

A human commons cell atlas reveals cell type specificity for OAS1 isoforms

Ángel Galvez-Merchán^{1*}, A. Sina Boeshaghi^{2*} & Lior Pachter^{3,4+}

1. Cellarity, Somerville, MA, USA
2. Department of Bioengineering, University of California Berkeley, Berkeley, CA, USA
3. Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA
4. Department of Computing & Mathematical Sciences, California Institute of Technology, Pasadena, CA, USA

*These authors contributed equally

+To whom correspondence should be addressed.

Abstract

We describe an open source Human Commons Cell Atlas comprising 2.9 million cells across 27 tissues that can be easily updated and that is structured to facilitate custom analyses. To showcase the flexibility of the atlas, we demonstrate that it can be used to study isoforms of genes at cell resolution. In particular, we study cell type specificity of isoforms of OAS1, which has been shown to offer SARS-CoV-2 protection in certain individuals that display higher expression of the p46 isoform. Using our commons cell atlas we localize the OAS1 p44b isoform to the testis, and find that it is specific to round and elongating spermatids. By virtue of enabling customized analyses via a modular and dynamic atlas structure, the commons cell atlas should be useful for exploratory analyses that are intractable within the rigid framework of current gene-centric cell atlases.

Introduction

The innate immune system plays a crucial role in defending the body against viruses (Takeuchi and Akira 2009; Koyama et al. 2008; Carty, Guy, and Bowie 2021). One of the key components of this system is Oligoadenylate synthetase 1 (OAS1), a protein that gets activated during viral infection through its binding to double-stranded RNA (Melchjorsen et al. 2009). Activated OAS1 produces 2',5'-oligoadenylates, promoting the activity of RNase L and triggering the degradation of cellular and viral RNAs to halt viral replication (Melchjorsen et al. 2009; Hovanessian and Justesen 2007). Importantly, the last exon of the human OAS1 gene undergoes alternative splicing, leading to the production of isoforms with unique C-terminal sequences (Di, Elbahesh, and Brinton 2020). Because of their distinct antiviral activities (Soveg et al. 2021), the differential expression of OAS1 isoforms correlates with susceptibility to certain viruses (Li et al. 2017). For example, the expression of OAS1-p46 has been shown to provide protection against viral infections such as West Nile virus (Lim et al. 2009), Dengue virus (Lin et al. 2009), Hepatitis C virus, and most recently, SARS-CoV-2 (Zhou et al. 2021). Given the clinical relevance, several studies have investigated the regulation of the expression of OAS1 isoforms, with a particular focus on the impact of various SNPs on the relative isoform abundance (Li et al. 2017). However, no study has explored whether this regulation is organ or cell-type specific. This question is significant: an OAS1 isoform can only protect against a virus if it is expressed in the cell-type the virus infects, and therefore an understanding of cell type specificity of OAS1 isoforms must accompany an analysis of viral tropism.

A single cell atlas is the ideal tool to study cell-type specific OAS1 isoform regulation. There are a number of available cell atlases that together contain data from most human organs, such as the Human Cell Atlas (Li et al. 2017; Rozenblatt-Rosen et al. 2017), the adult human cell atlas (S. He et al. 2020), Tabula Sapiens (Consortium* et al. 2022), the Human Cell Landscape (Han et al. 2020), Descartes (Cao et al. 2020), and Azimuth (Hao et al. 2021). However, our attempts to use these atlases for our research question uncovered three important limitations. First, current atlases exclusively provide gene-level expression data, and lack information on isoforms. This deficiency is not minor, as evidenced by the growing body of literature supporting the critical role of isoform expression in major biological processes (E. T. Wang et al. 2008; Chaponnier and Gabbiani 2004; Warren et al. 2003; Mincarelli et al. 2023), as well as in the identification and definition of cell-types (Booeshaghi et al. 2021). Second, current cell atlas projects depend on assay-specific preprocessing tools that can introduce computationally-induced batch effects that can be challenging to identify and correct. Finally, cell atlases are static objects that cannot be easily updated or re-processed to facilitate interpretation of the stream of new findings emerging from single-cell studies. Our atlas and the infrastructure it is built on address these limitations, and we demonstrate its potential by showing how it can shed light on OAS1 isoform expression at single cell resolution.

Results

Building the Humans Commons Cell Atlas

To study the cell-type specificity of OAS1 isoforms in humans, we created a human cell atlas using our recently developed Commons Cell Atlas (CCA) infrastructure. The Human CCA comprises over 2.9 million cells from 525 publicly available scRNA-seq datasets across 27 organs (**Figure 1A**).

We started by compiling a list of publicly available single-cell RNA-seq datasets deposited across GEO/SRA/ENA/DDBJ (Barrett et al. 2011; Leinonen, Sugawara, et al. 2011; Leinonen, Akhtar, et al. 2011; Ogasawara et al. 2020). This collection of data consisted of 147 billion sequencing reads (20.74 TB of FASTQs), **Supplementary Table 1**. We chose to start with raw FASTQs, instead of the gene count matrices, which is crucial for 1. ensuring a uniform read alignment strategy and barcode error correction, and 2. enabling isoform quantification, which is lost when counts are aggregated at the gene level. Data and metadata were downloaded and organized by “observation” which maps to GEO sample accessions (GSM) that group a set of FASTQ files with metadata.

Atlas building requires uniform processing and sequencing read quantification to minimize computational variability and apply concordant read alignment, barcode error correction, and read counting. To that end, we leveraged the recently developed kallisto bustools (kb-python) programs (Melsted et al. 2021; Sullivan et al. 2023) to generate all 525 gene count matrices. From the raw sequencing data, the Human Commons Cell atlas was built in about two weeks (305 hours) (**Figure S1**) with less than 8GB of memory usage. Reads were pseudoaligned to the human transcriptome, cell barcodes were corrected within hamming-1 distance of a barcode “onlist”, and naïve UMI collapsing was performed to generate gene counts (see Methods, **Supplementary Knee**).

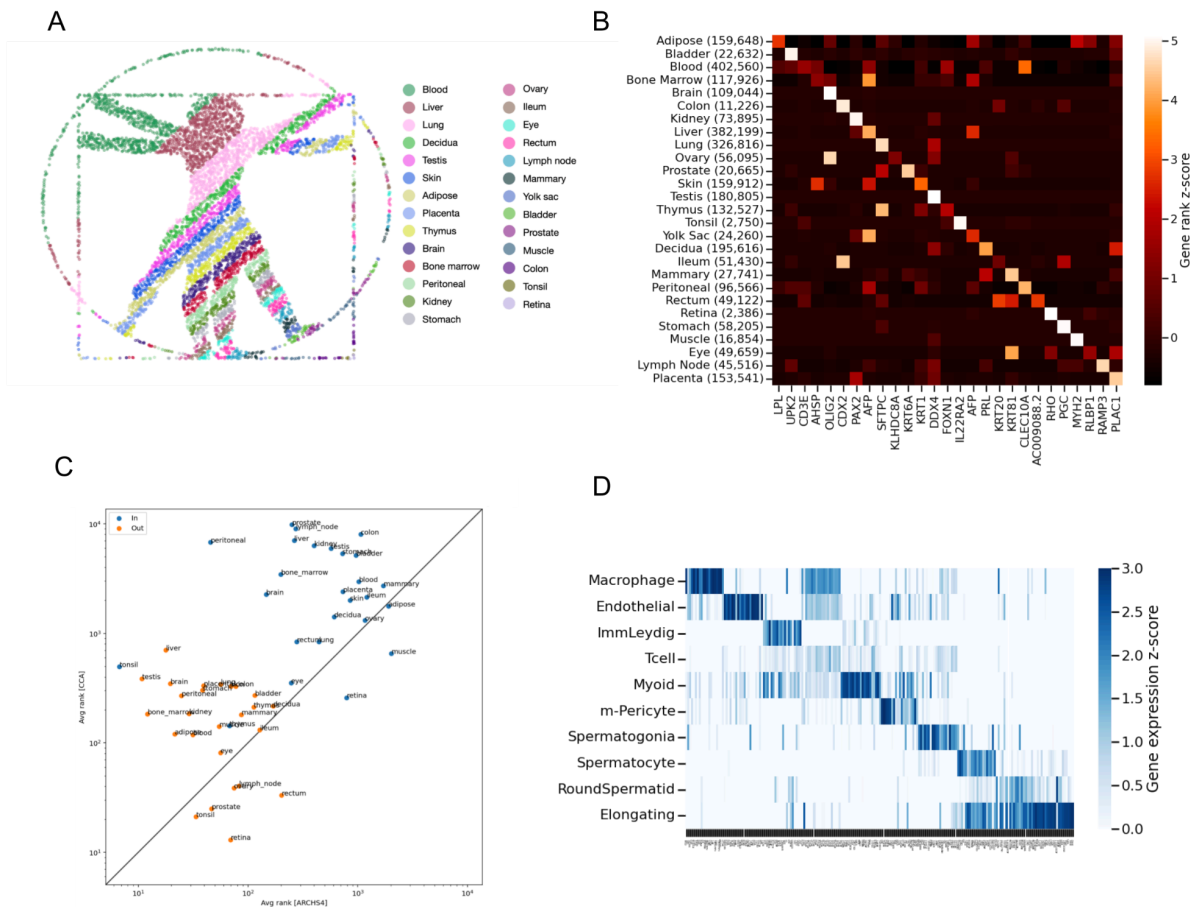


Figure 1. The Human Commons Cell Atlas. (A) 2D representation of the atlas. Cells were downsampled to match the number of coordinates of the Vitruvian man maintaining the original proportions by organ. (B) Heatmap displaying the z-score of the average rank of select organ marker genes calculated across organs. Elements along the diagonal indicate genes (columns) that mark major cell-types in the organ (rows). (C) Correlation between ranks of marker genes in the CCA and ARCHS4 computed on organ-specific marker genes (in, blue) and non-organ specific marker genes (out, red). (D) Z-score of average expression of testis marker genes by cell-type.

