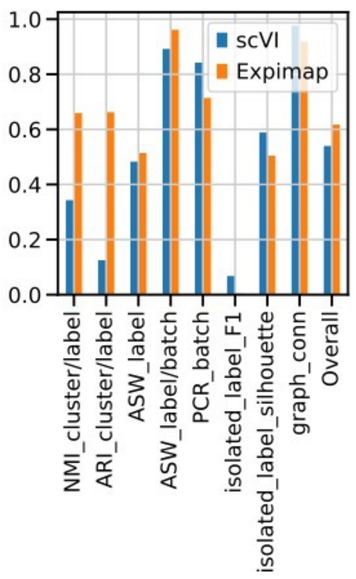
Expimap outperforms unconstrained models



Over-correction loses separation of organoid secretory cells

Successful separation of organoid populations which do not have *in vivo* counterparts

scIB[7] benchmarking



- in vivo_SOX9
- in vivo_SOX9_LGR5
- organoid_Estrogen Induced
- in vivo_Pre-ciliated
- in vivo_Ciliated
- organoid_Pre-ciliated
- organoid_Ciliated
- in vivo_Glandular_secretory
- in vivo_Glandular
- organoid_Secretory
- organoid_Inflamatory
- organoid_NH_Day0

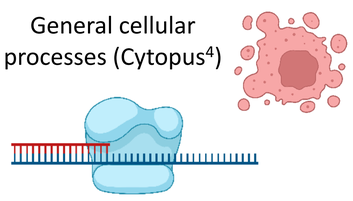
Future directions

- Extension to scRNA-seq datasets from organoids cultured according to different protocols
- Extension to multiple modalities (eg chromatin accessibility)
- Comparison to attention-based methods
- Perturbing differentially active pathways *in silico* and experimentally

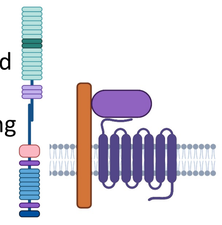
Building the Expimap model

Curated gene programs in the bottleneck layer:

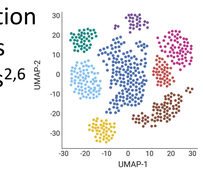
General cellular processes (Cytopus[4])

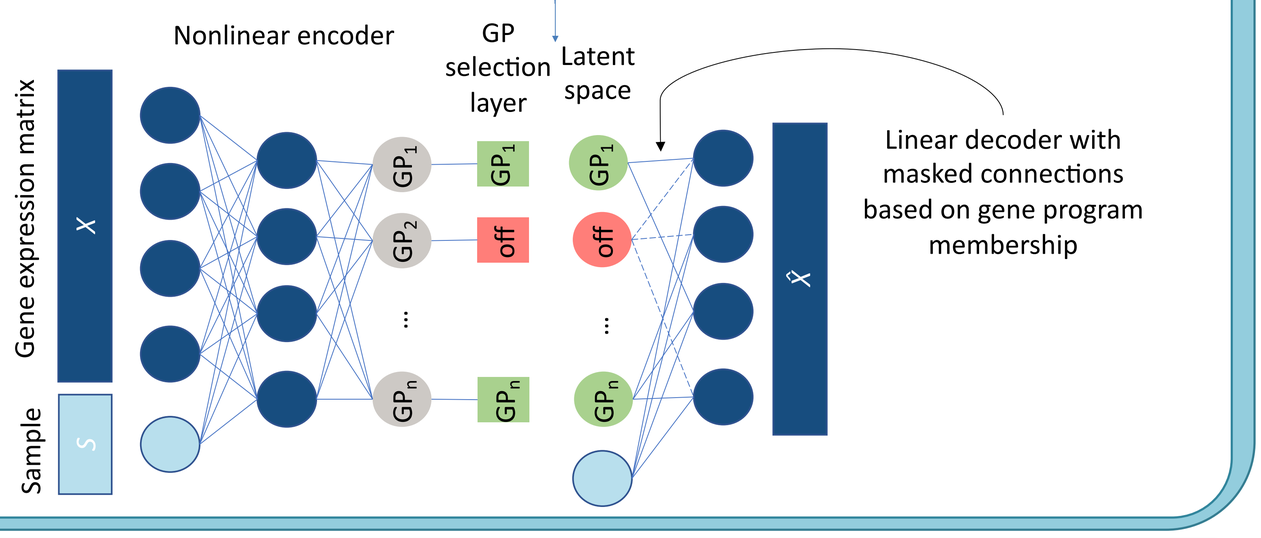


Pathways related to organoid culture conditions, Wnt, Notch and hormonal signaling (from Reactome[5])

