

Chromatin mapping and single-cell immune profiling define the temporal dynamics of ibrutinib response in CLL

André F. Rendeiro^{1*}, Thomas Krausgruber^{1*}, Nikolaus Fortelny¹, Fangwen Zhao^{1,2}, Thomas Penz¹, Matthias Farlik¹, Linda C. Schuster¹, Amelie Neme¹, Szabolcs Tasnády³, Marienn Réti³, Zoltán Mátrai³, Donat Alpar^{1,4†}, Csaba Bödör^{4†}, Christian Schmidl^{1,7†}, Christoph Bock^{1,2,5,6†}

¹ CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria

² Ludwig Boltzmann Institute for Rare and Undiagnosed Diseases, Vienna, Austria

³ Department of Haematology and Stem Cell Transplantation, Central Hospital of Southern Pest, National Institute of Hematology and Infectious Diseases, Budapest, Hungary

⁴ MTA-SE Lendület Molecular Oncohematology Research Group, 1st Department of Pathology and Experimental Cancer Research, Semmelweis University, Budapest, Hungary

⁵ Department of Laboratory Medicine, Medical University of Vienna, Vienna, Austria

⁶ Max Planck Institute for Informatics, Saarland Informatics Campus, Saarbrücken, Germany

⁷ Current address: Regensburg Center for Interventional Immunology (RCI), Regensburg, Germany

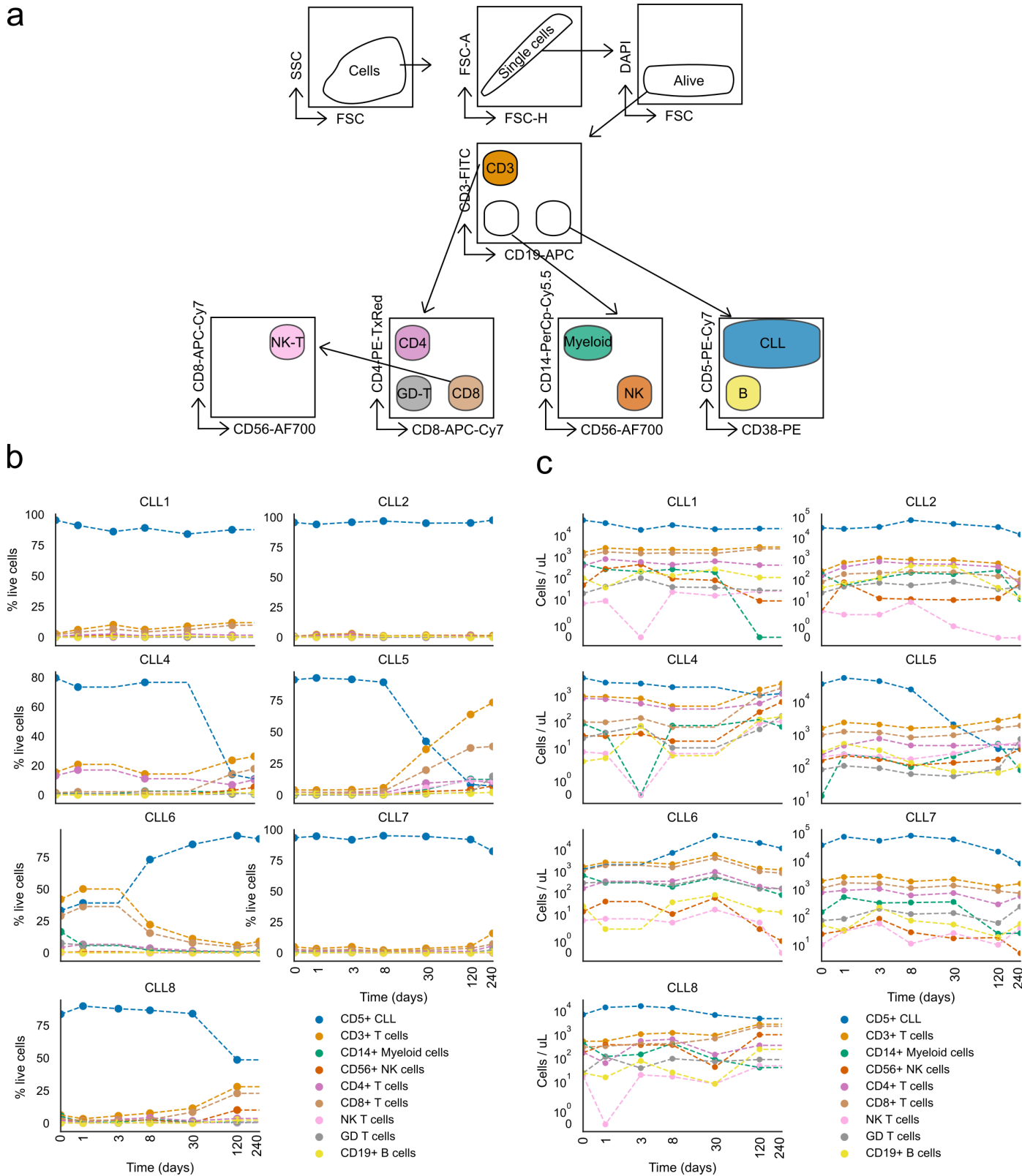
* These authors contributed equally to this work

† Co-last author / These authors jointly directed this work

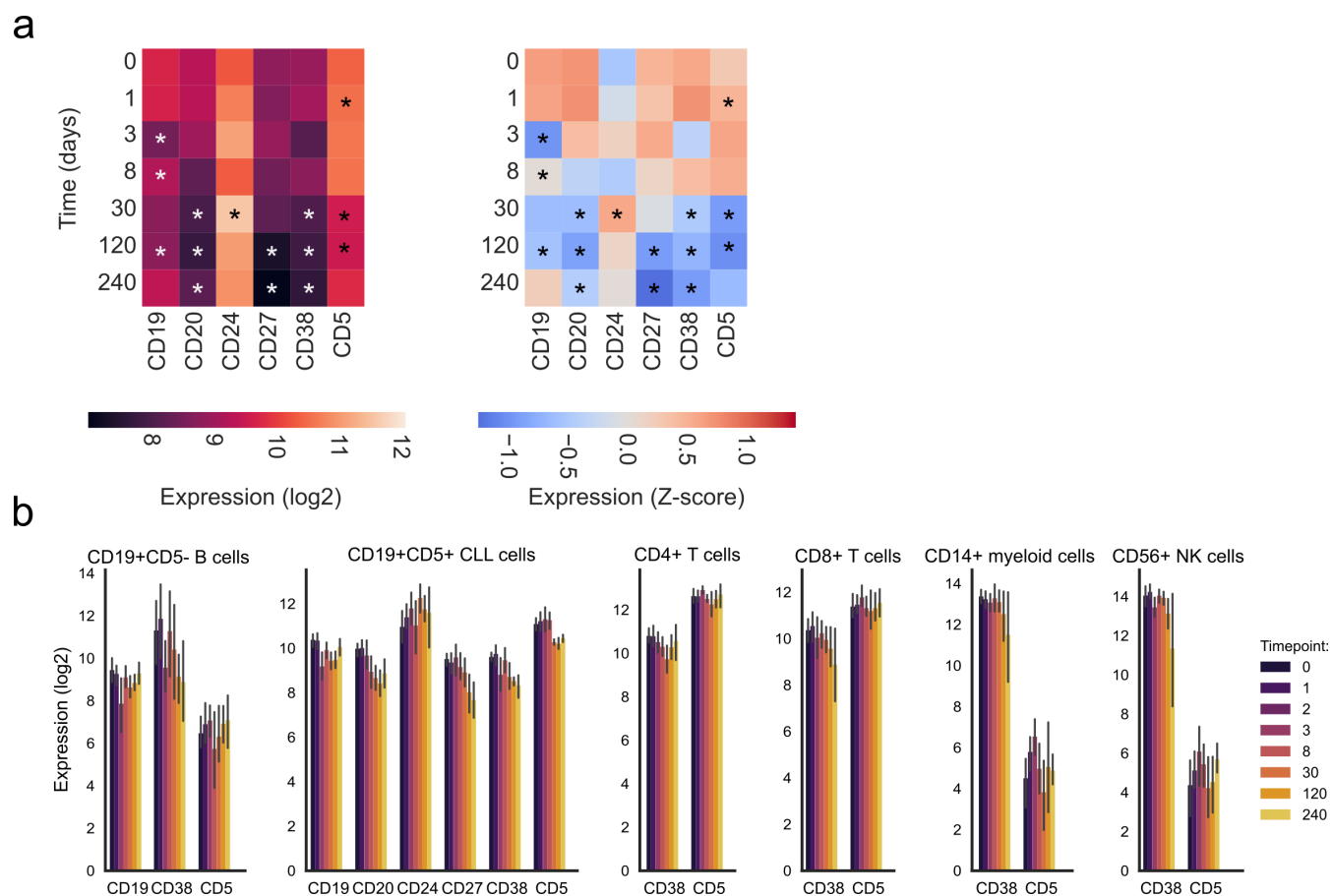
Correspondence: Christoph Bock (cbock@cemm.oeaw.ac.at)

