



# Single-cell multi-modal data handling in R/Bioconductor using TENxMultiomeTools package

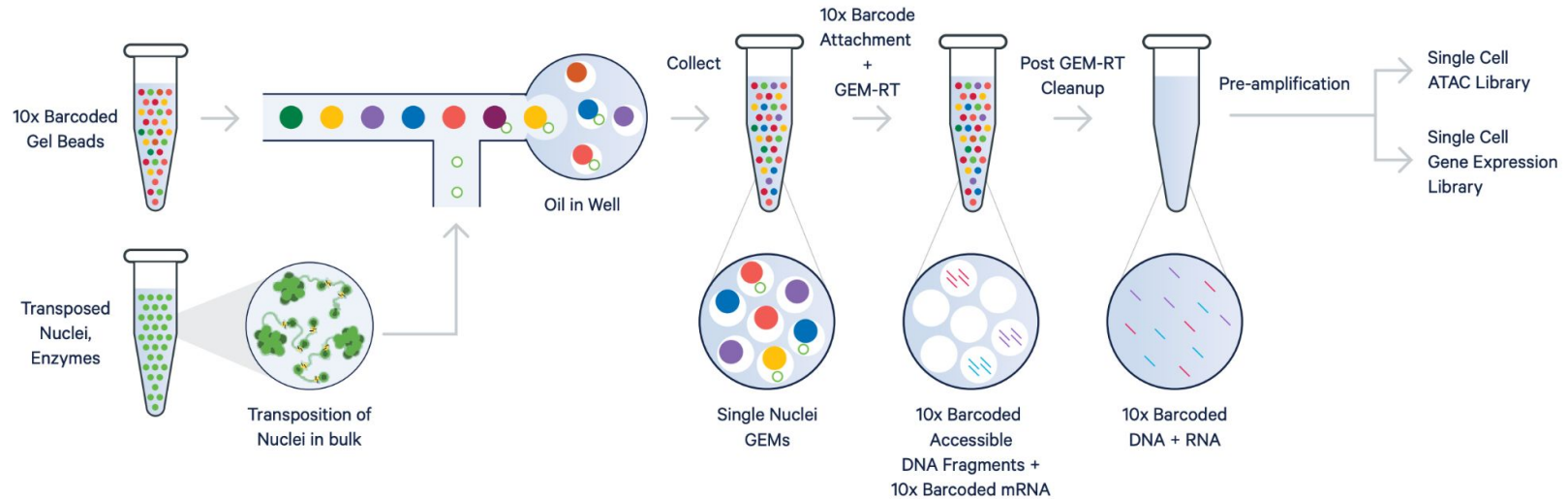
Dario Righelli, PhD

PostDoc - Risso's Lab

Department of Statistical Sciences, University of Padova, Padua, Italy

dario.righelli@gmail.com

# The 10X Multiome protocol



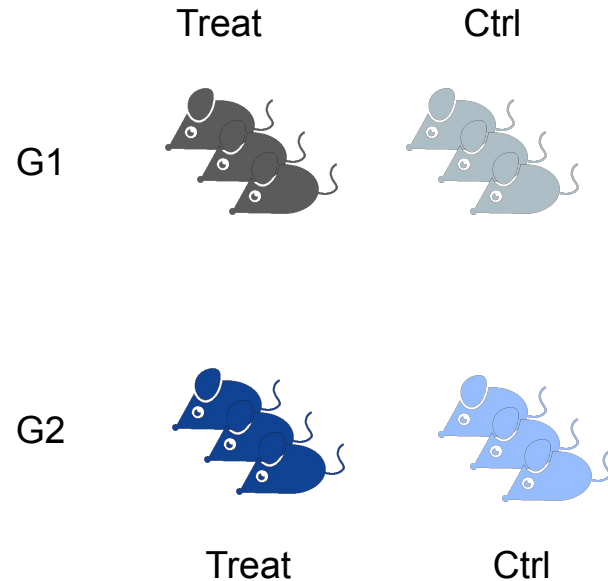
**Figure 2. Efficient and robust workflow.** Access a unified view of transcription and the chromatin landscape by combining gene expression and ATAC-seq data from the same single cell with a simple, streamlined workflow. Starting with a single nuclei suspension, transposition is performed in bulk before individual nuclei are captured in GEMs (Gel Bead-in-emulsion), where DNA fragments and the 3' ends of mRNA are barcoded. Generate two complementary libraries from each sample, and link gene expression and open chromatin profiles back to the same cell with certainty.

