Vylyny Chat
Applied Sequencing Informatics
May 10, 2020

Impacts of the PI3K/AKT and MRN inhibitors on superior cervical ganglia's cell population: A novel scRNA-seq approach

Abstract:

The PI3K/AKT pathway and MRN complex are involved in various cellular functions ranging from regulating cell proliferation to maintaining genomic integrity. Previous research reported strong associations of this crucial pathway and protein complex with major diseases such as diabetes, cancer or neuronal disorders. However, little is known about their biological relevance at single cell resolution. In this study, with the power of novel scRNA-seq approach, we aim to characterize cell subtypes in a superior cervical ganglia (SCG) of embryonic rats, and investigate how LY294002 and Mirin (inhibitors of the PI3K/AKT and MRN, respectively) affect different cell populations. We observed five major cell clusters in the SCG: satellite glial, endothelial, macrophages, mature neurons and Schwann cells. Comparing against mocked-control tissue isolates, we reported that Mirin modulated a large number of different genes across the five cell subtypes and reduced endothelial cell abundance, while LY294002 preferentially perturbed the transcriptional signatures of the mature neurons and depleted macrophage population. Taken together, our study did not only shed light on the presence of different cell subtypes in the SCG, but also identified relevant cell populations and biological pathways that were disproportionately modulated by LY294002 and Mirin inhibitors. This knowledge holds clinical significance for the PI3K/AKT and MRN associated disorders, as it may point to novel potential treatments that selectively target only the relevant cell types/pathways while leaving others unscathed.

Keywords: PI3K/AKT; MRN complex; LY294002; Mirin; single cell RNA-seq; Superior cervical ganglia

Introduction:

The PI3K/AKT signaling pathway (Phosphoinositide 3-kinase) is essential in various cellular functions including cell growth, proliferation and survival (1, 2). The MRN complex comprising of Mre11, Rad50 and Nbrs1 plays a crucial role in sensing and signaling DNA-double strand breaks (3). This complex recruits ATM and the recruited ATM, then, phosphorylates various downstream tumor suppressor genes such as p53 and BRCA1, together maintaining genome integrity and stability during DNA replication and cell cycle (3, 4). Prior research reported associations between the PI3K/AKT pathway and MRN-ATM complex with major diseases such as cancers, diabetes and neurological disorders (5-7). While there has been an in-depth understanding about this important pathway and complex, little research has explored their biological relevance at a granular single cell resolution. Being able to disentangle certain cell types these biological pathways perturb in various diseases may point to specific treatment targets, resulting in reduced off-target side effects towards other unrelated cell populations.

LY294002 and Mirin are commonly used biological inhibitors to study the role of PI3K/AKT and MRN, respectively (1, 4). LY294002 has been shown to affect many downstream biological processes such as metabolism, transcription or protein trafficking (8). This drug was reported to target specifically class I&III PI3K isoforms and potentially other PI3K-like kinases (1). Mirin was discovered as an inhibitor of the MRN complex through a forward chemical genetic screen (4). Mirin specifically targets the exonuclease activity of the Mre11, while leaving the ATM kinase activity and other MRN functions intact. Due to its specificity and decreased impacts on cell viability, this MRN complex inhibitor emerges as an essential tool to study biological underpinnings of MRN-related disorders (4).

In this study, we harvested the sympathetic neurons of the prenatal rat superior cervical ganglia (SCG) and treated these tissue isolates with a PI3K/AKT inhibitor (LY294002) and an MRN inhibitor (Mirin). SCG is part of the peripheral nervous system and plays an important role to innervate different organs throughout the head (9). We employed an ex-vivo system in this study because these drugs may be too toxic for live animals (10) . The SCG is dissected from the sympathetic chain of the embryonic rat at pregnancy day 21. We cultured a mixture of these neuronal cells for 5 days and then exposed them to either LY29002 or Mirin. With advent of a scRNA-seq technology, we sequenced and analyzed our treated tissue samples along with mocked controls. The study aims to characterize cellular compositions of the SCG and to further investigate how these inhibitors of the two crucial biological processes (PI3K/AKT and MRN complex) modulate transcriptomic signatures, biological pathways and cellular abundance of the superior cervical ganglia at a single cell resolution.

Results:

Deconvoluting cell types with scRNA-seq

We collected tissue samples from the SCG of embryonic mice (mocked-control and drug-treated) for sequencing with scRNA-seq 10X chromium droplet technology (details in Methods). Three sets of the samples (Exp4-control, Ly-treated and Mirin-treated) were sequenced at the same time, while Exp1-control was processed separately. Raw sequences were aligned to the *Rattus norvegicus(Rnor)*genome. Exp4-control had the highest number of cells and genes, followed by Exp1-control, Mirin-treated and Ly-treated samples (**Table 1**).

To ensure high quality of cells in our final analysis, we performed QC filtering to remove cells with extremely high numbers of features/genes and enriched proportion of mitochondria. Cells with excess gene features usually suggest doublets in 10X droplets, while high percent mitochondria indicate cell deaths. We used interquartile ranges to identify outlier cells (details in Methods) and filtered them out. After QC, approximately 45-55% of the cell population remained in the datasets. Ex4-control still had the highest number of cells (N=5,688) and Ly-treated samples had the lowest (N=2,283). We did not filter our data based on gene quality since feature selection of the top 2000 genes was later performed, resulting in the inclusions of only the most variable and relevant genes for downstream analyses.

