Science Advances
AAAS

Supplementary Materials for

Severe hematotoxicity after CD19 CAR-T therapy is associated with suppressive immune dysregulation and limited CAR-T expansion

Kai Rejeski *et al.*

Corresponding author: Kai Rejeski, kai.rejeski@med.uni-muenchen.de; Michael D. Jain, michael.jain@moffitt.org; Marion Subklewe, marion.subklewe@med.uni-muenchen.de

Sci. Adv. **9**, eadg3919 (2023)
DOI: 10.1126/sciadv.adg3919

This PDF file includes:

Supplementary Material

Supplementary Tables

Neutrophil Recovery Phenotype Group	Quick (n=136)	Intermittent (n=144)	Aplastic (n=64)	*p*
Supportive therapies – n (%)				
Granulocyte colony stimulating factor (G-CSF) use	15 (11%)	101 (70%)	45 (70%)	**<0.001**
Thrombopoetin (TPO) agonist use	0 (0%)	1 (0.7%)	5 (8%)	**<0.001**
CD34+ Stem cell boost	0 (0%)	0 (0%)	5 (8%)	**<0.001**

Supplemental Table 1: Management of CAR T-cell related hematotoxicity
Overview of management strategies during the first 100 days after CAR T-cell infusion stratified by neutrophil recovery phenotype. P-values were determined by Chi-squared (X^2) test.

Neutrophil Recovery Phenotype Group	All patients (n=344)	Quick (n=136)	Intermittent (n=144)	Aplastic (n=64)	*p*
Supportive therapies – n (%)					
Received broad-spectrum IV antibiotics during the first 10 days post CAR-T infusion (e.g. piperacillin/tazobactam, carbapenems, 3./4. Generation cephalosporine)	298 (87%)	108 (79%)	132 (93%)	58 (91%)	**0.006**
Received PO antibiotic prophylaxis (e.g. ciprofloxacin, levofloxacin)	218 (63%)	100 (74%)	75 (52%)	43 (67%)	**<0.001**
Maximal Infection Severity, any infection type – n (%)					
0 (no infection)	180 (52%)	83 (61%)	74 (51.5%)	23 (36%)	**0.002**
1 (mild)	28 (8%)	13 (10%)	9 (6%)	6 (9%)	
2 (moderate)	54 (16%)	18 (13%)	29 (20%)	7 (11%)	
3 (severe)	58 (17%)	17 (13%)	23 (16%)	18 (28%)	
4 (life-threatening)	15 (4%)	3 (2%)	7 (5%)	5 (8%)	
5 (fatal)	9 (3%)	2 (1%)	2 (1.5%)	5 (8%)	
Maximal Infection Severity, bacterial infections – n (%)					
0 (no infection)	249 (72%)	106 (78%)	101 (70%)	42 (66%)	**0.006**
1 (mild)	17 (5%)	8 (6%)	7 (5%)	2 (3%)	
2 (moderate)	30 (9%)	12 (9%)	16 (11%)	2 (3%)	
3 (severe)	29 (8.5%)	6 (4%)	12 (8%)	11 (17%)	
4 (life-threatening)	15 (4%)	3 (2%)	8 (6%)	4 (6%)	
5 (fatal)	4 (1%)	1 (1%)	0 (0%)	3 (5%)	

Supplemental Table 2: Antimicrobial strategies and infection severity by neutrophil recovery phenotype
Overview of management strategies during the first 100 days after CAR T-cell infusion stratified by neutrophil recovery phenotype. P-values were determined by Chi-squared (X^2) test.

Neutrophil Recovery Phenotype Group	Quick (n=136)	Intermittent (n=144)	Aplastic (n=64)	*p*
CRS maximum grade – n (%)				0.22
0	23 (17%)	13 (9%)	11 (17%)	
1 – 2	102 (75%)	118 (82%)	45 (70%)	
3 – 5	11 (8%)	13 (9%)	8 (13%)	
ICANS maximum grade – n (%)				**0.002**
0	83 (61%)	65 (45%)	33 (52%)	
1 – 2	41 (30%)	50 (35%)	16 (25%)	
3 – 4	12 (9%)	39 (20%)	15 (23%)	
Treatment for toxicity – n (%)				
Steroid use	44 (32%)	75 (52%)	29 (45%)	**0.004**
Received high-dose corticosteroids ≥9 days	10 (7%)	29 (20%)	14 (22%)	**0.004**
Tocilizumab use	58 (43%)	91 (63%)	41 (64%)	**<0.001**
Anakinra use	0 (0%)	4 (3%)	5 (8%)	**0.005**
ICU admission	16 (12%)	19 (13%)	16 (25%)	**0.02**

Supplemental Table 3: Coincident immunotoxicity and management
Severity of CRS and ICANS and distribution of toxicity management by neutrophil recovery phenotype. High-dose corticosteroids defined as ≥9 days of dexamethasone 10 mg or equivalent (cutoff based on Rejeski et al JITC 2022). P-values were explored using Chi squared test for categorical variables.

