Multimodal blood based proteomic profiling reveals insights into mechanisms of immunotherapy resistance

Supplementary Notes

Refitting linear mixed effects models with subtype as a covariate, using cutaneous-only melanoma samples and with stable disease as responders

We refit the linear mixed regression models using subtype as an additional covariate and noted that most time point-associated proteins remained the same suggesting that this variable did not account for the time point-associated proteomic changes. For response-associated proteins, 8 of the 27 remained significant (CDH17, ITGB6, BMP4, CPA2, MYOC, ITGAV, LEFTY2, CD1C, HMCN2), while for non-response-associated proteins, only 2 of 58 remained significant (STC1, MIA) suggesting that there may be more distinct mechanisms of nonresponse across histologies. The 137 proteins that were associated with subtypes included IGF1R, GDF15, TYRO3, BCAN and PODXL2 (**Fig. S4, Table S7-8**).

We also refit the linear mixed regression models with only cutaneous samples, and interestingly no proteins had significant response effects (**Table S9**). We do note, however, that the proteins which had significant response effects in the original models have highly correlated slopes with the same proteins in the cutaneous-only models (the cutaneous-only model response slopes tended to be lower in magnitude) (**Fig. S5a**), indicating that there is likely some association between the abundances of these proteins and response that fails to reach the level of significance. When comparing the skewness of these protein distributions in responder samples in the overall setting to the cutaneous-only setting, we note that these are similar across response proteins (**Fig. S5b**). However, we noticed that the skewness in the overall setting was higher than the skewness in the cutaneous-only (**Fig. S5c**). In other words, an arm of the protein distribution of non-responder samples that differentiates it from responder samples is lost when non-cutaneous samples are removed from the models.

There is a precedent for including stable disease with responders, so the linear mixed regression models were refitted using SD patients as responders instead of non-responders, and again no proteins had significant response effects (**Table S10**). Again we see a high correlation between the response effect slopes of the response-associated proteins in the original model and the response effect slopes of these proteins in the SD responder models (**Fig. S6a**) Among the responder samples, the skewness in the distribution of these proteins is essentially even (**Fig. S6b**), but in non-responder samples the skewness in the distribution of these proteins is in general slightly higher in the original models, compared to the SD responder models (**Fig. S6c**). This again brings up the possibility that the protein abundance distribution in non-responders is losing an arm that differentiates it from responder samples when SD samples are moved from the non-responder to the responder group.

External validation of proteomics analysis results

To validate our treatment- and response-associated plasma proteomics signatures, we re-analyzed orthogonal data sets using our linear mixed effects model methodology. We found publicly available data sets with similar structures involving plasma sampling of melanoma immunotherapy cohorts pre- and on-treatment with Babačić et al.[21] and Nuñez et al.[36]. We also utilized an internal plasma proteomics data set from a cohort of gastric cancer patients undergoing sequential chemo-immunotherapy (samples obtained from 46 patients pre-treatment, after one cycle of chemotherapy and after 7 cycles of combination chemo-immunotherapy). While the Babačić et al. and Nuñez et al. data sets included a panel of around 90 proteins, the gastric data set included approximately 3000 proteins, which is in the range of what our melanoma immunotherapy data set had.

There are 47 significant proteins from the original analysis that are present in the Babačić et al. data (**Table S12**, **Fig. S8a**). There were 15 proteins in the Babačić et al. data with significant effects. 1 with an interaction effect (with a negative coefficient), 5 with response effects (all with positive response effects), and 9 with timeline effects (all with positive timeline coefficients). IL6 was more abundant in non-responders in our data and more abundant in responders in Babačić et al. CD244, CRTAM, CXCL11, CXCL9, GZMH, CXCL10, CXCL13 and KLRD1 were more abundant post-ICB in both data sets. Babačić et al. called SD responders in their analysis, and we decided to change this to non-responders in our validation analysis to match the approach with our data. On-treatment samples were also collected at 3 weeks while ours were collected at 6 weeks, which may help explain limited overlap between significant effects, especially treatment-associated proteins.

There are 38 significant proteins from the original analysis that are present in the Nuñez et al. data (**Table S13**, **Fig. S8b**). There were 16 proteins in the Nuñez et al. data with significant effects. 13 with interaction effects (2 with negative coefficient, 11 with positive coefficient), 8 with response effects (7 negative coefficients, 1 positive coefficient), and 6 with timeline effects (3 with negative coefficient, 3 with positive coefficient). 2 of the significant Nuñez proteins were also significant in the original analysis in the same direction (KITLG was higher in responders in both cases and CXCL9 had positive timeline effects in both cases). However, there were some proteins with positive interaction effects in the Nuñez data that had positive timeline effects in our data (CXCL9, CXCL11, CXCL10, MMP10, TNF, IFNG, CCL20, IL10RA, IL17A, CDCP1, IL10RA). There were four more significant proteins from Nuñez et al. that overlapped with our significant proteins but that didn't have the same type of effect. The structure of the data was different because there were 4 sampling time points centered around detection of Tox (pre-ICI, d7-d14, 1 month pre-Tox, and Tox), which may help explain limited overlap between our results and theirs.

80 proteins that were significant for timeline effects and more abundant on-treatment were shared between the gastric and melanoma data sets (**Table S14**, **Fig. S8c**). GO term enrichment for those 80 overlapping proteins showed that they are enriched in inflammatory processes (**Fig. S8d**), which indicates that the upregulation of these proteins may be a common feature of systemic response in circulation to therapeutics in cancer (including immunotherapy).

There was limited overlap between proteins with other effects, which may indicate that global host response to immunotherapy may be shared across cancers, response- and non-response-associated proteins may differ from disease to disease. CXCL16 and PAEP were both more abundant pre-treatment in both data sets. The melanoma-only timeline proteins were most enriched in other immune-related processes, especially myeloid and cytokine signaling (**Fig. S8e**), while the gastric-only timeline proteins are most enriched in a lot of adhesion/ECM-related processes (**Fig. S8f**). This may imply that immunotherapy elicits specifically T cell activation/proliferation related changes in circulation, while melanoma may enrich for additional immune processes and gastric may enrich for EMT-related processes.

Analysis of circulating immune cell proportions data

To leverage our relative circulating immune cell proportions data, we correlated modules of proteins from our linear mixed effects models (response proteins, non-response proteins, and four hierarchically clustered post-ICB protein modules shown in **Fig. 2c**) using cosine similarity with the relative immune cell proportions from our PBMC data. We noticed specific patterns of correlation between the protein modules and immune proportions, such as CD8 T effector memory cells and CD8 TEMRA cells correlating with the response proteins, and Tregs correlating with the non-response proteins (**Fig. S9a**). As both of these data derive from peripheral blood, apparent correlations between immune cell proportions and proteins in circulation may be due system drivers of both. As such, it is important to acknowledge that relationships between immune cell proportions and proteins does not imply that one inherently causes the other but instead that these may be co-regulated changes reflective of systemic response to immunotherapy.

We additionally fit linear mixed effects models on the relative immune proportions from our PBMC data using the same strategy as with the Olink NPX values, in which we used patient response, time point and an interaction term to predict each relative immune proportion (**Fig. S9b, Table S15**). While no proportions had terms that reached the level of significance following multiple hypothesis correction, some proteins had uncorrected p-values below 0.05 and thus showed a trend toward significance. Interestingly, response was significantly associated with naive CD8 T cells, indicating that the abundance of this subset in circulation may have some bearing on favorable melanoma ICB outcomes.

Ki67 levels in PD-1+CD8+ cells have been reported to be associated with response to anti-PD1 therapy[18], and as such we examined our immune cell proportion data to determine whether there was a correlation between Ki67 levels in PD-1+ CD8+ cells in post-ICB samples. Normalizing cell subpopulation fractions by tumor burden has been used as an approach to find correlations with response[18]. We elected to perform this normalization using lactate dehydrogenase (LDH), which is a more clinically actionable measurement of tumor bulk. Indeed, we noticed that there was a significant correlation between the baseline patient composite score (calculated using scanpy score_genes function) of these circulating melanoma-specific proteins (**Methods**) and baseline LDH level in these patients, and the melanoma protein score had a less constrained range than the LDH level, signifying that this may be a more nuanced reflection

of tumor bulk (**Fig. S9c**). This approach recapitulated PD1+ CD8+ TEMRA cells having significantly different proportions of Ki67+ cells between responders and non-responders (**Fig. S9d)**, and additionally showed a strong trend of responders having overall PD1+CD8+ cells having higher proportions of Ki67+ cells compared to non-responders (**Fig. S9e**).

