Generation and analysis of context-specific genome-scale metabolic models derived from single-cell RNA-Seq data

Johan Gustafsson[a,b] (iD), Mihail Anton[c] (iD), Fariba Roshanzamir[a], Rebecka Jörnsten[d], Eduard J. Kerkhoven[a] (iD), Jonathan L. Robinson[a,e] (iD), and Jens Nielsen[a,b,e,1] (iD)

Single-cell RNA sequencing combined with genome-scale metabolic models (GEMs) has the potential to unravel the differences in metabolism across both cell types and cell states but requires new computational methods. Here, we present a method for generating cell-type-specific genome-scale models from clusters of single-cell RNA-Seq profiles. Specifically, we developed a method to estimate the minimum number of cells required to pool to obtain stable models, a bootstrapping strategy for estimating statistical inference, and a faster version of the task-driven integrative network inference for tissues algorithm for generating context-specific GEMs. In addition, we evaluated the effect of different RNA-Seq normalization methods on model topology and differences in models generated from single-cell and bulk RNA-Seq data. We applied our methods on data from mouse cortex neurons and cells from the tumor microenvironment of lung cancer and in both cases found that almost every cell subtype had a unique metabolic profile. In addition, our approach was able to detect cancer-associated metabolic differences between cancer cells and healthy cells, showcasing its utility. We also contextualized models from 202 single-cell clusters across 19 human organs using data from Human Protein Atlas and made these available in the web portal Metabolic Atlas, thereby providing a valuable resource to the scientific community. With the ever-increasing availability of single-cell RNA-Seq datasets and continuously improved GEMs, their combination holds promise to become an important approach in the study of human metabolism.

GEM | single-cell | RNA-Seq | modeling

Genome-scale metabolic models (GEMs) have been extensively used to further our understanding of metabolism in both unicellular organisms such as yeast and bacteria (1–3) and multicellular species such as humans (4–6). For multicellular species, the existence of many different cell types and tissues poses a challenge for metabolic modeling since the full reaction network encoded by the genome is typically not present in individual tissues or cell types. To remedy this, several methods have been developed that utilize RNA sequencing or proteomics data to estimate the active subnetwork in a sample (7–9), such as the task-driven integrative network inference for tissues (tINIT) algorithm. Such methods start with a full model and generate context-specific models, containing only the active portion of the network within a given tissue or cell type.

Each tissue in the human body contains many cell types and cell subtypes, where each of these often has several transcriptional states. Bulk RNA-Seq measurements are useful for generating context-specific models that describe the collective metabolism of the cell types in a tissue. Moreover, if used with for example fluorescence-activated cell sorting (FACS), bulk measurements can be used to target individual cell types. However, the technique is limited to cell types and states that can be separated by cell surface markers, which must be decided beforehand. The availability of single-cell RNA-Seq (scRNA-Seq) presents a new opportunity to generate context-specific models at the level of individual cell types and cell states.

Obtaining a representative gene expression profile for a cell type can be challenging. Due to technical variation in the data (10, 11), data sparsity in particular is a substantial challenge when generating context-specific models from scRNA-Seq data. The variation in the data, particularly in single-cell data from droplet-based methods, is dominated by sampling effects, often requiring averaging (pooling) the individual profiles of hundreds or thousands of cells to obtain the same expected variation as observed between bulk RNA-Seq samples (12). Previously reported methods for generating context-specific models for single cells either focus on small simplified models targeting highly expressed enzymes (13) or use different strategies to integrate data from neighboring cells, also focusing on highly expressed pathways (14). While these methods are useful for finding differences in metabolism, they do not focus on capturing the entire metabolic network of a cell type, with the purpose of using these networks for further simulation. Others

Significance

The metabolic reactions active in a cell can vary substantially across cell types and organs in the body. While bulk measurements from biopsies do not provide resolution at cell-type level, the arrival of single-cell RNA-Seq holds promise to make this possible. Here, we developed a method to determine the active reactions in single-cell clusters and showed that the active reactions differed substantially across cell types, both in the brain and the tumor microenvironment. To reach a broader audience, we also made the active reaction networks of 202 single-cell clusters from 19 human organs available in the web portal Metabolic Atlas.

have generated context-specific models from pooled scRNA-Seq data (15, 16), but do not fully address the statistical uncertainties introduced by the data sparsity. Therefore, new methods should aim to address these shortcomings.

In this work, we developed methods for generating context-specific GEMs from pools (often clusters) of scRNA-Seq profiles. The methods include an estimation of the required pool size and a bootstrapping strategy to estimate uncertainties in the ensuing reaction subnetwork. Since the bootstrapping strategy requires the generation of many models, we developed a new optimized version of tINIT called fast tINIT (ftINIT), which is substantially faster than the previous versions. We applied our methods on a mouse brain scRNA-Seq dataset, showcasing the ability of the methods to identify differences in metabolic capabilities across neurons. Furthermore, we used our methods to investigate a dataset from the tumor microenvironment of lung cancer and found unique metabolic capabilities of cells known to be associated with cancer. Finally, we extended the web portal Metabolic Atlas with the ability to visualize enzyme presence for 202 human cell populations from 19 organs, based on context-specific models generated using ftINIT.

Results

Generation of Cell-Type-Specific Models. To investigate the difference in active metabolic network between cell types, we generated context-specific genome-scale models by reducing the generic GEM Human1 (4) based on scRNA-Seq data (Fig. 1*A*). The process starts with generating clusters of single cells by cell type. To enable comparison across cell types, it is desirable to estimate the uncertainty in modeling results and apply statistical inference. For scRNA-Seq data, we propose to generate GEMs from multiple bootstrapped (randomly sampled) cell populations from each cluster to assess the robustness of the modeling results. This procedure is required since the total number of unique

molecular identifiers (UMIs) or reads is usually too small to apply statistics across models generated from separate biological samples (*SI Appendix, Note S1*). Each bootstrapped cell population is then pooled into an RNA-Seq profile, and context-specific models are generated for each such profile, henceforth called bootstrap models. Further analyses, such as the evaluation of metabolic tasks, are performed for each individual model and statistical methods can be applied across cell types, where each cell type is represented by a group of bootstrap models.