Table1: Descriptive statistics of number of cells and features before and after QC filtering across datasets

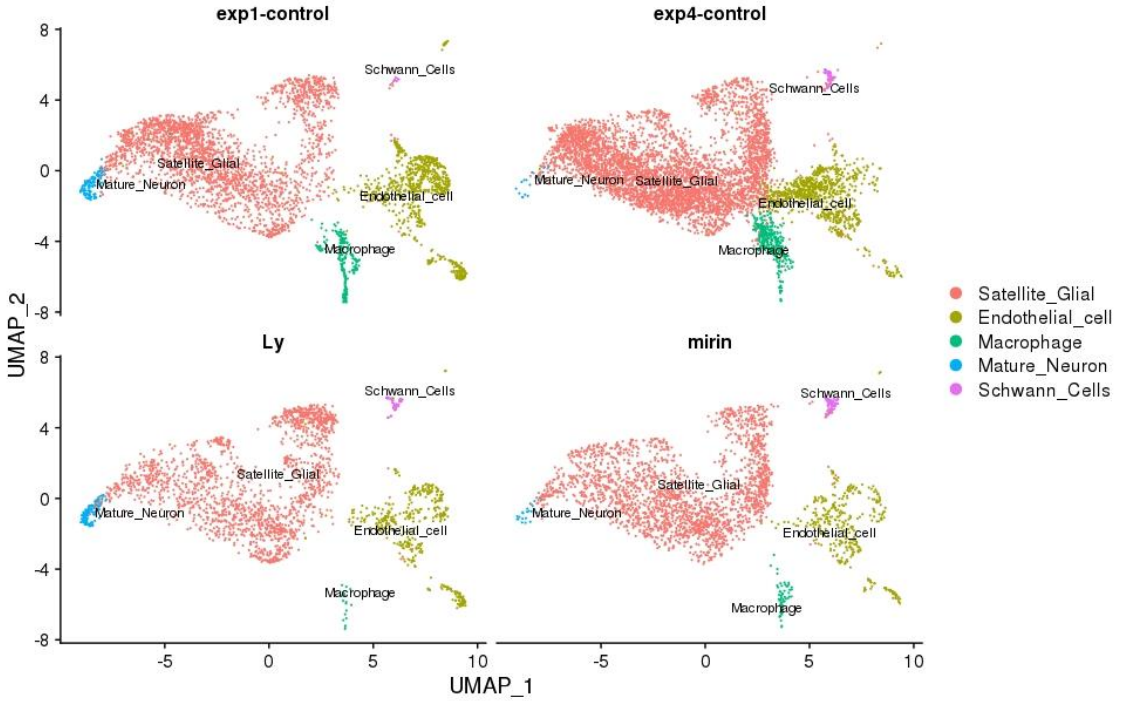
Characteristics	**Exp1-Control**	**Exp4-Control**	**Ly-treated**	**Mirin-treated**
Before filtering				
Number of cells	6,552	11,733	4,418	5,901
Number of genes	15,504	15,733	15,037	15,431
QC filtering thresholds				
Number of features[a]	>5,970	>794.375	>8,294.25	>1,793.5
Percent mitochondria	>10%	>10%	>10%	>10%
After filtering				
Number of cells	3,537	5,688	2,283	2,625
Number of features/genes	15,504	15,733	15,037	15,431
Integrating into one dataset[b]				
Number of cells	3,537	5,688	2,283	2,625
Number of features	2,000	2,000	2,000	2,000

[a] Thresholds for number of features were determined with formula to identify outliers: Q3/Q1 ±1.5*IQR (interquartile range)
[b] These are numbers of cells and genes/features after integrating the four datasets into one. Feature selection of top 2000 most variable genes was performed.

We integrated all four datasets into one combined dataset for final analyses and investigated the presence of different cell types. Data was log-normalized and scaled to remove unwanted source of variation (eg. Batch effects) prior to performing PCA (Principle component analysis). Top 15 principle components, as

determined by the plateau of an elbow plot, were used to run UMAP (Uniform Manifold Approximation and Projection) for a non-linear dimensionality reduction and data visualization (11). Graph-based clustering approach was then performed to partition different cells that belong to the same subtype/community. With a resolution parameter 0.04 (details in Methods), we observed 5 distinct clusters of cell population across our datasets. Capitalizing on a novel single cell annotation tool (*SCfind*) (12) and previously known gene markers of cells in the peripheral nervous system (13), we were able to annotate the clusters to satellite glial, endothelial, macrophage, mature neurons and Schwann cells (**Figure 1**).

Figure1: UMAP plots of cell clusters detected in all four datasets. There were five cell types present in the SCG of embryonic rats in this study, regardless of treatments (mock-control and Ly/Mirin-treated). Each cell cluster is labeled as satellite glial, endothelial, macrophage, mature neuron and Schwann cell.