Supplementary Data Legends

Supplementary Data 1. Clinical metadata of patient samples across plasma proteomics cohort.

Supplementary Data 2. Information on TMB testing in patients for which proteomic data was available.

Supplementary Data 3. Plasma proteomics normalized proteomic expression (NPX) matrix across patient samples.

Supplementary Data 4. Relative immune proportions (mass cytometry) profiled from PBMC samples across patient samples.

Supplementary Data 5. Anatomic sites of metastases for patients for which proteomic data was available.

Supplementary Data 6. Proteins that had at least one significant linear mixed effect model covariate. A one is indicated in the appropriate column if that protein was significant for a particular term. Slope coefficients are shown for each term in the rightmost columns. P-values shown are associated with the F-statistic for the fixed effect of the corresponding variable in the mixed-effect model. Benjamini-Hochberg was used for false discovery control.

Supplementary Data 7. Proteins that had at least one significant covariate in the linear mixed effects model with the added subtype term. A one is indicated in the appropriate column if that protein was significant for a particular term. Slope coefficients are shown for each term in the rightmost columns. P-values shown are associated with the F-statistic for the fixed effect of the corresponding variable in the mixed-effect model. Benjamini-Hochberg was used for false discovery control.

Supplementary Data 8. Comparison of proteins significant for response in the original linear mixed effects models and in the linear mixed effects models with the added subtype term.

Supplementary Data 9. Proteins that had at least one significant covariate in the cutaneous-only linear mixed effects models. P-values are indicated in the appropriate column for each covariate. Slope coefficients are shown for each term in the rightmost columns. P-values shown are associated with the F-statistic for the fixed effect of the corresponding variable in the mixed-effect model. Benjamini-Hochberg was used for false discovery control.

Supplementary Data 10. Proteins that had at least one significant covariate in the SD responder linear mixed effects models. P-values are indicated in the appropriate column for each covariate. Slope coefficients are shown for each term in the rightmost columns. P-values shown are associated with the F-statistic for the fixed effect of the corresponding variable in the mixed-effect model. Benjamini-Hochberg was used for false discovery control.

Supplementary Data 11. Top 30 genes in each post-ICB protein non-negative factorization module with six components.

Supplementary Data 12 Proteins that had at least one significant linear mixed effect model covariate in the Babačić et al. data. A one is indicated in the appropriate column if that protein was significant for a particular term. Slope coefficients are shown for each term in the rightmost columns. P-values shown are associated with the F-statistic for the fixed effect of the corresponding variable in the mixed-effect model. False discovery control was not performed.

Supplementary Data 13 Proteins that had at least one significant linear mixed effect model covariate in the Nuñez et al. data. A one is indicated in the appropriate column if that protein was significant for a particular term. Slope coefficients are shown for each term in the rightmost columns. P-values shown are associated with the F-statistic for the fixed effect of the corresponding variable in the mixed-effect model. False discovery control was not performed.

Supplementary Data 14 Proteins that had at least one significant linear mixed effect model covariate in the gastric data. A one is indicated in the appropriate column if that protein was significant for a particular term. Slope coefficients are shown for each term in the rightmost columns. P-values shown are associated with the F-statistic for the fixed effect of the corresponding variable in the mixed-effect model. Benjamini-Hochberg was used for false discovery control.

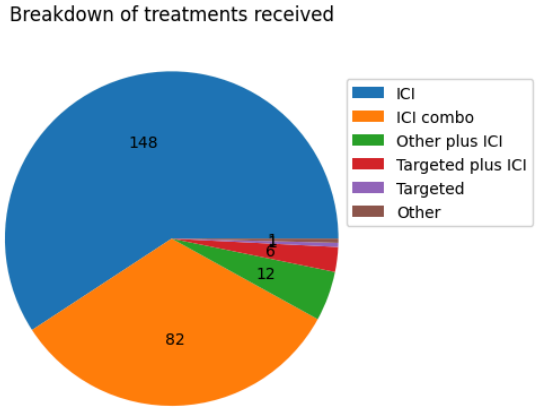
Supplementary Data 15. P-values and slope coefficients for linear mixed effects models predicting PBMC relative immune proportions. P-values shown are associated with the F-statistic for the fixed effect of the corresponding variable in the mixed-effect model. False discovery control was not performed.

Supplementary Data 16. CellPhoneDB output file showing cell-cell interactions with significant means.

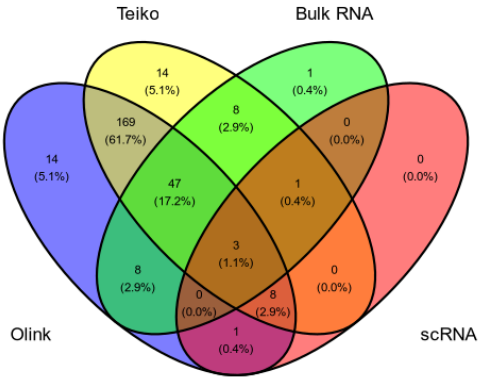
Supplementary Data 17. Pancreatitis incidence by Patient ID in the cohort, with a 1 designating patients with pancreatitis and a 0 designating patients without pancreatitis.

Supplementary Figures

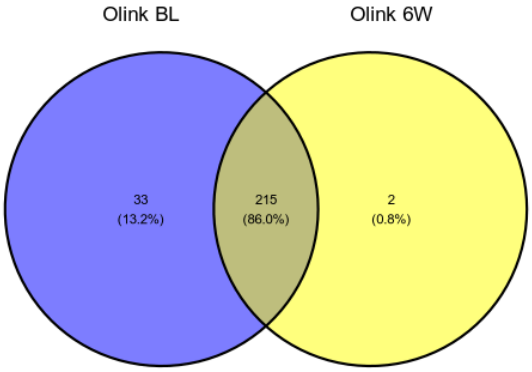
A



B



C



D

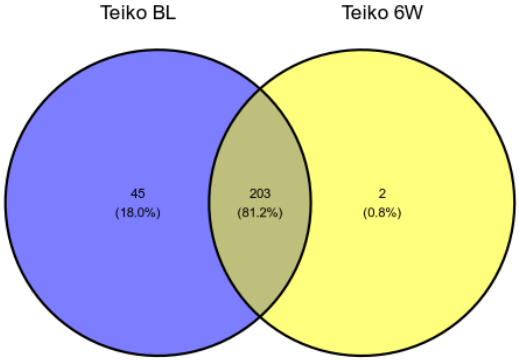


Figure S1. Cohort summary and comparison between proteomics and circulating relative immune proportions analysis. **a,** Breakdown of treatment type received by each patient whose plasma underwent proteomic profiling. (ICI, immune checkpoint inhibition) **b,** Number of patients in the melanoma immunotherapy cohort for which Olink (plasma proteomics), Teiko (PBMC relative immune cell proportions), Bulk RNA-seq and single-cell RNA-seq were available. Numbers in overlapping segments represent quantities of patients for which more than one data modality was available. **c,** Number of patients from the plasma proteomic cohort for which one or both of pre-treatment baseline (BL) and 6 weeks on-treatment (6W) time points were available. **d,** Number of patients from the PBMC relative immune proportions (Teiko) cohort for which one or both of pre-treatment baseline (BL) and 6 weeks on-treatment (6W) time points were available.