Supplementary Figures

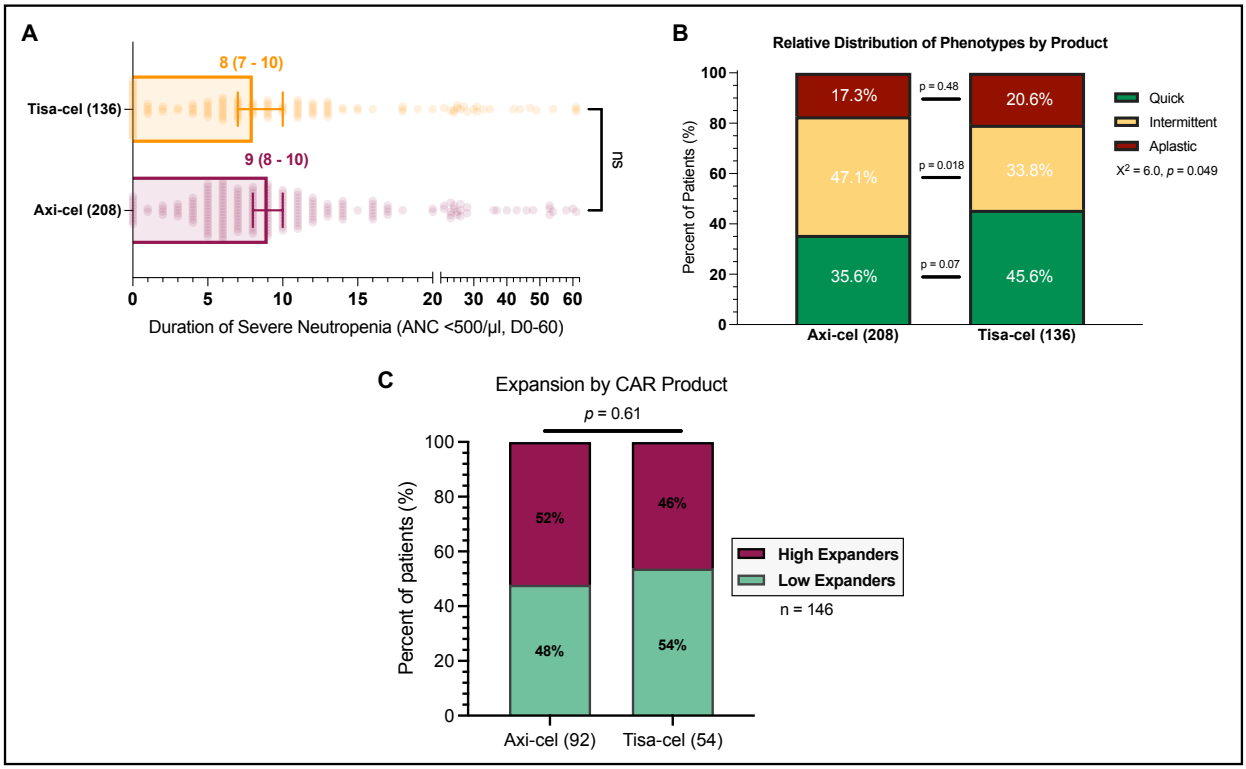


Figure S1: Cytopenia incidence rates and peak CAR expansion by CAR product

A Cumulative duration of severe neutropenia (ANC <500/µL) in days measured during the first 60 days after CAR-T infusion comparing patients treated with Axi-cel (n=208) vs. Tisa-cel (n=136). Box and whiskers indicate the median and 95% confidence interval. The p-value was determined by Mann-Whitney test. **B** Distribution of neutrophil recovery phenotypes by CAR product. Results of Chi-squared (X^2) test and individual t tests are provided. **C** Comparison of high CAR expanders (above median) vs. low expanders (below median) by CAR-T product. P-value determined by Fisher's exact test.