Gene-level Human CCA

In order to study the expression of OAS1 we first sought to assess the suitability and robustness of the Human CCA for computationally identifying tissue-level marker genes. We first performed differential expression between 28 tissues on the rank of all genes and identified tissue-level markers from the list of DE genes. We identified genes that are highly and specifically expressed in tissues, most of them representing bona-fide markers for each organ. The list of genes include the LPL gene (encoding Lipoprotein Lipase) in adipose tissue (Zechner et al. 2000), the RLBP1 gene (encoding Retinaldehyde Binding Protein 1) in eye (Eichers et al. 2002), and the SFPTC gene (encoding pulmonary-associated surfactant protein C) in lung (Tredano et

al. 2004) (**Figure 1B**). These findings serve as a positive control on the atlas's ability to join tissue-level groupings and prior known marker genes.

We then assessed the concordance of gene-level pseudobulk quantifications produced with CCA to bulk gene-level bulk quantifications in ARCHS4 (Lachmann et al. 2018) across 27 organs. We found a higher concordance of ranked- pseudobulk profiles for those genes that mark organs, than those that do not (Methods). This demonstrates that pseudobulk qualifications for organ-specific marker genes computed from the CCA are consistent with those found in ARCHS4 (**Figure 1C**).

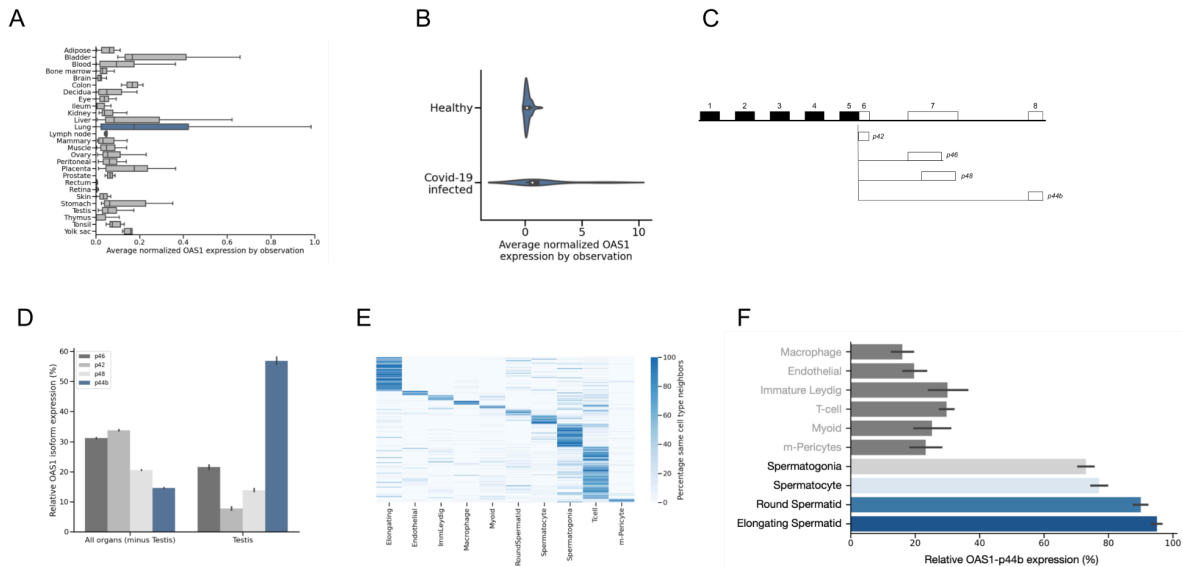


Figure 2. Cell-type specificity of OAS1 isoforms. (A) Normalized OAS1 gene expression by organ. Results for the lung, the organ with highest expression, are highlighted in blue. (B) Normalized OAS1 expression in Lung datasets by health status of the individual. (C) Diagram of the 4 main OAS1 isoforms we found in our data. (D) Relative OAS1 isoform expression in all organs except testis, and testis. OAS1-p44b is highly and specifically expressed in testis. (E) Validation of testis cell-type assignments. The clusters indicate that cells from the same cell-type are frequently neighbors in the K-Nearest Neighbor Graph. (F) OAS1-p44b is the main OAS1 isoform in spermatogenesis cells.

We then sought to identify if the OAS1 gene exhibits tissue-level specificity. As expected given its crucial function, OAS1 was detected in most tissues, with no significant enrichment in any of them (**Figure 2A**). Notably, OAS1 was upregulated in lung samples from COVID-19 infected individuals, which is consistent with OAS1 being a type I interferon (IFN)-induced gene (Melchjorsen et al. 2009) (**Figure 2B**).

Isoform-level Human CCA

The expression of isoforms within genes can vary greatly, even when the overall gene expression remains unchanged (Booeshaghi et al. 2021). This information is lost in currently published atlases, which fail to quantify transcript isoforms. We hypothesized that OAS1

isoforms exhibited tissue-level specificity. To test this hypothesis, we rebuilt our atlas at the isoform level leveraging transcript compatibility counts (See Methods) (Booeshaghi et al. 2021; Ntranos et al. 2019). Since most of the publicly available single-cell data derives from 3' technologies, our quantification was limited to isoforms with distinct 3' ends.

OAS1 isoform cell-type specificity

We leveraged the unique 3'UTR of OAS1 isoforms to study their differential expression (**Figure 2C**). We found that, out of the 11 OAS1 isoforms annotated in the human transcriptome, only 4 of them were significantly expressed across the atlas: p46, p42, p48 and p44-b (**Figure S2**). We failed at detecting cell-type specificity of the protective isoform p46 (Zhou et al. 2021), which exhibited broad expression across all organs and was among the most highly expressed isoforms alongside p42. The least expressed isoform, p44-b, was responsible for only around 10% of the total OAS1 expression. This is in line with numerous studies consistently showing that amplifying the p44-b isoform by qPCR is difficult due to its very low or undetectable expression level (Noguchi et al. 2013; Iida et al. 2021). Interestingly, we found that this trend was reversed in testis, with p44b accounting for almost 60% of total OAS1 expression and being the predominant isoform (**Figure 2D**).

To investigate if OAS1-p44b expression was cell-type specific, we assigned cell-types using `mx assign`. We validated the output of `mx assign` by calculating, for each cell, the percentage of cells belonging to the same cell-type within their 20-nearest neighbors (Z. Zhang 2016). Cells from the same cell-type clustered together, validating our assignment results (**Figure 2E**). We observed that the high OAS1-p44b expression was specific to germ cells undergoing spermatogenesis, where p44b represented over 80% of total OAS1 expression (**Figure 2F**). To discard any possible artifacts caused by pseudoalignment, we visualized the alignments of one of the testis samples (GSM3302525) and observed high density of reads mapping to OAS1's exon 8, which is unique to the isoform p44b (**Figure S3**). This result was not sample or paper-dependent, with Round Spermatids across all testis samples expressing high levels of OAS1-p44b (**Figure S4**).

Screen for cell-type specific isoform switching

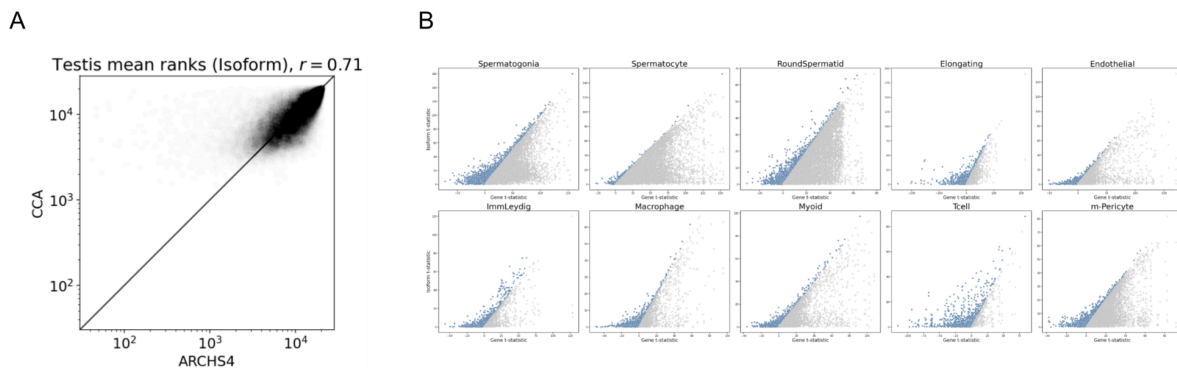


Figure 3. Quantifying isoforms with distinct 3' UTRs. (A) Scatterplot of isoform ranks in the testis as measured according to CCA and ARCHS4. (B) Examples of cell-type specific isoform switches.