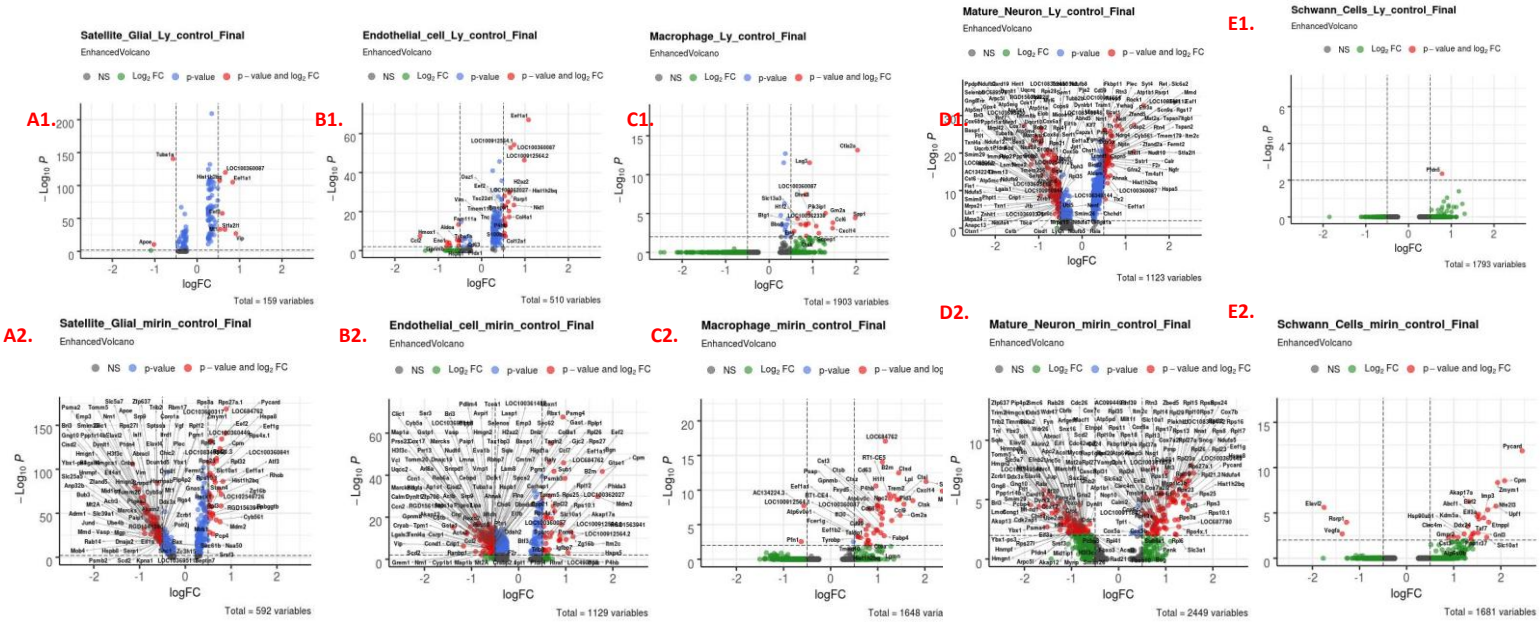


Differentially expressed genes (DE) across treatment conditions

To investigate the effects of treatments on cells in the SCG at a single cell resolution, we performed DE analyses using non-parametric Wilcoxon Rank Sum tests. Two comparison groups were explored for each cell type separately: Ly-treated vs all controls (exp1 and exp4 combined), and Mirin-treated vs. all controls (exp1 and exp4 combined). Given low mRNA content captured in a single cell, scRNA-seq data are much noisier compared to bulk mRNA-seq (14). Therefore, to be able to capture meaningful biological DE genes and prevent false negative, we relaxed our thresholds of significance to adjusted-p value of 0.05 and absolute log fold-change value of 0.5.

Across all five identified cell subtypes in our studies, Mirin seemed to modulate more number of genes than Ly did, compared to mocked controls (**Figure 2**). For examples, for satellite glial cells in particular, Mirin modulated 127 genes, while Ly only affected expressions of 9 genes. We further performed a secondary analysis to explore DE genes that were truly specific to Mirin in satellite glial cells. Interestingly, when we compared gene expressions for Mirin-treated vs. Ly-treated, we only observed three genes (Stfa2l1,Vip and Atf3) that were significantly modulated for satellite glials, despite the overwhelmingly high number of DE genes observed when comparing to controls. This suggests that Ly may have also modulated many different genes like Mirin did, yet in a moderate fashion which was not extreme enough to be captured as statistically significant in our DE analyses when compared against mocked controls.

Figure2: Volcano plots depicting differentially expressed genes (DE) for each cell type across treatment conditions. The upper panels A1-B1-C1-D1 and E1 were DE genes comparing Ly-treated with mocked controls as reference, while the lower panels A2-B2-C2-D2 and E2 depicted DE genes comparing Mirin-treated with mocked-controls for each cell subtype. Log fold change of 0.5 and adjusted p-value 0.05 were used as thresholds of biological and statistical significance. Genes surpassing these significant thresholds were labelled in each volcano plot. Transcriptional profiles of each cell type tended to be more affected by Mirin more than by Ly for SCG tissue isolates in this study.

