**Fitting the topic model**

We used fastTopics [countClust2017, fastTopics2022] to fit a (multinomial) topic model to the (UMI?) counts, with *K* = 16 topics. fastTopics implemennts the following simple strategy to fit the topic model: first, fit a non-negative matrix factorization based on a *Poisson model* (“Poisson NMF”) [hien2021]; second, recover maximum-likelihood estimates (MLEs) of the topic model parameters by a simple reparameterization [fastTopics2022].

In detail, we fit the *K* = 16 topic model to the UMI counts by taking the following steps. First, we removed genes with very low expression (total UMI count ≤ 20). Therefore, UMI counts for 364 samples and 28,209 genes were used to estimate the parameters of the topic model. Second, we ran 20 expectation maximization (EM) updates, without extrapolation, to get close to a MLE solution (“prefitting phase”). This prefitting phase was implemented in R by calling fit\_poisson\_nmf from fastTopics with the following settings: numiter = 20, method = "em", init.method = "random", control = list(nc = 8). Third, we performed an additional 180 coordinate descent (CD) updates, with extrapolation, to improve the fit (“refinement phase”). This refinement phase was implemented by calling fit\_poisson\_nmf with the following settings: method = "scd", numiter = 180, control = list(numiter = 4, nc = 8, extrapolate = TRUE), in which the model fit was initialized using the fit obtained from the prefitting phase. Finally, the topic model was recovered from the Poisson NMF model by calling function poisson2multinom. The convergence diagnostics suggested that, after a total of 200 combined iterations of the Poisson NMF optimization, the parameter estimates were very close to a MLE: the change in log-likelihood between successive iterations was less than 1 × 10-5, and the largest residual in the first-order (“Karush-Kuhn-Tucker”) conditions was less than 1.

Reassuringly, the estimated topics capture the predominant expression patterns, which largely coincide with the 13 tissues. (The one exception was iLN and SP) Two other topics (topics 1 and 6) capture variation specific to two tissues (LI and PBMC), and one topic (topic 9) captures changes in expression over time that are not specific to any one tissue.

*If necessary, we can add details here about fitting topic models separately to each tissue.*

**Visualizing the topic proportions**

The *n* × *K* matrix of topic proportions, *L*, where *n* denotes the number of RNA-seq samples and *K* is the number of topics, can be viewed as an embedding of the samples in a (*K* − 1)- dimensional space. A simple way to visualize this embedding in 2-d (or 3-d) is to apply a nonlinear dimensionality reduction technique such as *t*-SNE [tsne2008] to *L*. A more powerful approach, first suggested by [countClust2017], is to visualize all *K* − 1 dimensions simultaneously using a Structure plot [rosenberg2002]. The Structure plot is essentially a stacker bar chart, in which the bars correspond to samples (rows of *L*) and bar heights (in different colors, one for each topic) are determined by the topic proportions. To arrange the samples in the Structure plot, we first grouped the RNA-seq samples by tissue, then we ordered them within each tissue by time point.

**Differential expression analysis allowing for grades of membership**

To annotate the topics, we used the grade-of-membership differential expression (GoM DE) analysis methods developed in [GoM-DE2022] and implemented in the de\_analysis function in the fastTopics package. In brief, the GoM DE analysis is conceptually similar to standard DE analysis methods, but extends the idea of comparing expression between groups by allowing the cells to have *partial membership* to multiple groups (*i.e.*, topics). We called the de\_analysis function with the following settings: shrink.method = "ash", pseudocount = 0.1 and control = list(ns = 1e5, nc = 8). We also performed a second DE analysis, with the same settings, after merging topics 1 and 5 (capturing expression in PBMC), 2 and 13 (BM) and topics 6 and 14 (LI). The GoM DE analysis produces, for each gene *j* and topic *k*, estimates of expression differences and statistics quantifying support for these differences. Expression differences are defined by the “least extreme” log-fold change (l.e. LFC), which is defined for gene *j* and topic *k* as the log-fold change that is the smallest (in magnitude) among all pairs of topics (*k*, *l*). The GoM DE analysis (with shrink.method = “ash”) performs an adaptive shrinkage step [ash2016] to stabilize the l.e. LFC estimates, separately for each topicc. We use the posterior mean estimates, posterior standard errors, posterior z-scores (posterior mean/posterior s.e.) and local false sign rates (*lfsr*) produced by the adaptive shrinkage step to report results of the GoM DE analysis. Note that the *lfsr* can be interpreted similarly to the *q*-value, or *p*-value adjusted for multiple testing, but the *lfsr* is a more conservative measure [ash2016].

**Computing environment for topic modeling analysis**

Most computations on real data sets were performed in R 3.5.1 [R2018], linked to the OpenBLAS 0.2.19 optimized numerical libraries, on Linux machines (Scientific Linux 7.4) with Intel Xeon E5-2680v4 (“Broadwell”) processors. For performing the Poisson NMF optimization and DE analysis, which included some multithreaded computations, as many as 8 CPUs and 16 GB of memory were used. More details about the computing environment, including the R packages used, are recorded in the workflowr [workflowr2019] pages in the companion code repository [TO DO: create Zenodo repository].

[ash2016]: Stephens, M.: False discovery rates: a new deal. *Biostatistics* 18(2), 041 (2016)

[cistopic2019] González-Blas, C., Minnoye, L., Papasokrati, D., Aibar, S., Hulselmans, G., Christiaens, V., Davie, K., Wouters, J., Aerts, S.: cisTopic: cis-regulatory topic modeling on single-cell ATAC-seq data. *Nature Methods* 16(5), 397–400 (2019)

[countClust2017] Dey, K.K., Hsiao, C.J., Stephens, M.: Visualizing the structure of RNA-seq expression data using grade of membership models. *PLoS Genetics* 13(3), 1006599 (2017)

[fastTopics2022] Carbonetto, P., Sarkar, A., Wang, Z., Stephens, M.: Non-negative matrix factorization algorithms greatly improve topic model fits. *arXiv* 2105.13440 (2021).

[GoM-DE2022] Carbonetto, P., Luo, K., Sarkar, A., Hung, A., Pott, S., Stephens. M.:

Interpreting structure in sequence count data with differential expression analysis allowing for

grades of membership. *Manuscript in preparation.*

[hien2021] Hien, L.T.K., Gillis, N.: Algorithms for nonnegative matrix factorization with the Kullback–Leibler divergence. *Journal of Scientific Computing* 87(3), 93 (2021)

[R] R Core Team: R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2018). R Foundation for Statistical Computing. https://www.R-project.org

[rosenberg2002] Rosenberg, N.A.: Genetic structure of human populations. *Science* 298(5602), 2381–2385 (2002)

[tsne2008] van der Maaten, L., Hinton, G.: Visualizing data using t-SNE. *Journal of Machine Learning Research* 9(86), 2579–2605 (2008)

[workflowr2019] Blischak, J.D., Carbonetto, P., Stephens, M.: Creating and sharing reproducible research code the workflowr way [version 1; peer review: 3 approved]. *F1000Research* 8(1749) (2019)