To confirm that the atlas had accurate isoform information, compared isoforms quantified with ARCHS4 to those quantified with CCA that: i) derived from genes with more than one isoform, ii) had reads in testis samples that mapped uniquely to it and iii) had a minimum average

normalized expression of 0.002 per cell. 20,768 isoforms passed these filters in testis, and were therefore amenable to quantification. We performed a bulk isoform-to-isoform comparison between ARCHS4 and CCA (Methods) for cells assayed from the testis and found a Spearman correlation of 0.71 pointing to concordance of pseudobulk single-cell profiles to bulk profiles (**Figure 3A**).

After confirming that the human CCA can accurately quantify isoforms with distinct 3' UTRs, we decided to test how many isoforms were amenable to this analysis in our data. Approximately 27% of genes with more than one isoform had at least 80% of their counts in single equivalence classes (**Figure S5**), suggesting that for a relatively large fraction of genes, isoform quantifications can be meaningfully utilized. Indeed, performing differential expression of both isoforms and genes in the cell-types of the testis revealed a number of isoforms whose DE signal was stronger than that of the gene it belonged to (**Figure 3B**). We decided to leverage this additional layer of information and screen for cell-type specific isoform switching within testis. We found genes with differential isoform usage in all cell-types except for T-cells. Some of the hits, such as IFT27 in Spermatocytes (Y. Zhang et al. 2017), have been shown to have essential roles within the testis (**Figure S6**). Our results suggest that some of these roles may be carried out by specific isoforms, and that isoform switching may be an important regulatory mechanism in a number biological processes.

Discussion

The question of how to organize a single-cell atlas is complex, and there is little agreement on even the most basic questions, such as what constitutes a “cell-type”. In addition to conceptual problems, engineering challenges abound. Single-cell genomics datasets are growing in both number and size, and it is non-trivial to engineer atlases so that they can be updated when new datasets or biology are discovered. As we have demonstrated, the Commons Cell Atlas concept offers a solution to many of these problems by virtue of reframing single-cell atlases as a dynamic collection of data and, crucially, tools for processing, querying and interpreting the data. Our isoform analysis was possible thanks to this design principle; in order to obtain isoform quantifications we needed to reprocess the data several times. Other questions may demand alternative atlas processing that, with the Commons Cell Atlas architecture, should be tractable to implement. Most importantly, the Commons Cell Atlas principle dictates that there is no definitive Commons Cell Atlas, but rather numerous Commons Cell Atlases that are customized and specific to the questions being explored.

One of the main aims of cell atlases is to provide a comprehensive characterization and classification of cells (Rozenblatt-Rosen et al. 2017), which relies heavily on identifying their cell-type. The cell-type assignments within the Commons Cell Atlas can be constantly updated, thereby enabling continual refinement of the derived results as new information emerges. To show that this is a feasible strategy, we obtained all marker genes from single cell publications whose data was not used to build the atlas. Despite data and marker genes coming from different sources, we were able to successfully assign cells from most tissues in the atlas (**Figure S7**), as measured by the percentage of same-cell-type neighbors (as shown in **Figure 2** for testis). We expect that these assignments will change as we learn more about each tissue, enabled by the dynamism and efficiency of the Common Cell Atlas infrastructure.

The pivotal function of OAS1 in innate immunity has been well-established, and recent studies have demonstrated that differential expression of its isoforms can affect susceptibility to viruses, including SARS-CoV-2 (Zhou et al. 2021). The work described here enables the study of this differential expression across organs and cell-types, providing a valuable resource for our

understanding of viral immunity. We have used our Human Commons Cell Atlas to discover previously unknown cell-type specificity for an OAS1 isoform. Interestingly, this isoform was deemed undetectable across many studies, and our atlas provides an explanation for these results. Our finding is preliminary, and a comprehensive assessment of p44b function and activity in the testis is beyond the scope of our paper. But our discovery highlights the utility of a Commons Cell Atlas, and more generally, points towards the importance of carefully assessing isoform cell-type specificity.

The Human Commons Cell Atlas is composed of 27 tissues, 526 cell-types and 3,554 marker genes. The whole atlas is hosted on GitHub, and it is therefore readily available to download, inspect, modify and use. We envision that the Human Commons Cell Atlas will evolve as we increase our knowledge on tissues and cell-types, moving away from the idea of achieving a "final" or "complete" atlas. For as long as we continue discovering new cell-types, developing new single cell technologies, and gathering new single cell data, the Human Commons Cell Atlas and all its derived results can continue to grow and improve.

Acknowledgments

We thank Tara Chari for assistance in producing **Figure 1A**. The authors acknowledge the Howard Hughes Medical Institute for funding A.S.B. through the Hanna H. Gray Fellows program. Thanks to the Caltech Bioinformatics Resource Center for assisting with pre-processing the data.

Methods

Downloading data

Accession ids for scRNA-seq datasets were obtained using (Svensson, da Veiga Beltrame, and Pachter 2020). Metadata and links to raw data were collected using the `ffq` program (Gálvez-Merchán et al. 2022) version 0.2.1 (available at <https://github.com/pachterlab/ffq>) by running '`ffq DATASETID`' where DATASETID is the sample ID (GSM, SRS, etc.).

Preprocessing data

Reads from each dataset were pseudoaligned to the human transcriptome, which was obtained by running '`kb ref -i index.idx -g t2g.txt -d human`'. Reads were uniformly processed using the `kallisto | bustools` python wrapper `kb-python`, running the command '`kb count -i index.idx -g t2g.txt -x [technology]`'. The datasets used in the atlas are listed below:

GSE68596 (Kind et al. 2015),
GSE130636 (Voigt et al. 2019),
GSE96583 (Kang et al. 2018)
GSE112013 (Guo et al. 2022)
GSE130228 (Xu et al. 2022)
GSE130973 (Solé-Boldo et al. 2020)
GSE129363 (Vijay et al. 2020)
SRP126175 (Sivakamasundari et al. 2017)
GSE131391 (Duclos et al. 2019)
GSE125527 (Boland et al. 2020)
GSE165860 (King et al. 2021)
GSE143704 (Fitzgerald et al. 2023)
GSE114297 (Sui et al. 2021)
GSE148963 (Leir et al. 2020)
GSE129363 (Vijay et al. 2020)

GSE131685 (Su et al. 2021)
GSE157526 (Fiege et al. 2021)
PRJCA002413 (Wen et al. 2020)
GSE128889 (Merrick et al. 2019)
GSE129845 (Yu et al. 2019)
GSE104600 (Das et al. 2018)
GSE107747 (D'Avola et al. 2018)
GSE111014 (Rendeiro et al. 2020)
GSE115189 (Freytag et al. 2018)
GSE119594 (Schiroli et al. 2019)
GSE124334 (Galinato et al. 2019)
GSE124898 (Borcherding et al. 2023)
GSE128066 (Sun et al. 2019)
GSE130117 (Richer et al. 2020)
GSE130430 (Yang et al. 2019)
GSE117824 (Nam et al. 2019)
PRJNA610059 (Turner et al. 2020)
GSE115149 (Jitschin et al. 2019)
GSE120446 (Oetjen et al. 2018)
GSE130430 (Yang et al. 2019)
GSE108571: No associated publication
GSE119212 (Zhong et al. 2020)
GSE130238 (Trujillo et al. 2019)
EMTAB6701 (Vento-Tormo et al. 2018)
GSE145926 (Liao et al. 2020)
GSE146188 (van Zyl et al. 2020)
GSE112570 (Lindström et al. 2018)
GSE114530 (Mircea et al. 2022)
GSE117211 (Kumar et al. 2019)
GSE119561 (Howden et al. 2019)
GSE130073 (Ouchi et al. 2019)
GSE103918 (McCauley et al. 2018)
GSE102592 (C. Wang et al. 10 2018)
GSE121600 (Ruiz García et al. 2019)
GSE135893 (Habermann et al. 2020)
GSE134174 (Goldfarbmuren et al. 2020)
GSE124494 (Takeda et al. 2019)
GSE11472 (Azizi et al. 2018)
GSE123926 (Merino et al. 2019)
GSE118127 (Fan et al. 2019)
GSE130888 (Kind et al. 2015)
PRJNA492324: No associated publication
GSE130318: (McCray et al. 2019)
GSE128066 (Sun et al. 2019)
GSE109037 (Hermann et al. 2018)
GSE124263 (Sohni et al. 2019)
GSE130151 (Laurentino et al. 2019)
GSE119506 (Durand et al. 2019)
GSE139522 (Durante et al. 2020)
EMTAB7407 (Popescu et al. 2019)
GSE120508 (Guo et al. 2022)
GSE115469 (Sigiel et al. 1978)
EMTAB8581 (Park et al. 2020)
GSE125970 (Y. Wang et al. 2020)

Filtering matrices

To filter the gene count matrices, barcodes with low UMIs were filtered with '*mx filter*', which uses a derivation of the knee plot approach. Barcodes with more than 40% mitochondrial genes were discarded.

Normalizing matrices

Gene count matrices were normalized using '*mx norm*', which uses the log1pPF method (Sina Boeshaghi et al. 2022).