The tINIT (9) algorithm was previously developed to generate context-specific GEMs based on either transcriptomic or proteomic data. A drawback of the method is the computation time, which for more complex models such as Human1 can range from 15 min to 3 h on a standard laptop computer for a single model. Since the bootstrapping strategy requires generation of large quantities of context-specific GEMs, we sought to optimize the method, yielding ftINIT (*SI Appendix, Note S2*). The results from ftINIT are different from that of previous versions. For example, ftINIT employs a different strategy for reactions lacking gene associations, where if desired many such reactions can be included rather than excluded, and the ftINIT optimization is divided into two steps to reduce computation time. ftINIT supports two modes: Mode "1+0" only runs the first step, with the result that most reactions without Gene-Protein-Reaction (GPR) associations are included in the final model, while the mode "1+1" runs both steps. The "1+0" mode is suitable for structural comparisons, while the "1+1" mode can be useful to generate a smaller model. We evaluated the performance of the previous and new version of tINIT using gene essentiality analysis on 891 cell lines from DepMap (17, 18), which showed a similar ability of the produced models to predict gene essentiality (Fig. 1*B*). We compared ftINIT with the previous version and models generated from transcriptomic data from the Genotype-Tissue Expression (GTEx) project (19) grouped by tissue type in a comparable way as the original tINIT method
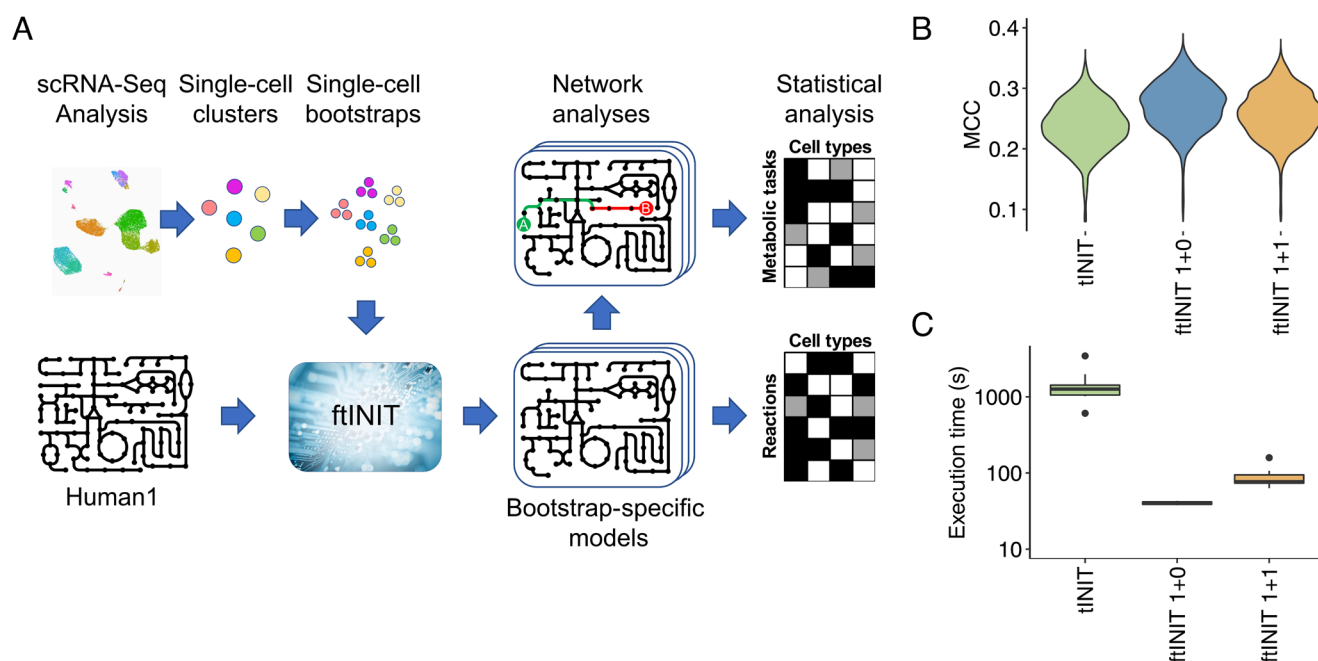

Fig. 1. Generation of context-specific models from scRNA-Seq data. *A.* Overview of model generation and analysis. Cells are first clustered in the scRNA-Seq data. Bootstraps of single cells are then generated from each cluster, followed by pooling the single cells to form a transcriptomic profile, which together with the template model Human1 is used as input to ftINIT to generate context-specific models for each bootstrap of each cell type. Network analyses in the form of metabolic task analysis are then performed for each bootstrap model, and statistical analysis is applied across bootstraps to decide if a reaction or metabolic task is available (black), unavailable (white), or uncertain (gray) in each cell type. *B.* Evaluation of the ability to predict essential genes by the models created using tINIT and ftINIT (run in the two modes "1+0" and "1+1"), respectively, for 891 cell lines from DepMap. The performance was measured using the Matthews correlation coefficient (MCC). *C.* Execution times for tINIT and ftINIT applied on 10 samples from GTEx.

(*SI Appendix,* Figs. S1 and S2). We likewise evaluated the reduction in execution time, which exceeded an order of magnitude (Fig. 1*C*).

Due to performance reasons, tINIT needs to be run with parameters that simplify the original problem when using large models such as Human1, which can cause gaps in the extracted model. To evaluate this effect, we ran both tINIT and ftINIT on a small test model, resulting in undesired gaps when using tINIT (*SI Appendix,* Figs. S3 and S4), whereas ftINIT produced no such gaps, suggesting that ftINIT may generate models of higher quality (*SI Appendix,* Fig. S5). We further investigated model statistics for models generated by the different methods. ftINIT generated larger models, where the major increase in size can be attributed to the inclusion of more reactions without GPR associations (*SI Appendix,* Fig. S6). However, ftINIT has tunable parameters that can be used to reduce the model size should that be desired (*SI Appendix, Note S2*).

Technical Evaluation of Modeling from scRNA-Seq Data. To evaluate the technical limitations of scRNA-Seq data, we first investigated the reproducibility of context-specific GEMs generated from such data. Specifically, we compared models generated from non-overlapping randomly selected cell subpopulations from the same cell-type cluster (Fig. 2*A*). Surprisingly, thousands of cells were typically needed for droplet-based single-cell data to generate models with the same variation as observed between bulk samples, and increasing the pool size beyond 10,000 cells continues to reduce the variation of the cell-type-specific GEMs. Furthermore, the number of cells required for stable model generation varied across datasets, where datasets with more UMIs per cell generally required fewer cells, emphasizing the need to evaluate the required pool size per dataset, as illustrated by the cell population "HCA CB T" in Fig. 2*A*. Direct model comparison, as shown in Fig. 2*A*, is impractical due to the large computational cost required for such a method. We therefore investigated the use of our previously developed scRNA-Seq variation estimation method DownSAmpling based Variation Estimation (DSAVE) (12) to quantify the variation between pools of cells, which takes less than a minute to run on a standard laptop computer. DSAVE demonstrated reasonable agreement in the estimated required pool size (Fig. 2*B*). Based on our results, when generating context-specific models we recommend pooling at least the number of cells required to reach the DSAVE total variation score of the bulk reference.

The presence of misclassified cells is a common problem in scRNA-Seq data, especially when dividing the cells into cell subtype populations, and there are tools available for detecting such cells (12). We investigated to what extent such cells affected the generation of context-specific GEMs by comparing models generated from pure T cell populations to models generated from populations contaminated with varying fractions of cancer cells (Fig. 2*C*). Seemingly, a few percent of misclassified cells have only a negligible effect on model generation compared to other sources of variation (such as data sparsity), while levels of 10 to 20% of misclassified cells have a clear negative effect ($P = 2.0 \times 10^{-4}$ and $P = 3.0 \times 10^{-13}$, respectively, at 5,000 cells, Wilcoxon rank sum test).

A structural comparison of models generated for both bulk and scRNA-Seq profiles from different tissues showed good agreement between models originating from the same tissue and technology (Fig. 2*D*). However, models generated from similar tissues and different technologies only partly clustered together, suggesting that a combination of technical batch effects and differences in cell-type composition between single-cell and bulk have a substantial effect on model generation. Interestingly, immune cells

from single-cell lung datasets clustered with GTEx blood samples, which can be expected to have a high immune cell content. We quantified the differences within and across different groups of tissue and technology, which showed that both these variables have a substantial effect on model generation (Fig. 2*E*).

For practical reasons, context-specific models are often generated from bulk data normalized to transcript per million (TPM) since many other normalization methods are designed to operate on gene counts and hence do not compensate for gene length. For droplet-based single-cell data, this is not a problem, as such data do not need to be normalized by gene length, although such data is still normalized to counts per million (CPM) (11). We have previously shown that trimmed mean of M values (TMM) (20) can be applied on TPM data by scaling the TPM values to produce pseudo-counts (11), and we therefore investigated the impact of different normalization methods (Fig. 2*F*). While it is difficult to draw any general conclusions from just a few datasets, TMM normalization seems to generally have a small effect compared to library normalization. However, models seem to become more similar across technologies for both TMM and quantile normalization, and TMM may be a good option for such cases, since the samples still group as expected (*SI Appendix,* Fig. S7). While quantile normalization (21) yields models with even greater similarity, it worsens the grouping on tissue (*SI Appendix,* Fig. S8) and is therefore not recommended.