Keywords: Chronic lymphocytic leukemia, drug response profiling, ibrutinib therapy, chromatin mapping, single-cell RNA sequencing, time series analysis, machine learning, translational bioinformatics

Supplementary Figures

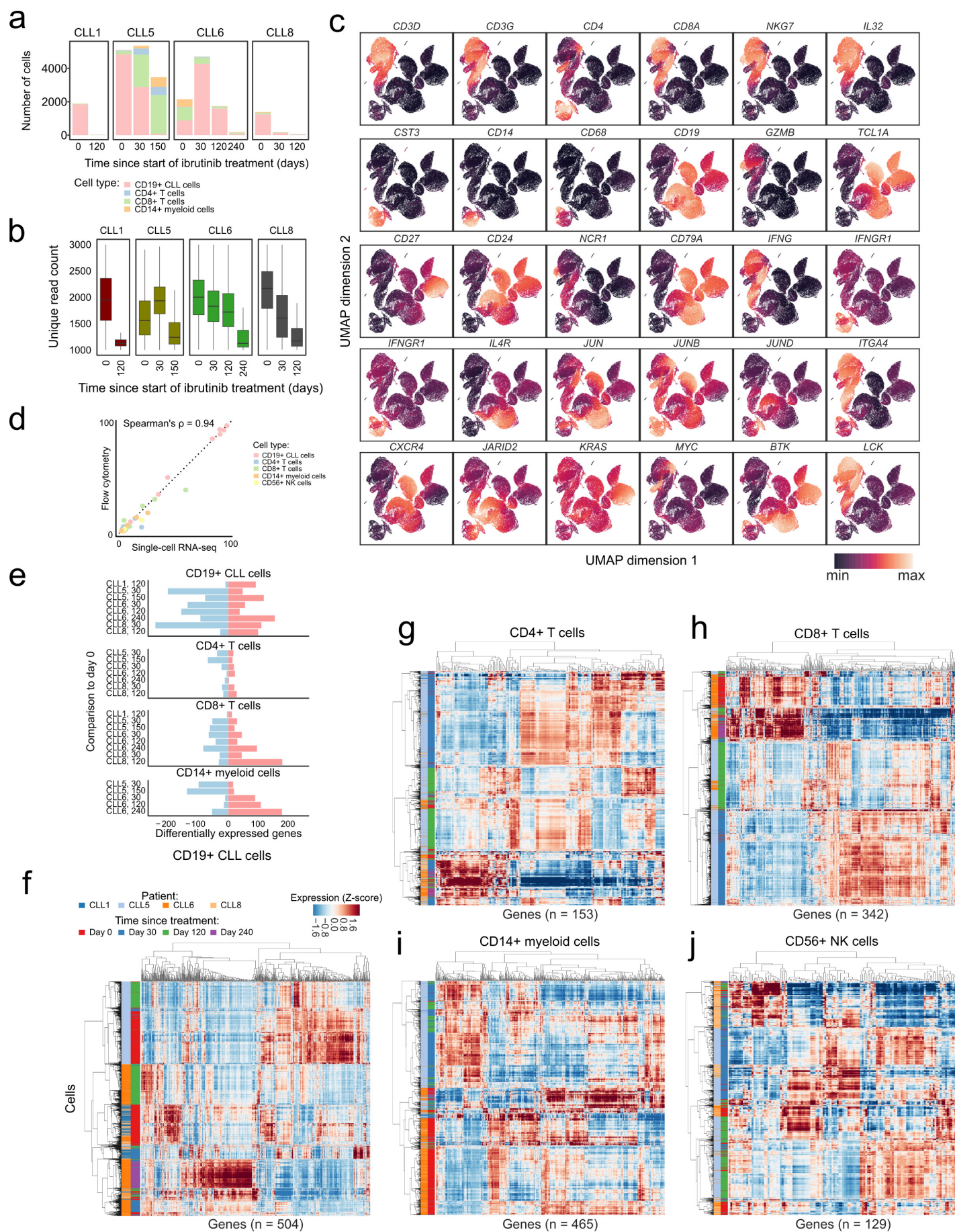


CD14+), NK cells (CD3-CD19-CD56+), NK-T cells (CD3+CD8+CD56+), and T cells (CD3+) from PBMCs of patients with CLL. This gating strategy was used to: (i) analyze the composition of the immune cell compartment (Fig. 1b-c); (ii) quantify surface marker expression (Fig. 1d); (iii) enrich cell populations for single-cell transcriptome analysis (Fig. 1e-i, Fig. 2d-f, Fig. 3e-f); and (iv) sort cell populations for bulk chromatin mapping (Fig. 2a-e, Fig. 3a-d). **b-c**) Flow cytometry quantification of the relative (b) or absolute (c) abundance of CLL cells and several non-malignant immune cell types in patients undergoing ibrutinib therapy. In total, 45 samples obtained from seven patients were included in the analysis.