Marker genes curation

A list of marker genes for each organ was generated from supplementary tables of single cell publications containing differential expression information. Markers were selected applying filters to the corrected p-value, log fold change, and percentage of cells expressing the gene. A Google Colab notebook that downloads the supplementary table, filters it, and generates a markers file is available at the Human Cell Atlas repository for each organ. We curated markers for the following tissues:

Adipose: (Emont et al. 2022)
Bladder: (Yu et al. 2019)
Blood:(Jiang et al. 2023)
Bone: (J. He et al. 2021)
Bone marrow: (Jiang et al. 2023)
Brain: (Jiang et al. 2023)
Colon:(Jiang et al. 2023)
Decidua: (Jiang et al. 2023)
Eye: (Gautam et al. 2021)
Heart: (Tucker et al. 2020)
Ileum: (Tucker et al. 2020)
Kidney: (Wu et al. 2018)
Liver: (Guilliams et al. 2022)
Lung: (Adams et al. 2020)
Lymph node: (Jiang et al. 2023)
Mammary: (Jiang et al. 2023)
Muscle: (Jiang et al. 2023)
Ovary: (Wagner et al. 2020)
Pancreas: (Segerstolpe et al. 2016)
Peritoneal: (Jiang et al. 2023)
Placenta: (Liu et al. 2018)
Prostate: (Jiang et al. 2023)
Rectum: (Jiang et al. 2023)
Retina: (Jiang et al. 2023)
Skin: (Jiang et al. 2023)
Stomach: (Jiang et al. 2023)
Testis: (Shami et al. 2020)
Thymus: (Jiang et al. 2023)
Tonsil: (Jiang et al. 2023)
Uterus: (Legeth et al. 2021)
Yolk sac: (Jiang et al. 2023)

Cell-type assignment and validation

Cell-types were assigned running `mx assign` on each individual dataset using the marker gene file generated above. Assignments were validated by calculating, for each cell, the percentage of cells that belong to the same cell-type in the k-nearest neighbors (KNN) graph, with $k=20$, within each dataset. The KNN graph was calculated using the normalized expression values of the union of marker genes of the corresponding organ.

2D representation of the atlas

To create a 2D latent space in which cells from the same organ are neighbors (**Figure 1A**), we used MCML (Chari and Pachter 2023) with the `fracNCA` parameter set to 1 (this is, optimizing only the Neighborhood Component Analysis (NCA) loss). We then calculated the pairwise distances of each cell's 2D coordinates to the Vitruvian man 2D coordinates, using the `L_1` norm or Manhattan distance. The distances were used as input to the `scipy 'linear_sum_assignment'` function to map the 2D latent space to the 2D shape coordinates, assigning each cell coordinate to a shape coordinate while minimizing the total cost or distance (as per the distance matrix).

Organ-level markers

For each dataset, we calculated the average expression of each gene across all cells, and used that value to rank each gene in the dataset. The gene ranks of datasets from the same organ were averaged, resulting in an organ x gene matrix. Genes with high value in one organ and low in the others were selected, and the Z-scores across organs were plotted using a heatmap.

OAS1 gene expression by tissue

The average normalized OAS1 expression for each dataset was calculated across all cells. Each data point in **Figure 2A,B** corresponds to a different dataset.

Isoform quantification

A transcript Compatibility Counts (TCC) matrix for each sample was obtained by running `'bustools count'` without the `'--genecount'` option on the bus files generated after pseudoaligning the raw reads. Transcript abundances were quantified using the EM algorithm by running `'kallisto quant-tcc'` on the TCC matrices. The transcript abundance matrix of each sample was normalized within each cell-type using $\log_1 pPF$. The normalized matrix was then subsetted to isoforms that i) derived from genes with more than one isoform, ii) had reads in the samples that mapped uniquely to it and iii) had a minimum average normalized expression of 0.002 per cell.

Isoforms screen

We performed t-tests of the expression of isoforms and genes in each cell-type vs all other cells. Isoforms with higher t-statistics than its corresponding gene were interpreted as containing isoform-specific cell-type signal.

Percentage of genes with at least 80% counts in single equivalence classes

The raw TCC matrix of the testis sample GSM2928378 was filtered to counts that mapped to a single equivalence class, which were averaged across all cells. A isoform-to-gene ratio was calculated by dividing the result of each isoform by the averaged raw counts of the corresponding gene expression. The number of isoforms with a ratio higher than 0.8 was calculated.

CONCORDEX ratio

The metric was calculated as previously shown (Jackson et al. 2023) for each sample of the human CCA where gene markers were available.

ARCHS4 Gene Comparison

ARCHS4 quantifications were ranked across genes for each observation and the mean rank was computed across all observations. CCA pseudobulk profiles were computed across all cells in an observation for each organ and genes were ranked for each observation. The mean rank was computed across all observations for each organ. Ranks for marker genes and non-marker genes in CCA and ARCHS4 were plotted for each organ.

ARCHS4 Isoform Comparison

We first took ARCHS4 quantifications in testis across 695 samples and filtered out samples with less than 1e7 counts. We then ranked all of the isoforms for each sample across the remaining 18 samples and took the average rank across the samples. We performed a similar computation with all 25 samples from the testis for the CCA data but first generated pseudobulk isoform profiles for each sample from single cells. Then the isoforms were ranked for each sample and the mean taken across all samples. A Spearman correlation was computed.

Data and code availability

The code and data needed to reprocess the results of this manuscript can be found here https://github.com/pachterlab/GBP_2024/. The CCA GitHub can be found here <https://github.com/cellatlas/human/>. A summary of the datasets in the CCA can be found here <https://cellatlas.github.io/human/>.

Author contributions

The CCA atlas concept emerged from an initiative by ASB to uniformly preprocess the datasets in (Svensson, da Veiga Beltrame, and Pachter 2020). ÁGM conceived of the idea of examining the OAS1 isoforms at single-cell resolution across human tissues after the publication of (Zhou et al. 2021). ASB conceived the CCA structure and associated mx toolkit. ÁGM pre-processed the CCA datasets, and ASB wrote mx. ÁGM and ASB developed the CCA quality control. ÁGM led the OAS1 analysis, with help from ASB and LP. ASB developed the `mx assign` cell assignment approach, and ÁGM and ASB benchmarked it. ÁGM drafted the initial version of the manuscript, which was edited and reviewed by all authors.

Competing Interests

None.

References

- Adams, Taylor S., Jonas C. Schupp, Sergio Poli, Ehab A. Ayaub, Nir Neumark, Farida Ahangari, Sarah G. Chu, et al. 2020. "Single-Cell RNA-Seq Reveals Ectopic and Aberrant Lung-Resident Cell Populations in Idiopathic Pulmonary Fibrosis." *Science Advances* 6 (28): eaba1983.
- Azizi, Elham, Ambrose J. Carr, George Plitas, Andrew E. Cornish, Catherine Konopacki, Sandhya Prabhakaran, Juozas Nainys, et al. 2018. "Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment." *Cell* 174 (5): 1293–1308.e36.
- Barrett, Tanya, Dennis B. Troup, Stephen E. Wilhite, Pierre Ledoux, Carlos Evangelista, Irene F. Kim, Maxim Tomashevsky, et al. 2011. "NCBI GEO: Archive for Functional Genomics Data Sets--10 Years on." *Nucleic Acids Research* 39 (Database issue): D1005–10.
- Boland, Brigid S., Zhaoren He, Matthew S. Tsai, Jocelyn G. Olvera, Kyla D. Omilusik, Han G. Duong, Eleanor S. Kim, et al. 2020. "Heterogeneity and Clonal Relationships of Adaptive Immune Cells in Ulcerative Colitis Revealed by Single-Cell Analyses." *Science Immunology* 5 (50). <https://doi.org/10.1126/sciimmunol.abb4432>.
- Booeshaghi, A. Sina, Zizhen Yao, Cindy van Velthoven, Kimberly Smith, Bosiljka Tasic, Hongkui Zeng, and Lior Pachter. 2021. "Isoform Cell-Type Specificity in the Mouse Primary Motor Cortex." *Nature* 598 (7879): 195–99.
- Borcherding, Nicholas, Kevin J. Severson, Nicholas Henderson, Luana S. Ortolan, Allison C. Rosenthal, Andrew M. Bellizzi, Vincent Liu, Brian K. Link, Aaron R. Mangold, and Ali Jabbari. 2023. "Single-Cell Analysis of Sézary Syndrome Reveals Novel Markers and Shifting Gene Profiles Associated with Treatment." *Blood Advances* 7 (3): 321–35.
- Cao, Junyue, Diana R. O'Day, Hannah A. Pliner, Paul D. Kingsley, Mei Deng, Riza M. Daza, Michael A. Zager, et al. 2020. "A Human Cell Atlas of Fetal Gene Expression." *Science* 370 (6518). <https://doi.org/10.1126/science.aba7721>.
- Carty, Michael, Coralie Guy, and Andrew G. Bowie. 2021. "Detection of Viral Infections by Innate Immunity." *Biochemical Pharmacology* 183 (January): 114316.
- Chaponnier, Christine, and Giulio Gabbiani. 2004. "Pathological Situations Characterized by Altered Actin Isoform Expression." *The Journal of Pathology* 204 (4): 386–95.
- Chari, Tara, and Lior Pachter. 2023. "The Specious Art of Single-Cell Genomics." *PLoS Computational Biology* 19 (8): e1011288.
- Consortium*, The Tabula Sapiens, The Tabula Sapiens Consortium*, Robert C. Jones, Jim Karkanas, Mark A. Krasnow, Angela Oliveira Pisco, Stephen R. Quake, et al. 2022. "The Tabula Sapiens: A Multiple-Organ, Single-Cell Transcriptomic Atlas of Humans." *Science*. <https://doi.org/10.1126/science.abl4896>.
- Das, Rituparna, Noffar Bar, Michelle Ferreira, Aaron M. Newman, Lin Zhang, Jithendra Kini Bailur, Antonella Bacchiocchi, et al. 2018. "Early B Cell Changes Predict Autoimmunity Following Combination Immune Checkpoint Blockade." *The Journal of Clinical Investigation* 128 (2): 715–20.
- D'Avola, Delia, Carlos Villacorta-Martin, Sebastiao N. Martins-Filho, Amanda Craig, Ismail Labгаа, Johann von Felden, Allette Kimaada, et al. 2018. "High-Density Single Cell mRNA Sequencing to Characterize Circulating Tumor Cells in Hepatocellular Carcinoma." *Scientific Reports* 8 (1): 11570.
- Di, Han, Husni Elbahesh, and Margo A. Brinton. 2020. "Characteristics of Human OAS1 Isoform Proteins." *Viruses* 12 (2). <https://doi.org/10.3390/v12020152>.
- Duclos, Grant E., Vitor H. Teixeira, Patrick Autissier, Yaron B. Gesthalter, Marjan A. Reinders-Luinge, Robert Terrano, Yves M. Dumas, et al. 2019. "Characterizing Smoking-Induced Transcriptional Heterogeneity in the Human Bronchial Epithelium at Single-Cell Resolution." *Science Advances* 5 (12): eaaw3413.
- Durand, Mélanie, Thomas Walter, Tiphène Pirnay, Thomas Naessens, Paul Gueguen, Christel