# 10X Cell Ranger pipeline

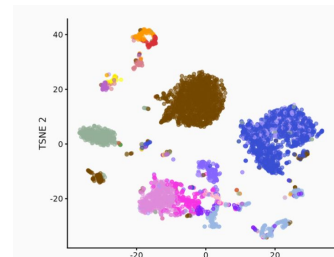
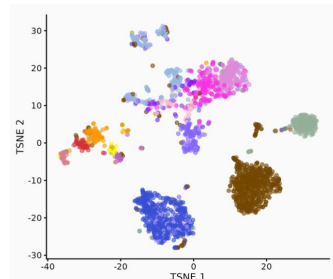
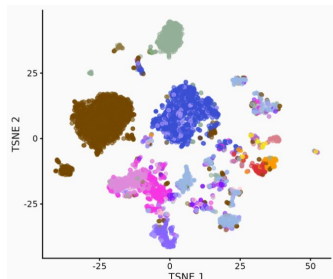
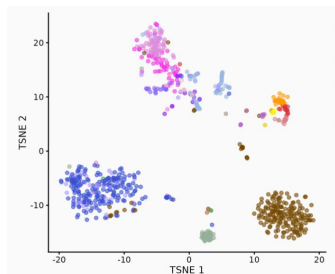
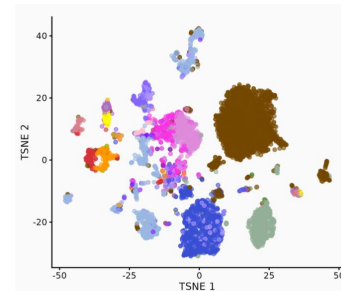
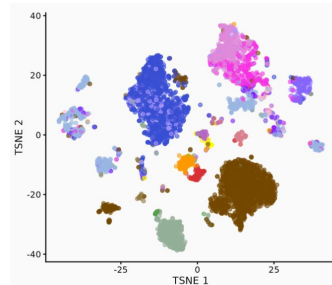
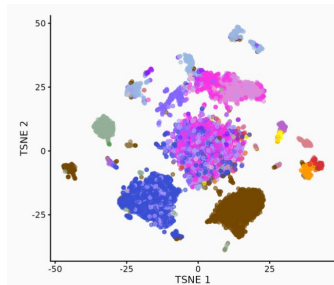
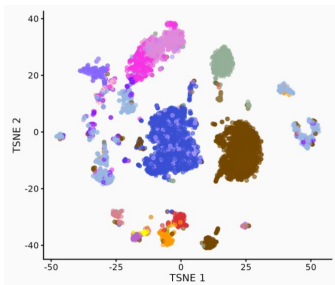
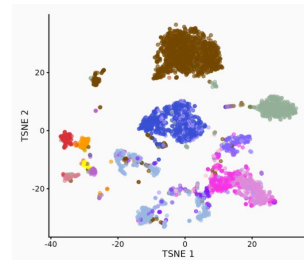
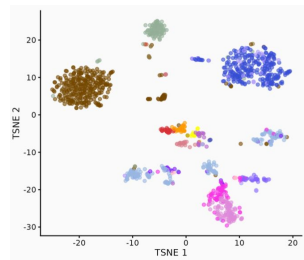
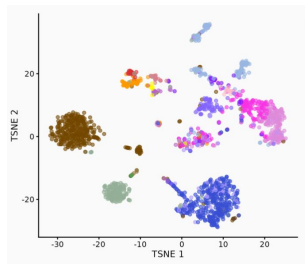
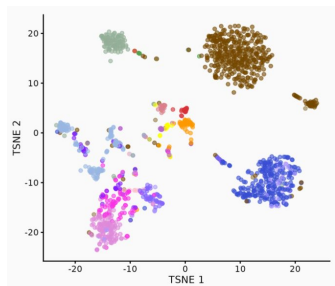
Output Files <a href="#">format details →</a>		
<a href="#">Loupe Browser visualization file (CLOUPE)</a>	1.40 GB	49d0621186974c368442ee10008de461
<a href="#">Secondary analysis outputs (DIR)</a>	383.16 MB	6c1b8f17fc38b457588b565846438597
<a href="#">Run summary (HTML)</a>	5.68 MB	36f5b9f0d27b12ae194476fea4f47ea7
<a href="#">Run summary metrics (CSV)</a>	2 KB	755d8e040aaf159e9d99aefcc52bc953
<a href="#">Per barcode summary metrics (CSV)</a>	85.51 MB	356b41d11800316657573798e48ec18f
<a href="#">Filtered feature barcode matrix MEX (DIR)</a>	396.43 MB	b99400e8000dec92c1b6c9b080e9b01a
<a href="#">Filtered feature barcode matrix (HDF5)</a>	158.62 MB	cee000a98c8a05d0456add3ff864cdb0
<a href="#">Raw feature barcode matrix MEX (DIR)</a>	483.21 MB	f1f78862adaccb6a614803239bd3bd70
<a href="#">Raw feature barcode matrix (HDF5)</a>	195.95 MB	8b27a5fff23f82b60b062399ca29917e
<a href="#">GEX Position-sorted alignments (BAM)</a>	31.76 GB	29c8674d7e45229f0a3eb777f3bfd78c
<a href="#">GEX Position-sorted alignments (BAM index)</a>	10.69 MB	d4e6b5b4720c75c14322ee262d9e16e5
<a href="#">GEX Per molecule information file (HDF5)</a>	182.94 MB	aadd4323de6f46f00076d4989fa2357a
<a href="#">ATAC Position-sorted alignments (BAM)</a>	57.87 GB	965f5f3df8167d342b55a9ad2e8f99bc
<a href="#">ATAC Position-sorted alignments (BAM index)</a>	8.47 MB	fa7f7688587338b9fce01cc845ff444a
<a href="#">ATAC Per fragment information file (TSV.GZ)</a>	3.30 GB	d0046f31a25b076ec389dcf9c017763b
<a href="#">ATAC Per fragment information index (TSV.GZ index)</a>	1.44 MB	7837beb5ed0dc2e31d972cc012d962cf
<a href="#">ATAC peak locations (BED)</a>	2.55 MB	22bf997edb4ce5c6e28490c632b3e2d4
<a href="#">ATAC smoothed transposition site track (BIGWIG)</a>	2.56 GB	6171046154084e2c8c76d16a8da0aff0
<a href="#">ATAC peak annotations based on proximal genes (TSV)</a>	5.90 MB	f6fd80ca3b5c33d50e9f1e25cad38b1b