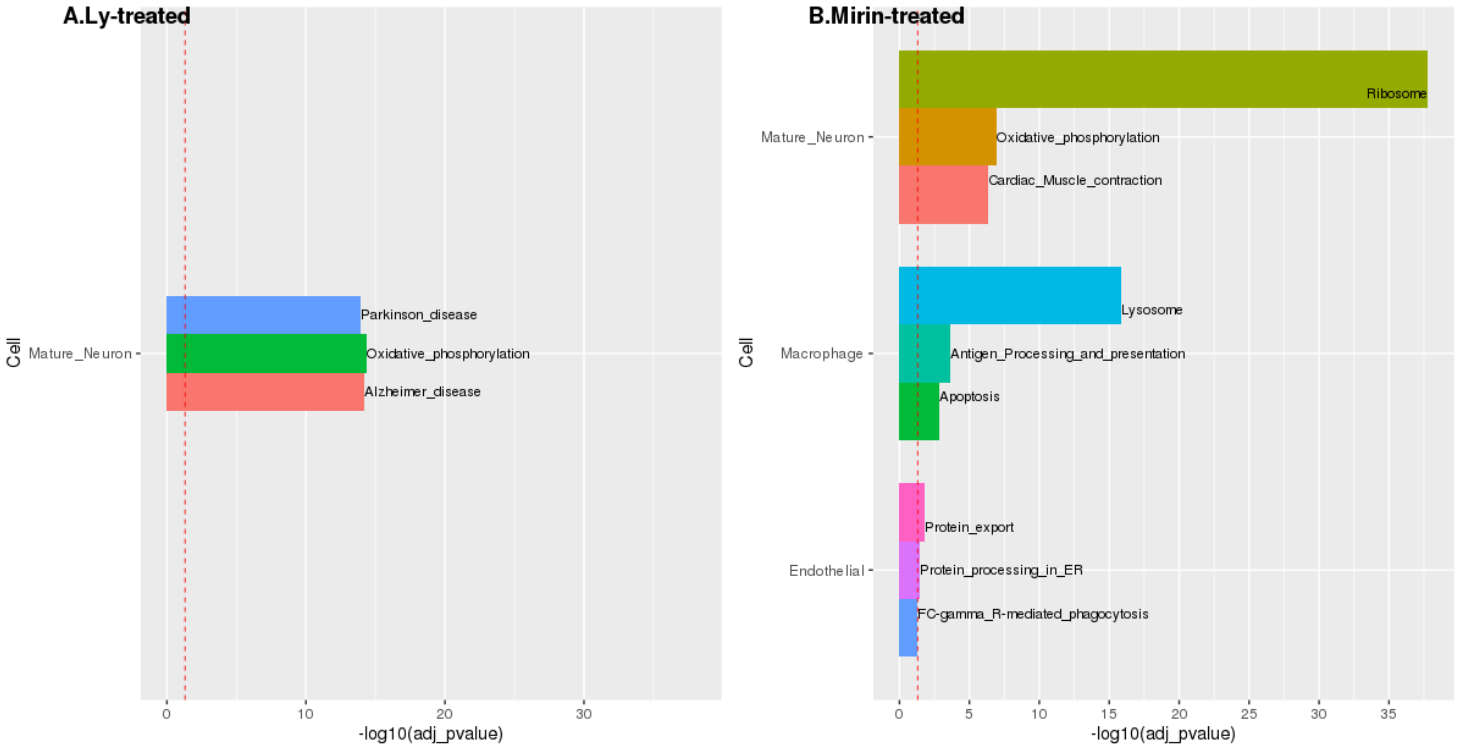


Biological pathways enriched for differentially expressed genes (DE)

We asked if DE genes identified earlier were enriched in specific biological pathways. To anwer this question, we took all the DE genes (up and down-regulated) in each cell type and performed pathway analyses with EnrichR (15) separately for each condition (LY294002 and Mirin). **Figure 3** presented the top three enriched pathways surpassing statistical significance (FDR <0.05) in one cell (Mature neurons)

and three cell types (Mature neurons, macrophages and endothelial cells) for LY294002 and Mirin-treated, respectively. For other cell types, no biological pathway was enriched, partly maybe due to low number of detected DE genes.

Figure3: Pathways enrichment analyses of DE genes in LY294002 (Panel A) and Mirin-treated (Panel B). Adjusted-p values were converted to –log10 scales so that the lowest adjusted p-values appeared on top. The vertical dashed red line demarcated threshold of significance at adjusted-p value 0.05. Only DE genes identified in mature neurons for Ly-treated sample were enriched for certain biological pathways (Parkinson disease, oxidative phosphorylation and Alzheimer disease). For Mirin-treated, enriched biological pathways were detected in three cell types (Mature neuron, macrophage and endothelial cells).