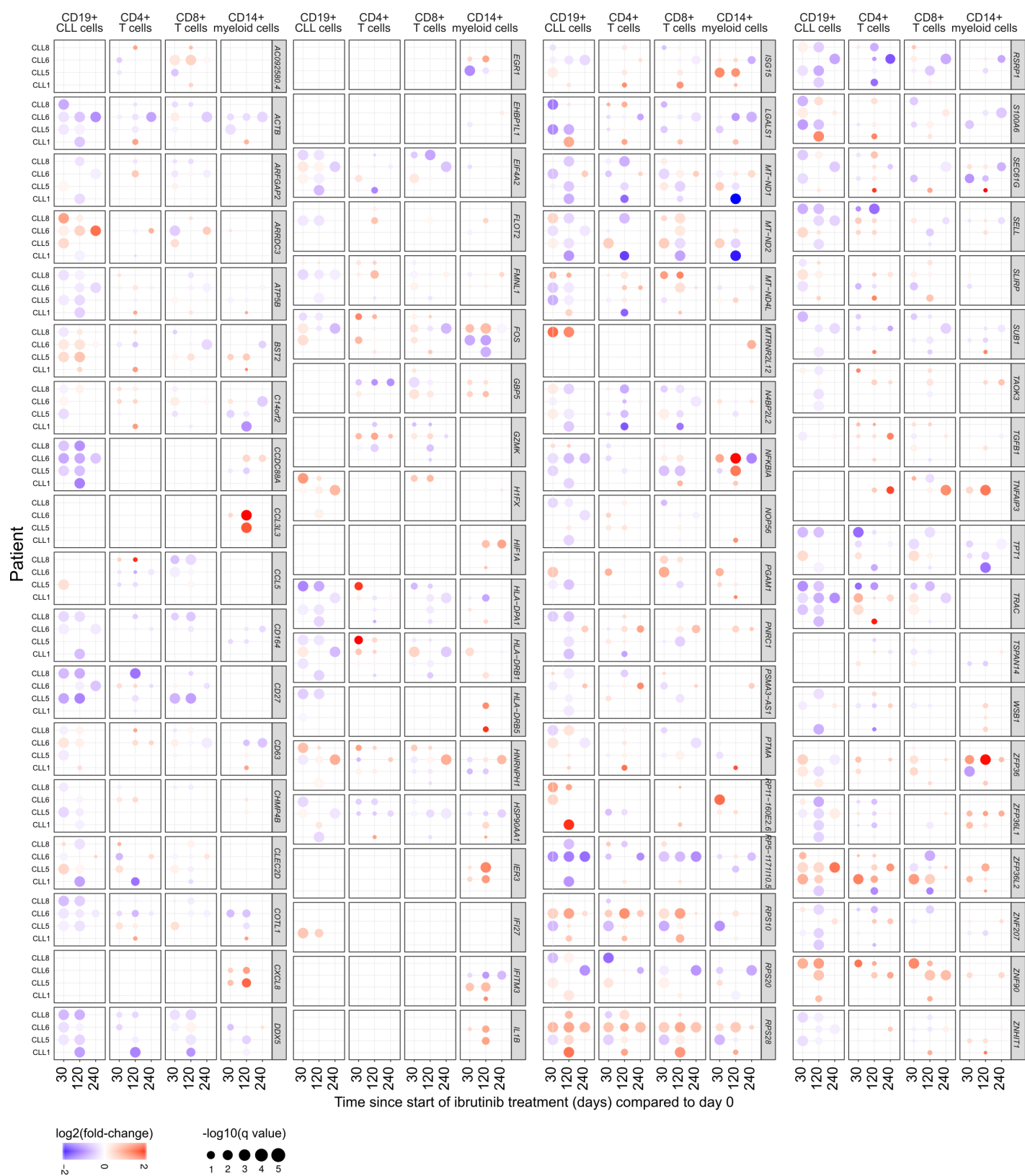
Supplementary Figure 2: Temporal dynamics of T cell subsets and surface marker expression

a) Mean expression of surface marker proteins in CD19⁺CD5⁺ CLL cells during ibrutinib treatment. The left panel displays absolute (log scale) expression, while the right panel displays column-wise Z-transformed values. Stars mark significant changes compared to time 0 (paired t-test, $p < 0.05$). **b)** Expression of surface marker proteins in immune cell subsets of patients with CLL as measured by flow cytometry (error bars represent 95 percent confidence intervals of the mean). In total, 45 samples obtained from seven patients were included in the analysis.



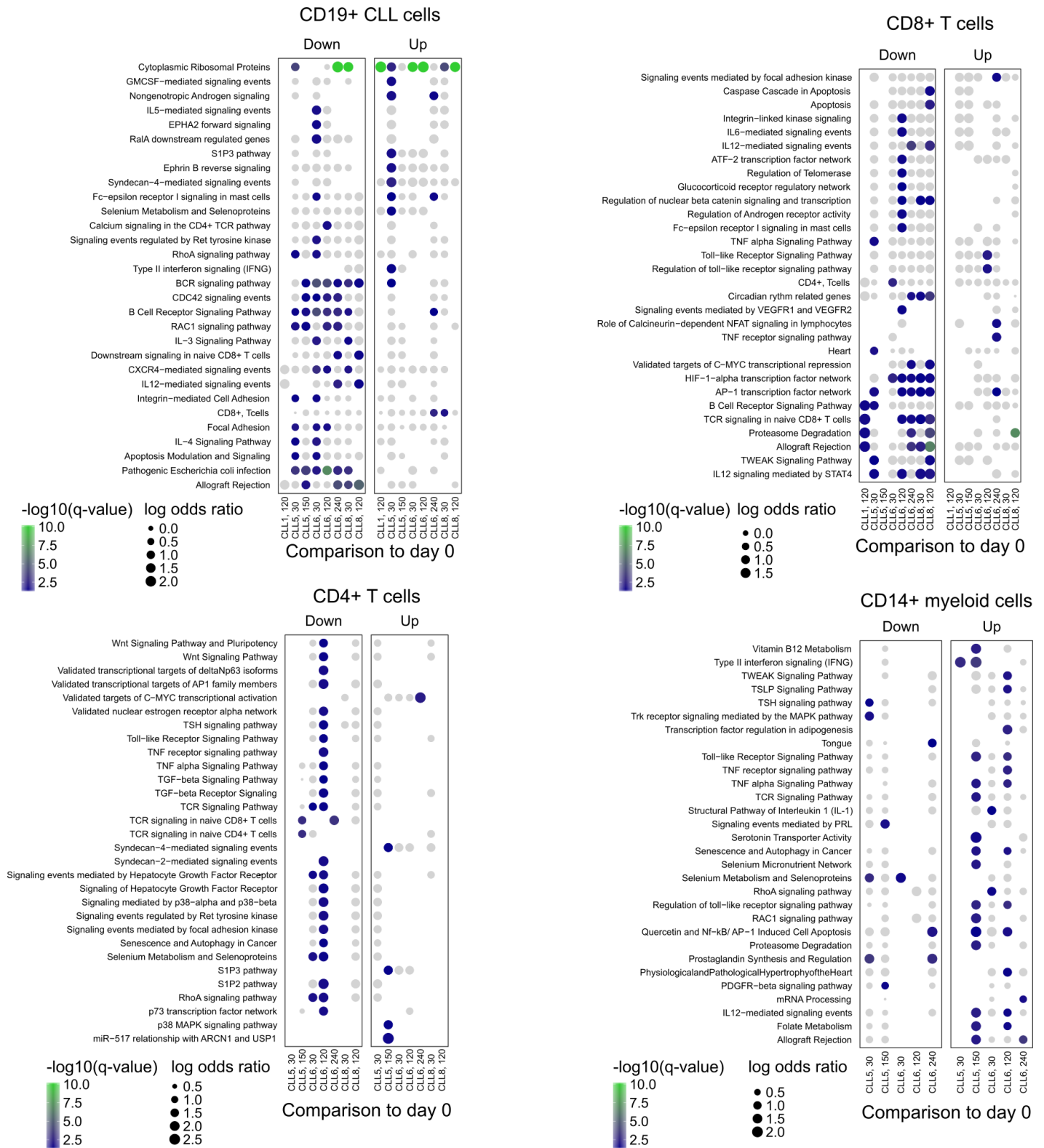
Supplementary Figure 3: Single-cell RNA-seq profiling over the ibrutinib time course

a) Bar plot displaying the number of single-cell transcriptome profiles that passed quality control, shown separately for each patient, cell type, and time point. **b)** Box plot displaying the number of unique molecular identifiers (UMIs) detected per single cell, shown separately for each patient, cell type, and time point. **c)** Heatmap showing mean expression levels of the marker genes that were used to assign the single-cell transcriptomes to defined cell types. Values represent expression levels (normalized UMI counts) scaled from minimum to maximum in each row. **d)** Scatterplot comparing the fraction of cells of each type based on single-cell RNA-seq versus flow cytometry across all patients and time points. **e)** Number of differentially expressed genes for each cell type, patient, and time point. **f-j)** Heatmaps of differentially expressed genes over the ibrutinib time course separately for each cell type.



