# Alexandria: A Single-Cell RNA-Seq and Analytics Platform for Global Health



Work-in-progress documentation for the Alexandria platform and all associated tool workflows.

The first tool Alexandria offers is the dropseq\_scCloud, a workflow that joins the functionalities of [Drop-seq pipeline](http://mccarrolllab.org/dropseq/) with [scCloud](https://sccloudpy.readthedocs.io/en/latest/).

### Background

Single-cell genomic profiling approaches are transforming basic and applied biological inquiry. These methods—and, in particular single-cell RNA-Seq (scRNA-Seq)—enable powerful, general frameworks for identifying cell types/states, circuits, and biomarkers by explicitly measuring single-cell properties (e.g., gene expression), rather than just the mean of a population of unknown complexity, as well as their co-variation across single cells. The exciting possibility of realizing unbiased “atomic” resolution cellular maps of biological specimens with scRNA-Seq has spurred efforts by the broader scientific community to begin formally constructing a Human Cell Atlas (HCA) as an essential resource for the community2. Unfortunately for the global health community, initial HCA efforts will focus on healthy human tissues alone. While the pipelines used by the HCA to preprocess data will be open source, a distinct effort is needed to bring their transformative potential to bear on datasets of relevance to global health. Moreover, as many global health studies rely on model organisms, such as non-human primates (NHP; e.g., Cynomolgus or Rhesus) or rodents (e.g., mice, rats, guinea pigs), where perturbation studies (e.g., challenge or vaccination) are more feasible and ethical, and more controlled experiments can be performed, similar approaches need to be available to analyze these datasets. Further, new integrative analytic pipelines and visualization, currently outside of the scope of HCA, need to be established to enable efficient comparisons across datasets, species, and diseases to facilitate rapid in silico hypothesis testing and the prioritization of follow-up experiments. Equally importantly, a dedicated front-end portal is needed to facilitate globally access and querying of these data resources to allow data inspection by all who have domain expertise and the potential to contribute.

Here, we propose to realize these goals via “Alexandria” – a single-cell portal and data resource for the global health community. Alexandria will use unified HCA pipelines to preprocess, store, and visualize datasets of interest to the global health community, enabling rapid realization of transformative insights and the prioritization of follow-ups. To maximize impact and utility, Alexandria will build upon existing efforts at the Broad which leverage HCA pipelines—namely Single Cell Portal—but will be further enhanced to enable queries across gene sets, cell types, and models as inspired by the types of data collected by the global health community. This will power vital cross-comparisons while simultaneously providing novel analytic capabilities for HCA community at large. Moreover, Alexandra will similarly empower the broader global research community—from individuals with limited experience in single-cell analysis to power users looking to more rapidly isolate specific subsets of data from several experiments—to examine and parse scRNA-Seq data, so that the insights and intuitions of the entire scientific community can be leveraged to enable rapid progress in fighting a variety of human maladies.