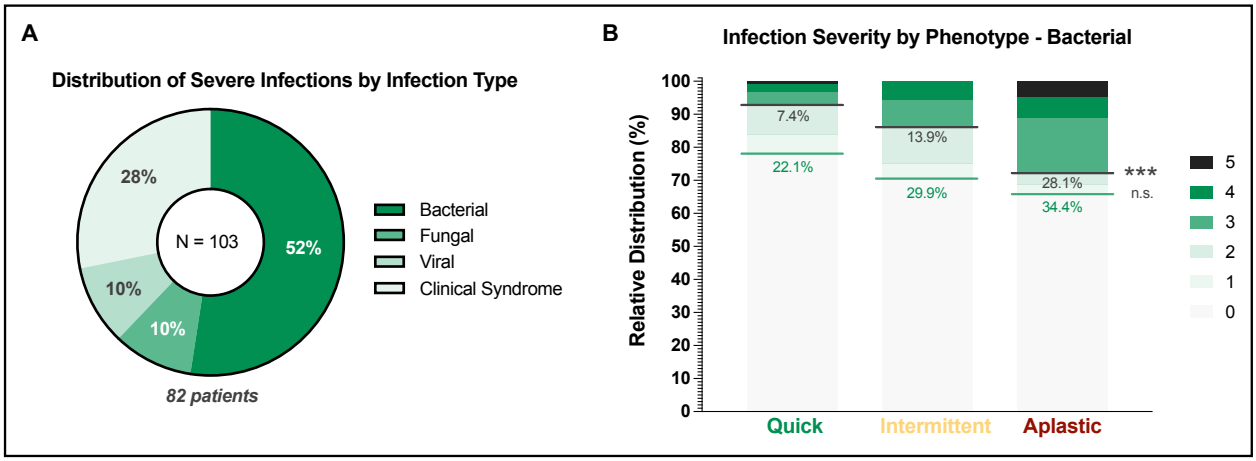


Figure S2: Types of severe infections and bacterial infections severity by neutrophil recovery phenotype

A Overview of infection subtype in the 103 severe infection events determined across 82 patients. An episode of infection was defined as bacterial, viral or fungal based on microbiologic or histopathologic data or as a clinical syndrome of infection (eg, pneumonia, cellulitis, cystitis) based on retrospective chart review. **B** Relative distribution of infection grades by phenotype (bacterial infections only). Infection grades (1-5°) are color-coded in shades of green with the connecting green and grey lines and percentage numbers comparing all-grade and grade ≥3 infections by phenotype (infection grading: Rejeski et al, JITC 2022). P-value determined by Chi-squared (X^2) test (*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001).

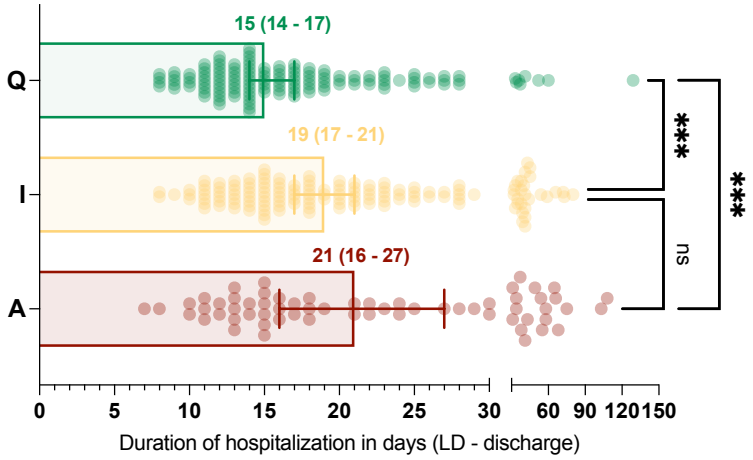


Figure S3: Duration of hospitalization by neutropenia recovery phenotype

Duration of hospitalization from start of lymphodepletion until first discharge in days. P-valuse determined by Kruskal-Wallis and Dunn's tests. Significance values: *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