Another source of variation in scRNA-Seq data that has recently received much attention is variation across samples (*SI Appendix, Note S1*). For example, differential expression analysis with single-cell data is known to produce false positives if the variation is measured across cells when not accounting for sample origin (22). In such an approach, the variation across samples is not accounted for, and pooling cells per sample to pseudo-bulk samples followed by applying methods such as DESeq2 (23), which was originally designed for bulk data, remedies the problem. The same problem is faced when trying to estimate the uncertainty in context-specific models generated from single-cell data. The variation across samples is high in single-cell data (*SI Appendix,* Fig. S9) and ideally procedures that estimate uncertainty should take this into account. In practice, datasets seldom have enough cells to generate reliable models per cell type and sample, and in such cases, we recommend our bootstrapping strategy, although it does not fully account for variation across samples.

Metabolism across Neuron Subtypes in the Mouse Cortex. To assess the utility of our method, we generated context-specific models for different neuron subtypes in the mouse primary motor cortex from a deeply sequenced publicly available dataset (24). Analysis with Seurat (25) yielded a good agreement between the cell subtype definition by Booeshaghi et al. and the Uniform Manifold Approximation and Projection (UMAP) projection (Fig. 3*A*). We selected 17 neuron subtypes for further analysis, each with more than 450 cells in the dataset, consistent with our recommendation based on the DSAVE total variation scores (*SI Appendix,* Figs. S10–S12). When using only metabolic genes, the UMAP was still able to separate the dataset per cell subtype (Fig. 3*B*), suggesting that the neuron subtypes exhibit distinct metabolic gene expression signatures that vary more across cell subtypes than within cells of the same subtype.

To investigate the metabolic networks of the neuron subtypes, we generated 100 bootstrapped single-cell populations from each neuron subtype and generated context-specific models for each bootstrap, yielding in total 1,700 models. Since the dataset contains mouse data, we used the Mouse-GEM, which is derived from Human1 by gene orthology (26). The bootstrap models were then
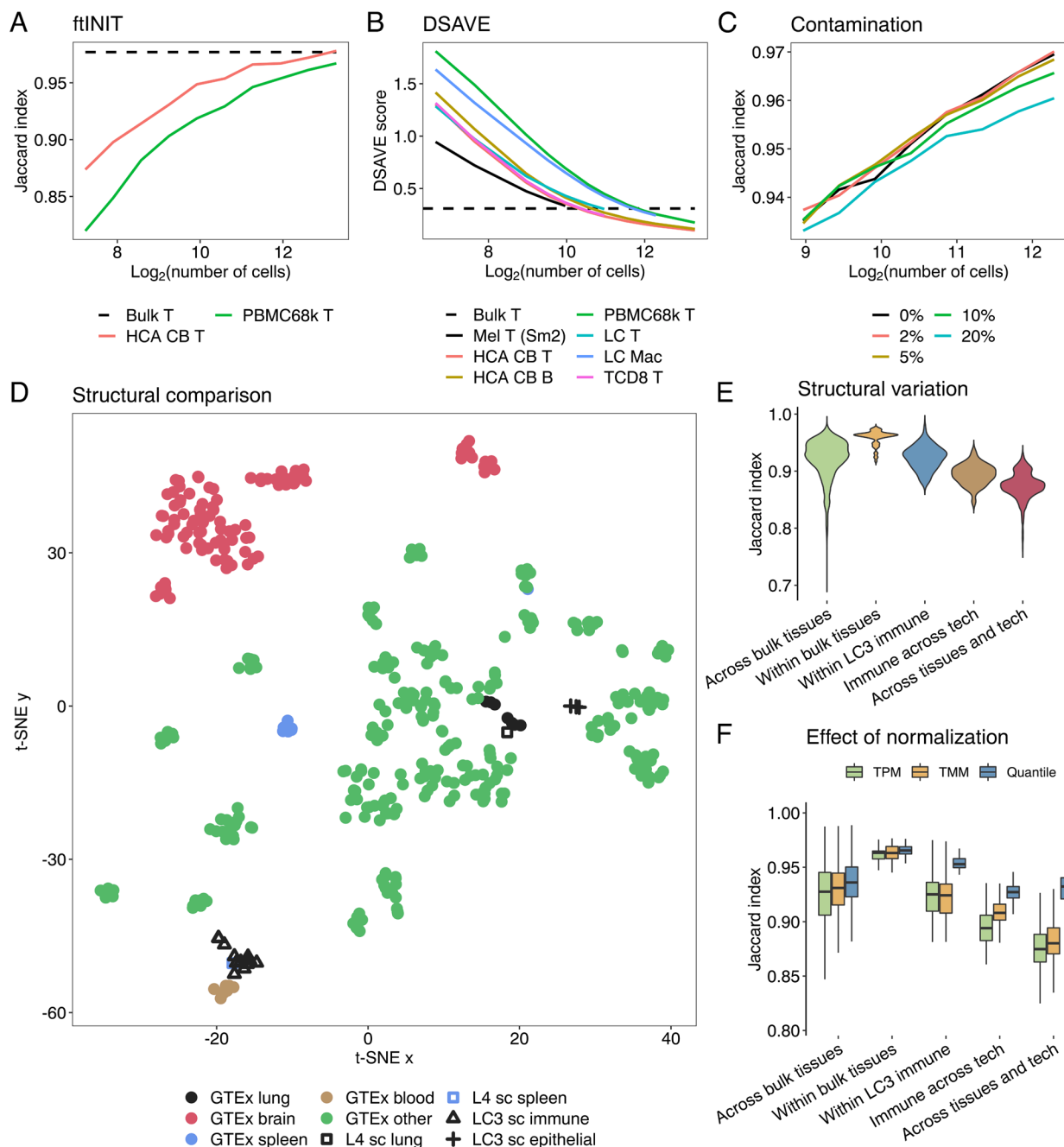
Fig. 2. Technical evaluation of generating context-specific models from scRNA-Seq data. *A*. Reproducibility of context-specific model generation per single-cell pool size, using ftINIT. Pairs of two non-overlapping sets of single cells were pooled from the same cell type (T cells) and dataset, followed by GEM generation by ftINIT and reaction content comparison (Jaccard index per pool pair, mean of 20 repetitions per pool size). The bulk reference value represents a comparison of bulk T cells (FACS-sorted). *B*. DSAVE total variation score per pool size. *C*. Structural variation of ftINIT-generated models from contaminating cell pools (T cells, LC dataset) to a varying degree with cells of another cell type (tumor cells, from the same patient). *D*. Structural comparison of models generated by ftINIT from various sources, both GTEx bulk samples from 53 tissue types (eight samples per tissue) and different single-cell datasets: 2 models (pooled cells from spleen and lung) from the L4 dataset and 16 different cell-type models from the LC3 dataset (10 from the tumor microenvironment and 6 from healthy tissue). "GTEx other" consist of the tissues from GTEx that were not expected to overlap with other datasets. Normalization: TPM/CPM. *E*. Investigation of structural variation across and within different model groups (*Materials and Methods*). *F*. The effect of different RNA-Seq normalization strategies on model similarity across and within model groups.

