Inference of cell cycle regulation between glioblastoma subpopulations in vivo to drive mathematical models

Nicholas Harbour[1], Lee Curtin[3], Sebastian Velez[3], Michael Chappell[2], Matthew Hubbard[1], Osama Al-Dalahmah[4], Peter Canoll[4], Kristin Swanson[3], Markus Owen[1]
[1]Centre for Mathematical Medicine and Biology, University Nottingham, UK, [2]Sir Peter Mansfield Imaging Centre, University Nottingham, UK,
[3]Mathematical NeuroOncology Lab, Mayo Clinic, AZ, USA, [4]Department of Pathology and Cell Biology, Columbia University Vagelos College of Physicians and Surgeons, NY, USA

INTRODUCTION

Glioblastoma (GBM) is the most aggressive and most common primary malignant brain tumor in adults, with a poor average survival time of 15 months. One of the key challenges in successfully treating GBM is its heterogeneity, with multiple distinct cellular subtypes that have been shown to occur on both inter- and intra-patient levels. We analyses a recently published single nucleus RNAseq data set that identified six subpopulations of GBM cells which are: gl_pro1, gl_pro2, gl_mes1, gl_mes2, gl_PN1, gl_PN2..

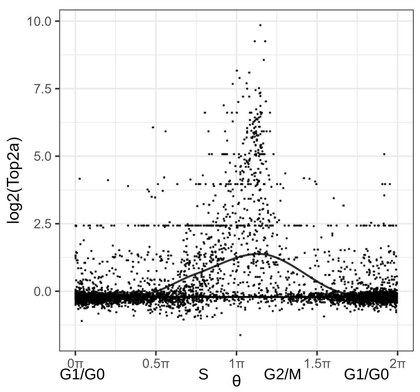
The cell cycle is a fundamental and highly conserved process that controls faithful division of cells; dysregulation of the cell cycle is known to be a key driver in many cancers. However, how the cell cycle is differently regulated between malignant GBM subtypes has not been well classified *in vivo.* We compare cell cycle regulation/dysregulation among these six subtypes, using Tricycle[2], an R/Bioconductor package that utilises principal component analysis and transfer learning to infer cell cycle stage (CCS) from any snRNAseq data set.

We find that as expected both proliferative subtypes have the highest proportion of actively dividing cells (cells in S/G2/M phases), while the other glioma subtypes have relatively low proportion of actively dividing cells. We compare the cell cycle phases between primary and recurrent samples and find that the mesenchymal subtype has a significant decrease in the number of actively proliferating cells in recurrent samples.

TOP2A EXPRESSION

We assign cell cycle position using Tricycle, this produces a cell cycle pseudotime upon which to order cells. TOP2A is a known cell cycle gene that can be used to verify the accuracy of the cell cycle assignment.

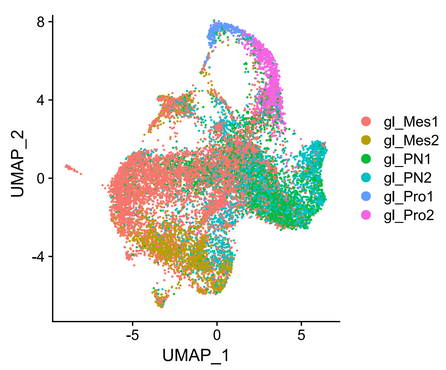
TOP2A expression along cell cycle pseudotime (θ)



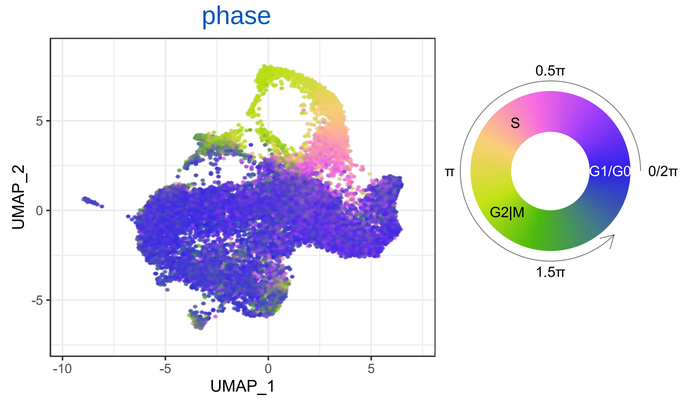
CELL CYCLE RESULTS

We visualise the snRNAseq data using UMAP and labelling it by both cell type and cell cycle position, this clearly shows that the gl_pro subtypes are highly proliferative (in S/G2/M phases) compared to the other subtypes.