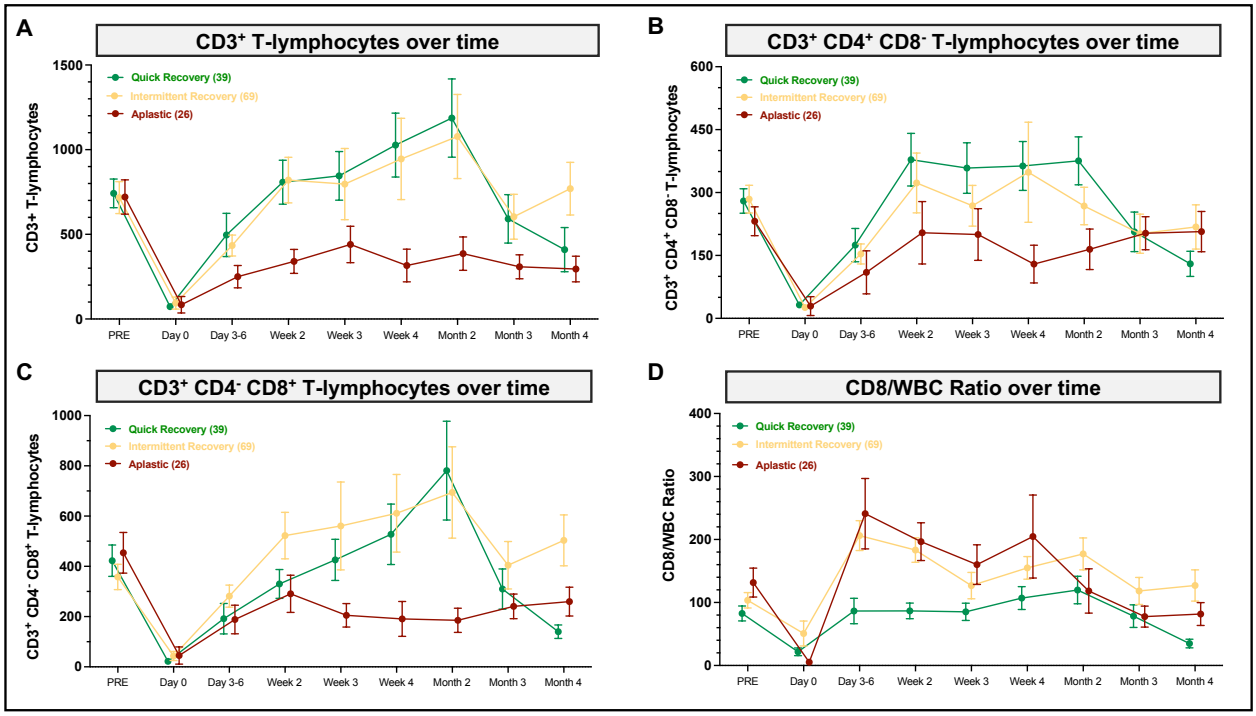


Figure S4: Endogenous T-cell subpopulations over time by phenotype

A-C Median absolute number of $CD3^+$ (**A**), $CD3^+/CD4^+/CD8^-$ (**B**), and $CD3^+/CD4^-/CD8^+$ (**C**) non-CAR bearing T-cells over time by neutrophil recovery phenotype. **D** Ratio of $CD3^+/CD4^-/CD8^+$ T-cells to white blood cell count (WBC) on the respective day of measurement. Box and whisker describe the median with the 95% confidence interval of the median.

