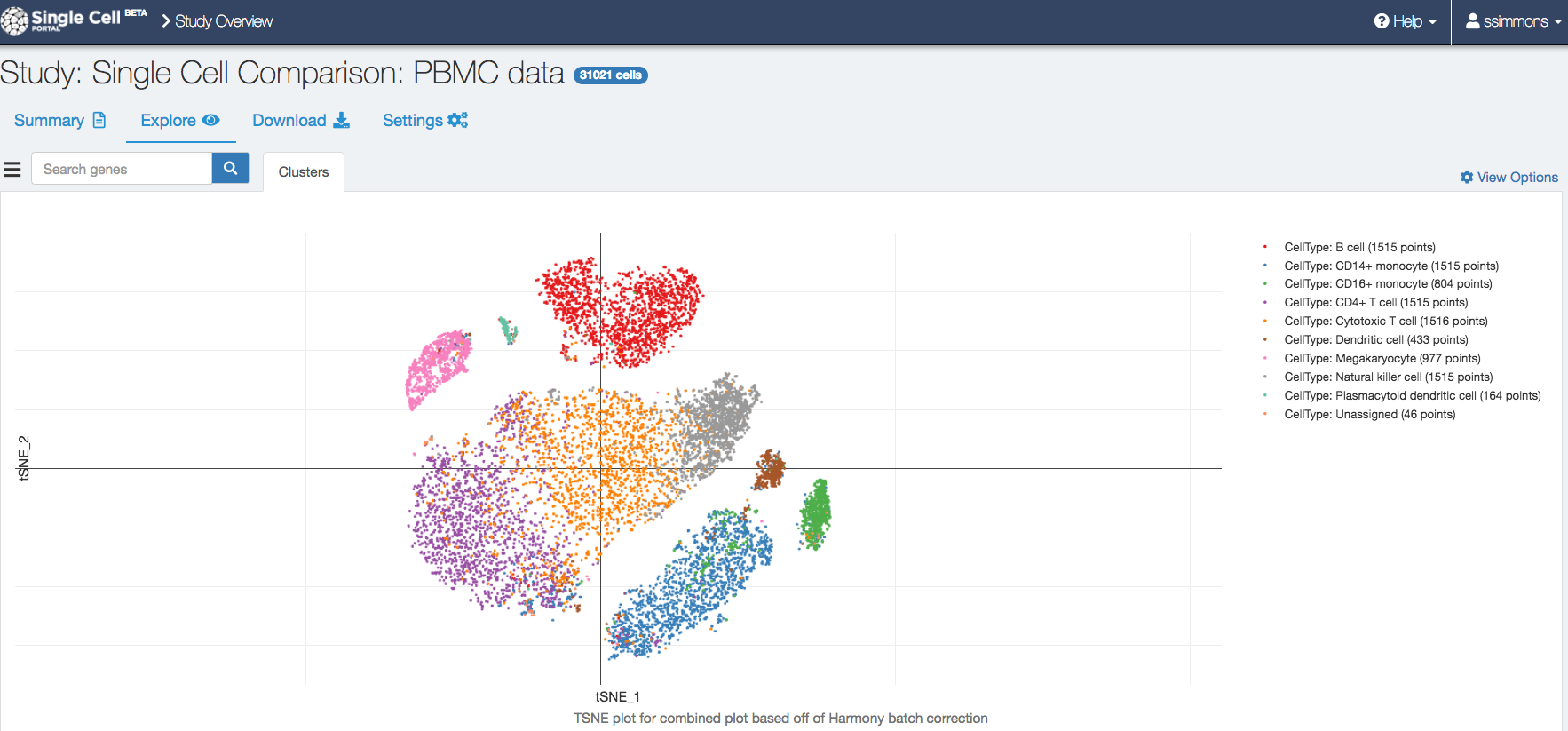
***Description of Single Cell Portal access for “Single Cell Comparison: PBMC data”***