UMAP of CNVpos cells

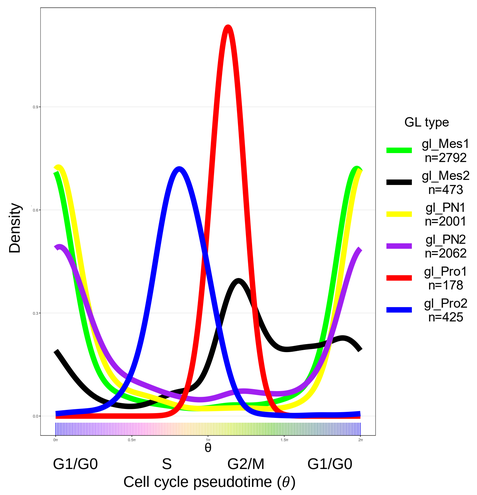


UMAP of CNVpos cells: colored by cell cycle phase

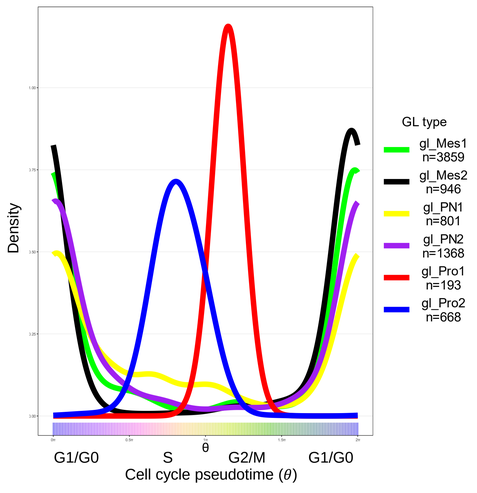


We separate the cells into primary and recurrent samples and plot a histogram of the cell cycle pseudotime. Interestingly the gl_Pro2 are predominantly in S phase while gl_Pro1 are mainly in G2/M phase. We also observe that gl_Mes2 are mainly in G2/M phase in primary tumors but are not cycling (G1/G0) in recurrent samples.

Histogram of cell cycle phase for primary subtypes

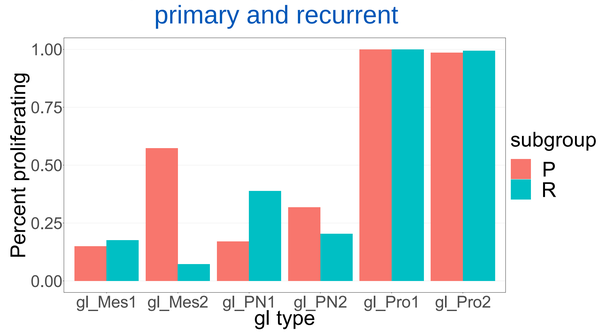


Histogram of cell cycle phase for primary subtypes Recurrent

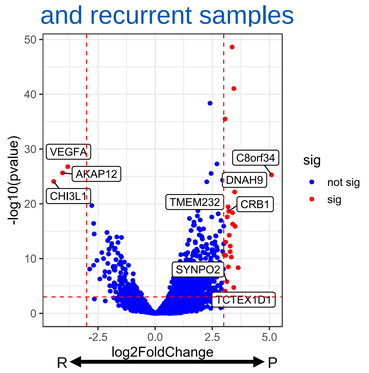


We bin the cells into either 'actively proliferating' (S,G2,M phases) and 'non-proliferative' (G1,G0 phases) and compare the GBM cell types between primary and recurrent samples. The largest difference between primary and recurrent samples is in the gl_Mes2 populations, to investigate this further we carry out differential expression analysis.

Comparison of actively proliferating cells between primary and recurrent



Differential expression analysis of gl_Mes2 between primary and recurrent samples



MATHEMATICAL MODEL

To further investigate cell cycle regulation we plan to develop an age structured PDE mathematical model. In general, our model for the cell cycle takes the form

$$\frac{\partial g}{\partial t} + c(t,\theta,g)\frac{\partial g}{\partial \theta} + d(t,\theta)g(t,\theta) = 0$$

$$g(t,\theta=0) = 2\int_0^{2\pi} K(t,\hat{\theta})g(t,\hat{\theta})\,d\,\hat{\theta}$$

where θ represents cell cycle phase, $c(t,\theta,g)$ is the speed of progression through the cell cycle, $d(t,\theta)$ is the death rate, and $K(t,\theta)$ the mitosis rate. We plan to parameterise such a model using the snRNAseq data. This will allow us to predict how changes at the cell cycle level can affect changes at the whole tissue level.

DISCUSSION

- The gl_Pro GBM subtype have by far the greatest number of actively dividing cells, while gl_PN and gl_Mes have relatively few cells actively cycling. This is shown by the high density of gl_Pro cells found between $0.5\pi - 1.5\pi$.
- The number of actively proliferating gl_Mes2 cells is significantly decreased in recurrent samples compared to primary.

FUTURE WORK

- Monocle Trajectory inference.
- CellChat to infer cell-cell interactions.
- Calibrate and build mathematical models to integrate bioinformatics results with mechanistic modelling.

REFERENCES

1. Al-Dalahmah et al., Nature Communications, 2023
2. Zheng et al., Genome Biology, 2022
3. Schwabe et al., Molecular systems Biology, 2020

nicholas.harbour@nottingham.ac.uk

University of Nottingham
UK | CHINA | MALAYSIA