- Goudot, Sonia Lameiras, et al. 2019. "Human Lymphoid Organ cDC2 and Macrophages Play Complementary Roles in T Follicular Helper Responses." *The Journal of Experimental Medicine* 216 (7): 1561–81.
- Durante, Michael A., Stefan Kurtenbach, Zoukaa B. Sargi, J. William Harbour, Rhea Choi, Sarah Kurtenbach, Garrett M. Goss, Hiroaki Matsunami, and Bradley J. Goldstein. 2020. "Single-Cell Analysis of Olfactory Neurogenesis and Differentiation in Adult Humans." *Nature Neuroscience* 23 (3): 323–26.
- Eichers, Erica R., Jane S. Green, David W. Stockton, Christopher S. Jackman, James Whelan, J. Arch McNamara, Gordon J. Johnson, James R. Lupski, and Nicholas Katsanis. 2002. "Newfoundland Rod-Cone Dystrophy, an Early-Onset Retinal Dystrophy, Is Caused by Splice-Junction Mutations in RLBP1." *American Journal of Human Genetics* 70 (4): 955–64.
- Emont, Margo P., Christopher Jacobs, Adam L. Essene, Deepti Pant, Danielle Tenen, Georgia Colletuori, Angelica Di Vincenzo, et al. 2022. "A Single-Cell Atlas of Human and Mouse White Adipose Tissue." *Nature* 603 (7903): 926–33.
- Fan, X., M. Bialecka, I. Moustakas, E. Lam, V. Torrens-Juaneda, N. V. Borggreven, L. Trouw, et al. 2019. "Single-Cell Reconstruction of Follicular Remodeling in the Human Adult Ovary." *Nature Communications* 10 (1): 3164.
- Fiege, Jessica K., Joshua M. Thiede, Hezkiel Arya Nanda, William E. Matchett, Patrick J. Moore, Noe Rico Montanari, Beth K. Thielen, et al. 2021. "Single Cell Resolution of SARS-CoV-2 Tropism, Antiviral Responses, and Susceptibility to Therapies in Primary Human Airway Epithelium." *PLoS Pathogens* 17 (1): e1009292.
- Fitzgerald, Gillian, Guillermo Turiel, Tatiane Gorski, Inés Soro-Arnaiz, Jing Zhang, Nicola C. Casartelli, Evi Masschelein, et al. 2023. "MME+ Fibro-Adipogenic Progenitors Are the Dominant Adipogenic Population during Fatty Infiltration in Human Skeletal Muscle." *Communications Biology* 6 (1): 111.
- Freytag, Saskia, Luyi Tian, Ingrid Lönnstedt, Milica Ng, and Melanie Bahlo. 2018. "Comparison of Clustering Tools in R for Medium-Sized 10x Genomics Single-Cell RNA-Sequencing Data." *F1000Research* 7 (August): 1297.
- Galinato, Melissa, Kristen Shimoda, Alexis Aguiar, Fiona Hennig, Dario Boffelli, Michael A. McVoy, and Laura Hertel. 2019. "Single-Cell Transcriptome Analysis of CD34+ Stem Cell-Derived Myeloid Cells Infected With Human Cytomegalovirus." *Frontiers in Microbiology* 10 (March): 577.
- Gálvez-Merchán, Ángel, Kyung Hoi (Joseph) Min, Lior Pachter, and A. Sina Boeshaghi. 2022. "Metadata Retrieval from Sequence Databases with Ffq." *bioRxiv*. <https://doi.org/10.1101/2022.05.18.492548>.
- Gautam, Pradeep, Kiyofumi Hamashima, Ying Chen, Yingying Zeng, Bar Makovoz, Bhav Harshad Parikh, Hsin Yee Lee, et al. 2021. "Multi-Species Single-Cell Transcriptomic Analysis of Ocular Compartment Regulons." *Nature Communications* 12 (1): 5675.
- Goldfarbmuren, Katherine C., Nathan D. Jackson, Satria P. Sajuthi, Nathan Dyjack, Katie S. Li, Cydney L. Rios, Elizabeth G. Plender, et al. 2020. "Dissecting the Cellular Specificity of Smoking Effects and Reconstructing Lineages in the Human Airway Epithelium." *Nature Communications* 11 (1): 2485.
- Guilliams, Martin, Johnny Bonnardel, Birthe Haest, Bart Vanderborght, Camille Wagner, Anneleen Remmerie, Anna Bujko, et al. 2022. "Spatial Proteogenomics Reveals Distinct and Evolutionarily Conserved Hepatic Macrophage Niches." *Cell* 185 (2): 379–96.e38.
- Guo, Jie, Shuang Wang, Zhenzhen Jiang, Le Tang, Zhizhong Liu, Jian Cao, Zhaolan Hu, Xiao Chen, Yanwei Luo, and Hao Bo. 2022. "Long Non-Coding RNA RFPL3S Functions as a Biomarker of Prognostic and Immunotherapeutic Prediction in Testicular Germ Cell Tumor." *Frontiers in Immunology* 13 (May): 859730.
- Habermann, Arun C., Austin J. Gutierrez, Linh T. Bui, Stephanie L. Yahn, Nichelle I. Winters, Carla L. Calvi, Lance Peter, et al. 2020. "Single-Cell RNA Sequencing Reveals Profibrotic