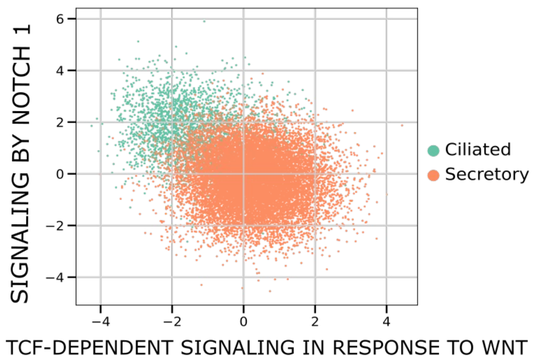


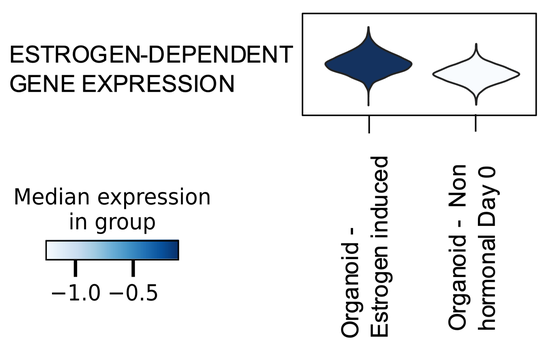
Cell-type specific markers and transcription factors regulons[2,6]





Expimap recovers expected biological patterns



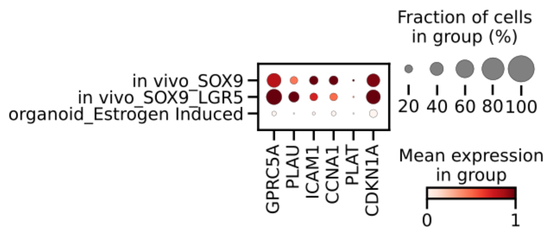


Expimap identifies *in vivo/in vitro* differences

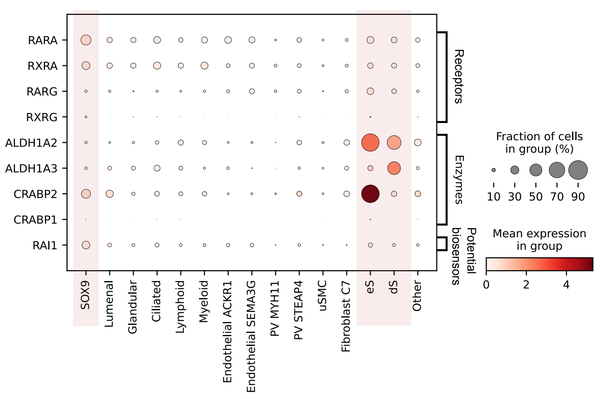
Differentially active GP are linked to cell division and proliferation, as well as organoid culture conditions and absence of stromal cell interactions

Retinoic acid

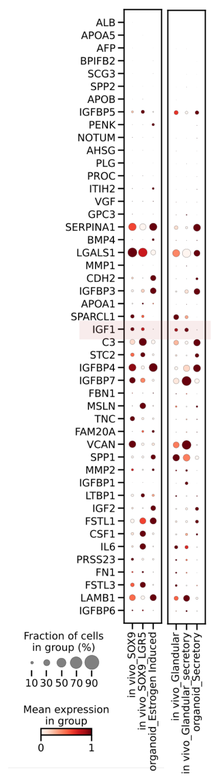
- Increased expression of retinoic acid receptor α target genes *in vivo*



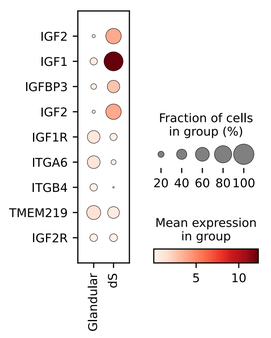
- *In vivo*, stromal cells signal to epithelial cells via RA



Insulin growth factor



- Differential activity of genes involved in regulation of IGF transport and uptake by IGFBP5
- Stromal cells signal to epithelial cells via IGF *in vivo :*



➢ Do these differences remain in organoids cultured with media containing vitamin A and insulin?

References:

1. Turco MY, Gardner L, Hughes J, et al. Long-term, hormone-responsive organoid cultures of human endometrium in a chemically defined medium. Nature Cell Biology. 2017;19(5):568-77.
2. Garcia-Alonso L, Handfield L-F, Roberts K, et al. Mapping the temporal and spatial dynamics of the human endometrium in vivo and in vitro. Nature Genetics. 2021;53(12):1698-711.
3. Lotfollahi M, Rybakov S, Hrovatin K, et al. Biologically informed deep learning to query gene programs in single-cell atlases. Nature Cell Biology. 2023;25(2):337-50.
4. Walle T. wallet-maker/cytopus: Cytopus v1.21. Zenodo; 2022. DOI:10.5281/zenodo.7306238
5. Gillespie M, Jassal B, Stephan R, et al. The reactome pathway knowledgebase 2022. Nucleic Acids Research. 2021;50(D1):D687-D92.
6. Garcia-Alonso L, Holland CH, Ibrahim MM, Turei D, Saez-Rodriguez J. Benchmark and integration of resources for the estimation of human transcription factor activities. Genome Res. 2019;29(8):1363-75.
7. Luecken MD, Büttner M, Chaichoompu Ket al. Benchmarking atlas-level data integration in single-cell genomics. Nature Methods. 2022;19(1):41-50.
8. Boretto M, Cox B, Noben M, Hendriks N, et al. Development of organoids from mouse and human endometrium showing endometrial epithelium physiology and long-term expandability. Development. 2017;144(10):1775-86.