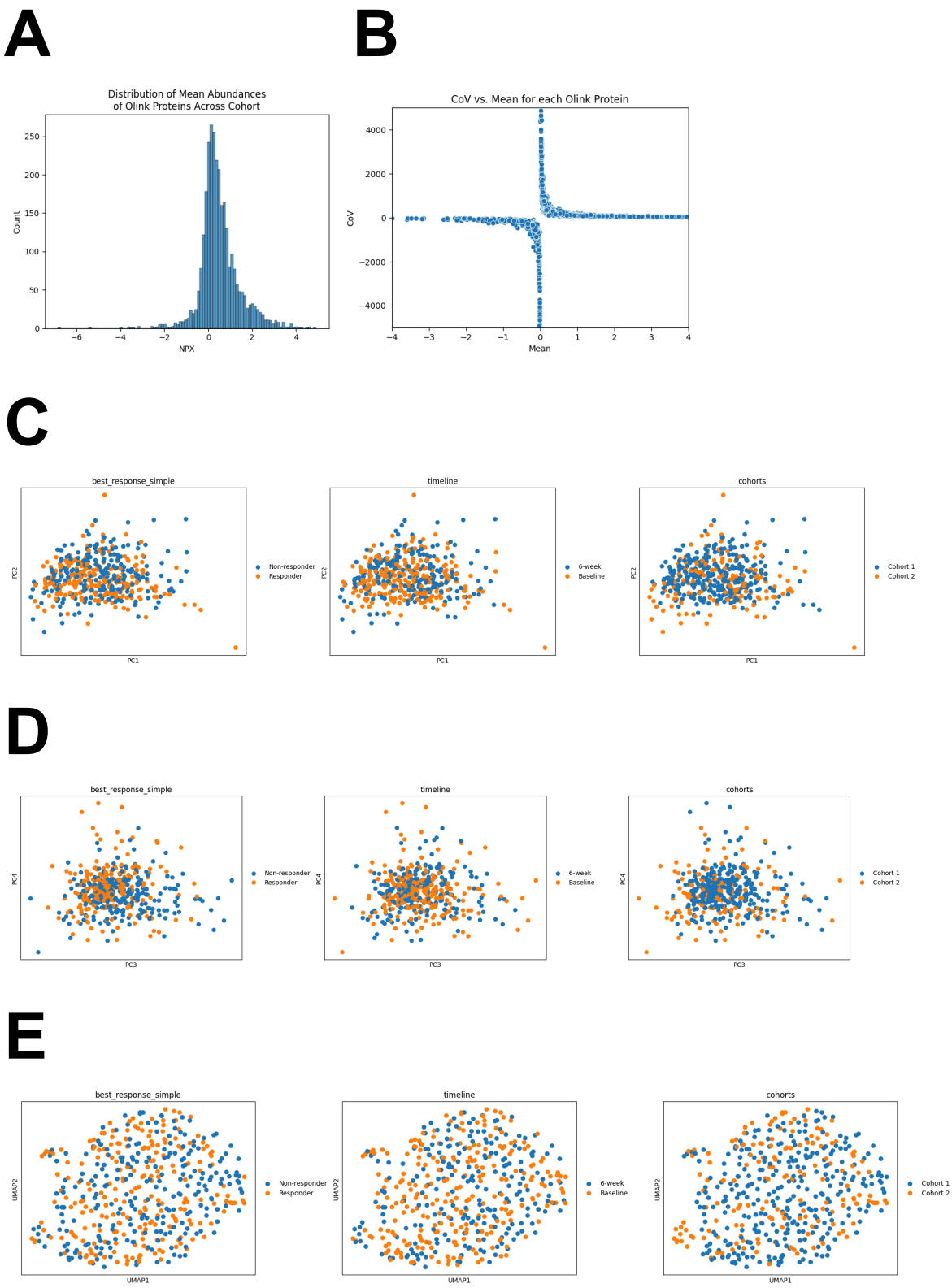


Figure S2. Plasma proteomics data quality analysis. **a,** Distribution of normalized protein expression (NPX) means across samples. **b,** Coefficient of variation (CoV) for each protein

plotted against its mean. **c-e**, PC (**c,d**) and UMAP (**e**) clustering of proteomics samples, colored by response (left), treatment time point (middle) and batch (right).

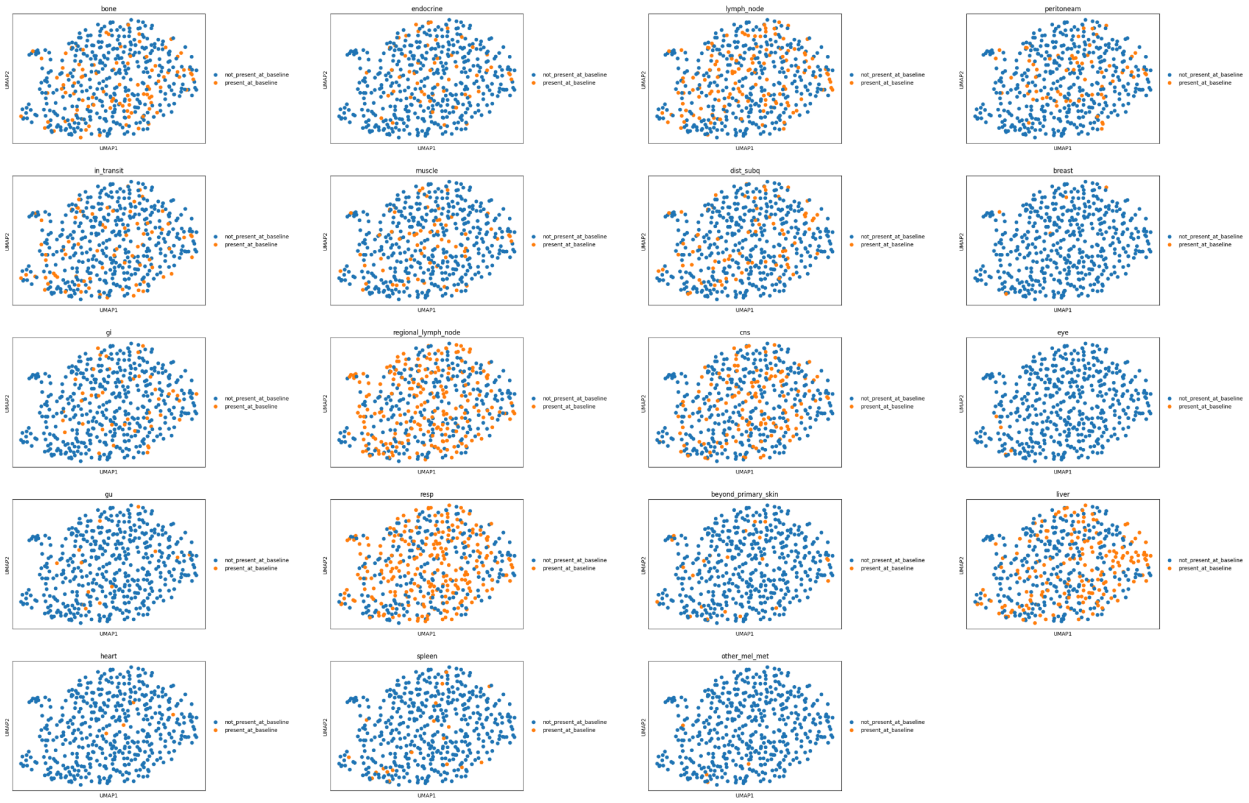


Figure S3. Clustering of plasma proteomics samples on the basis of metastatic site. UMAP clustering of proteomics samples, colored by whether a patient from which the sample was derived had metastasis at a specific anatomic site.