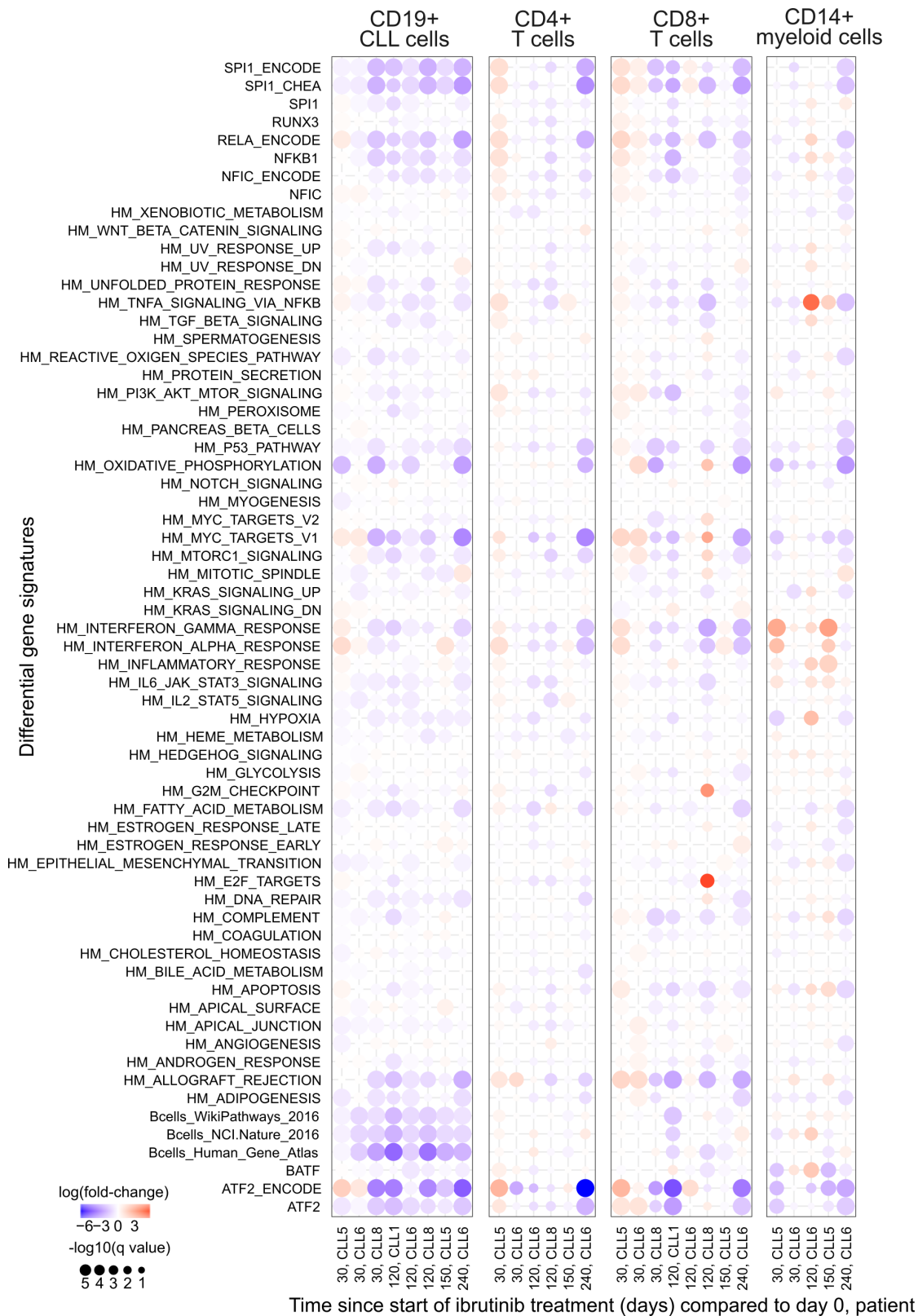
Supplementary Figure 4: Transcriptional changes of differentially expressed genes upon ibrutinib treatment

Differential gene expression for those genes that were significantly differentially expressed (absolute log fold change higher than one in more than two patients) over the course of ibrutinib therapy, shown separately for each cell type and comparing each sample to the pre-treatment (day 0) sample from the same patient.



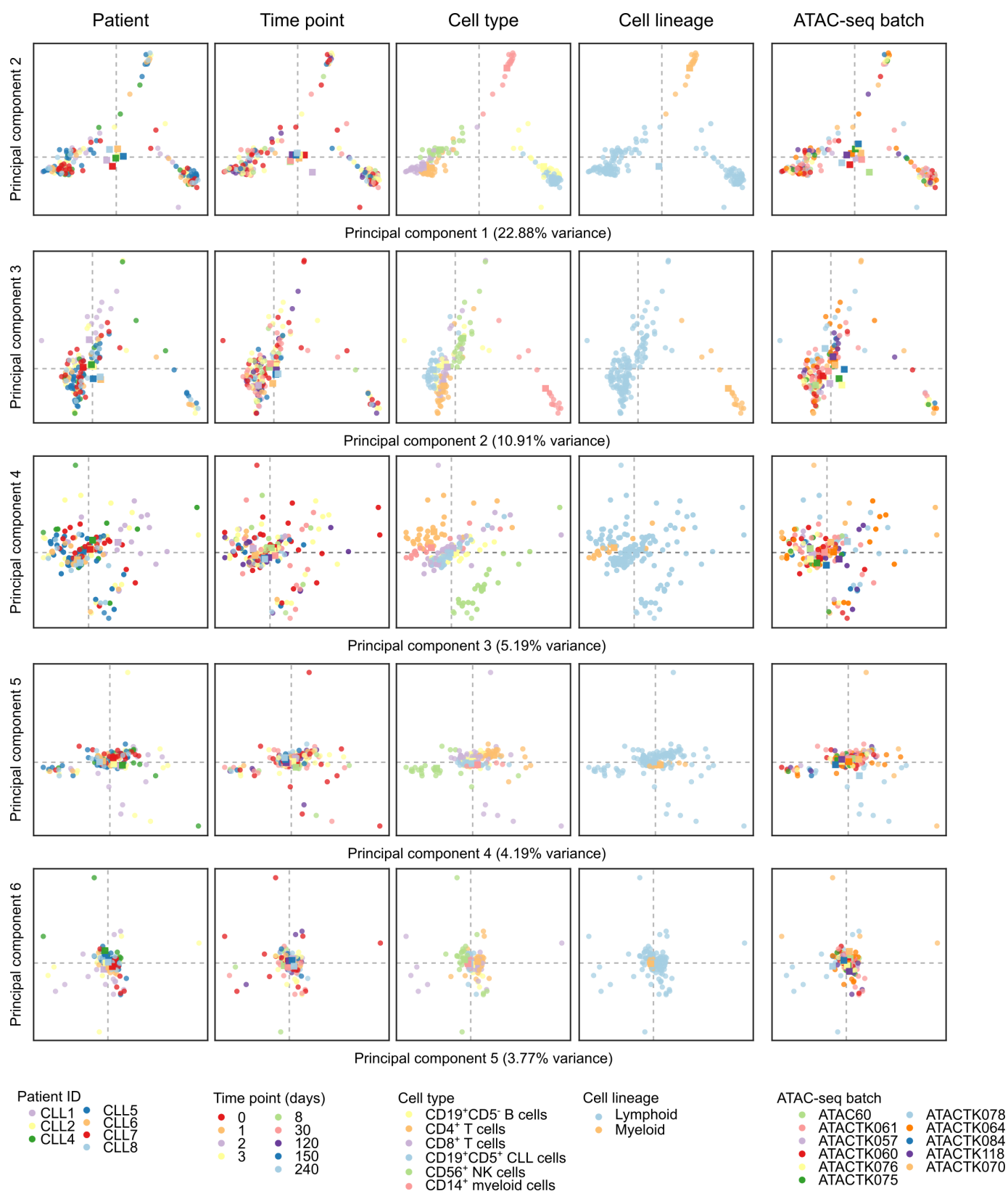
Supplementary Figure 5: Gene set enrichments for single-cell RNA-seq over the ibrutinib time course

Enrichment of differentially expressed genes over the course of ibrutinib therapy for gene sets and biological processes involved in immune regulation, calculated separately for each cell type.