# A real 10x Multiome Experiment

- Brain sections of twelve mice
  - Two different genotypes
  - Two different conditions
    - Treatments and controls
  - Three replicates per condition
- 10x Multiome per each mouse
  - ATAC + RNA
- We are interested in the differences between conditions across genotypes
- We can build a SingleCellExperiment per each mouse
  - concatenate them as a list to execute the  
TENxMultiomeTools pipeline

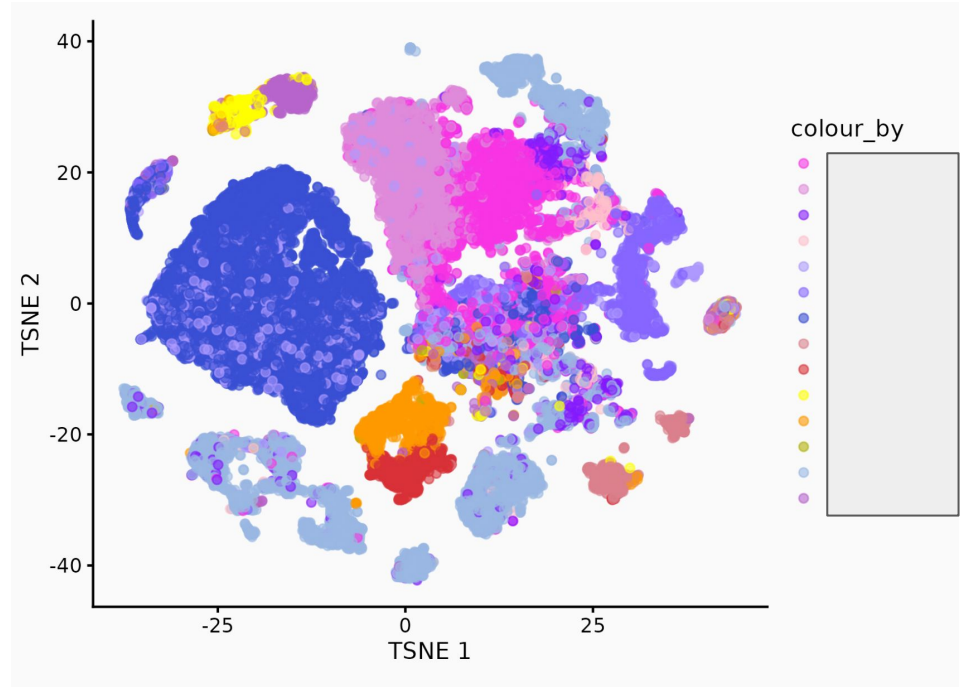


# t-sne for each mouse

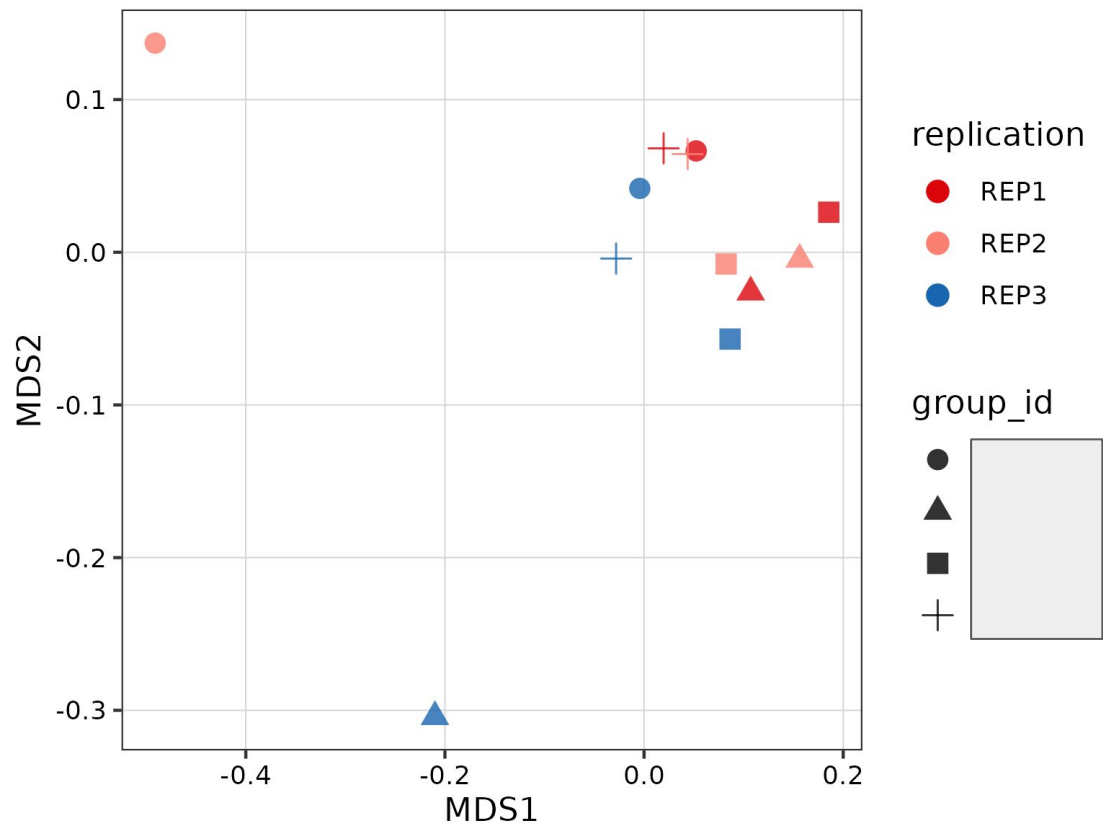


# t-sne MNN for visualization all mice together

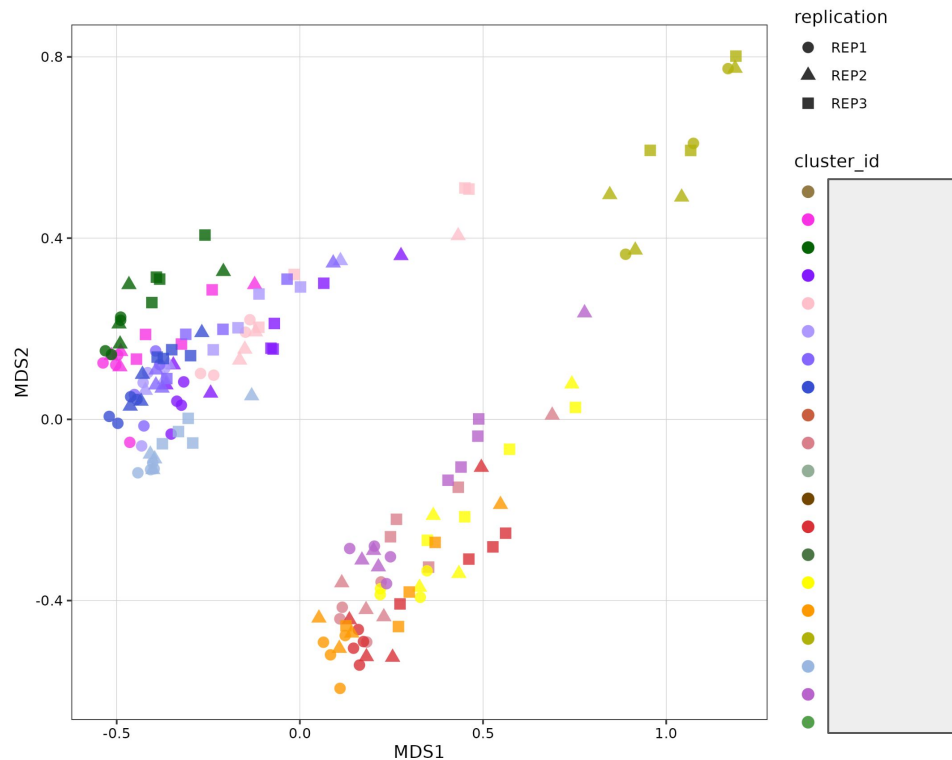
- Harmonize data with MNN
  - using *batchelor::fastMNN*
- See results with a simple t-sne
  - *plotTSNE* function