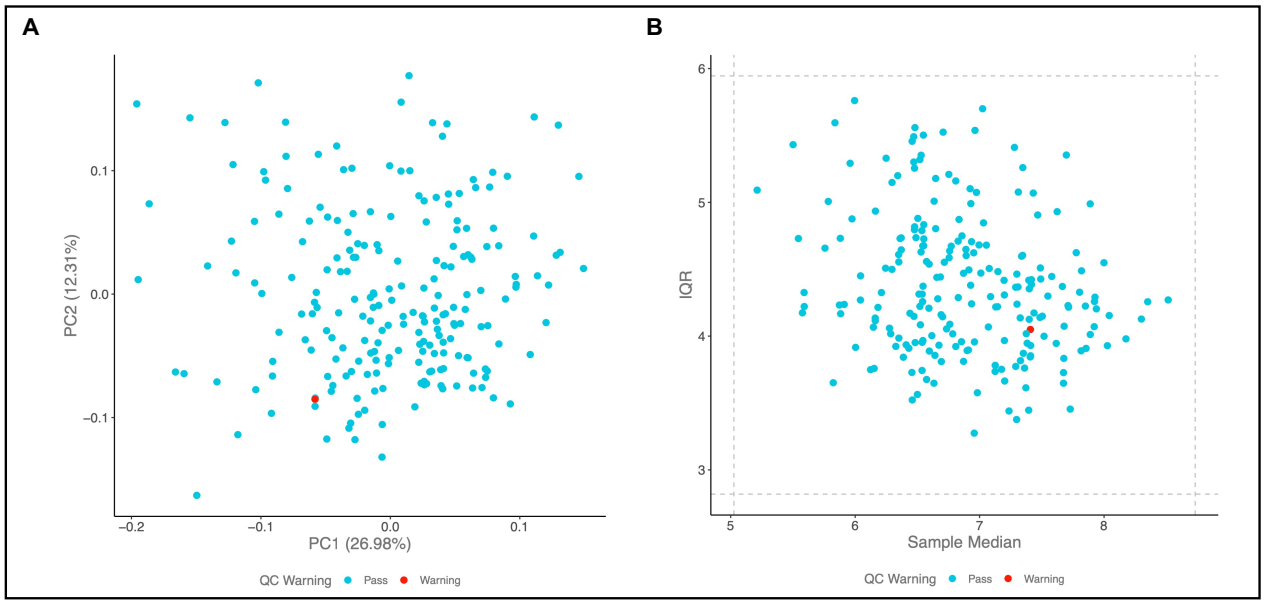


Figure S5: Quality control measures – Olink serum proteome analysis

A PCA projection of PC2 vs. PC1. Each point corresponds to one sample from the whole data set. Samples that were flagged with a QC warning on one of the two panels are indicated in red, and samples with no QC warnings are blue. **B** The median NPX of each sample (x-axis) plotted versus the inter