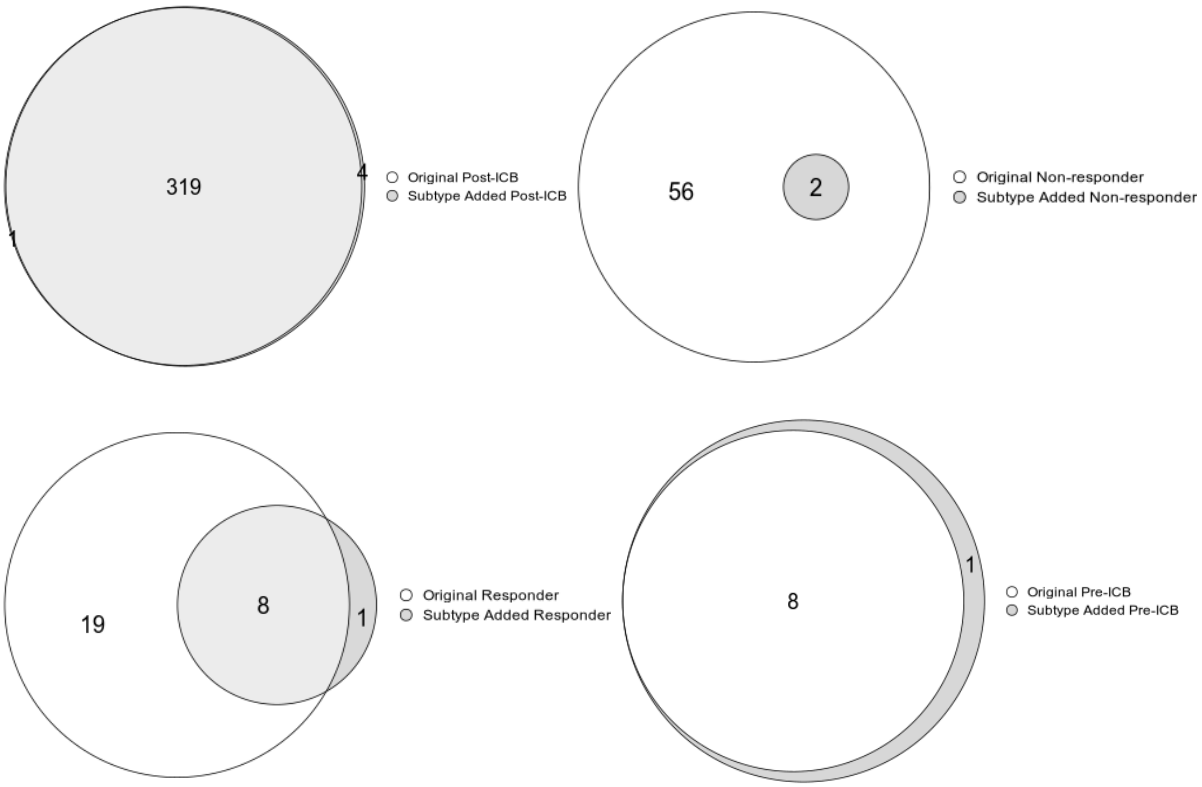


Figure S4. Overlap of proteins with significant response or time point effects in original linear mixed effects models vs. linear mixed effects models with an additional subtype covariate. Top left, overlapping post-ICB proteins; bottom left, overlapping pre-ICB proteins; top right, overlapping non-responder proteins; bottom right, overlapping responder proteins.

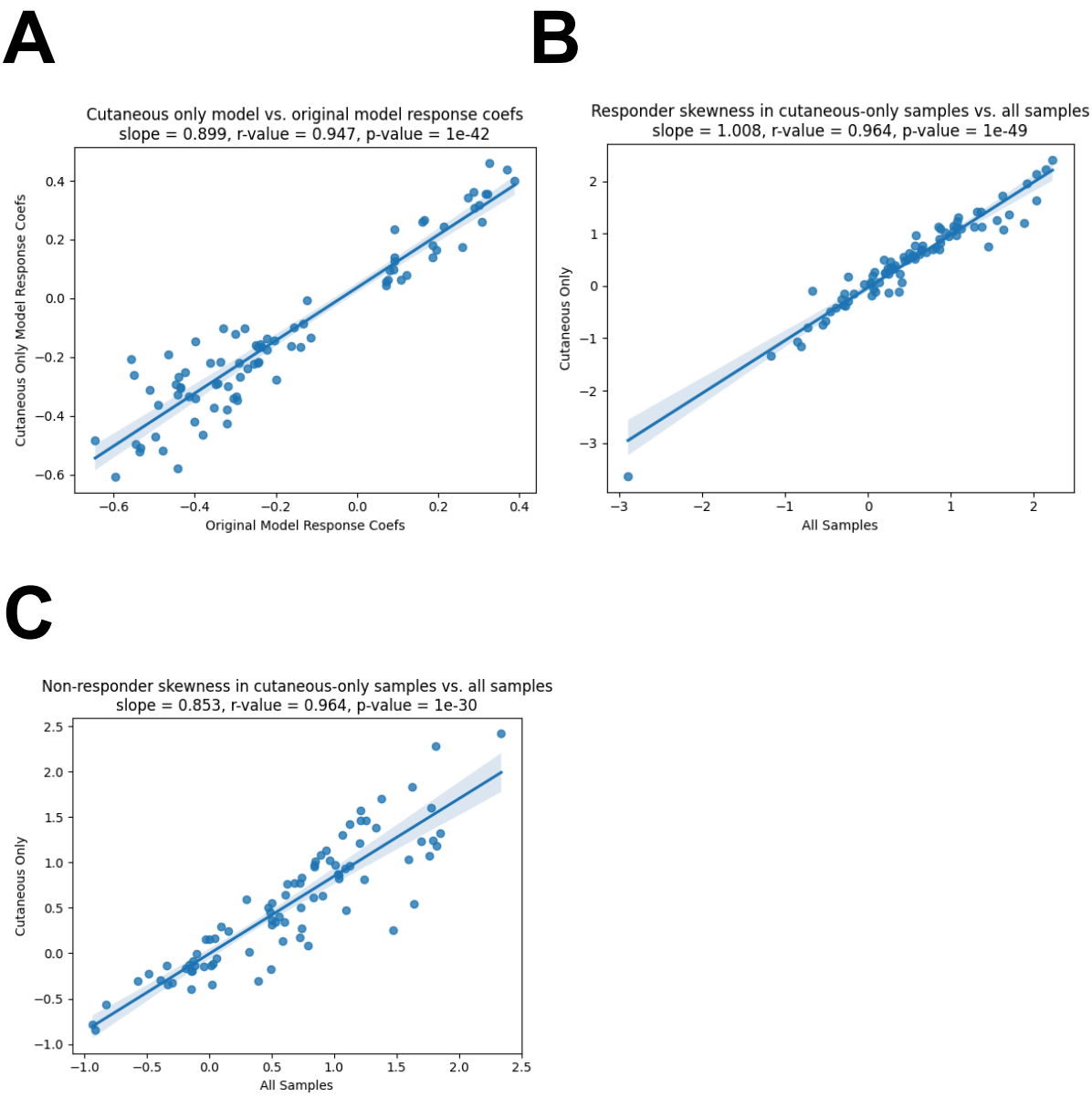


Figure S5. Linear mixed effects model comparisons between original and cutaneous-only models. **a,** Comparison of response coefficients in original and cutaneous-only linear mixed effects models, among proteins from original linear mixed effects models that had significant response effects. **b,** Comparison of responder distribution skewness in total and cutaneous-only sample groups among proteins from original linear mixed effects models that had significant response effects. **c,** Comparison of non-responder distribution skewness in total and cutaneous-only sample groups among proteins from original linear mixed effects models that had significant response effects.