- Roles of Distinct Epithelial and Mesenchymal Lineages in Pulmonary Fibrosis.” *Science Advances* 6 (28): eaba1972.
- Han, Xiaoping, Ziming Zhou, Lijiang Fei, Huiyu Sun, Renying Wang, Yao Chen, Haide Chen, et al. 2020. “Construction of a Human Cell Landscape at Single-Cell Level.” *Nature* 581 (7808): 303–9.
- Hao, Yuhan, Stephanie Hao, Erica Andersen-Nissen, William M. Mauck 3rd, Shiwei Zheng, Andrew Butler, Maddie J. Lee, et al. 2021. “Integrated Analysis of Multimodal Single-Cell Data.” *Cell* 184 (13): 3573–87.e29.
- He, Jian, Jing Yan, Jianfang Wang, Liangyu Zhao, Qian Xin, Yang Zeng, Yuxi Sun, et al. 2021. “Dissecting Human Embryonic Skeletal Stem Cell Ontogeny by Single-Cell Transcriptomic and Functional Analyses.” *Cell Research* 31 (7): 742–57.
- Hermann, Brian P., Keren Cheng, Anukriti Singh, Lorena Roa-De La Cruz, Kazadi N. Mutoji, I-Chung Chen, Heidi Gildersleeve, et al. 2018. “The Mammalian Spermatogenesis Single-Cell Transcriptome, from Spermatogonial Stem Cells to Spermatids.” *Cell Reports* 25 (6): 1650–67.e8.
- He, Shuai, Lin-He Wang, Yang Liu, Yi-Qi Li, Hai-Tian Chen, Jing-Hong Xu, Wan Peng, et al. 2020. “Single-Cell Transcriptome Profiling of an Adult Human Cell Atlas of 15 Major Organs.” *Genome Biology* 21 (1): 294.
- Hovanessian, Ara G., and Just Justesen. 2007. “The Human 2'-5'-oligoadenylate Synthetase Family: Unique Interferon-Inducible Enzymes Catalyzing 2'-5' instead of 3'-5' Phosphodiester Bond Formation.” *Biochimie* 89 (6-7): 779–88.
- Howden, Sara E., Jessica M. Vanslambrouck, Sean B. Wilson, Ker Sin Tan, and Melissa H. Little. 2019. “Reporter-Based Fate Mapping in Human Kidney Organoids Confirms Nephron Lineage Relationships and Reveals Synchronous Nephron Formation.” *EMBO Reports* 20 (4). <https://doi.org/10.15252/embr.201847483>.
- Iida, Kei, Masahiko Ajiro, Yukiko Muramoto, Toru Takenaga, Masatsugu Denawa, Ryo Kurosawa, Takeshi Noda, and Masatoshi Hagiwara. 2021. “Switching of OAS1 Splicing Isoforms Mitigates SARS-CoV-2 Infection.” *bioRxiv*. <https://doi.org/10.1101/2021.08.23.457314>.
- Jackson, Kayla, A. Sina Boeshaghi, Ángel Gálvez-Merchán, Lambda Moses, Tara Chari, and Lior Pachter. 2023. “Quantitative Assessment of Single-Cell RNA-Seq Clustering with CONCORDEX.” *bioRxiv*. <https://doi.org/10.1101/2023.06.28.546949>.
- Jiang, Shuai, Qiheng Qian, Tongtong Zhu, Wenting Zong, Yunfei Shang, Tong Jin, Yuansheng Zhang, et al. 2023. “Cell Taxonomy: A Curated Repository of Cell Types with Multifaceted Characterization.” *Nucleic Acids Research* 51 (D1): D853–60.
- Jitschin, R., M. Böttcher, D. Saul, S. Lukassen, H. Bruns, R. Loschinski, A. B. Ekici, A. Reis, A. Mackensen, and D. Mougiakakos. 2019. “Inflammation-Induced Glycolytic Switch Controls Suppressivity of Mesenchymal Stem Cells via STAT1 Glycosylation.” *Leukemia* 33 (7): 1783–96.
- Kang, Hyun Min, Meena Subramaniam, Sasha Targ, Michelle Nguyen, Lenka Maliskova, Elizabeth McCarthy, Eunice Wan, et al. 2018. “Multiplexed Droplet Single-Cell RNA-Sequencing Using Natural Genetic Variation.” *Nature Biotechnology* 36 (1): 89–94.
- Kind, Jop, Ludo Pagie, Sandra S. de Vries, Leila Nahidiazar, Siddharth S. Dey, Magda Bienko, Ye Zhan, et al. 2015. “Genome-Wide Maps of Nuclear Lamina Interactions in Single Human Cells.” *Cell* 163 (1): 134–47.
- King, Hamish W., Kristen L. Wells, Zohar Shipony, Arwa S. Kathiria, Lisa E. Wagar, Caleb Lareau, Nara Orban, et al. 2021. “Integrated Single-Cell Transcriptomics and Epigenomics Reveals Strong Germinal Center-Associated Etiology of Autoimmune Risk Loci.” *Science Immunology* 6 (64): eabh3768.
- Koyama, Shohei, Ken J. Ishii, Cevayir Coban, and Shizuo Akira. 2008. “Innate Immune Response to Viral Infection.” *Cytokine* 43 (3): 336–41.

- Kumar, Santhosh V., Pei X. Er, Kynan T. Lawlor, Ali Motazedian, Michelle Scurr, Irene Ghobrial, Alexander N. Combes, et al. 2019. "Kidney Micro-Organoids in Suspension Culture as a Scalable Source of Human Pluripotent Stem Cell-Derived Kidney Cells." *Development* 146 (5). <https://doi.org/10.1242/dev.172361>.
- Lachmann, Alexander, Denis Torre, Alexandra B. Keenan, Kathleen M. Jagodnik, Hoyjin J. Lee, Lily Wang, Moshe C. Silverstein, and Avi Ma'ayan. 2018. "Massive Mining of Publicly Available RNA-Seq Data from Human and Mouse." *Nature Communications* 9 (1): 1366.
- Laurentino, Sandra, Laura Heckmann, Sara Di Persio, Xiaolin Li, Gerd Meyer Zu Hörste, Joachim Wistuba, Jann-Frederik Cremers, et al. 2019. "High-Resolution Analysis of Germ Cells from Men with Sex Chromosomal Aneuploidies Reveals Normal Transcriptome but Impaired Imprinting." *Clinical Epigenetics* 11 (1): 127.
- Legeth, Oscar, Johan Rodhe, Stefan Lang, Parashar Dhapola, Mattias Wallergård, and Shamit Soneji. 2021. "CellexVR: A Virtual Reality Platform to Visualize and Analyze Single-Cell Omics Data." *iScience* 24 (11): 103251.
- Leinonen, Rasko, Ruth Akhtar, Ewan Birney, Lawrence Bower, Ana Cerdeno-Tárraga, Ying Cheng, Iain Cleland, et al. 2011. "The European Nucleotide Archive." *Nucleic Acids Research* 39 (Database issue): D28–31.
- Leinonen, Rasko, Hideaki Sugawara, Martin Shumway, and International Nucleotide Sequence Database Collaboration. 2011. "The Sequence Read Archive." *Nucleic Acids Research* 39 (Database issue): D19–21.
- Leir, Shih-Hsing, Shiyi Yin, Jenny L. Kerschner, Wilmel Cosme, and Ann Harris. 2020. "An Atlas of Human Proximal Epididymis Reveals Cell-Specific Functions and Distinct Roles for CFTR." *Life Science Alliance* 3 (11). <https://doi.org/10.26508/lsa.202000744>.
- Liao, Mingfeng, Yang Liu, Jing Yuan, Yanling Wen, Gang Xu, Juanjuan Zhao, Lin Cheng, et al. 2020. "Single-Cell Landscape of Bronchoalveolar Immune Cells in Patients with COVID-19." *Nature Medicine* 26 (6): 842–44.
- Li, He, Tove Ragna Reksten, John A. Ice, Jennifer A. Kelly, Indra Adrianto, Astrid Rasmussen, Shaofeng Wang, et al. 2017. "Identification of a Sjögren's Syndrome Susceptibility Locus at OAS1 That Influences Isoform Switching, Protein Expression, and Responsiveness to Type I Interferons." *PLoS Genetics* 13 (6): e1006820.
- Lim, Jean K., Andrea Lisco, David H. McDermott, Linda Huynh, Jerrold M. Ward, Bernard Johnson, Hope Johnson, et al. 2009. "Genetic Variation in OAS1 Is a Risk Factor for Initial Infection with West Nile Virus in Man." *PLoS Pathogens* 5 (2): e1000321.
- Lindström, Nils O., Guilherme De Sena Brandine, Tracy Tran, Andrew Ransick, Gio Suh, Jinjin Guo, Albert D. Kim, et al. 2018. "Progressive Recruitment of Mesenchymal Progenitors Reveals a Time-Dependent Process of Cell Fate Acquisition in Mouse and Human Nephrogenesis." *Developmental Cell* 45 (5): 651–60.e4.
- Lin, Ren-Jye, Han-Pang Yu, Bi-Lan Chang, Wei-Chun Tang, Ching-Len Liao, and Yi-Ling Lin. 2009. "Distinct Antiviral Roles for Human 2',5'-Oligoadenylate Synthetase Family Members against Dengue Virus Infection." *Journal of Immunology* 183 (12): 8035–43.
- Liu, Yawei, Xiaoying Fan, Rui Wang, Xiaoyin Lu, Yan-Li Dang, Huiying Wang, Hai-Yan Lin, et al. 2018. "Single-Cell RNA-Seq Reveals the Diversity of Trophoblast Subtypes and Patterns of Differentiation in the Human Placenta." *Cell Research* 28 (8): 819–32.
- McCauley, Katherine B., Konstantinos-Dionysios Alysandratos, Anjali Jacob, Finn Hawkins, Ignacio S. Caballero, Marall Vedaie, Wenli Yang, et al. 2018. "Single-Cell Transcriptomic Profiling of Pluripotent Stem Cell-Derived SCGB3A2+ Airway Epithelium." *Stem Cell Reports* 10 (5): 1579–95.
- McCray, Tara, Daniel Moline, Bethany Baumann, Donald J. Vander Griend, and Larisa Nonn. 2019. "Single-Cell RNA-Seq Analysis Identifies a Putative Epithelial Stem Cell Population in Human Primary Prostate Cells in Monolayer and Organoid Culture Conditions." *American Journal of Clinical and Experimental Urology* 7 (3): 123–38.