quantile range (IQR) (y-axis). Each dot corresponds to one sample. Samples that are flagged with a QC warning are indicated in red, and samples with no QC warnings are blue. Horizontal dashed lines indicate ±3 standard deviations from the mean IQR. Vertical dashed lines indicate ±3 standard deviations from the mean sample median.

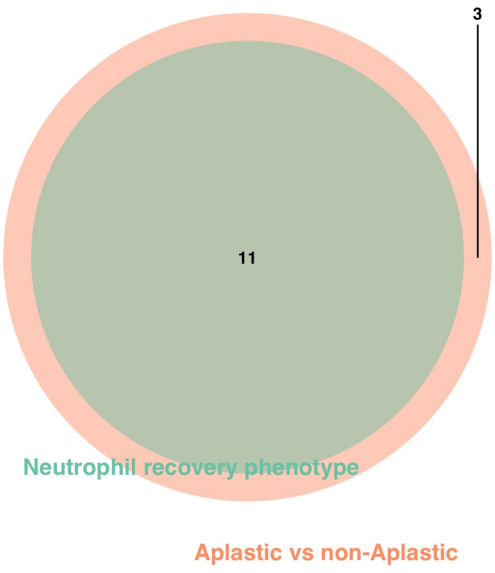
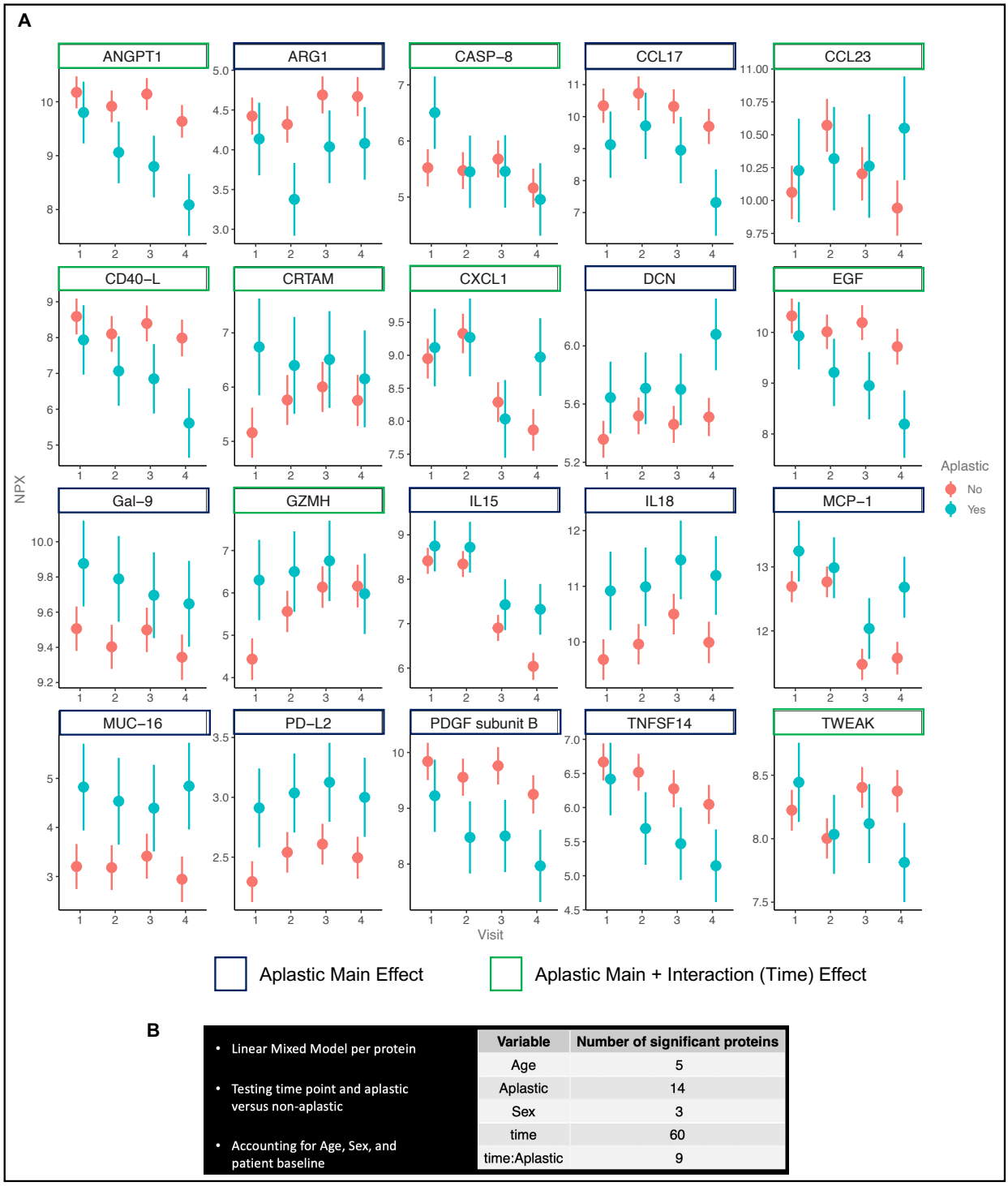