pooled together for structural comparison, where each reaction was scored between 0 and 100 representing the number of bootstrap models in which the reaction was present. A Principal Component Analysis (PCA) revealed structural grouping of neuron subtypes (inferior temporal (IT), near-projecting (NP), corticothalamic (CT), and Lamp5-expressing neurons) when using Principal Component (PC) 1 and PC3 (Fig. 3*C* and *SI Appendix*, Fig. S13*A*). We could not find any clear grouping of cell types from PC 1 and PC 2, and we could not identify any other cell-type

property that was related to PC 2 (explained variance 16.6%) (*SI Appendix*, Fig. S13*B*). We also could not see any clear grouping from cortex layer (L2, L5, or L6) (*SI Appendix*, Fig. S13 *C* and *D*), suggesting that the neuron metabolism is likely defined more by cell function than location, although PC2 may also represent an important factor. To quantify the number of reactions that were present in some subtypes but not in others, we defined reactions to be "on" in a subtype if it was present in at least 99 out of 100 bootstrap models, and likewise to be "off" if it was missing in at
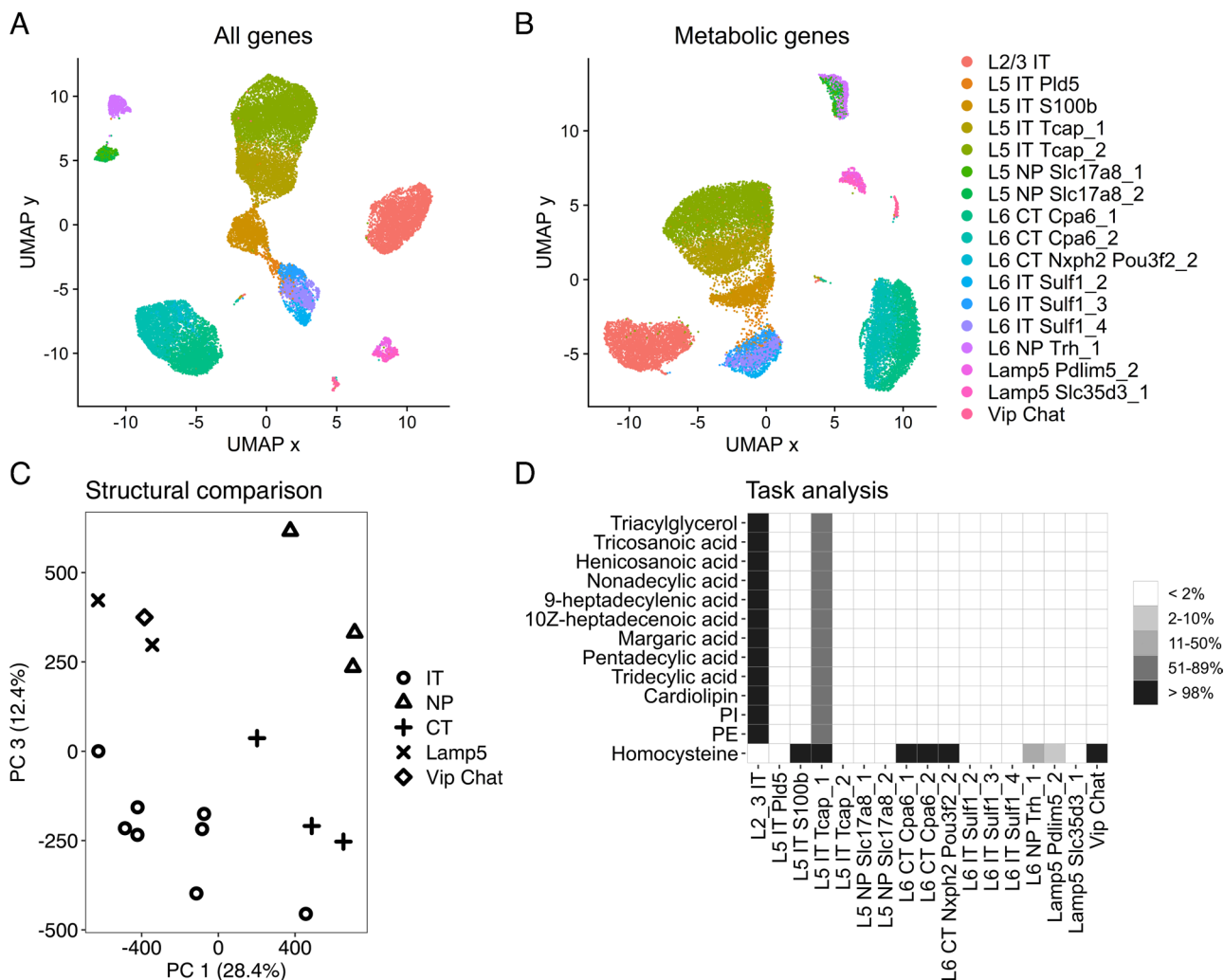
Fig. 3. Generation of context-specific GEMs for mouse primary motor cortex cell types. *A*. Single-cell UMAP projection using all genes, colored by neuron subtype classifications published together with the data. The data displayed are a subset of all cells; only the selected clusters are shown, and only for one batch of the data (with date 4/26/2019). *B*. Similar to *A*, but only using the subset of genes present in the Mouse-GEM metabolic model. *C*. Structural comparison of the context-specific models derived from each neuron subtype. Each reaction is scored based on its presence in 100 bootstrap models, which is used as input to the PCA. *D*. Metabolic task analysis of 100 bootstrap models from each cell subtype. The colors indicate the fraction of the bootstrap models that could perform each task. Only tasks where at least one cell type had more than 98% success rate and at least one had less than 2% such rate are shown. All tasks presented here represent de novo synthesis of the compounds.

least 99 out of 100 bootstraps ($P < 2.2 \times 10^{-16}$ against the null hypothesis that two reactions, where one is considered on and the other off, should be equally available; exact Fisher's test. For statistical considerations regarding multiple testing, see *SI Appendix, Note S1*). A total of 387 reactions (out of 10,376 total reactions) were defined as on in at least one cell subtype, and at the same time off in at least one other, suggesting a clear distinction between the available reaction networks in the different neuron subtypes. It is also possible to pairwise compare if a reaction statistically has a higher tendency to be on in one cell type compared to another, even for reactions that are not considered on or off (*SI Appendix, Note S1*).

What metabolic capabilities are available to a cell is an interesting property of a cell that can be evaluated by its ability to carry out different metabolic tasks such as de novo synthesis or catabolism of important metabolites. We again used our bootstrapped models to perform an analysis of 257 tasks defined in Human1, where we similarly defined on if at least 99 out of 100 bootstraps successfully completed the task and off if 99 out of 100 models failed. We found a total of 13 tasks that were considered on for at least one cell subtype while off for another (Fig. 3*D*). Most of these differentiating

tasks were related to de novo synthesis of fatty acids, phospholipids (phosphoinositides (PI) and phosphatidylethanolamines (PE)), and cardiolipin. Interestingly, the importance of fatty acids as signaling molecules in neurons has recently been emphasized, and deficiencies in lipid metabolism have been associated with cognitive problems and neurodegenerative diseases (27). The variation of homocysteine synthesis capabilities across neuron subtypes is also interesting. High homocysteine levels in blood are associated with neurological disorders (28, 29), and although homocysteine regulation is mainly managed by the liver (28), the ability of some neuron subtypes to synthesize this metabolite suggests that they could play a role in neurological disease. The diversity in homocysteine production capacity among neurons has not been studied, and potential dysregulation of this biosynthetic pathway could therefore be of interest to investigate further.