- Melchjorsen, Jesper, Helle Kristiansen, Rune Christiansen, Johanna Rintahaka, Sampsa Matikainen, Søren R. Paludan, and Rune Hartmann. 2009. "Differential Regulation of the OASL and OAS1 Genes in Response to Viral Infections." *Journal of Interferon & Cytokine Research: The Official Journal of the International Society for Interferon and Cytokine Research* 29 (4): 199–207.
- Melsted, Páll, A. Sina Boeshaghi, Lauren Liu, Fan Gao, Lambda Lu, Kyung Hoi Joseph Min, Eduardo da Veiga Beltrame, Kristján Eldjárn Hjörleifsson, Jase Gehring, and Lior Pachter. 2021. "Modular, Efficient and Constant-Memory Single-Cell RNA-Seq Preprocessing." *Nature Biotechnology* 39 (7): 813–18.
- Merino, D., T. S. Weber, A. Serrano, F. Vaillant, K. Liu, B. Pal, L. Di Stefano, et al. 2019. "Barcoding Reveals Complex Clonal Behavior in Patient-Derived Xenografts of Metastatic Triple Negative Breast Cancer." *Nature Communications* 10 (1): 766.
- Merrick, David, Alexander Sakers, Zhazira Irgebay, Chihiro Okada, Catherine Calvert, Michael P. Morley, Ivona Percec, and Patrick Seale. 2019. "Identification of a Mesenchymal Progenitor Cell Hierarchy in Adipose Tissue." *Science* 364 (6438). <https://doi.org/10.1126/science.aav2501>.
- Mincarelli, Laura, Vladimir Uzun, David Wright, Anita Scoones, Stuart A. Rushworth, Wilfried Haerty, and Iain C. Macaulay. 2023. "Single-Cell Gene and Isoform Expression Analysis Reveals Signatures of Ageing in Haematopoietic Stem and Progenitor Cells." *Communications Biology* 6 (1): 558.
- Mircea, Maria, Mazène Hochane, Xueying Fan, Susana M. Chuva de Sousa Lopes, Diego Garlaschelli, and Stefan Semrau. 2022. "Phiclust: A Clusterability Measure for Single-Cell Transcriptomics Reveals Phenotypic Subpopulations." *Genome Biology* 23 (1): 18.
- Nam, Anna S., Kyu-Tae Kim, Ronan Chaligne, Franco Izzo, Chelston Ang, Justin Taylor, Robert M. Myers, et al. 2019. "Somatic Mutations and Cell Identity Linked by Genotyping of Transcriptomes." *Nature* 571 (7765): 355–60.
- Noguchi, Satoshi, Emi Hamano, Ikumi Matsushita, Minako Hijikata, Hideyuki Ito, Takahide Nagase, and Naoto Keicho. 2013. "Differential Effects of a Common Splice Site Polymorphism on the Generation of OAS1 Variants in Human Bronchial Epithelial Cells." *Human Immunology* 74 (3): 395–401.
- Ntranos, Vasilis, Lynn Yi, Páll Melsted, and Lior Pachter. 2019. "A Discriminative Learning Approach to Differential Expression Analysis for Single-Cell RNA-Seq." *Nature Methods* 16 (2): 163–66.
- Oetjen, Karolyn A., Katherine E. Lindblad, Meghali Goswami, Gege Gui, Pradeep K. Dagur, Catherine Lai, Laura W. Dillon, J. Philip McCoy, and Christopher S. Hourigan. 2018. "Human Bone Marrow Assessment by Single-Cell RNA Sequencing, Mass Cytometry, and Flow Cytometry." *JCI Insight* 3 (23). <https://doi.org/10.1172/jci.insight.124928>.
- Ogasawara, Osamu, Yuichi Kodama, Jun Mashima, Takehide Kosuge, and Takatomo Fujisawa. 2020. "DDBJ Database Updates and Computational Infrastructure Enhancement." *Nucleic Acids Research* 48 (D1): D45–50.
- Ouchi, Rie, Shodai Togo, Masaki Kimura, Tadahiro Shinozawa, Masaru Koido, Hiroyuki Koike, Wendy Thompson, et al. 2019. "Modeling Steatohepatitis in Humans with Pluripotent Stem Cell-Derived Organoids." *Cell Metabolism* 30 (2): 374–84.e6.
- Park, Jong-Eun, Rachel A. Botting, Cecilia Domínguez Conde, Dorin-Mirel Popescu, Marieke Lavaert, Daniel J. Kunz, Issac Goh, et al. 2020. "A Cell Atlas of Human Thymic Development Defines T Cell Repertoire Formation." *Science* 367 (6480). <https://doi.org/10.1126/science.aay3224>.
- Popescu, Dorin-Mirel, Rachel A. Botting, Emily Stephenson, Kile Green, Simone Webb, Laura Jardine, Emily F. Calderbank, et al. 2019. "Decoding Human Fetal Liver Haematopoiesis." *Nature* 574 (7778): 365–71.
- Rendeiro, André F., Thomas Krausgruber, Nikolaus Fortelny, Fangwen Zhao, Thomas Penz,

- Matthias Farlik, Linda C. Schuster, et al. 2020. "Chromatin Mapping and Single-Cell Immune Profiling Define the Temporal Dynamics of Ibrutinib Response in CLL." *Nature Communications* 11 (1): 577.
- Richer, Amanda L., Kent A. Riemondy, Lakotah Hardie, and Jay R. Hesselberth. 2020. "Simultaneous Measurement of Biochemical Phenotypes and Gene Expression in Single Cells." *Nucleic Acids Research* 48 (10): e59.
- Rozenblatt-Rosen, Orit, Michael J. T. Stubbington, Aviv Regev, and Sarah A. Teichmann. 2017. "The Human Cell Atlas: From Vision to Reality." *Nature* 550 (7677): 451–53.
- Ruiz García, Sandra, Marie Deprez, Kevin Lebrigand, Amélie Cavard, Agnès Paquet, Marie-Jeanne Arguel, Virginie Magnone, et al. 2019. "Novel Dynamics of Human Mucociliary Differentiation Revealed by Single-Cell RNA Sequencing of Nasal Epithelial Cultures." *Development* 146 (20). <https://doi.org/10.1242/dev.177428>.
- Schirotti, Giulia, Anastasia Conti, Samuele Ferrari, Lucrezia Della Volpe, Aurelien Jacob, Luisa Albano, Stefano Beretta, et al. 2019. "Precise Gene Editing Preserves Hematopoietic Stem Cell Function Following Transient p53-Mediated DNA Damage Response." *Cell Stem Cell* 24 (4): 551–65.e8.
- Segerstolpe, Åsa, Athanasia Palasantza, Pernilla Eliasson, Eva-Marie Andersson, Anne-Christine Andréasson, Xiaoyan Sun, Simone Picelli, et al. 2016. "Single-Cell Transcriptome Profiling of Human Pancreatic Islets in Health and Type 2 Diabetes." *Cell Metabolism* 24 (4): 593–607.
- Shami, Adrienne Niederriter, Xianing Zheng, Sarah K. Munyoki, Qianyi Ma, Gabriel L. Manske, Christopher D. Green, Meena Sukhwani, Kyle E. Orwig, Jun Z. Li, and Saher Sue Hammoud. 2020. "Single-Cell RNA Sequencing of Human, Macaque, and Mouse Testes Uncovers Conserved and Divergent Features of Mammalian Spermatogenesis." *Developmental Cell* 54 (4): 529–47.e12.
- Sigiel, M., M. C. Laxenaire, A. Girard, L. Picard, and P. Vieux-Rochat. 1978. "[Fractionated gas encephalography with nitrous oxide under general anesthesia]." *Annales de l'anesthesiologie française* 19 (6): 509–16.
- Sina Boeshaghi, A., Ingileif B. Hallgrímsdóttir, Ángel Gálvez-Merchán, and Lior Pachter. 2022. "Depth Normalization for Single-Cell Genomics Count Data." *bioRxiv*. <https://doi.org/10.1101/2022.05.06.490859>.
- Sivakamasundari, V., Mohan Bolisetty, Santhosh Sivajothi, Shannon Bessonett, Diane Ruan, and Paul Robson. 2017. "Comprehensive Cell Type Specific Transcriptomics of the Human Kidney." *bioRxiv*. <https://doi.org/10.1101/238063>.
- Sohni, Abhishek, Kun Tan, Hye-Won Song, Dana Burow, Dirk G. de Rooij, Louise Laurent, Tung-Chin Hsieh, et al. 2019. "The Neonatal and Adult Human Testis Defined at the Single-Cell Level." *Cell Reports* 26 (6): 1501–17.e4.
- Solé-Boldo, Llorenç, Günter Raddatz, Sabrina Schütz, Jan-Philipp Mallm, Karsten Rippe, Anke S. Lonsdorf, Manuel Rodríguez-Paredes, and Frank Lyko. 2020. "Single-Cell Transcriptomes of the Human Skin Reveal Age-Related Loss of Fibroblast Priming." *Communications Biology* 3 (1): 188.
- Soveg, Frank W., Johannes Schwerk, Nandan S. Gokhale, Karen Cerosaletti, Julian R. Smith, Erola Pairo-Castineira, Alison M. Kell, et al. 2021. "Endomembrane Targeting of Human OAS1 p46 Augments Antiviral Activity." *eLife* 10 (August). <https://doi.org/10.7554/eLife.71047>.
- Su, Cheng, Yufang Lv, Wenhao Lu, Zhenyuan Yu, Yu Ye, Bingqian Guo, Deyun Liu, et al. 2021. "Single-Cell RNA Sequencing in Multiple Pathologic Types of Renal Cell Carcinoma Revealed Novel Potential Tumor-Specific Markers." *Frontiers in Oncology* 11 (October): 719564.
- Sui, Lina, Yurong Xin, Qian Du, Daniela Georgieva, Giacomo Diedenhofen, Leena Haataja, Qi Su, et al. 2021. "Reduced Replication Fork Speed Promotes Pancreatic Endocrine