Fig. S6 Quality control measures – Overlap between aplastic vs. non-aplastic and phenotype proteome analysis

Venn diagram showing the number of proteins showing significant differences (main effect in the LMM analyzes) between the neutrophil recovery phenotypes and/or between Aplastic and non-Aplastic patients.



B

- Linear Mixed Model per protein
- Testing time point and aplastic versus non-aplastic
- Accounting for Age, Sex, and patient baseline

Variable	Number of significant proteins
Age	5
Aplastic	14
Sex	3
time	60
time:Aplastic	9

Figure S7: Protein-level differences between the aplastic and non-aplastic phenotype in the Olink proteomic analysis

A 20 proteins displayed significant serum level differences in the linear mixed model (LMM) analysis between the aplastic (blue) and non-aplastic (quick or intermittent, red) conditions. The proteins with an aplastic main effect are framed in blue, while proteins with both a main effect and interaction (or time) effect are framed in green. **B** Number of significant proteins by variable of the LMM analysis.

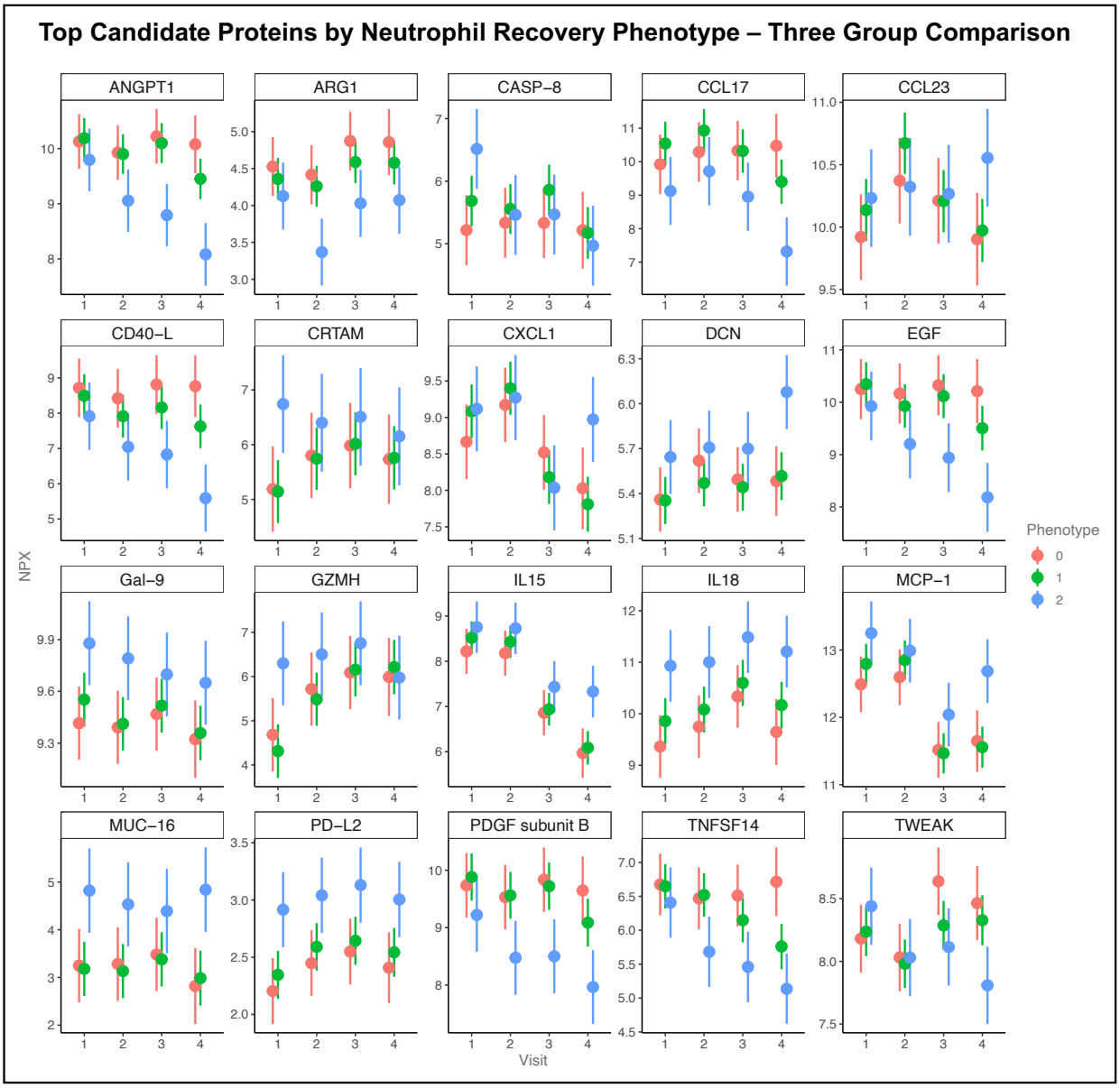


Figure S8: Top candidate proteins from LMM analysis in three-way comparison of neutrophil recovery phenotypes

A The 20 proteins with significant differences in the aplastic vs. non-aplastic LMM analysis (Fig. S4) are depicted in a three-way comparison of phenotypes (red = quick, green = intermittent, blue = aplastic).