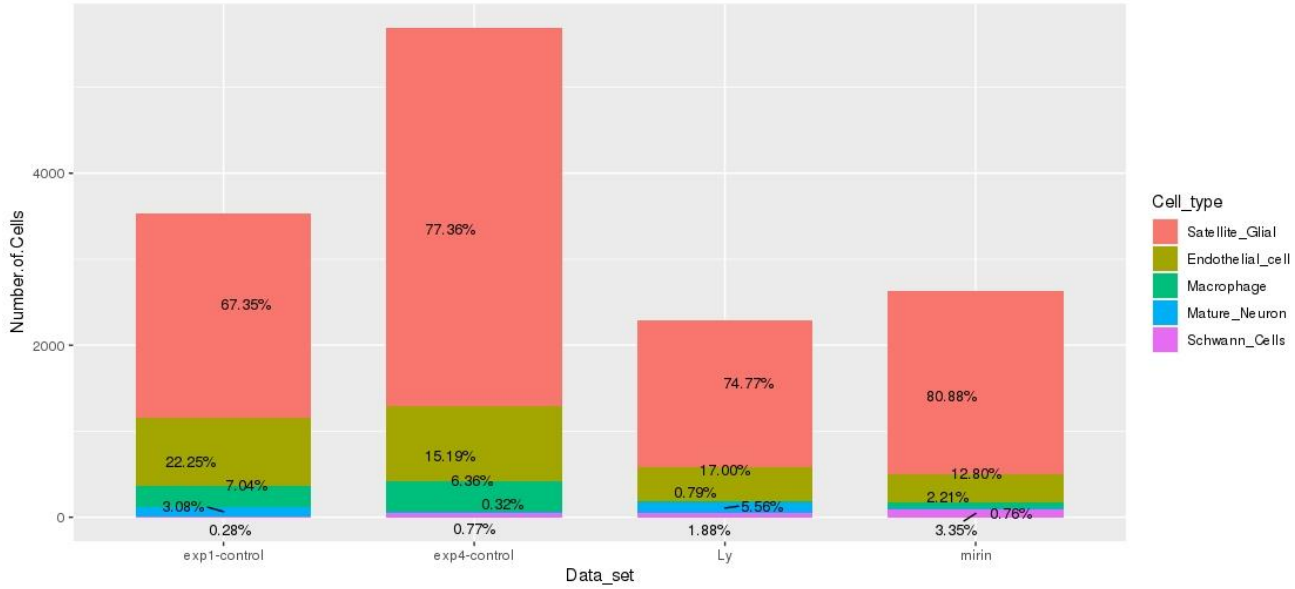


Impacts of LY294002 and Mirin on cell population

Prior studies reported potential impacts of LY294002 and Mirin on cell viability (1, 2, 4). In our studies, we investigated if there were any particular cell subtypes that these two drugs disproportionately affected. Satellite glial, mature neurons and Schwann cells seemed to be distributed comparably across conditions (**Figure 4**). Interestingly, only 0.79% of the whole cell population in LY294002-treated was annotated as macrophages while in controls, the proportion of macrophages were 6.7% (average of both exp1 and exp4 controls); test of differential proportions Chi-square one-sided p<2.2E-16. For endothelial cells, Mirin had reduced the proportion of these cells to 12.80%, compared to in controls (average of 18.72%); test of differential proportions Chi-square one-sided p=8.19E-09. However, we cannot rule out that this finding may be due to composition effect bias, in which an increase in abundance of one cell type (80.88% Mirin-

treated Satellite glial vs 77.36% in mocked controls) would, by default, result in a reduced abundance in another cell type (12.80% Mirin-treated endothelial vs. 18.72% in mocked controls).

Figure4: Proportions of cell population for each cell types across conditions. Mocked-controls had the highest absolute number of cells while Ly-treated had the lowest. Proportions of satellite glial, mature neurons and Schwann cells were comparable in both controls and drug-treated samples. Macrophages seemed to be depleted in Ly-treated while endothelial cells were reduced in Mirin-treated.



Cellular reprograming of macrophages

In some cell types, these drugs do not impact the total cell population per se, but they may be modulating transcriptional signatures. For examples, Mirin-treated macrophages made up 2.21% of the total cell population which was higher than the depleted proportion in Ly-treated (0.79%). Compared to mocked-controls (macrophages 6.7%), this proportion of 2.21% is relatively low, however not as extreme as in Ly-treated case (0.79%). Interestingly, our DE and pathway analyses suggested that Mirin had significantly modified transcriptional signatures of macrophages, compared to controls. Top three enriched biological pathways included lysosome, antigen processing/presentation and apoptosis. We took genes that drove the signals in these three pathways (N genes=15) and investigated if this subset of genes were sufficient to generate transcriptional signatures specific to the treated drugs. We used scaled expression data of these 15 genes and plotted a heat map. As shown in **Figure 5**, there were distinct signatures (visually) between mocked-control and drug-treated indicating that these 15 genes were more likely to be overexpressed in LY294002 or Mirin treated. Notably, many of these genes did not show up as DE genes in Ly-treated sample (**Figure 2**), which might be due to the small number of cells (Number

of macrophages in LY294002 was 18 vs. 58 in Mirin) resulting in reduced power to detect the statistical significance.