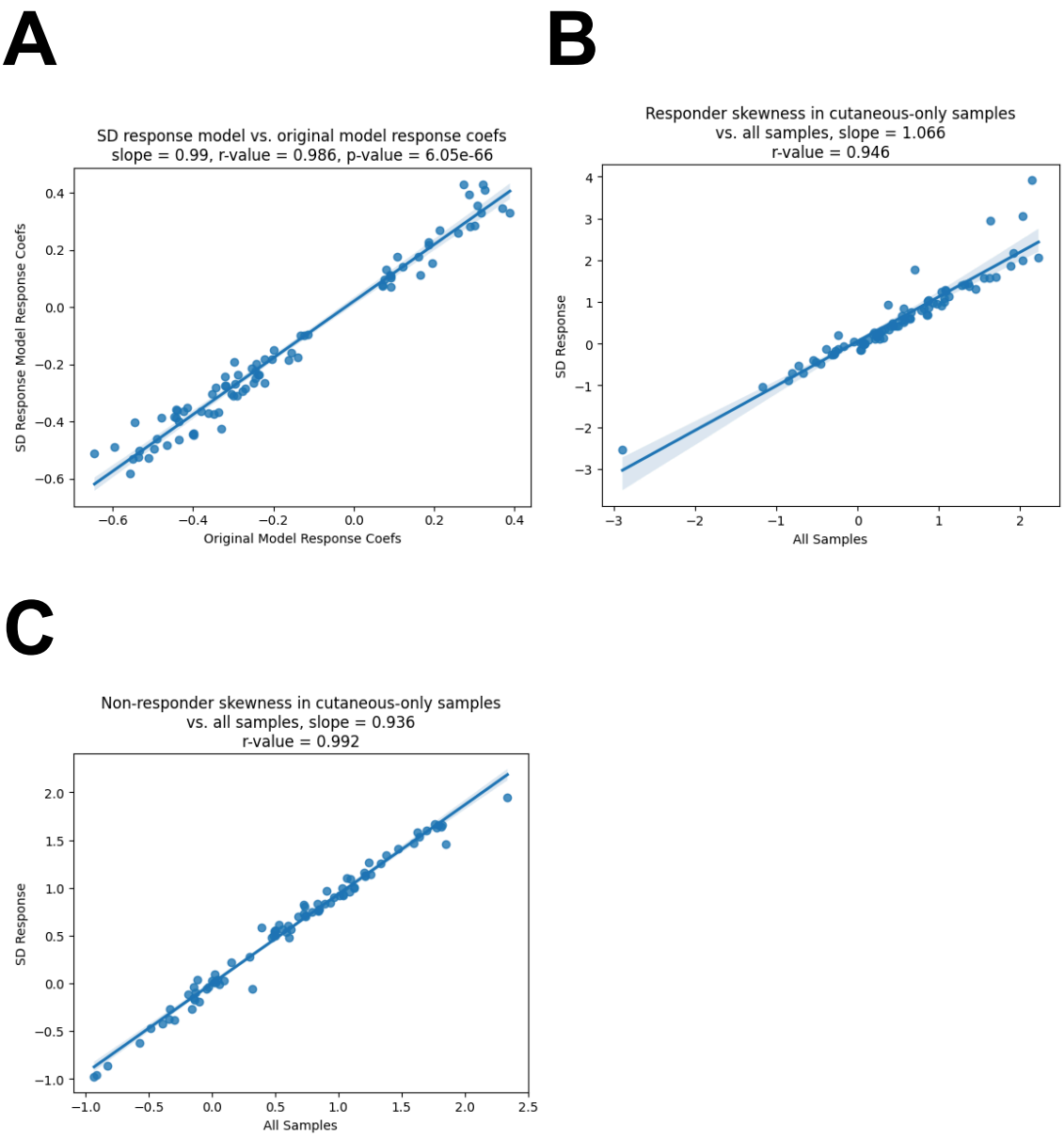


Figure S6. Linear mixed effects model comparisons between original and SD responder models. **a,** Comparison of response coefficients in original and SD responder linear mixed effects models, among proteins from original linear mixed effects models that had significant response effects. **b,** Comparison of responder distribution skewness in total and SD-responder sample groups among proteins from original linear mixed effects models that had significant response effects. **c,** Comparison of non-responder distribution skewness in total and SD-responder sample groups among proteins from original linear mixed effects models that had significant response effects.

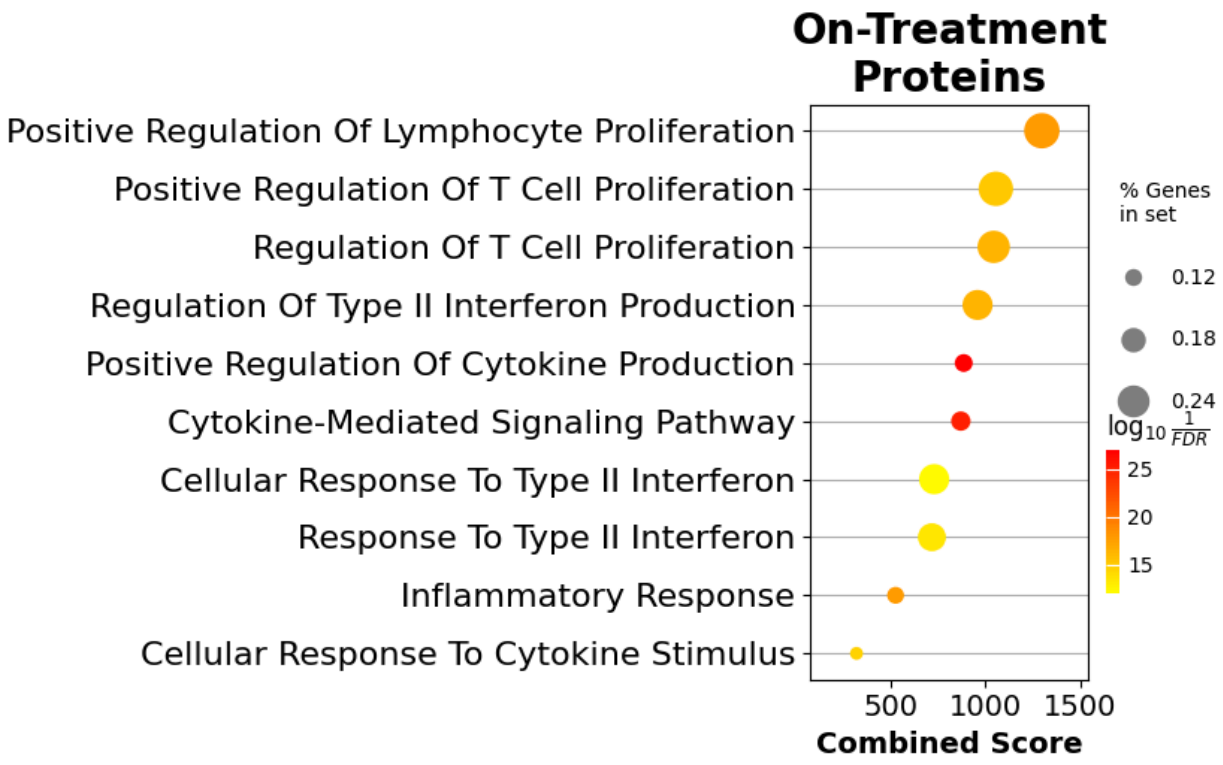


Figure S7. Top gene ontology (GO) enrichment terms on the on-treatment proteins using the GO Biological Processes 2023 library. GO score shown in x-axis, percent of on-treatment genes in the gene set shown by the size of each dot, -log(FDR) of enrichment shown by color of each dot (Benjamini Hochberg false discovery correction).