Metabolism across Cell Types in the Tumor Microenvironment. As a second application, we investigated the diversity in metabolism across cell types in the tumor microenvironment. We downloaded a publicly available lung adenocarcinoma dataset (30) containing RNA-Seq data from more than 200,000 cells from both healthy
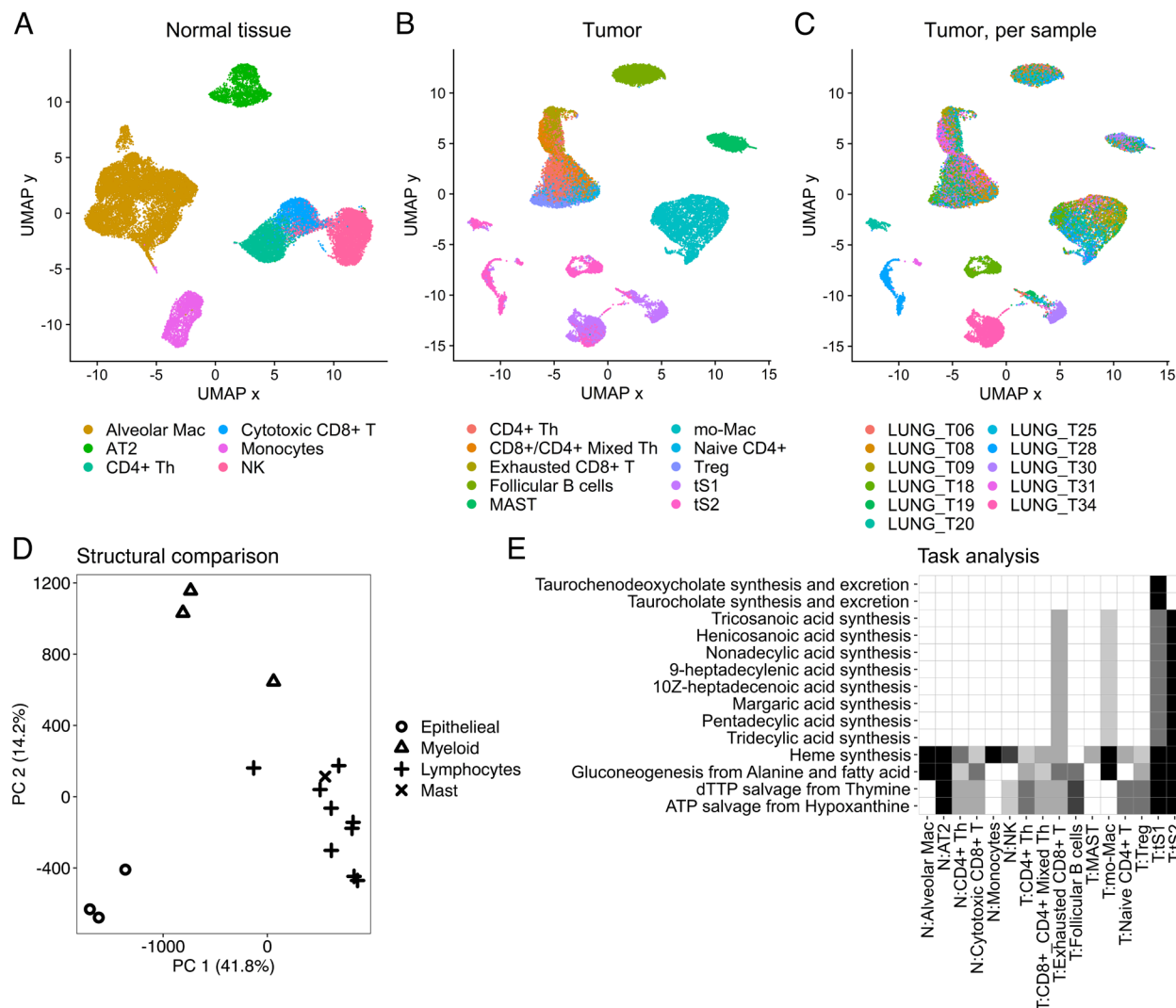
Fig. 4. Analysis of the cell types of the tumor microenvironment in lung cancer. *A*. UMAP projection of cells from healthy lung tissue. The cells originate from multiple patients and only cell clusters with at least 1,600 cells are included. The cell-type classification used was published together with the dataset. *B*. Similar to *A*, but for tumor tissue. *C*. Similar to *B* but showing sample origin per cell instead of cell subtype. *D*. Structural comparison of the context-specific models derived from clusters from both the cancer and healthy tissues. Each reaction was scored based on its presence in 100 bootstrap models, which was used as input to the PCA. The symbol indicates type of cell. *E*. Metabolic task analysis of 100 bootstrap models from each cluster. The colors indicate the fraction of the bootstrap models that could perform each task. Only tasks where at least one cell type had more than 98% success rate and at least one had less than 2% such rate are shown.

lung tissue and tumors originating from 44 patients. The data were first processed using Seurat and the UMAP projections matched well with the cell-type classifications provided with the dataset for both cells from healthy lung tissue (Fig. 4*A*) and cells originating from the tumor (Fig. 4*B*). The number of UMIs per cell varied substantially across the clusters (*SI Appendix*, Fig. S14*A*). Using DSAVE, we estimated the minimum required cluster size to be between 800 and 2,000 cells (*SI Appendix*, Fig. S14*B*) and therefore included the 16 clusters with more than 1,600 cells in the analysis. As expected, the cancer cells showed more diversity than the healthy cell types, since cancers from different patients can have varying transcriptional programs (Fig. 4*C*). Reprocessing the datasets using only metabolic genes yielded similar results, although slightly less separated per cell type, suggesting that each cell type has a unique metabolic program (*SI Appendix*, Fig. S15).

The diversity in metabolism across cell types was first investigated by a structural comparison (Fig. 4*D* and *SI Appendix*, Fig. S16). The cell types roughly clustered into a few groups: epithelial cells (alveolar cells and cancer cells), myeloid cells (macrophages and monocytes), and lymphocytes (T, natural killer

(NK), and B cells) together with mast cells, while we could not observe that cell types grouped by tissue of origin (tumor/healthy tissue). In total, 1,104 reactions were identified as on in at least one cell type and off in at least one other type, yielding a diverse set of metabolic networks.

To investigate the differences in metabolic capabilities between cell types, we performed a task analysis on the bootstrap models from all cell types, resulting in 14 tasks that could be confidently completed for at least one cell type while being absent in another (Fig. 4*E*). At least some tumor cells (tS2) had the ability to generate several types of fatty acids, which has been linked to tumor progression (31). In healthy tissues, de novo lipid production is normally limited to adipocytes and hepatocytes. However, cancer cells have been reported to be capable of lipogenesis of fatty acids from cytoplasmic acetyl-CoA (32). While even-chain fatty acids are produced from acetyl-CoA, the mechanism for production of odd-chain fatty acids (to which all fatty acids identified by the task analysis belong) was elongation of propionyl-CoA (verified in the model) but could also be supported by α-oxidation of even-chain fatty acids (33–35).

The capacity of the cancer cells to synthesize heme is another interesting observation. Cancer cells have been shown to display high heme levels, increased activity of heme containing proteins, and enhanced expression of heme exporters (36–41), suggesting that the entire heme biosynthetic pathway is frequently expressed in tumors. However, the reason why tumors enhance heme synthesis is largely unknown. One possibility is that heme together with iron–sulfur complexes is needed for oxygen-utilizing hemoproteins (e.g., mitochondrial cytochromes), which are essential for both the tricarboxylic acid (TCA) cycle and the electron transport chain (37, 41–43). Controversially, some studies showed that suppression of oxidative phosphorylation (OXPHOS) and enhanced glycolysis in tumors could be associated with increased heme biosynthesis, suggesting that heme can mediate additional functions in cancer (39). For example, heme synthesis followed by heme degradation and secretion of bilirubin provides means to dispose of succinyl-CoA from mitochondria, and this pathway was proven essential for cell lines with dysfunctional fumarate hydratase, where it can be used to keep part of the TCA cycle running (42, 44). In addition, a recent study showed that heme synthesis and export regulate the TCA cycle and OXPHOS in proliferating cells with high-energy demand (39).