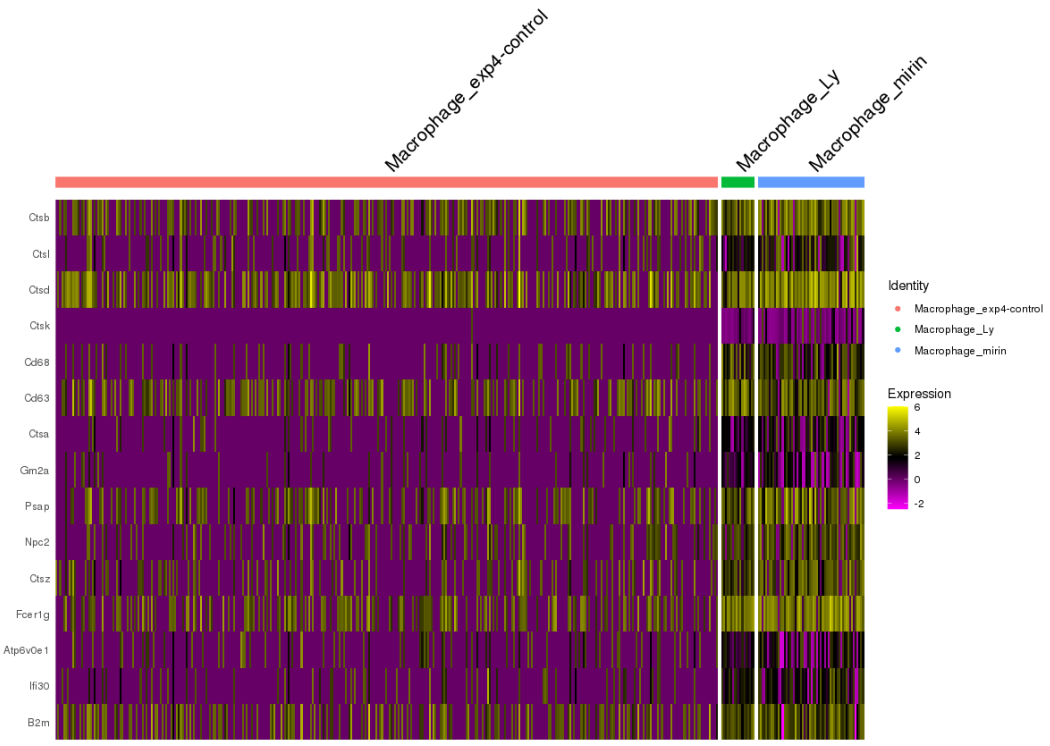


Figure5: Transcriptional heterogeneity of macrophages in mocked-control, LY294002 and Mirin-treated samples. Fifteen DE genes of the top 3 enriched biological pathways for Mirin-treated macrophages were selected (details in Methods) and their scaled expression values were plotted as a heat map. Yellow indicated high expressions, while pink corresponded to low expressions.

Discussion:

With a novel scRNA-seq approach, we deconvoluted five cell subtypes (Satellite, endothelial, macrophage, mature neurons and Schwann cells) in the superior cervical ganglia (SCG) of embryonic rats. For the first time, we reported that these SCG cell types were preferentially modulated by LY294002 and Mirin inhibitors. LY294002, a PI3K-AKT inhibitor, affected gene expressions largely in the mature neurons and depleted macrophage populations when compared against mocked controls. An MRN complex inhibitor, Mirin, on the other hand, modified transcriptional signatures for almost all five SCG subtypes, particularly for Satellite, endothelial and mature neurons; however, we observed reduced cell population only for endothelial cells upon Mirin treatment.

Pathway enrichment analyses of the detected DE genes revealed interesting findings about the mechanisms underlying the effects of LY294002 and Mirin on the SCG cell populations. Despite a high numbers of DE genes observed in all cell types, we only found these genes to be enriched for certain biological pathways in mature neurons when treated with LY294002 and only in mature neuron, macrophages and endothelial cells when treated with Mirin. Notably, the top three enriched pathways for Ly-treated mature neurons pointed to neurological disorders such as Parkinson and Alzheimer diseases, which have been previously reported to be associated with defects in the PI3K-AKT pathway (5). In the Mirin-treated samples, we found different biological mechanisms for the SCG subtypes. For mature neurons, Mirin modulated ribosomal and oxidative phosphorylation pathways, potentially reflecting the involvement of the MRN complex in cell cycle and growth (3, 16). In macrophages and endothelial cells, however, Mirin seemed to be affecting the antigen processing/presenting and phagocytosis, which may highlight another important role of the MRN complex in DNA damage and immune response (17, 18).

In addition to modulating transcriptional signatures of the SCG cell subtypes, LY294002 and Mirin also preferentially influenced the abundance of cellular composition. We reported that LY294002 significantly reduced macrophage population. In fact, prior study in mutant mice showed that constitutive activation of PI3K/AKT (a target of LY294002 inhibitor) to phosphorylate downstream transcription factor ets-2 and a colony-stimulating factor 1 (CSF-1) was essential in promoting macrophage survival (19). Interestingly, Mirin exposure did not affect macrophage population as much; yet, this MRN complex inhibitor did significantly alter the transcriptional profile of macrophages. Using the top 15 DE genes driving the signals of the enriched biological pathways, we were able to reconstruct a unique signature of this cell type that were distinct from mocked-controls. This finding was in line with previous research reporting that Mirin may not affect cell survival per se, but it may modulate macrophage cellular functions (4, 20).

In summary, our study emphasized the power of transcriptomic profiling at a single cell resolution to characterize cell subtypes in the SCG and to investigate how these cells could be affected differentially or similarly by inhibitors of the PI3K/AKT (LY294002) pathway and the MRN complex (Mirin). Upon validations, our findings hold clinical potentials to identify specific cell subtypes/pathways as novel treatment targets for neuronal diseases such as Parkinson/Alzheimer or other disorders associated with defects in the PI3K/AKT pathway and the MRN complex.