# Pseudo-Bulk MDS highlighted by replicates

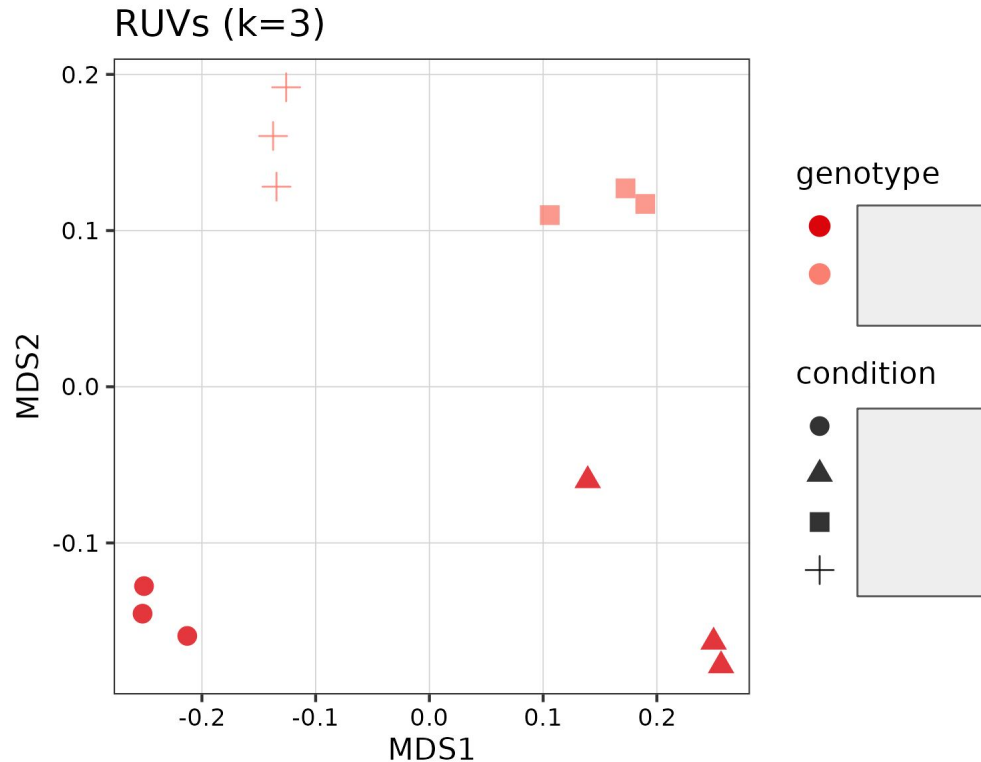


# Pseudo-Bulk MDS highlighted by cell types and replicates

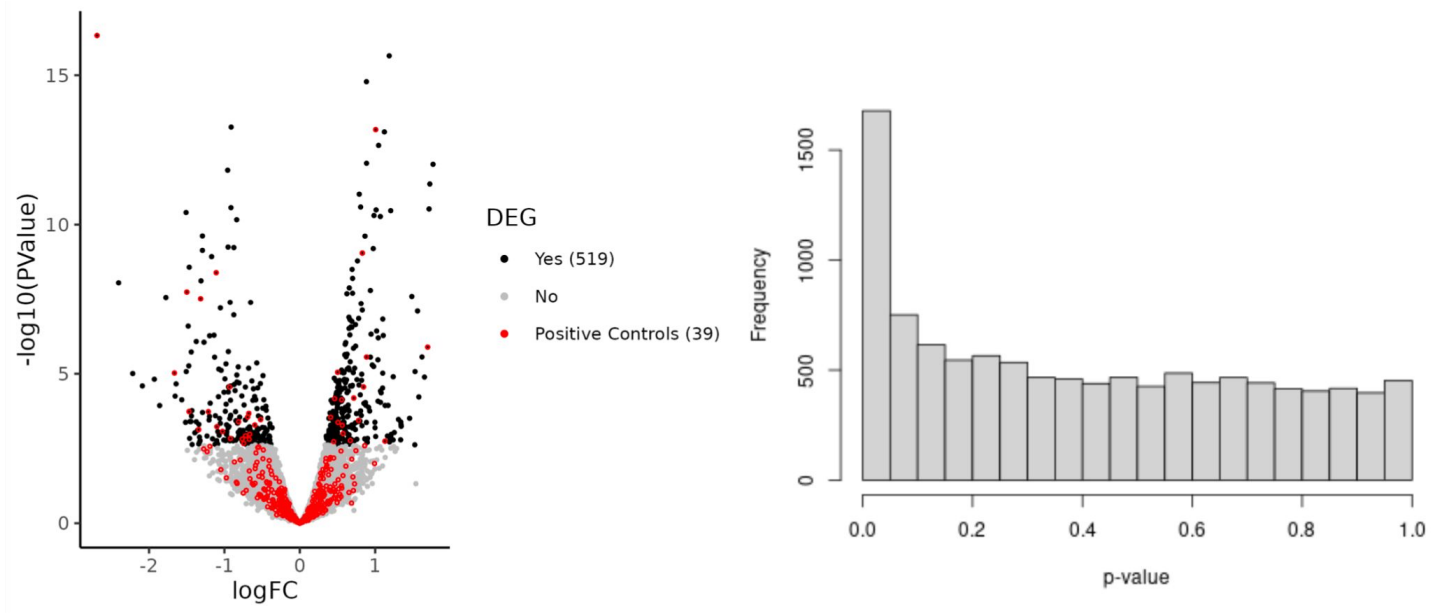




# A normalized example for a specific cell-type



# Two conditions DEGs with edgeR



# Future Plans

- Facilitate the execution of the package with multiple replicates
  - list of SingleCellExperiments or single object with all replicates
- Reduce the number of external dependencies for tracks generation
- Add other packages for better comparison during the analysis
  - azimuth for the cell type labeling
  - doublets identification both on RNA and ATAC
  - ...
- submit package to bioconductor

## Additional links

- package: <https://github.com/drighelli/TENxMultiomeTools>
- analysis rmd/html file: <https://github.com/drighelli/EuroBioc2022Resources>
- 10x multiome dataset from official website [here](#)

# Acknowledgements

Risso Lab @ University of Padova

Dr. Davide Risso

Elena Zuin

Peixoto Lab @ Washington State University

Dr. Lucia Peixoto

Ford Kaitlyn

Muheim Christine

Popescu Alexander

Medina Elizabeth

Schoch Hannah

Singletary Kristan

Pombo Lab @ Max Delbrück Center, Berlin

Dr. Ana Pombo

Dominika Vojtasova

Chan  
Zuckerberg  
Initiative 