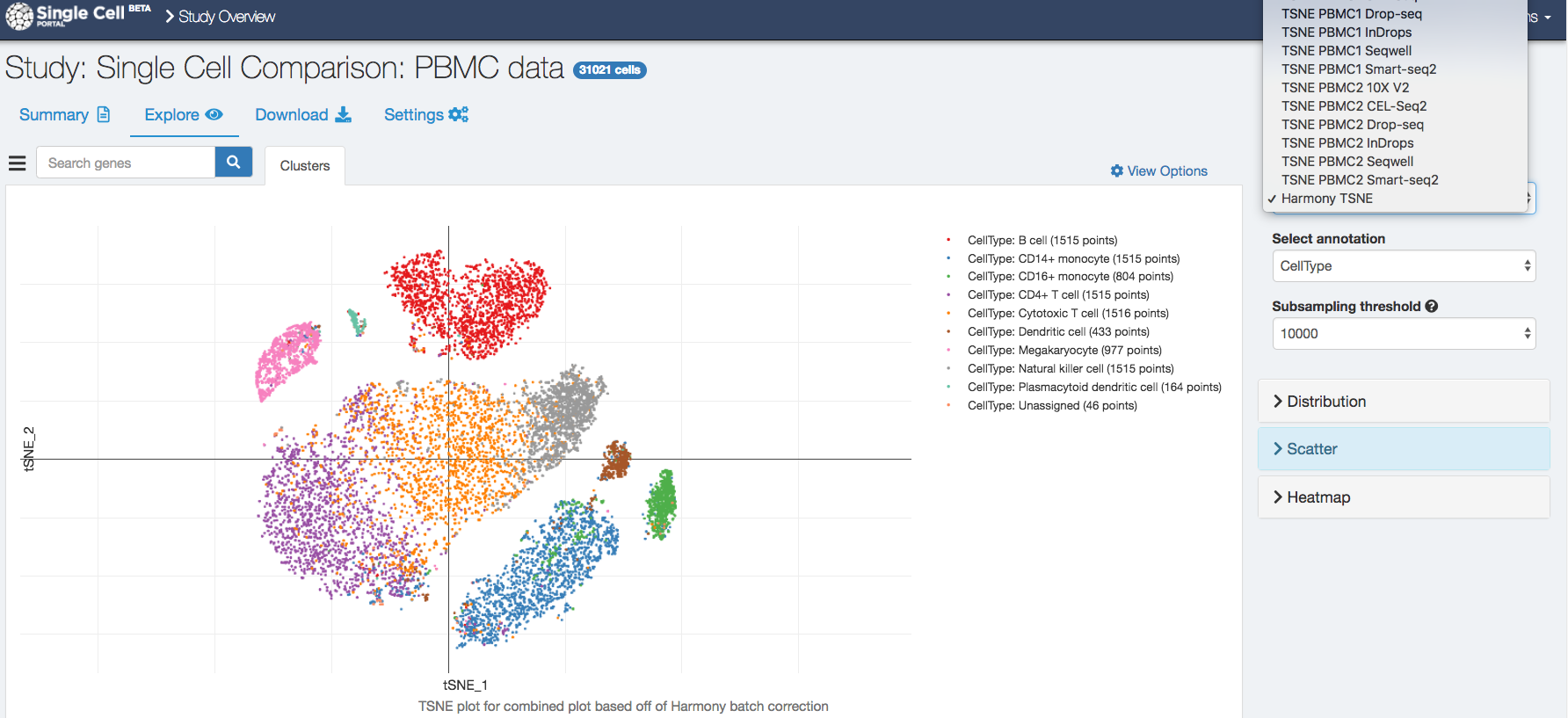
This single cell portal study corresponds to Peripheral Blood Mononuclear Cell (PBMC) single cell RNA-seq (scRNA-seq) data from the recent manuscript “Systematic comparative analysis of single cell RNA-sequencing methods” (see our preprint at: <https://www.biorxiv.org/content/10.1101/632216v1)>.

In particular, this study contains data from 6 scRNA-seq methods (Smart-seq2, CEL-Seq2, 10x Chromium, Drop-seq, Seq-Well, and inDrops). Each method was applied to two replicates (PBMC1 and PBMC2).