References:

1. Engelman JA, Luo J, Cantley LCJNRG. The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. 2006;7(8):606-19.
2. Hemmings BA, Restuccia DFJCSHpib. Pi3k-pkb/AKT pathway. 2012;4(9):a011189.
3. Rupnik A, Grenon M, Lowndes N. The MRN complex. Current Biology. 2008;18(11):R455-R7.
4. Dupré A, Boyer-Chatenet L, Sattler RM, Modi AP, Lee J-H, Nicolette ML, et al. A forward chemical genetic screen reveals an inhibitor of the Mre11–Rad50–Nbs1 complex. Nature chemical biology. 2008;4(2):119.
5. Xu F, Na L, Li Y, Chen L. Roles of the PI3K/AKT/mTOR signalling pathways in neurodegenerative diseases and tumours. Cell & Bioscience. 2020;10:1-12.
6. Huang X, Liu G, Guo J, Su Z. The PI3K/AKT pathway in obesity and type 2 diabetes. International journal of biological sciences. 2018;14(11):1483.
7. Situ Y, Chung L, Lee CS, Ho V. MRN (MRE11-RAD50-NBS1) complex in human cancer and prognostic implications in colorectal cancer. International journal of molecular sciences. 2019;20(4):816.
8. Wang Y, Kuramitsu Y, Baron B, Kitagawa T, Tokuda K, Akada J, et al. PI3K inhibitor LY294002, as opposed to wortmannin, enhances AKT phosphorylation in gemcitabine-resistant pancreatic cancer cells. 2017;50(2):606-12.
9. Li C, Horn JP. Physiological classification of sympathetic neurons in the rat superior cervical ganglion. Journal of neurophysiology. 2006;95(1):187-95.
10. Kobayashi M, Kim J-Y, Camarena V, Roehm PC, Chao MV, Wilson AC, et al. A primary neuron culture system for the study of herpes simplex virus latency and reactivation. JoVE (Journal of Visualized Experiments). 2012(62):e3823.
11. McInnes L, Healy J, Melville J. Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:180203426. 2018.
12. Lee JTH, Patikas N, Kiselev VY, Hemberg M. Fast searches of large collections of single cell data using scfind. bioRxiv. 2019:788596.
13. Abcam. Neural markers [Available from: https://www.abcam.com/neuroscience/neural-markers-guide.
14. Chen G, Shi T. Single-cell RNA-seq technologies and related computational data analysis. Frontiers in genetics. 2019;10:317.
15. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic acids research. 2016;44(W1):W90-W7.
16. Ashton TM, McKenna WG, Kunz-Schughart LA, Higgins GS. Oxidative phosphorylation as an emerging target in cancer therapy. Clinical Cancer Research. 2018;24(11):2482-90.
17. Rosales C, Uribe-Querol E. Phagocytosis: a fundamental process in immunity. BioMed research international. 2017;2017.
18. Bansal A, Neuhaus R, Izquierdo-Alvarez E, Nolte H, Vorholt D, Feldkoetter H, et al. ATM-mediated DNA damage response in macrophages primes phagocytosis and immune checkpoint regulation. bioRxiv. 2020.
19. Smith JL, Schaffner AE, Hofmeister JK, Hartman M, Wei G, Forsthoefel D, et al. ets-2 Is a Target for an AKT (Protein Kinase B)/Jun N-Terminal Kinase Signaling Pathway in Macrophages of motheaten-viable Mutant Mice. Molecular and cellular biology. 2000;20(21):8026-34.
20. Colonna M. DNA damage response impacts macrophage functions. Blood, The Journal of the American Society of Hematology. 2015;126(22):2440-2.
21. Kim JY, Shiflett LA, Linderman JA, Mohr I, Wilson AC. Using homogeneous primary neuron cultures to study fundamental aspects of HSV-1 latency and reactivation. Herpes Simplex Virus: Springer; 2014. p. 167-79.

22. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nature biotechnology. 2018;36(5):411-20.
23. Zappia L, Oshlack A. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. Gigascience. 2018;7(7):giy083.

Methods:

Isolating and seeding of SCG neurons from embryonic rats

We dissected superior cervical ganglia (SCG) from prenatal rats at embryonic day 21. Detailed protocols of the procedure are mentioned elsewhere (10, 21). In brief, a pregnant rat was sacrificed in a CO_2 chamber at pregnancy day 21. Embryonic rats were removed and both SCG of each pub were isolated. We filtered and extracted the neuronal cells and dispensed about 5000-6000 total live cells into a 96-well plate. The culture plate was incubated for five days with 5 μM aphidicolin and 20 μM 5-fluorouracil (anti-mitotic agents) to deplete fast-growing non neuronal cells. After DIV 5 (day in vitro), we exposed our cultured neuronal cells to either Mirin or LY294002. Mocked controls (no treatment exposed), Mirin and LY294002 tissue isolates were, then, harvested and analyzed.

Droplet-based single cell-sequencing

To generate single-cell gel bead-in-emulsions (GEMs), we used droplet-based 10X chromium v2 3' technology and loaded our tissue isolates onto the Chromium controller. We followed the manufacturer protocol accessible at https://bit.ly/3dUNOLZ. We then prepared our sequencing library and performed RNA sequencing with Illumina sequencing HiSeq/NextSeq.