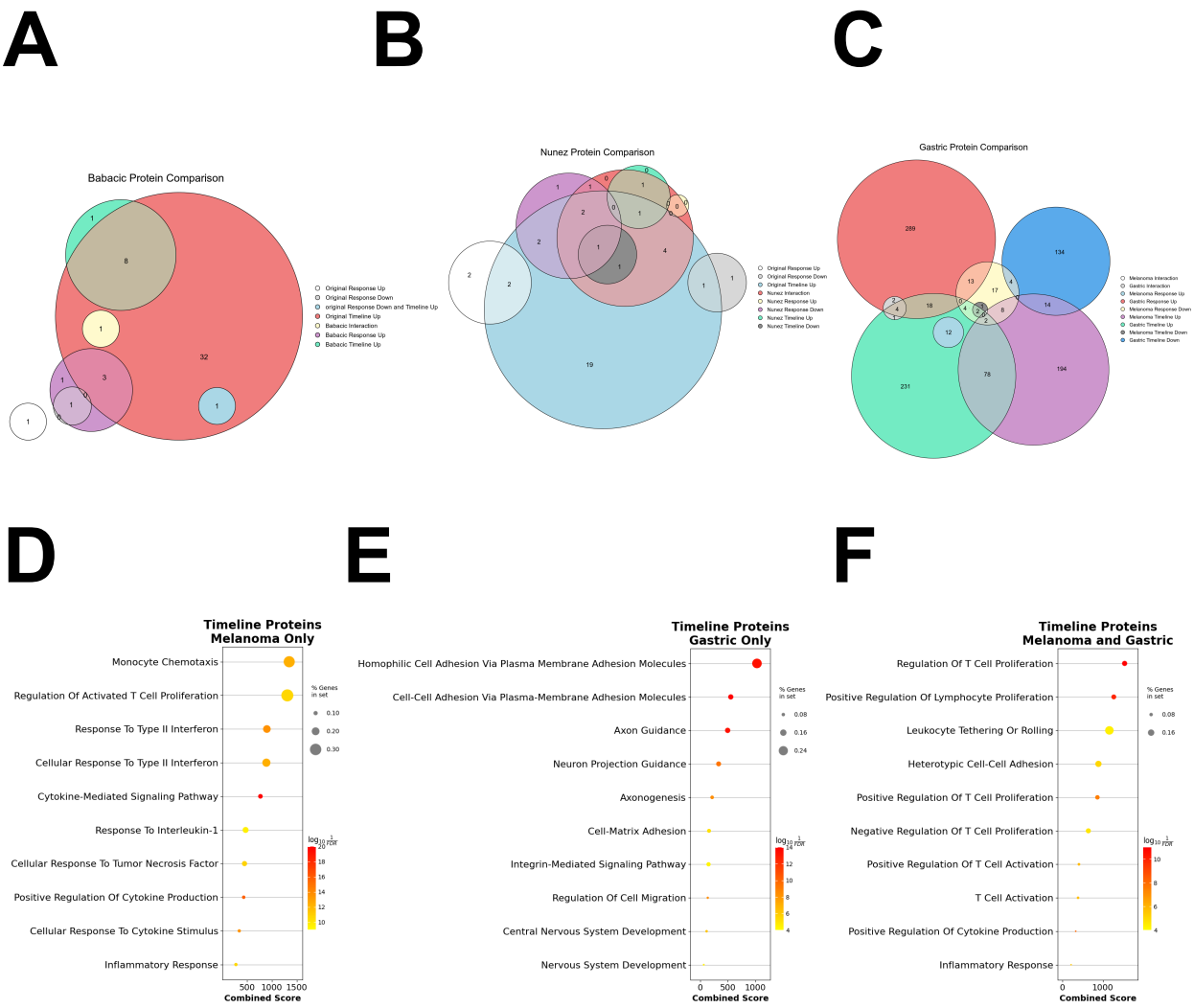


Figure S8 Comparison of original linear mixed effects model with linear mixed effects models on external data. **a-c,** Overlap of significant proteins in the melanoma plasma proteomics analysis and the (**a**) Babačić et al. data, (**b**) Nuñez et al. data and (**c**) gastric chemo-immunotherapy data, with overlapping segments of Venn Diagrams representing shared proteins between the indicated categories in the legend. **d-f,** Gene ontology enrichment using GO Biological Processes library of post-ICB proteins that were only found in the (**a**) melanoma or (**b**) gastric data sets, or that were (**c**) shared between both data sets.

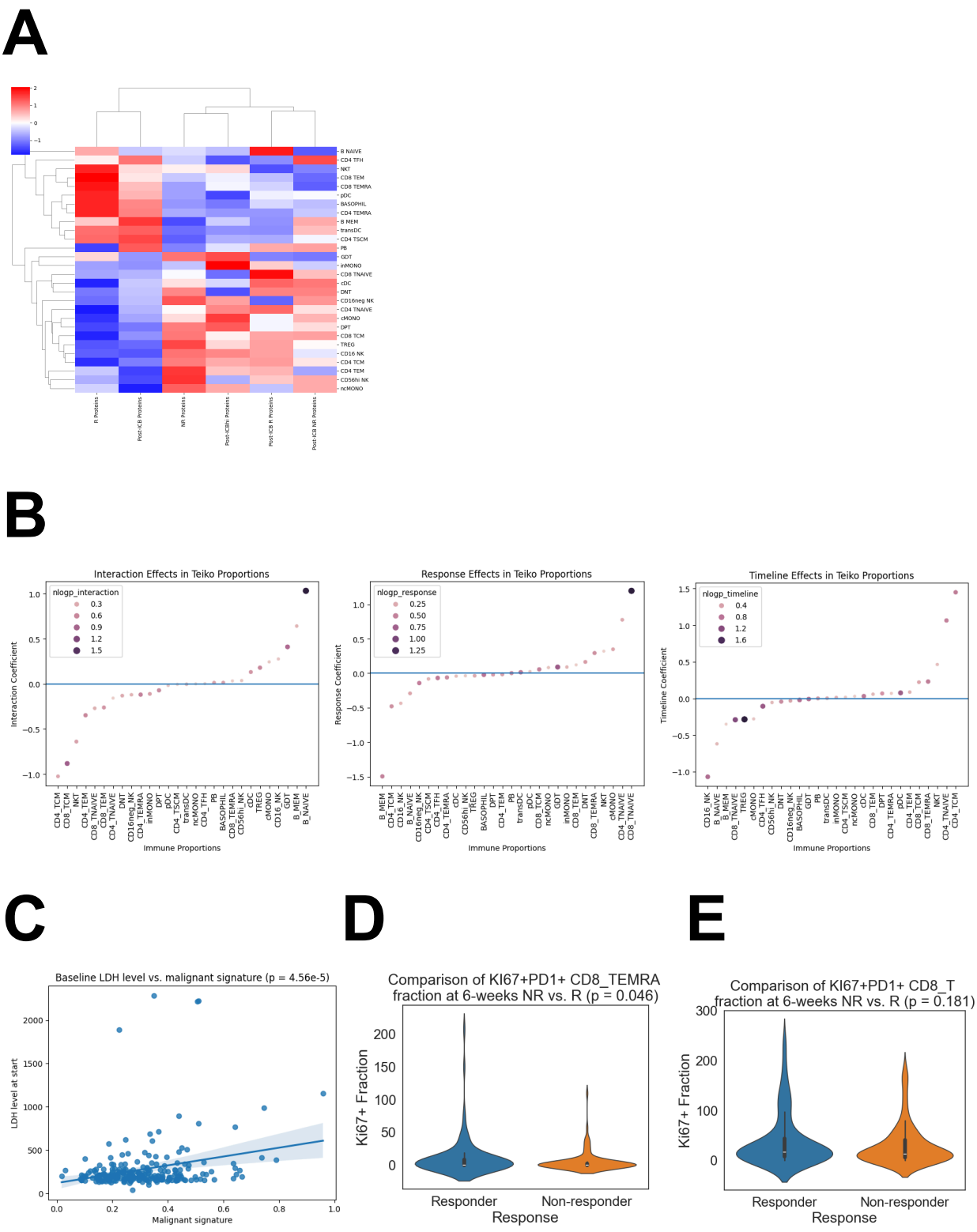


Figure S9. Analysis of circulating relative immune proportions data. **a,** Hierarchically-clustered heatmap of cosine similarity between composite protein group scores from linear mixed effects models (and hierarchical clustering of on-treatment proteins) and PBMC relative immune

proportions. **b,** PBMC relative immune proportions linear mixed effects model coefficients, ordered by magnitude and sign, with -log(p-value) shown by color of dot. Left, response effects. Middle, time point effects. Right, interaction effects. **c,** Correlation of LDH level and composite malignant proteomic score across samples. **d,** Association of KI67+PD1+ CD8 TEMRA cells with treatment response on-treatment, normalized by LDH level. Wilcoxon rank-sum test, p = 0.046. **e,** Association of overall KI67+PD1+ CD8 T cells with treatment response on-treatment, normalized by LDH level. Wilcoxon rank-sum test, p = 0.181.