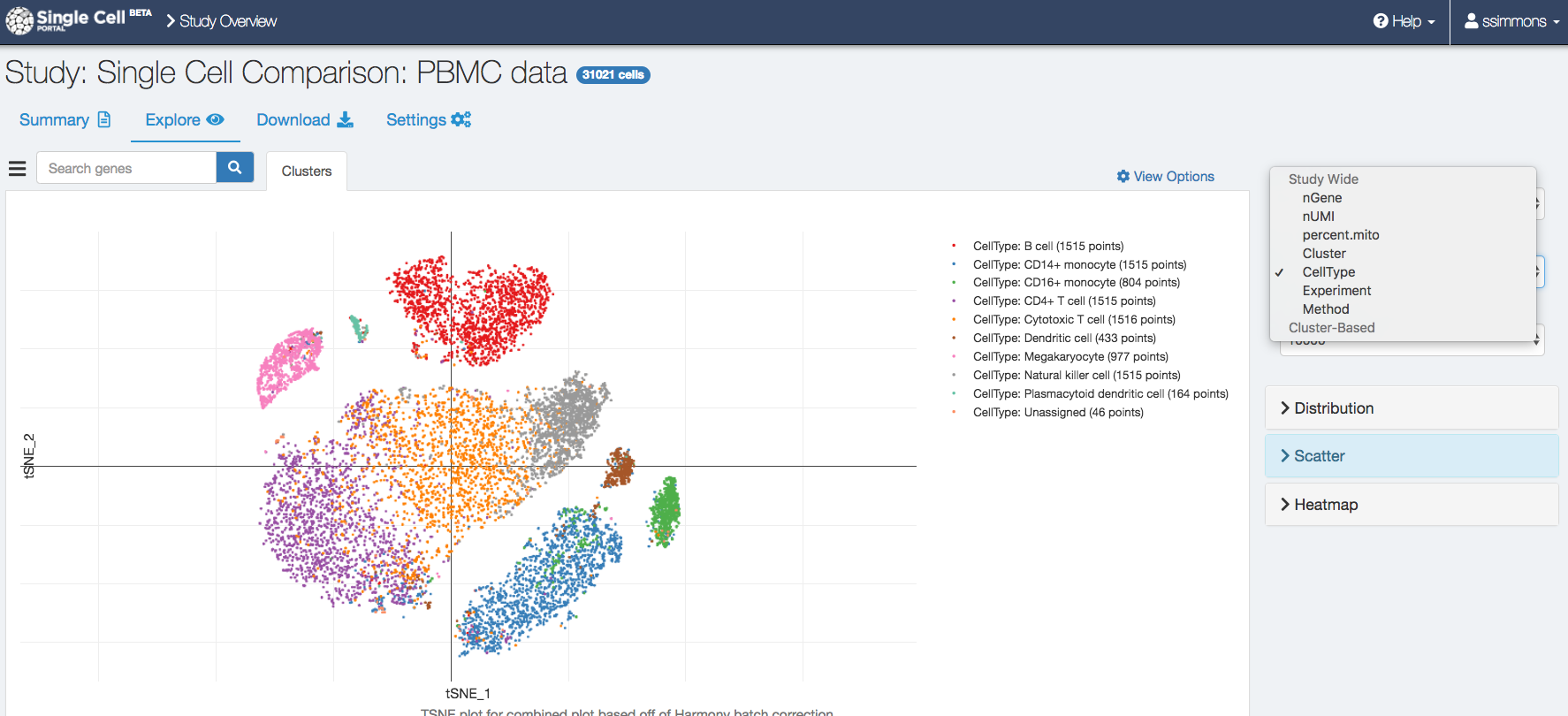
The portal contains log normalized UMI counts per 10,000 data from all these methods, except for Smart-seq2 with log normalized read counts per 10,000 data. Clicking on the explore tab shows a t-SNE plot of all the data combined (using Harmony), colored by cell type (as identified in our automated pipeline).



To view one dataset/method at a time, click the view options tab and choose another option: the sample name (PBMC1 or PBMC2) and followed by method name, which will show the t-SNE calculated on just those cells.



Similarly, one can switch from the select annotation to another annotation you want displayed, including number of genes, number of UMIs (=number of reads for Smart-seq2), batch (PBMC1 vs PBMC2), etc.



One can use the download tab to download the data. This includes the combined log counts per 10,000 matrix (gene\_sorted-Expression.txt in MM format, with barcodes.txt the cell names and genes.txt the gene information), the raw UMI count and Read count data (in compressed MM format, with counts.umi.txt.gz and counts.reads.txt.gz, with the gen names genes.umi.txt and genes.read.txt, and cell names cells.umi.txt and cells.reads.txt), as well as an annotation file describing what experiment/method each cell in the count data comes from (Count.cell.annotation.txt). Various t-SNE plots and the metadata can be downloaded as well.

A full list of files and short descriptions are available under the Download tab.

***Expression data description:***

**Description:** UMI count data, used for QC metrics in paper for UMI based methods, in MM format.

**Matrix:** counts.umi.txt.gz

**Colnames (cells):** cells.umi.new.txt

**Rownames (genes):** genes.umi.txt

**Metadata:** meta.counts.new.txt

**Description:** Read count data, used for Smart-seq2 QC metrics in paper, in MM format.

**Matrix:** counts.read.txt.gz

**Colnames (cells):** cells.read.new.txt

**Rownames (genes):** genes.read.txt

**Metadata:** meta.counts.new.txt

**Description:** Log counts per 10K (reads for Smart-seq2, UMIs for the rest), used for clustering/t-SNE in paper (except for Smart-seq2 which uses TPM from RSEM), used for all methods in the Harmony clustering, and for looking up expression in the portal, in MM format.

**Matrix:** gene\_sorted-Expression.txt

**Colnames (cells):** barcodes.txt

**Rownames (genes):** genes.txt

**Metadata:** meta.new.txt

**Description:** RSEM TPM for PBCM1, used for clustering/t-SNE in paper for Smart-seq2 in PBMC1, stored as tsv.

**Matrix:** RSEM.TPM.SM2.pbmc1.txt.gz

**Description:** RSEM TPM for PBCM2, used for clustering/t-SNE in paper for Smart-seq2 in PBMC2, stored as tsv.

**Matrix:** RSEM.TPM.SM2.pbmc2.txt.gz