Interestingly, the two different transcriptional states tS1 and tS2 of the tumor cells exhibited a distinct difference in bile acid metabolism (taurochenodeoxycholate and taurocholate synthesis and excretion), despite that each state was composed of cells from different patients with substantial transcriptional differences. The importance of bile acid metabolism is a topic of recent investigation (45), but its role in lung cancer is not clear and may be of interest for further research.
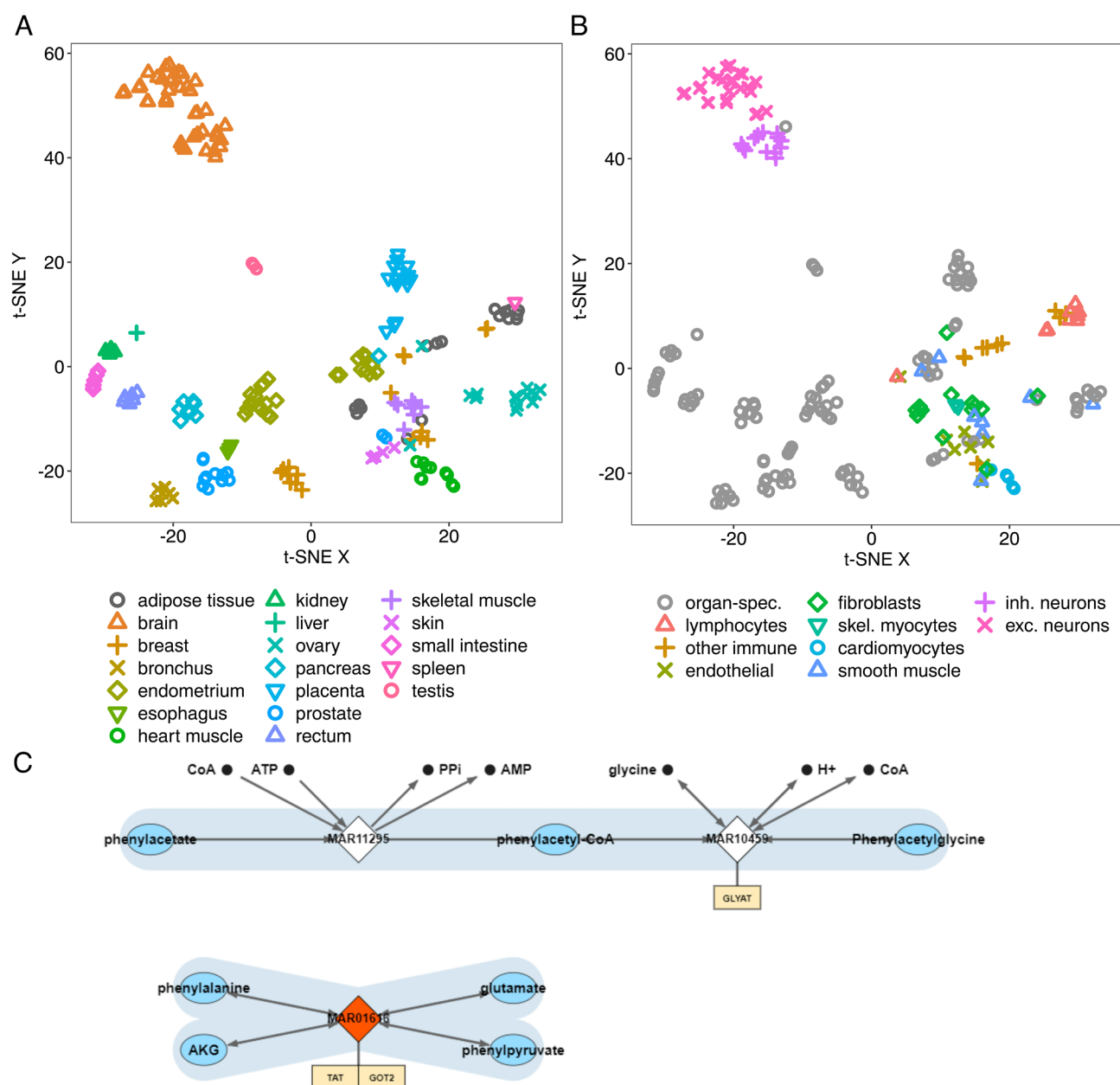

Fig. 5. Presentation of context-specific cell population models from the human body in Metabolic Atlas. *A.* t-distributed stochastic neighbor embedding (t-SNE) projection of 202 cell population reaction networks colored by tissue of origin. *B.* t-SNE projection of 202 cell population reaction networks colored by cell type. *C.* Visualization of reaction presence in the 2-dimensional maps in Metabolic Atlas, here showing phenylalanine metabolism in mitochondria for excitatory neurons (cluster 0). The color of each reaction indicates for a certain cell population the number of bootstrap models in which the reaction is present, where white represents reaction presence in zero bootstrap models and red represents presence in all models.

Presentation of Reaction Availability in Cell Types in Metabolic Atlas. Human Protein Atlas (46) has recently been extended with single-cell data (47), and we used this resource to generate context-specific models for the cell populations that were estimated to have enough cells (using DSAVE). We generated context-specific models for 202 different cell populations from 19 tissues where 100 bootstrap models were generated for each cell population, resulting in 20,200 context-specific models. The models from the cell populations cluster to some extent by organ (Fig. 5A) and by cell type (Fig. 5B). The data contain clusters from different datasets, and it is therefore difficult to determine if differences between organs originate from technical batch effects or actual differences between organs. Methods for correction of batch effects usually require cell-type overlaps between datasets, which is largely not available here, and we therefore did not apply any such method. However, while some similarities within organs can likely be attributed to batch effects, it is biologically feasible that different cell types within an organ share properties that can be attributed to an adaptation to the metabolic environment in that organ. Furthermore, cell types that exist in multiple organs, such as immune cells and endothelial cells, tend to cluster together.

Metabolic Atlas (https://metabolicatlas.org/) is a visualization web portal for GEMs. To make the reaction network of individual cell types easily accessible, we added functionality for visualizing reaction presence in context-specific models (Fig. 5C). Metabolic Atlas enables switching between cell types, making it possible to investigate differences in metabolism between cell types and organs.

Discussion

In this work, we developed methods for generating reliable context-specific models from cell populations of scRNA-Seq data. Specifically, we developed a method to estimate the required number of cells per population, a bootstrapping strategy to assess modeling results statistically, and a substantially faster version of tINIT to facilitate the bootstrapping strategy. In addition, we evaluated the effect of normalization methods for the RNA-Seq data and differences in models generated from single-cell and bulk RNA-Seq data. We found that metabolism differs substantially across cell types and subtypes, motivating our approach, and supporting that our methods were useful for finding differences across cell types and could identify metabolic properties known to be associated with the phenotype of interest.

There are many possible ways to investigate metabolism from scRNA-Seq data using GEMs. Reporter metabolites (48) is one method, whereby gene sets are defined based on which metabolites participate in the encoded reaction(s) and used in gene set analysis (GSA). The input to the GSA can for example be P-values obtained from differential expression analysis between clusters of single cells as input. Another approach is to penalize reactions based on gene expression and for different cell populations estimate the total penalty for carrying flux through the reaction network, which is implemented in the COMPASS method (14). The gene expression in COMPASS is estimated per cell by integration over nearby cells, which makes it possible to detect a metabolic switch in either a cell continuum or between clusters. While these methods were proven useful (14, 48, 49), they are designed to directly detect up-regulated pathways, while our method generates a model that can be used for further simulations, which enables the investigation of other questions. In this study, we showcased our method using metabolic task analysis, but it also allows for more advanced modeling approaches. Such methods could involve the use of metabolite uptake constraints (e.g., based on diffusion), constraints on

enzyme usage, or simulations involving the interplay between several cell types (50, 51).