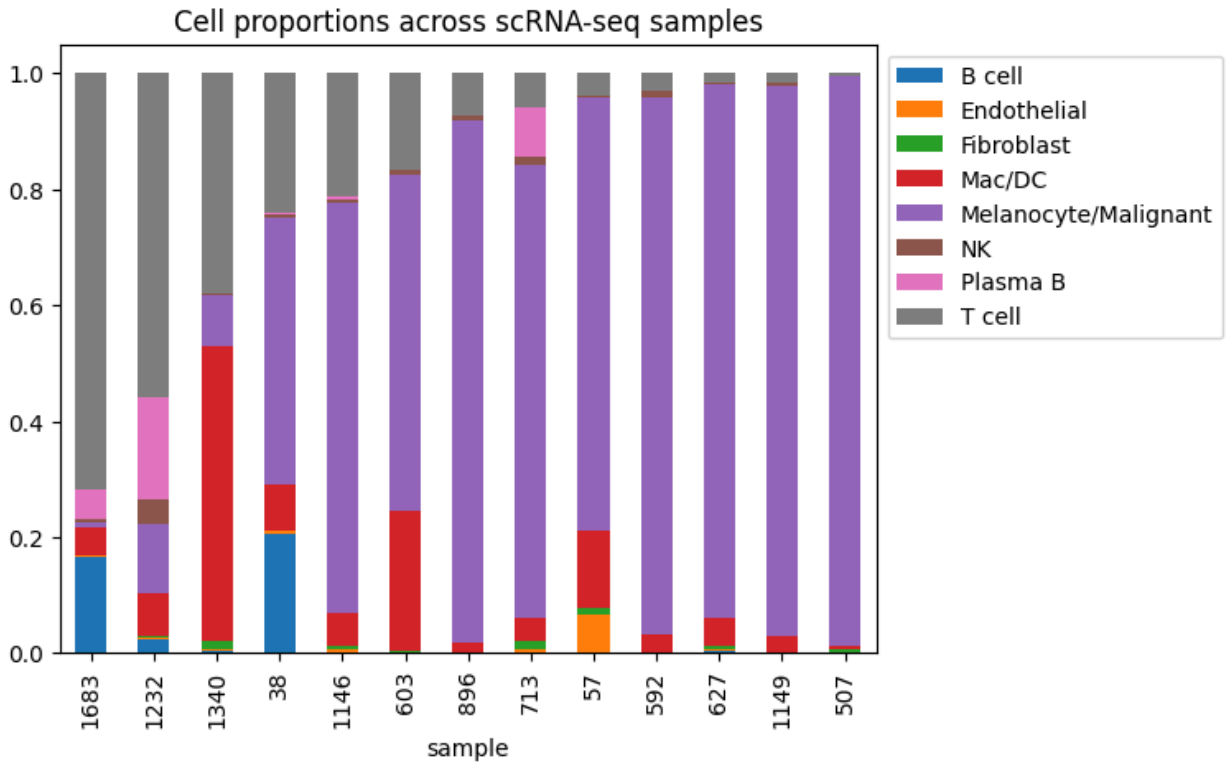
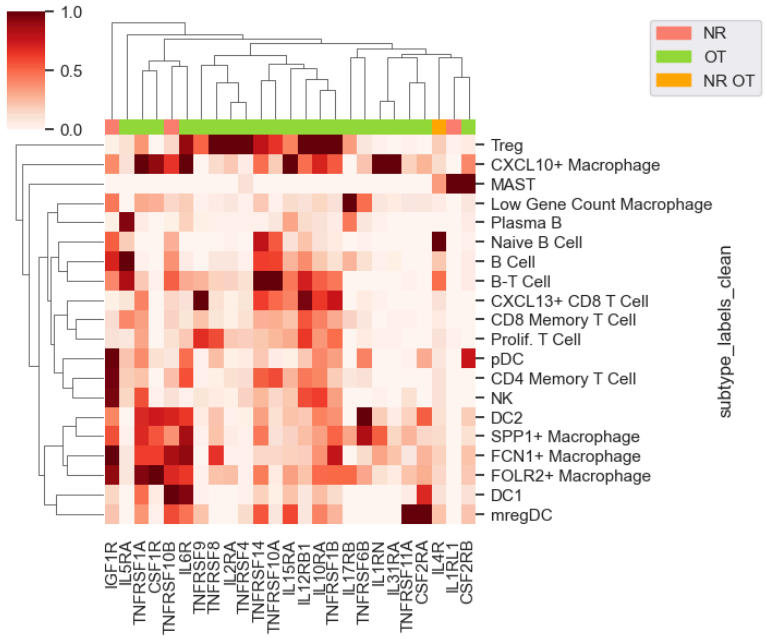


Figure S10. Cell type composition in each single-cell RNA-seq sample.

A



B

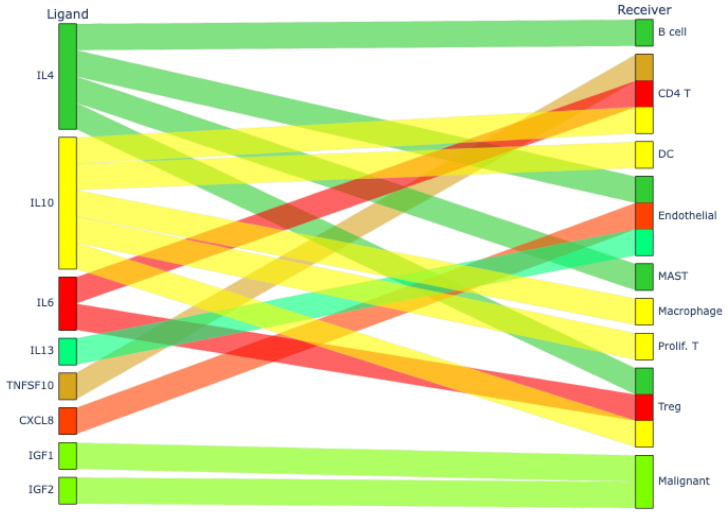


Figure S11. Gene expression analysis across single-cell RNA-sequencing cells for cytokines and chemokines among the circulating proteins of interest. **a,** Heatmap visualization of expression of receptors for cytokines, chemokines and other secreted immune-related genes across single-cell subtypes. **b,** Flow diagram showing potential interactions involving cytokines

and chemokines among non-responder proteins based on minimum expression of only receptors in receiver cells.

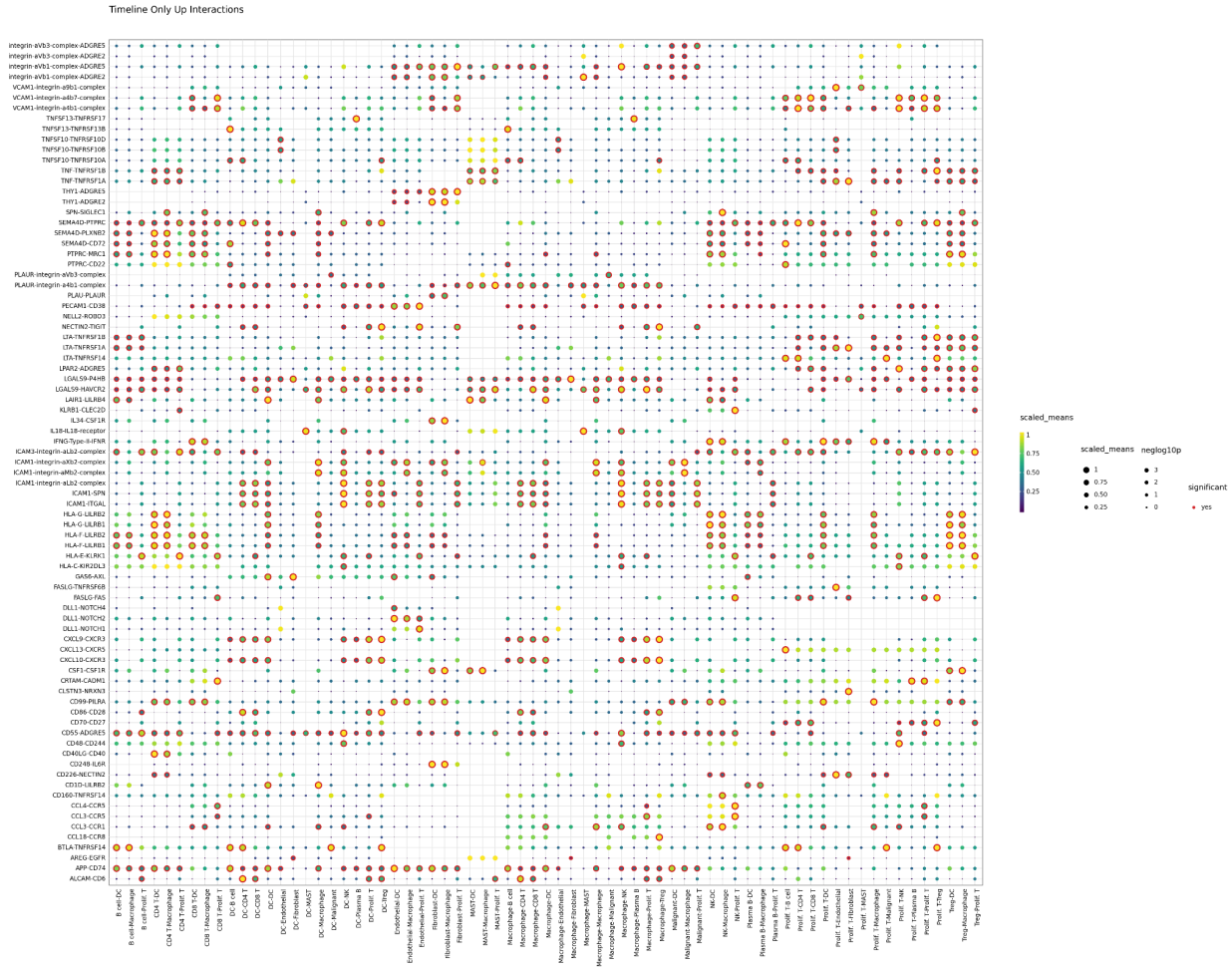


Figure S12. Ligand-receptor analysis of genes corresponding to circulating proteins enriched on treatment. CellPhoneDB output plot showing interaction score (shown by color), -log(p-value) (shown by size) and significance (whether the dot is circled in red) across cell type pairs for interactions involving post-ICB proteins.