Alignment and generation of read counts

We aligned and generated read counts with Cell Ranger pipelines as per instructions on the 10x Genomics website. Initially, we built genomic reference for *Rattus norvegicus(Rnor)*, a model organism used in our study. The *Rnor*_6 fasta file and the corresponding gtf file were downloaded from the Ensemble webpage. We used command *cellranger mkgtf* to filter only attributes of the gtf file required for Cell Ranger. The *Rnor* version 6 genomic reference was finally built with *cellranger mkref*. With this pre-built *Rnor* genome as a reference, we ran *cellranger count* to align our fastq files and generated single cell feature counts for downstream analyses.

scRNA-seq analyses with Seurat

We loaded all four outputs (Exp1-control, Exp4-control, LY294002-treated and Mirin treated samples) from Cell Ranger and performed scRNA-seq analyses with an R-package Seurat v3.0 (22). We created Seurat objects for each dataset and performed quality control (QC) separately. For all datasets, we removed cells with high proportion of mitochondrial genes (>10%), which often indicated cell deaths. To identify cell outliers that expressed excessively high number of genes (a marker of doublets) or extremely low gene numbers (a marker of empty droplets), we identified thresholds for each dataset, separately,

using non-parametric approach Q3/Q1 (75/25 percentiles) ±1.5IQR (interquartile range). We only retained cells whose number of expressed genes fell within this range.

After filtering, we integrated all four datasets into a single Seurat object. With *NormalizeData* function, we normalized to 10,000 counts per cell (scale factor) and log-transformed our feature expression matrix. We, then, performed feature selections with vst selection method modeling the mean-variance relationship in order to identify the top 2,000 most variable genes across cells. This step also worked to remove any genes of poor quality (present in small number of cells) and reduced noise in downstream analyses. We scaled our data using *ScaleData* function to regress out any unwanted variation such as batch effect before performing dimensionality reduction (PCA: principle component analysis).

Generating and Annotating different clusters

With scaled expression matrix of the selected 2000 genes, we conducted linear dimensional reduction PCA. An elbow plot ranking percent of variance explained by each PC was generated to facilitate the selection of PC numbers sufficient to capture the true signal in the dataset. The plot plateau at PC 15; thus, fifteen top PCs were used in our downstream clustering analyses. We used *FindNeighbors* and *FindClusters* functions to find different underlying clusters in our data. This clustering algorithm used a K-nearest neighbor graph-based approach (KNN) based on the Euclidean distance of the chosen 15 PCs in order to partition cells into 'communities'. We have tried different values of the hyper parameter *resolution* (0,0.02,0.04,0.06,0.08 and 0.1); this parameter influences the number of clusters formed. Using *Clustree* R package (23), we built a tree of possible clusters generated from all resolution values. We found that resolution = 0.04 gave the most stable clusters with distinct independent cell subtypes. This value was then chosen to generate cell clusters in our study.

We used a non-linear dimensional reduction method (UMAP) to visualize and investigate different clusters. We observed five clusters across our four datasets. To annotate each of these cell subtypes, we used *FindConservedMarkers* function. This function identified canonical cell markers of each cluster that were conserved across datasets regardless of treatments. For cluster 2, 3 and 4, we were able to use canonical neuronal gene markers from *abcam* (13) to annotate. Cluster 2 with marker gene CD68 was identified as macrophages; cluster 3 with MAP2 was annotated as mature neurons; while cluster 4 with GAP43 as Schwann cells. For cluster 0 and cluster 1, we leveraged a new tool called *SCfind* (12). We input top 30 conserved genes and these genes were found to be enriched for satellite glial cells for cluster 0 and endothelial cells for cluster 1.

Differential gene expression (DE) and pathway analyses

We performed differential gene expression analyses in two scenarios: LY294002 vs. controls and Mirin vs. controls. We combined both mocked controls as one reference dataset. Wilcoxon ranked sum test, a default setting in Seurat, was used with *FindMarkers* function. We ran these two separate comparisons in each of the identified cell subtypes. The thresholds of significant DE genes were set at absolute logFC > 0.5 and adjusted p-value <0.05. We relaxed these thresholds to combat highly noisy nature of scRNA-seq.

In each cell type, we took all the DE genes and performed pathway analyses using an online tool EnrichR (15). We only focused on KEGG 2019 mouse database as this was closest to the *Rattus norvegicus* in our study. Pathways with FDR < 0.05 were considered statistically significant.

Proportion of cell subtype population and heat map of macrophage signatures

We used *prop.table* function to generate a proportion table of each cell subtype across the datasets. The two control samples were combined using average values. One-sided test of proportion with 1 degree of freedom Chi-square statistic was used to compare different proportion of cells in the mocked control vs. drug-treated. We used one-sided test, instead of two-sided because we believed that if these cytotoxic drugs were to have any effects on the cell populations, it would be to reduce and not to increase cell numbers.

To investigate transcriptional heterogeneity of macrophages across conditions, we selected all genes that were driving the signals of the top three enriched pathways (Lysosome, Antigen processing/presentation and apoptosis) in Mirin-treated macrophages. These 15 genes included: *CTSB, CTSL, CTSD, CTSK, CD68, CD63, CTSA, GM2A, PSAP, NPC2, CTSZ, FCER1G, ATP6VOC, IFI30, B2M and LGLMN*. We used scaled data of their expressions and created a heat map with *DoHeatmap* function in our Seurat subset for only macrophage cluster in Exp-4 control, LY294002 and Mirin datasets.