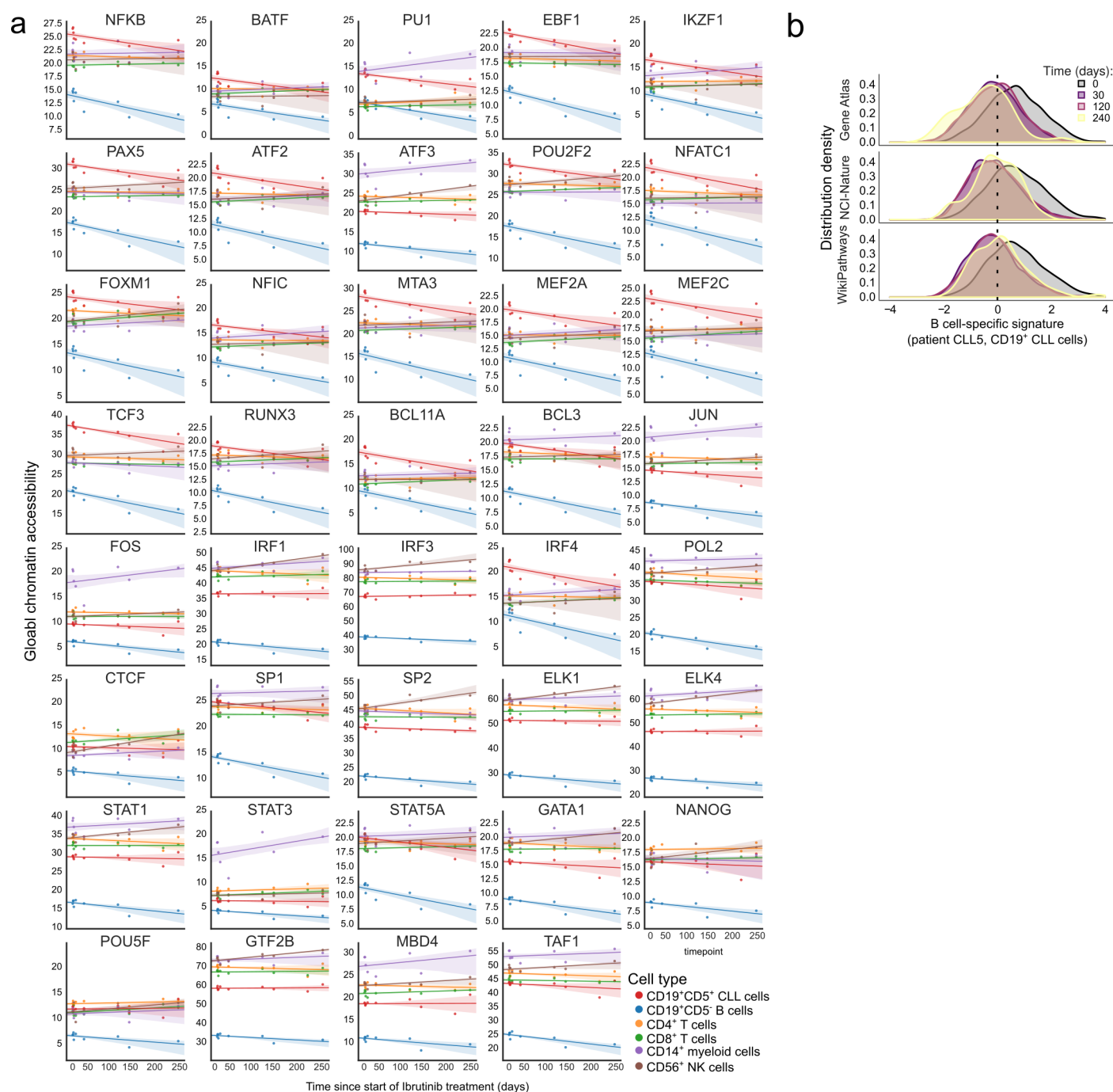
Supplementary Figure 6: Transcriptional changes of selected gene signatures upon ibrutinib treatment

Aggregate gene expression of selected gene signatures plotted over the ibrutinib time course. The list of gene signatures includes the hallmark signatures from MSigDB⁵⁹ (indicated by HM prefix) as well as sets of target genes for selected transcription factors (obtained from various sources).



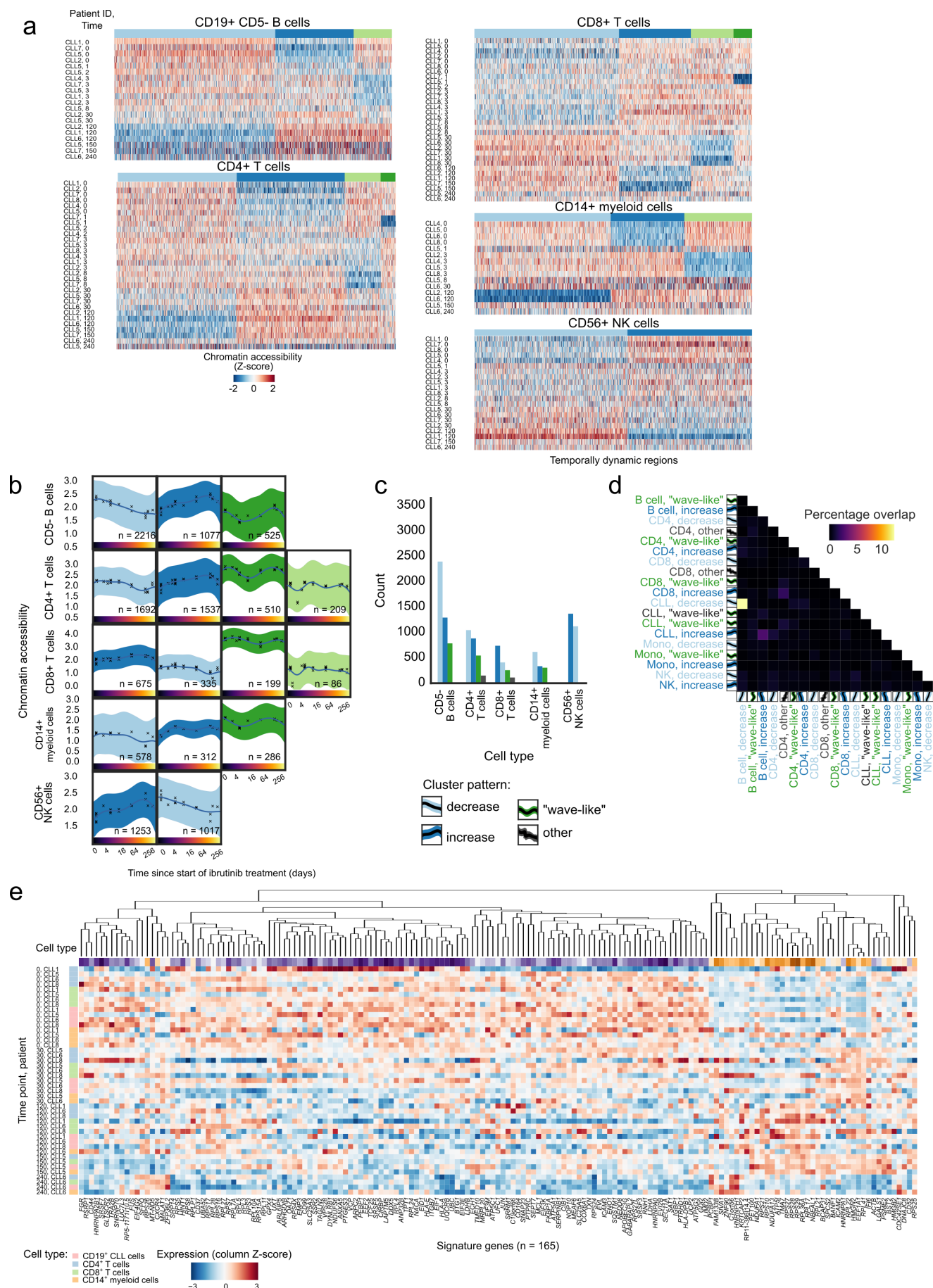
Supplementary Figure 7: Unsupervised analysis of chromatin accessibility over the ibrutinib time course