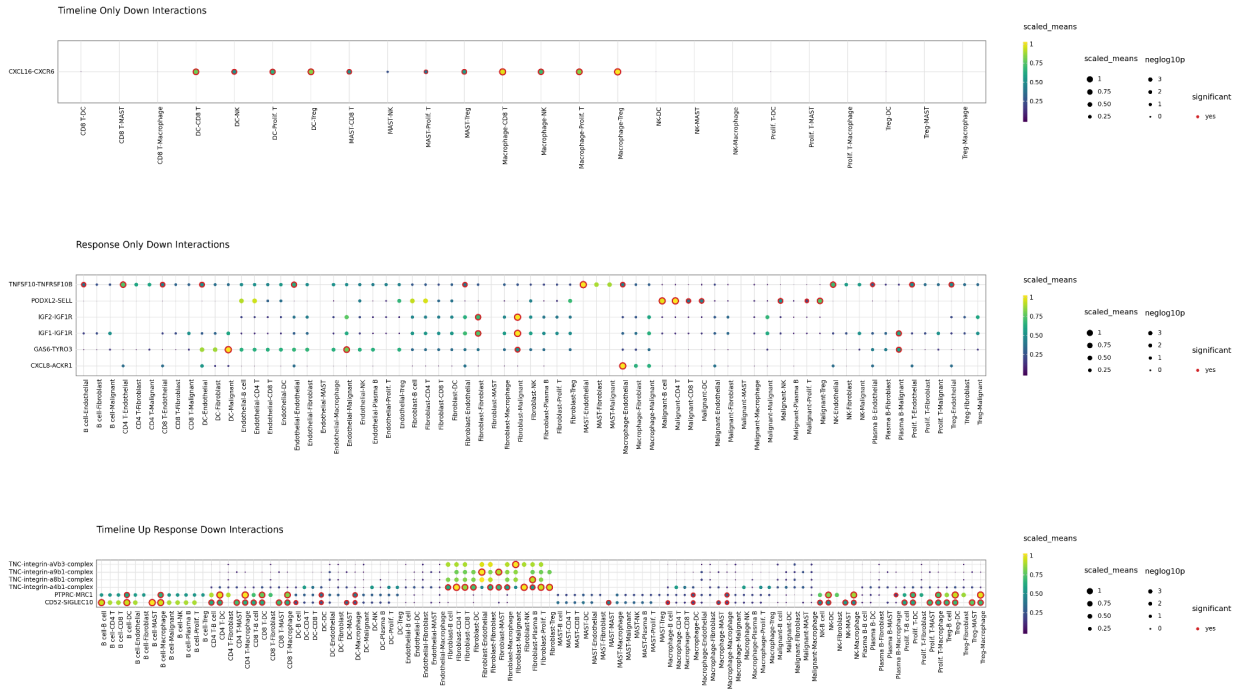


Figure S13. Ligand-receptor analysis of genes corresponding to circulating proteins of interest enriched pre-ICB and in non-responders. CellPhoneDB output plot showing interaction score (shown by color), -log(p-value) (shown by size) and significance (whether the dot is circled in red) across cell type pairs for interactions involving pre-ICB proteins (top), non-responder proteins (middle), and non-responder, post-ICB proteins (bottom).

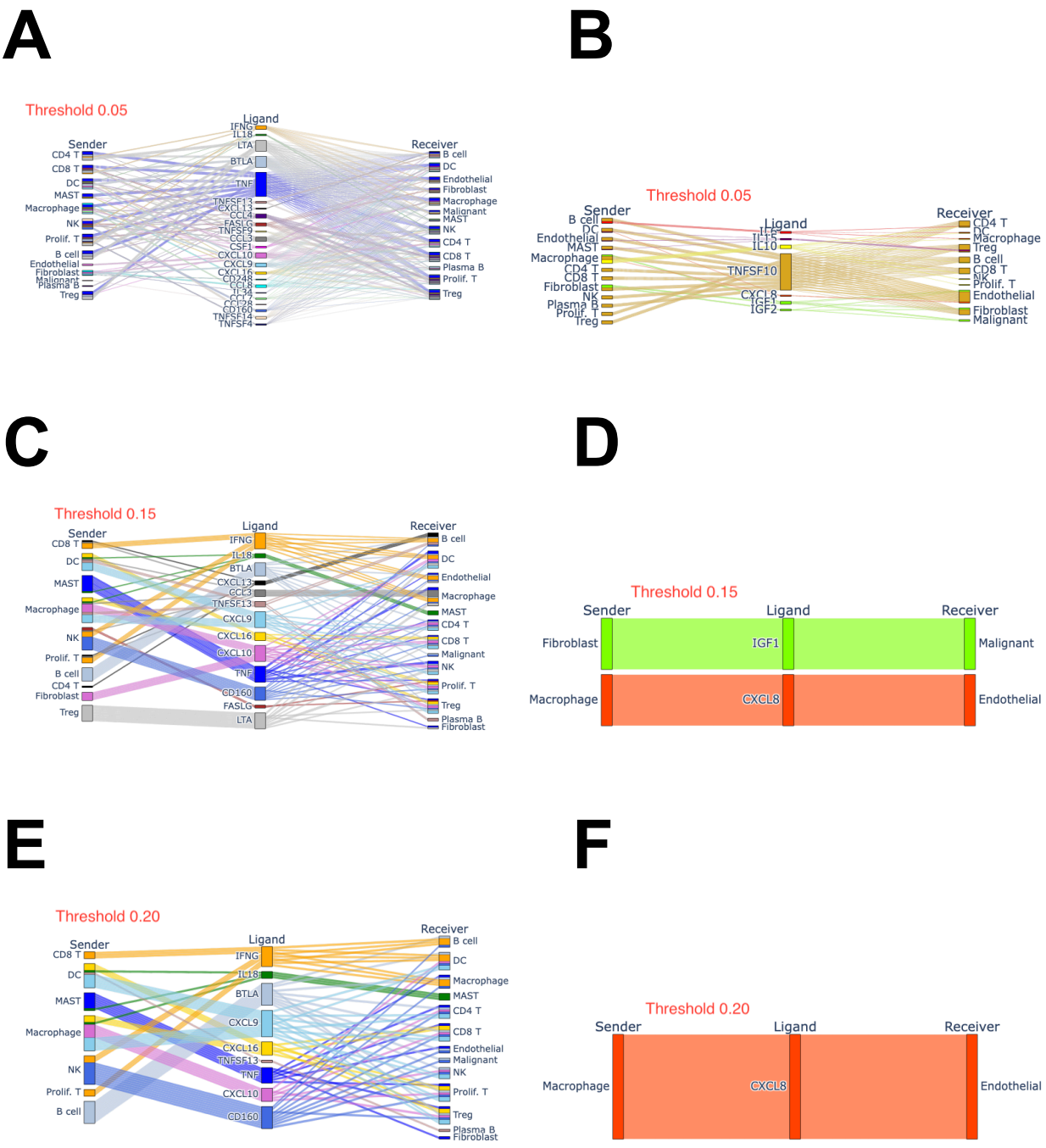


Figure S14. Sensitivity analysis of ligand-receptor pair gene expression. **a-f,** Flow diagram showing potential interactions involving cytokines and chemokines among post-ICB (**a, c, e**) proteins and non-responder proteins (**b, d, f**) based on minimum expression of ligands and receptors in sender and receiver cells, respectively. Minimum expression threshold set at 0.05 (**a,b**), 0.15 (**c,d**) and 0.20 (**e,f**).