- Differentiation and Controls Graft Size.” *JCI Insight* 6 (5).
<https://doi.org/10.1172/jci.insight.141553>.
- Sullivan, Delaney K., Kyung Hoi Joseph Min, Kristján Eldjárn Hjörleifsson, Laura Luebbert, Guillaume Holley, Lambda Moses, Johan Gustafsson, et al. 2023. “Kallisto, Bustools, and Kb-Python for Quantifying Bulk, Single-Cell, and Single-Nucleus RNA-Seq.” *bioRxiv : The Preprint Server for Biology*, November. <https://doi.org/10.1101/2023.11.21.568164>.
- Sun, Zhe, Li Chen, Hongyi Xin, Yale Jiang, Qianhui Huang, Anthony R. Cillo, Tracy Tabib, et al. 2019. “A Bayesian Mixture Model for Clustering Droplet-Based Single-Cell Transcriptomic Data from Population Studies.” *Nature Communications* 10 (1): 1649.
- Svensson, Valentine, Eduardo da Veiga Beltrame, and Lior Pachter. 2020. “A Curated Database Reveals Trends in Single-Cell Transcriptomics.” *Database: The Journal of Biological Databases and Curation* 2020 (November). <https://doi.org/10.1093/database/baaa073>.
- Takeda, Akira, Maija Hollmén, Denis Dermadi, Junliang Pan, Kevin Francis Brulois, Riina Kaukonen, Tapio Lönnberg, et al. 2019. “Single-Cell Survey of Human Lymphatics Unveils Marked Endothelial Cell Heterogeneity and Mechanisms of Homing for Neutrophils.” *Immunity* 51 (3): 561–72.e5.
- Takeuchi, Osamu, and Shizuo Akira. 2009. “Innate Immunity to Virus Infection.” *Immunological Reviews* 227 (1): 75–86.
- Tredano, Mohammed, Matthias Griese, Frank Brasch, Silja Schumacher, Jacques de Blic, Stéphanie Marque, Claude Houdayer, Jacques Elion, Rémy Couderc, and Michel Bahuau. 2004. “Mutation of SFTPC in Infantile Pulmonary Alveolar Proteinosis with or without Fibrosing Lung Disease.” *American Journal of Medical Genetics. Part A* 126A (1): 18–26.
- Trujillo, Cleber A., Richard Gao, Priscilla D. Negraes, Jing Gu, Justin Buchanan, Sebastian Preissl, Allen Wang, et al. 2019. “Complex Oscillatory Waves Emerging from Cortical Organoids Model Early Human Brain Network Development.” *Cell Stem Cell* 25 (4): 558–69.e7.
- Tucker, Nathan R., Mark Chaffin, Stephen J. Fleming, Amelia W. Hall, Victoria A. Parsons, Kenneth C. Bedi Jr, Amer-Denis Akkad, et al. 2020. “Transcriptional and Cellular Diversity of the Human Heart.” *Circulation* 142 (5): 466–82.
- Turner, Jackson S., Julian Q. Zhou, Julianna Han, Aaron J. Schmitz, Amena A. Rizk, Wafaa B. Alsoussi, Tingting Lei, et al. 2020. “Human Germinal Centres Engage Memory and Naive B Cells after Influenza Vaccination.” *Nature* 586 (7827): 127–32.
- Vento-Tormo, Roser, Mirjana Efremova, Rachel A. Botting, Margherita Y. Turco, Miquel Vento-Tormo, Kerstin B. Meyer, Jong-Eun Park, et al. 2018. “Single-Cell Reconstruction of the Early Maternal-Fetal Interface in Humans.” *Nature* 563 (7731): 347–53.
- Vijay, Jinchu, Marie-Frédérique Gauthier, Rebecca L. Biswell, Daniel A. Louiselle, Jeffrey J. Johnston, Warren A. Cheung, Bradley Belden, et al. 2020. “Single-Cell Analysis of Human Adipose Tissue Identifies Depot and Disease Specific Cell Types.” *Nature Metabolism* 2 (1): 97–109.
- Voigt, A. P., S. S. Whitmore, M. J. Flamme-Wiese, M. J. Riker, L. A. Wiley, B. A. Tucker, E. M. Stone, R. F. Mullins, and T. E. Scheetz. 2019. “Molecular Characterization of Foveal versus Peripheral Human Retina by Single-Cell RNA Sequencing.” *Experimental Eye Research* 184 (July): 234–42.
- Wagner, Magdalena, Masahito Yoshihara, Iyadh Douagi, Anastasios Dandimopoulos, Sarita Panula, Sophie Petropoulos, Haojiang Lu, et al. 2020. “Single-Cell Analysis of Human Ovarian Cortex Identifies Distinct Cell Populations but No Oogonial Stem Cells.” *Nature Communications* 11 (1): 1147.
- Wang, Chaoqun, Nabora S. Reyes de Mochel, Stephanie A. Christenson, Monica Cassandras, Rebecca Moon, Alexis N. Brumwell, Lauren E. Byrnes, et al. 10 2018. “Expansion of Hedgehog Disrupts Mesenchymal Identity and Induces Emphysema Phenotype.” *The Journal of Clinical Investigation* 128 (10): 4343–58.

- Wang, Eric T., Rickard Sandberg, Shujun Luo, Irina Khrebtkova, Lu Zhang, Christine Mayr, Stephen F. Kingsmore, Gary P. Schroth, and Christopher B. Burge. 2008. "Alternative Isoform Regulation in Human Tissue Transcriptomes." *Nature* 456 (7221): 470–76.
- Wang, Yalong, Wanlu Song, Jilian Wang, Ting Wang, Xiaochen Xiong, Zhen Qi, Wei Fu, Xuerui Yang, and Ye-Guang Chen. 2020. "Single-Cell Transcriptome Analysis Reveals Differential Nutrient Absorption Functions in Human Intestine." *The Journal of Experimental Medicine* 217 (2). <https://doi.org/10.1084/jem.20191130>.
- Warren, Chad M., Maria C. Jordan, Kenneth P. Roos, Paul R. Krzesinski, and Marion L. Greaser. 2003. "Titin Isoform Expression in Normal and Hypertensive Myocardium." *Cardiovascular Research* 59 (1): 86–94.
- Wen, Wen, Wenru Su, Hao Tang, Wenqing Le, Xiaopeng Zhang, Yingfeng Zheng, Xiuxing Liu, et al. 2020. "Immune Cell Profiling of COVID-19 Patients in the Recovery Stage by Single-Cell Sequencing." *Cell Discovery* 6 (May): 31.
- Wu, Haojia, Andrew F. Malone, Erinn L. Donnelly, Yuhei Kirita, Kohei Uchimura, Sai M. Ramakrishnan, Joseph P. Gaut, and Benjamin D. Humphreys. 2018. "Single-Cell Transcriptomics of a Human Kidney Allograft Biopsy Specimen Defines a Diverse Inflammatory Response." *Journal of the American Society of Nephrology: JASN* 29 (8): 2069–80.
- Xu, Heng, Ying Hu, Xinyu Zhang, Bradley E. Aouizerat, Chunhua Yan, and Ke Xu. 2022. "A Novel Graph-Based K-Partitioning Approach Improves the Detection of Gene-Gene Correlations by Single-Cell RNA Sequencing." *BMC Genomics* 23 (1): 35.
- Yang, Chao, Jason R. Siebert, Robert Burns, Zachary J. Gerbec, Benedetta Bonacci, Amy Rymaszewski, Mary Rau, et al. 2019. "Heterogeneity of Human Bone Marrow and Blood Natural Killer Cells Defined by Single-Cell Transcriptome." *Nature Communications* 10 (1): 3931.
- Yu, Zhenyuan, Jinling Liao, Yang Chen, Chunlin Zou, Haiying Zhang, Jiwen Cheng, Deyun Liu, et al. 2019. "Single-Cell Transcriptomic Map of the Human and Mouse Bladders." *Journal of the American Society of Nephrology: JASN* 30 (11): 2159–76.
- Zechner, R., J. Strauss, S. Frank, E. Wagner, W. Hofmann, D. Kratky, M. Hiden, and S. Levak-Frank. 2000. "The Role of Lipoprotein Lipase in Adipose Tissue Development and Metabolism." *International Journal of Obesity and Related Metabolic Disorders: Journal of the International Association for the Study of Obesity* 24 Suppl 4 (November): S53–56.
- Zhang, Yong, Hong Liu, Wei Li, Zhengang Zhang, Xuejun Shang, David Zhang, Yuhong Li, et al. 2017. "Intraflagellar Transporter Protein (IFT27), an IFT25 Binding Partner, Is Essential for Male Fertility and Spermiogenesis in Mice." *Developmental Biology* 432 (1): 125–39.
- Zhang, Zhongheng. 2016. "Introduction to Machine Learning: K-Nearest Neighbors." *Annals of Translational Medicine* 4 (11): 218.
- Zhong, Suijuan, Wenyu Ding, Le Sun, Yufeng Lu, Hao Dong, Xiaoying Fan, Zeyuan Liu, et al. 2020. "Decoding the Development of the Human Hippocampus." *Nature* 577 (7791): 531–36.
- Zhou, Sirui, Guillaume Butler-Laporte, Tomoko Nakanishi, David R. Morrison, Jonathan Afilalo, Marc Afilalo, Laetitia Laurent, et al. 2021. "A Neanderthal OAS1 Isoform Protects Individuals of European Ancestry against COVID-19 Susceptibility and Severity." *Nature Medicine* 27 (4): 659–67.
- Zyl, Tavé van, Wenjun Yan, Alexi McAdams, Yi-Rong Peng, Karthik Shekhar, Aviv Regev, Dejan Juric, and Joshua R. Sanes. 2020. "Cell Atlas of Aqueous Humor Outflow Pathways in Eyes of Humans and Four Model Species Provides Insight into Glaucoma Pathogenesis." *Proceedings of the National Academy of Sciences of the United States of America* 117 (19): 10339–49.