Principal component analysis of all chromatin accessibility profiles, highlighting biological and technical annotations of potential relevance. Samples are shown as circles, color-coded according to annotations, and the centroid for each annotation is shown as a color-coded square. In total, 155 samples obtained from seven patients were included in the analysis.



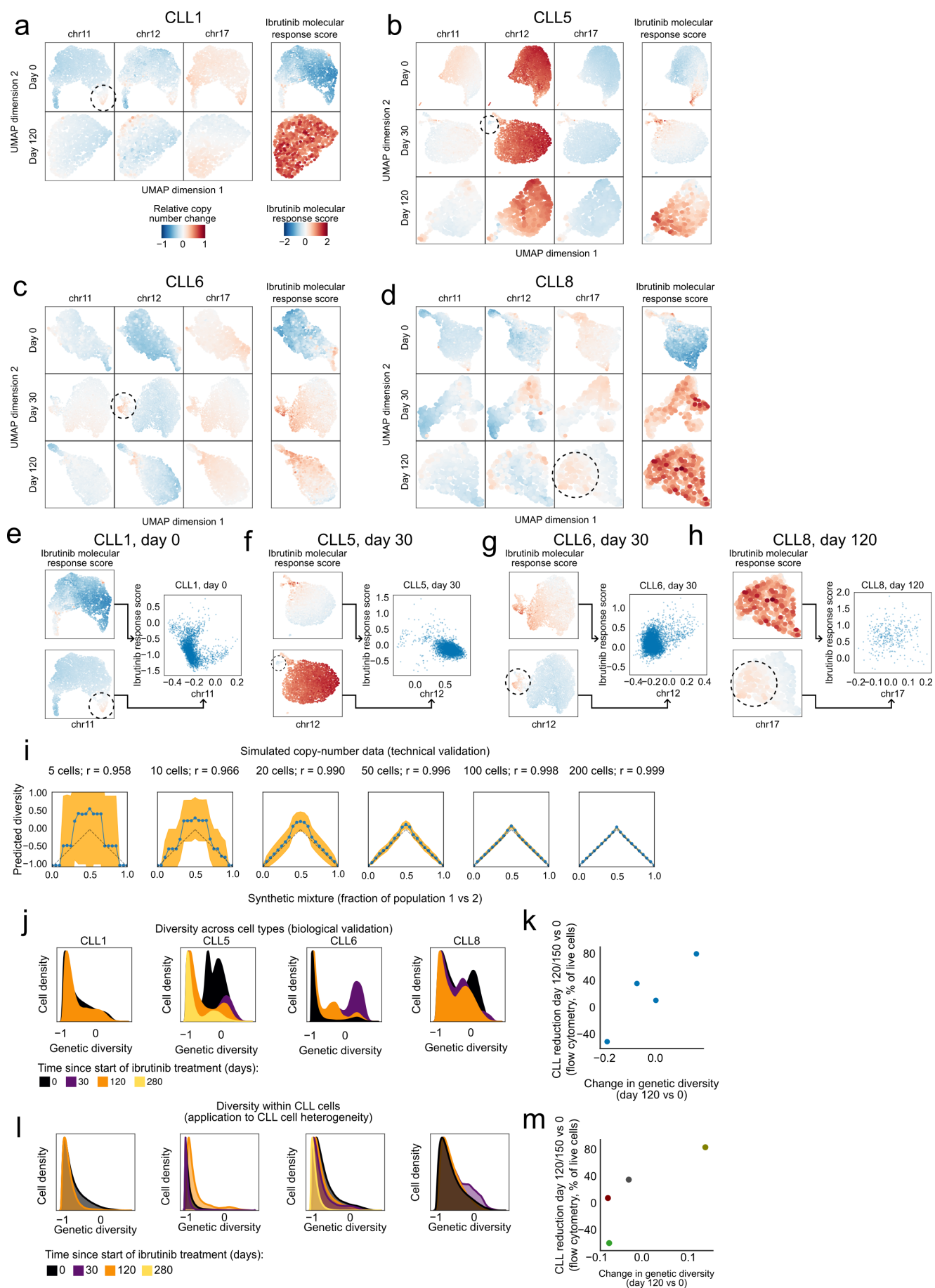
Supplementary Figure 8: Changes in transcription regulation and cell state over the ibrutinib time course

a) Line plots showing mean chromatin accessibility of regulatory regions overlapping putative binding sites of the respective transcription factors (based on publicly available ChIP-seq data) for each cell type and time point. Colored areas indicate 95 percent confidence intervals calculated over 1,000 bootstrap runs. 155 samples obtained from seven patients were included in the analysis, comprising CLL cells as well as non-malignant immune cells. **b)** Gene expression histogram across CLL cells in one patient, demonstrating the decline of B cell-specific expression signature (three alternative signatures are shown) over the time course of ibrutinib treatment. For illustration, the patient with most time points in the single-cell RNA-seq analysis (CLL5) is displayed.



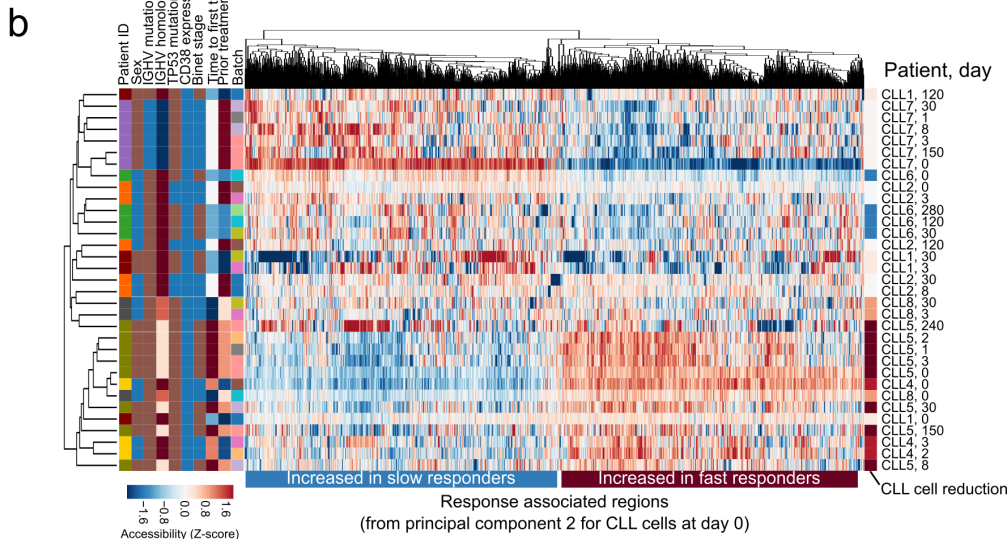
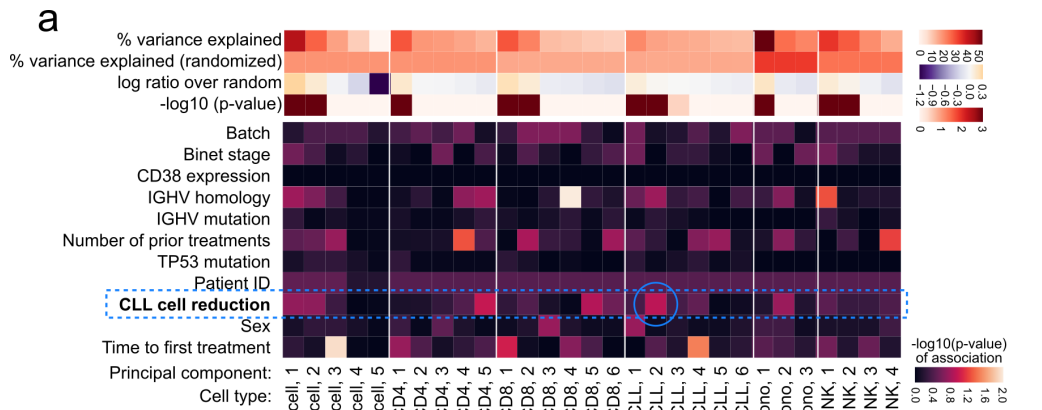
Supplementary Figure 9: Cluster analysis of regulatory regions in non-malignant immune cell types