Statistical inference is often a challenge when using scRNA-Seq data. Single-cell datasets often do not contain enough samples, or enough cells per sample, to apply statistics in a similar way as for bulk RNA-Seq samples. While our method partly suffers from the same weakness, our bootstrapping approach provides some statistical assurance, although subject to certain assumptions (*SI Appendix, Note S1*). It is important to realize that applying our method over cell populations collected from several samples requires that the cell-type proportions are reasonably similar across samples, since batch effects between patients could otherwise bias the results, and a prefiltering of cells to ensure cell-type proportions may be necessary.

Although some methods have recently emerged (14, 16), the use of GEMs together with scRNA-Seq data to study disease is still in its infancy. However, such analyses hold great potential—scRNA-Seq enables characterization of all cell types in the human body (52). In addition, each cell type can come in various states, and sometimes continuums, and such aspects are difficult to capture in bulk data, even for FACS-sorted cell populations. We have shown that the metabolic transcriptional program varies substantially across cell types, suggesting that study of individual cell types will provide further detail when studying metabolism in complex organs. With further development of both scRNA-Seq and GEMs, the combination of the two holds promise for a substantial contribution in unraveling the key metabolic features in human health and disease.

Materials and Methods

Datasets. We downloaded eight different scRNA-Seq datasets (30, 52–58), bulk RNA-Seq data from GTEx (19) and DepMap (also including gene essentiality data) (17, 18), and FACS-sorted T cell samples from the BLUEPRINT epigenome project (12, 59) previously assembled for DSAVE. Single-cell data from Human Protein Atlas were used for presentation of reaction presence for cell types in Metabolic Atlas (47). Some of the datasets were accessed through SingleCellToolbox (https://github.com/SysBioChalmers/SingleCellToolbox) (12), which provides single-cell datasets with cell-type classifications through a MATLAB-friendly interface. Detailed information about the datasets is available in *SI Appendix*, Table S1.

ftINIT. The ftINIT method is described in detail in *SI Appendix, Note S2*. In short, ftINIT runs in two steps: 1) Simplified run where many reactions without gene associations are omitted from the problem. 2) Run where the reactions turned on in step 1 are treated as essential and all reactions are included in the problem. Step 2 is optional and was omitted for the generation of all models used in this work except for the data presented in Metabolic Atlas. The "rxns to ignore mask" in step 1 was set to [1 1 1 1 1 1 1 0], which effectively means that a collection of reactions without gene rules, including spontaneous reactions, exchange reactions, transport reactions, and custom reactions are omitted from the optimization problem and are always included in the final model (except in the Metabolic Atlas data). The custom reactions were in this study selected to include reactions for protein generation and reactions that pool metabolites, in total 52 reactions. None of the custom reactions had gene rules.

ftINIT is designed to work with the Gurobi solver (60).

Like its predecessor tINIT, ftINIT supports both a fixed gene expression threshold value for all genes or individual gene thresholds that can be based for example on the mean expression of the gene in all samples. We did not seek to specifically evaluate gene expression thresholds here since such parameters have previously been evaluated (61) and therefore settled on a threshold value of 1 TPM/CPM.

ftINIT was implemented in Reconstruction, Analysis and Visualization of Metabolic Networks (RAVEN) Toolbox (62). All figures were generated using the version implemented in version 2.7.4.

Detailed Methods Description. A detailed description of the methods used to produce the results in this work is available in *SI Appendix, Note S3*.

Software. The data were analyzed using MATLAB R2019b and R version 4.1.1. To ensure the quality of our analyses, we verified and validated the code using a combination of test cases, reasoning around expected outcome of a function, and code review. The details of this activity are available in the verification matrix available with the code.

Data, Materials, and Software Availability. All the datasets used in this manuscript are available in public repositories, and references are given Table S1 in *SI Appendix* . The ftINIT method is implemented in RAVEN Toolbox (62) and is open source and publicly available on GitHub (https://github.com/SysBioChalmers/RAVEN). The Human1 model together with model-specific code for ftINIT is available on GitHub (https://github.com/SysBioChalmers/Human-GEM). The processed data and source code, including the used versions of RAVEN and SingleCellToolbox, are available on Zenodo (https://doi.org/10.5281/zenodo.7469969). The source code is also available on GitHub (https://github.com/SysBioChalmers/SingleCellModeling). Instructions on how to use ftINIT together with single-cell data can be found at https://sysbiochalmers.github.io/Human-GEM-guide/. The ftINIT code is planned to be moved to a new GitHub location, at which it should be accessed for future use at https://github.com/SysBioChalmers/ftINIT.

Author affiliations: [a]Department of Biology and Biological Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden; [b]Wallenberg Center for Protein Research, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden; [c]Department of Biology and Biological Engineering, National Bioinformatics Infrastructure Sweden, Science for Life Laboratory, Chalmers University of Technology, SE-41258 Gothenburg, Sweden; [d]Mathematical Sciences, University of Gothenburg and Chalmers University of Technology, SE-412 96 Gothenburg, Sweden; and [e]BioInnovation Institute, DK-2200 Copenhagen, Denmark