a) Heatmaps showing chromatin accessibility at dynamically changing regulatory regions for five FACS-purified non-malignant immune cell types collected over the ibrutinib time course. Values represent column Z-scores of normalized ATAC-seq signal strength. 122 samples obtained from seven patients were included in the analysis, comprising only non-malignant immune cells. **b)** Mean chromatin accessibility across patients plotted over the ibrutinib time course for each cluster of dynamically changing regulatory regions in each cell type. Each cross represents a single sample from a single patient at a specific time point, and 95% confidence intervals are shown as colored shapes. **c)** Absolute number of dynamic regulatory regions for each cell type and cluster. **d)** Pairwise overlap of dynamic regulatory regions between cell types and clusters. **e)** Clustered heatmap showing patient-specific gene expression levels for the quiescence-like gene expression signature (**Figure 3e**), based on the single-cell RNA-seq data over the ibrutinib time course. Values represent column Z-scores of gene expression. 12 independent patient samples were included in the analysis.

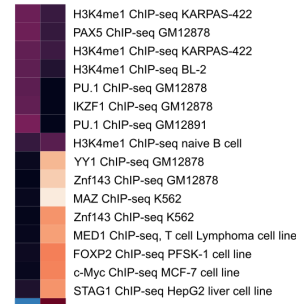


Supplementary Figure 10: Analysis of copy number and genetic diversity over the ibrutinib time course

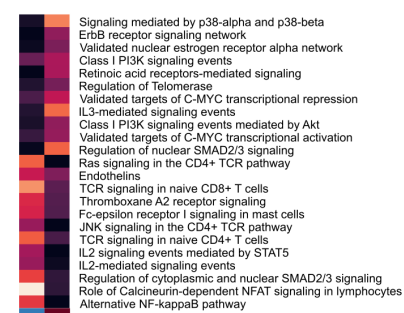
a-d) Two-dimensional similarity map (UMAP projection) based on DNA copy number profiles for single cells inferred from the single-cell RNA-seq data. These maps were calculated separately for each patient and time point. Color-coding indicates the relative copy number change for three chromosomal aberrations common in CLL (left) and for the ibrutinib molecular response score (right), i.e., the change in the CLL cell percentage on day 120/150 of ibrutinib treatment compared to day 0 as measured by flow cytometry. Genetically distinct subclones are highlighted by dashed circles. **e-h)** Scatterplots comparing selected subclonal copy number aberrations (x-axis and UMAP plots on the bottom left) with the ibrutinib molecular response score (y-axis and UMAP plots on the top left) across single cells in individual patients and time points. **i)** Accuracy of the computational approach for quantifying genetic diversity benchmarked on simulated copy number profiles that were combined at defined percentages (x-axis). Dashed lines indicate expected values (based on the simulation's known ground truth), blue lines indicate inferred values, and yellow areas represent 95th confidence intervals for the inferred values. Correlation coefficients quantify the overall agreement between expected and inferred values. **j)** Histograms showing the change in genetic diversity across all cells (i.e., CLL cells and immune cells) in the single-cell RNA-seq dataset. **k)** Scatterplot showing the correlation between the change in genetic diversity across all cells (x-axis) between time points and the cellular response to ibrutinib treatment (y-axis). **l)** Histograms showing the change in genetic diversity specifically for CLL cells. **m)** Scatterplot showing the correlation between changes in genetic diversity specifically in CLL cells (x-axis) and the cellular response to ibrutinib treatment (y-axis). Panel m is a reproduction of Figure 4b for consistency with panels j-k.



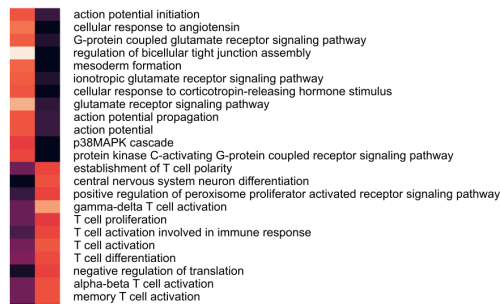
C Location overlap analysis (LOLA)



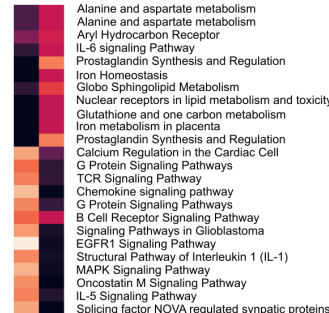
NCI-Nature



Biological Process (GO)

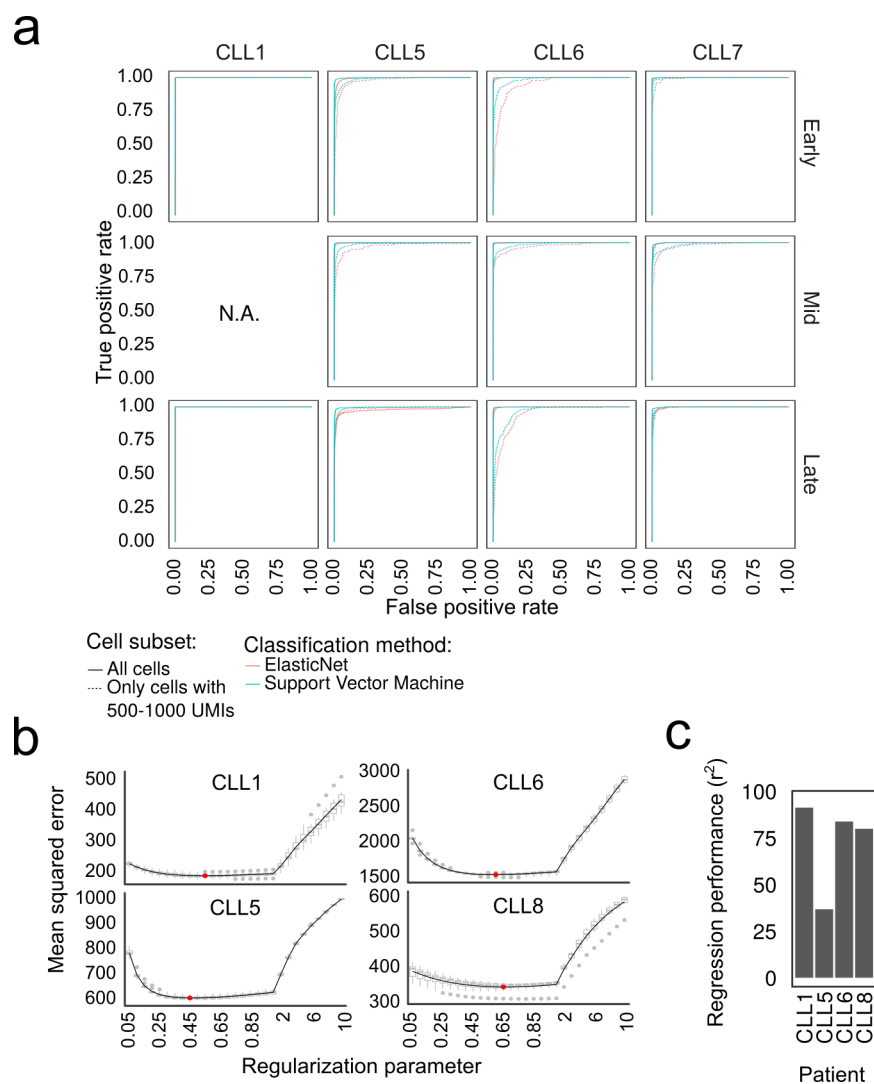


WikiPathways



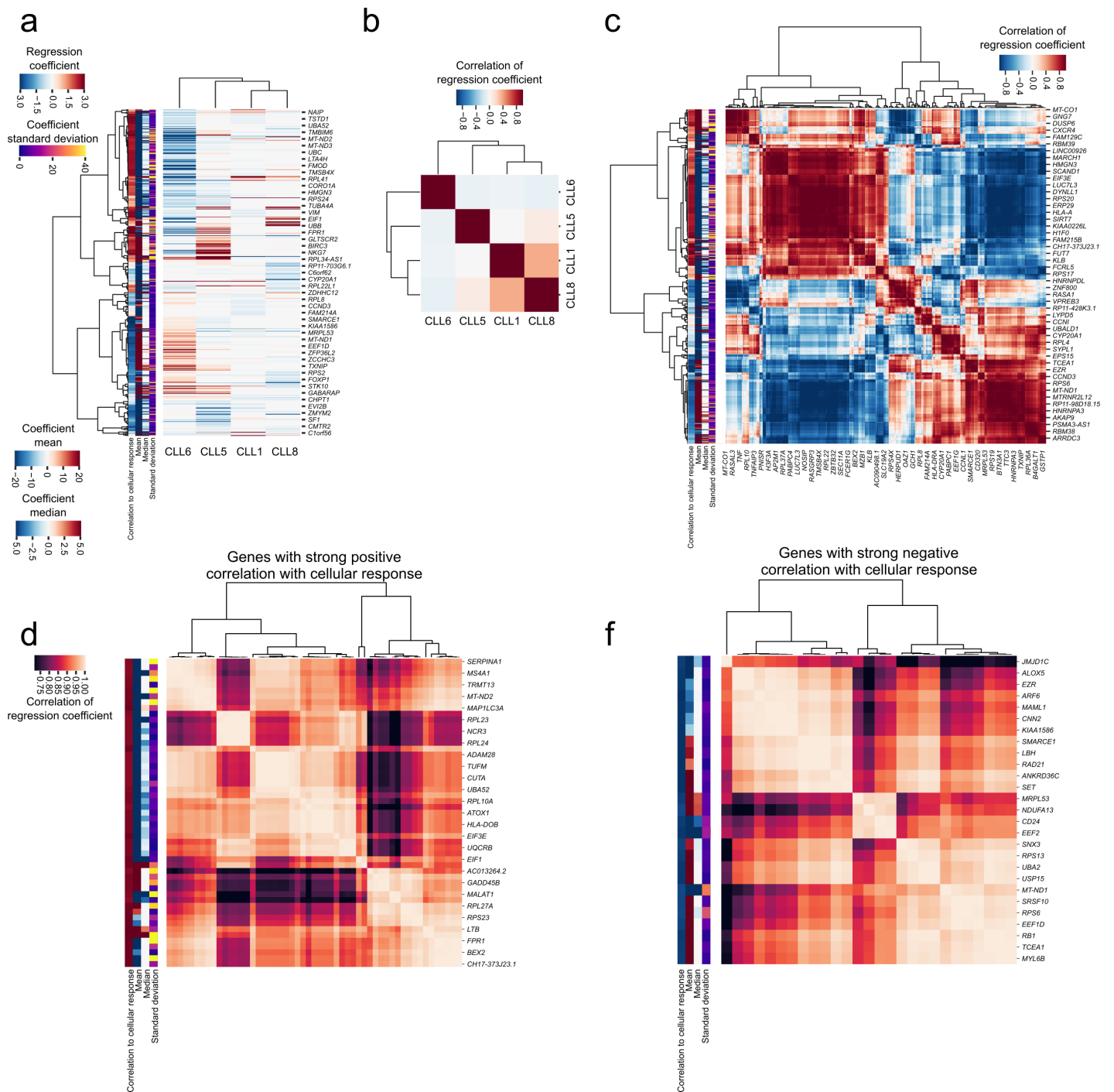
Supplementary Figure 11: Analysis of chromatin profiles and their association with the response to ibrutinib

a) Heatmaps showing the association of various clinical annotations with the principal components of the cell type specific chromatin accessibility profiles of different cell types prior to the start of ibrutinib treatment. The blue circle highlights the association between the second principal component for CLL cells and the change in the CLL cell percentage on day 120/150 of ibrutinib treatment compared to day 0, as measured by flow cytometry (y-axis). 7 independent patient samples were included in the analysis. **b)** Clustered heatmap showing patient-specific chromatin profiles for genomic regions associated with the second principal component (from panel a). 33 samples obtained from seven patients were included in the analysis. **c)** Enrichment analysis for genomic regions associated with the second principal component, separately for regions associated with a slow versus a fast response to ibrutinib treatment.



Supplementary Figure 12: Prediction of the time point of sample collection from single-cell transcriptomes

a) ROC curves showing the cross-validated test set performance of classifiers predicting the time point of sample collection based on single-cell transcriptome profiles, using two different machine learning methods (logistic regression with elastic net regularization and support vector machines) and two different thresholds for single-cell RNA-seq data quality (all cells vs. only cells with 500 to 1,000 UMIs). **b)** Optimization of the regularization parameter (lambda) for predicting the time since the start of ibrutinib treatment using elastic net regularized linear regression. Red dots indicate the chosen parameter for each patient. **c)** Cross-validated test set performance of the regression models (coefficient of determination) for predicting the time since the start of ibrutinib treatment for each patient.



Supplementary Figure 13: Analysis of regression coefficients for gene expression upon ibrutinib treatment

a) Clustered heatmap of regression coefficients estimated for each patient and each gene over the course of the ibrutinib time series. For each gene, the heatmap shows a set of summary statistics calculated across patients (median, mean, standard deviation) as well as the Pearson correlation to the reduction of CLL cells at day 120 of ibrutinib treatment. **b)** Pairwise Pearson correlation of regression coefficients between patients, aggregated across all genes. **c)** Pairwise Pearson correlation of regression coefficients, showing genes separately (as in panel a). **d-e)** Pairwise Pearson correlation of regression coefficients (as in panel c), shown separately for genes that are highly correlated with slow (panel d) or fast (panel e) response to ibrutinib, defined by an absolute Pearson correlation above 0.9. Due to space constraints, an equidistant subset of genes is shown in the heatmaps of panel a, c, and d.