1. C. Lieven *et al.*, MEMOTE for standardized genome-scale metabolic model testing. *Nat. Biotechnol.* **38**, 272–276 (2020).
2. D. S. Weaver, I. M. Keseler, A. Mackie, I. T. Paulsen, P. D. Karp, A genome-scale metabolic flux model of Escherichia coli K–12 derived from the EcoCyc database. *BMC Syst. Biol.* **8**, 79 (2014).
3. O. Oftadeh *et al.*, A genome-scale metabolic model of Saccharomyces cerevisiae that integrates expression constraints and reaction thermodynamics. *Nat. Commun.* **12**, 4790 (2021).
4. J. L. Robinson *et al.*, An atlas of human metabolism. *Sci. Signal.* **13**, eaaz1482 (2020).
5. E. Brunk *et al.*, Recon3D: A resource enabling a three-dimensional view of gene variation in human metabolism. *Nat. Biotechnol.* **36**, 272–281 (2018).
6. J. E. Lewis, T. E. Forshaw, D. A. Boothman, C. M. Furdui, M. L. Kemp, Personalized genome-scale metabolic models identify targets of redox metabolism in radiation-resistant tumors. *Cell Syst.* **12**, 68–81.e11 (2021).
7. A. Schultz, A. A. Qutub, Reconstruction of tissue-specific metabolic networks using CORDA. *PLoS Comput. Biol.* **12**, e1004808 (2016).
8. S. A. Becker, B. O. Palsson, Context-specific metabolic networks are consistent with experiments. *PLoS Comput. Biol.* **4**, e1000082 (2008).
9. R. Agren *et al.*, Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling. *Mol. Syst. Biol.* **10**, 721 (2014).
10. L. M. McIntyre *et al.*, RNA-seq: Technical variability and sampling. *BMC Genomics* **12**, 293 (2011).
11. J. Gustafsson *et al.*, Sources of variation in cell-type RNA-Seq profiles. *PLoS One* **15**, e0239495 (2020).
12. J. Gustafsson *et al.*, DSAVE: Detection of misclassified cells in single-cell RNA-Seq data. *PLoS One* **15**, e0243360 (2020).
13. N. Alghamdi *et al.*, A graph neural network model to estimate cell-wise metabolic flux using single-cell RNA-seq data. *Genome Res.* **31**, 1867–1884 (2021).
14. A. Wagner *et al.*, Metabolic modeling of single Th17 cells reveals regulators of autoimmunity. *Cell* **184**, 4168–4185.e21 (2021).
15. L. S. Yilmaz *et al.*, Modeling tissue-relevant Caenorhabditis elegans metabolism at network, pathway, reaction, and metabolite levels. *Mol. Syst. Biol.* **16**, e9649 (2020).
16. Y. Zhang, M. S. Kim, E. Nguyen, D. M. Taylor, Modeling metabolic variation with single-cell expression data. bioRxiv [Preprint] (2020). https://doi.org/10.1101/2020.01.28.923680 (Accessed 6 June 2022).
17. M. Ghandi *et al.*, Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**, 503–508 (2019).
18. R. M. Meyers *et al.*, Computational correction of copy number effect improves specificity of CRISPR–Cas9 essentiality screens in cancer cells. *Nat. Genet.* **49**, 1779–1784 (2017).
19. L. J. Carithers *et al.*, A novel approach to high-quality postmortem tissue procurement: The GTEx project. *Biopreserv. Biobank.* **13**, 311–319 (2015).
20. M. D. Robinson, A. Oshlack, A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).
21. J. K. Pickrell *et al.*, Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature* **464**, 768–772 (2010).
22. J. W. Squair *et al.*, Confronting false discoveries in single-cell differential expression. *Nat. Commun.* **12**, 5692 (2021).
23. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
24. A. S. Booeshaghi *et al.*, Isoform cell type specificity in the mouse primary motor cortex. *Nature* **598**, 195–199 (2021).
25. Y. Hao *et al.*, Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
26. H. Wang *et al.*, Genome-scale metabolic network reconstruction of model animals as a platform for translational research. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2102344118 (2021).
27. L. J. Falomir-Lockhart, G. F. Cavazzutti, E. Giménez, A. M. Toscani, Fatty acid signaling mechanisms in neural cells: Fatty acid receptors. *Front. Cell. Neurosci.* **13**, 162 (2019).
28. R. Moretti, P. Caruso, The controversial role of homocysteine in neurology: From labs to clinical practice. *Int. J. Mol. Sci.* **20**, 231 (2019).
29. W. Herrmann, R. Obeid, Homocysteine: A biomarker in neurodegenerative diseases. *Clin. Chem. Lab. Med.* **49**, 435–441 (2011).
30. N. Kim *et al.*, Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nat. Commun.* **11**, 2285 (2020).
31. F. Röhrig, A. Schulze, The multifaceted roles of fatty acid synthesis in cancer. *Nat. Rev. Cancer* **16**, 732–749 (2016).
32. C. M. Metallo *et al.*, Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. *Nature* **481**, 380–384 (2012).
33. K. Weitkunat *et al.*, Effects of dietary inulin on bacterial growth, short-chain fatty acid production and hepatic lipid metabolism in gnotobiotic mice. *J. Nutr. Biochem.* **26**, 929–937 (2015).
34. L. Guo, D. Zhou, K. M. Pryse, A. L. Okunade, X. Su, Fatty acid 2-hydroxylase mediates diffusional mobility of raft-associated lipids, GLUT4 level, and lipogenesis in 3T3-L1 adipocytes*. *J. Biol. Chem.* **285**, 25438–25447 (2010).
35. B. Jenkins, J. A. West, A. Koulman, A review of odd-chain fatty acid metabolism and the role of pentadecanoic acid (C15:0) and heptadecanoic acid (C17:0) in health and disease. *Molecules* **20**, 2425–2444 (2015).
36. Y. Fukuda *et al.*, Upregulated heme biosynthesis, an exploitable vulnerability in MYCN-driven leukemogenesis. *JCI Insight* **2**, e92409 (2017).
37. J. Hooda *et al.*, Enhanced heme function and mitochondrial respiration promote the progression of lung cancer cells. *PLoS One* **8**, e63402 (2013).
38. M. Y. Hsu, E. Mina, A. Roetto, P. E. Porporato, Iron: An essential element of cancer metabolism. *Cells* **9**, 2591 (2020).
39. C. Peng *et al.*, FLVCR1 promotes the proliferation and tumorigenicity of synovial sarcoma through inhibiting apoptosis and autophagy. *Int. J. Oncol.* **52**, 1559–1568 (2018).
40. V. Fiorito *et al.*, The heme synthesis-export system regulates the tricarboxylic acid cycle flux and oxidative phosphorylation. *Cell Rep.* **35**, 109252 (2021).
41. S. Sohoni *et al.*, Elevated heme synthesis and uptake underpin intensified oxidative metabolism and tumorigenic functions in non-small cell lung cancer cells. *Cancer Res.* **79**, 2511–2525 (2019).
42. V. Fiorito, D. Chiabrando, S. Petrillo, F. Bertino, E. Tolosano, The multifaceted role of heme in cancer. *Front. Oncol.* **9**, 1540 (2020).
43. H. J. Kim, O. Khalimonchuk, P. M. Smith, D. R. Winge, Structure, function, and assembly of heme centers in mitochondrial respiratory complexes. *Biochim. Biophys. Acta* **1823**, 1604–1616 (2012).
44. C. Frezza *et al.*, Haem oxygenase is synthetically lethal with the tumour suppressor fumarate hydratase. *Nature* **477**, 225–228 (2011).
45. J. Fu, M. Yu, W. Xu, S. Yu, Research progress of bile acids in cancer. *Front. Oncol.* **11**, 778258 (2022).
46. M. Uhlén *et al.*, Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
47. M. Karlsson *et al.*, A single-cell type transcriptomics map of human tissues. *Sci. Adv.* **7**, eabh2169 (2021).
48. K. R. Patil, J. Nielsen, Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 2685–2689 (2005).
49. S. Lee *et al.*, Integrated network analysis reveals an association between plasma mannose levels and insulin resistance. *Cell Metab.* **24**, 172–184 (2016).
50. J. Gustafsson, F. Roshanzamir, A. Hagnestal, J. L. Robinson, J. Nielsen, Cellular limitation of enzymatic capacity explains glutamine addiction in cancers. bioRxiv [Preprint] (2022). https://doi.org/10.1101/2022.02.08.479584 (Accessed 8 July 2022).
51. B. J. Sánchez *et al.*, Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. *Mol. Syst. Biol.* **13**, 935 (2017).
52. O. Rozenblatt-Rosen, M. J. T. Stubbington, A. Regev, S. A. Teichmann, The Human Cell Atlas: From vision to reality. *Nat. News* **550**, 451 (2017).
53. B. Li, Census of immune cells. *Human Cell Atlas Data Portal* (2018). https://data.humancellatlas.org/explore/projects/cc95ff89-2e68-4a08-a234-480eca21ce79?catalog=dcp1. Accessed 19 February 2019.
54. G. X. Y. Zheng *et al.*, Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017).
55. I. Tirosh *et al.*, Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).
56. D. Lambrechts *et al.*, Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat. Med.* **24**, 1277–1289 (2018).
57. J. Chen *et al.*, PBMC fixation and processing for chromium single-cell RNA sequencing. *J. Transl. Med.* **16**, 198 (2018).
58. E. Madissoon *et al.*, scRNA-seq assessment of the human lung, spleen, and esophagus tissue stability after cold preservation. *Genome Biol.* **21**, 1 (2019).
59. "Blueprint Epigenome Project". https://www.blueprint-epigenome.eu/. Accessed 4 March 2019.
60. Gurobi Optimization, LLC, Gurobi optimizer reference manual (2022).
61. S. Opdam *et al.*, A systematic evaluation of methods for tailoring genome-scale metabolic models. *Cell Syst.* **4**, 318–329.e6 (2017).
62. H. Wang *et al.*, RAVEN 2.0: A versatile toolbox for metabolic network reconstruction and a case study on Streptomyces coelicolor. *PLOS Comput. Biol.* **14**, e1006541 (2018).