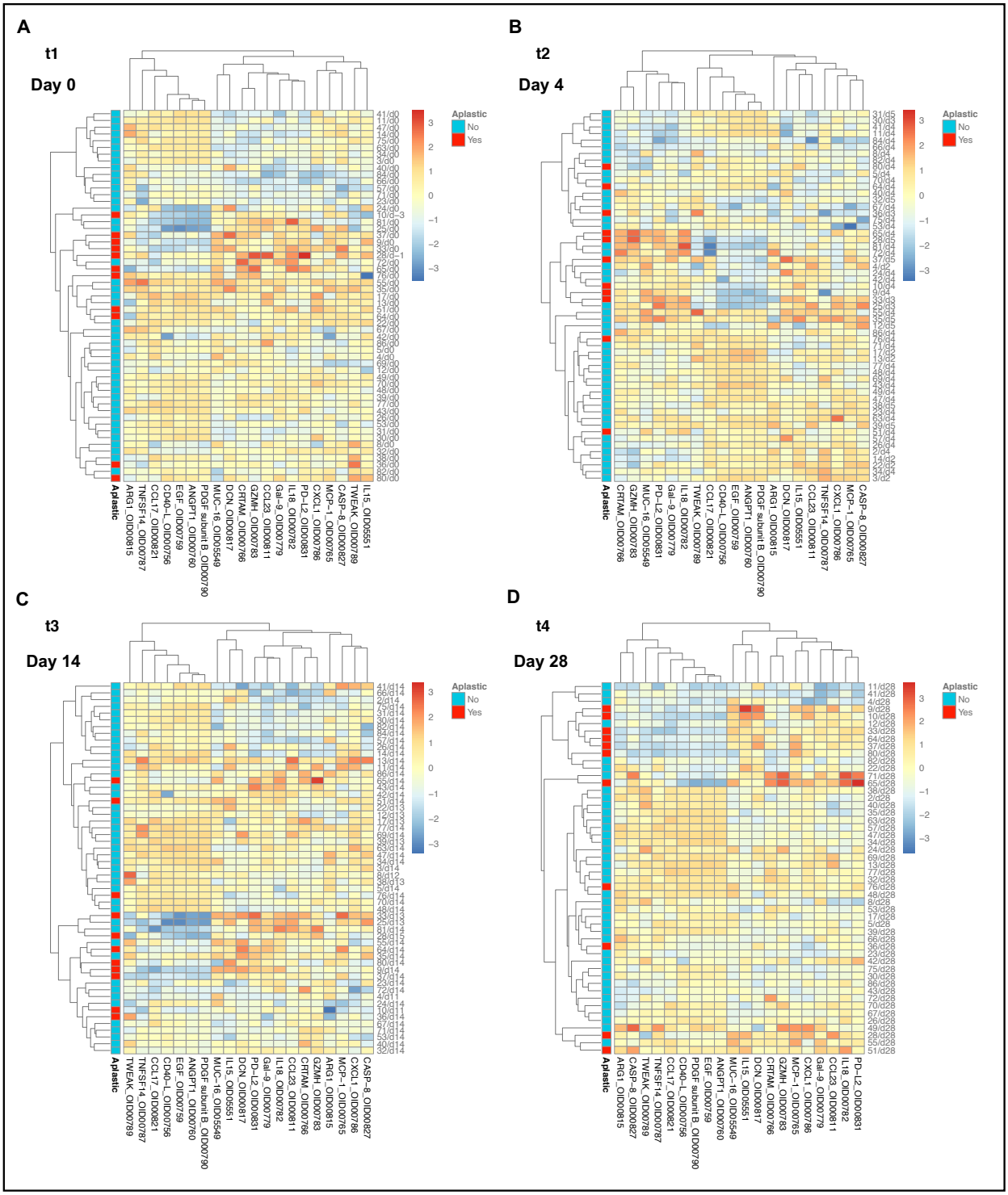


Figure S9: Hierarchical clustering analysis by aplastic vs. non-aplastic phenotype

Heatmaps depicting normalized and centered NPX values (color scale) of the 20 proteins that differed between aplastic and non-aplastic patients (significant main and/or interaction effect). Each row corresponds to a sample/time point, and each column to a protein. Both rows and columns are sorted by hierarchical clustering, with dendrograms showing the clustering result. Neutrophil recovery phenotype status is indicated on the leftmost columns. The individual heatmap is depicted for each time point: day 0 (**A**), day 4 (**B**), day 14 (**C**), and day 28 (**D**).

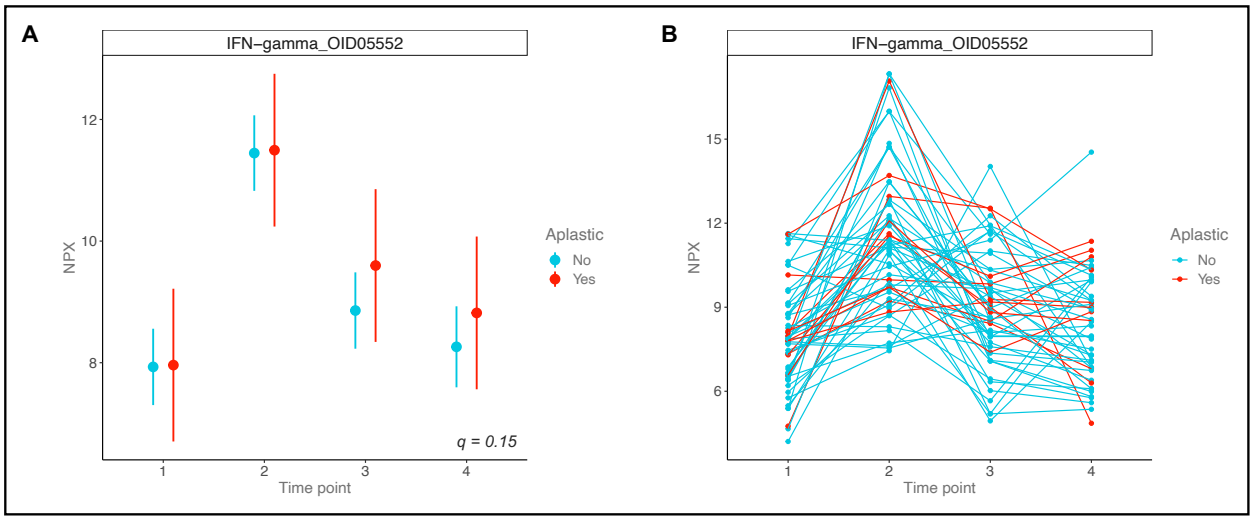


Figure S10: Interferon-gamma dynamics in Olink data set

A Serum level differences in IFN-gamma levels in the linear mixed model analysis between the aplastic (red) and non-aplastic (quick or intermittent, blue) conditions. The p-value corrected for multiple testing (q-value) is superimposed on the panel. As NPX is expressed on a log2 scale, an increase of 1 NPX corresponds to a doubling of the serum protein concentration. Timepoint 1-4: day 0, day 4-7, day 14, day 28. **B** Scatter plot of serum IFN-gamma levels over time with each continuous line depicting an individual patient. Aplastic patients are represented in red, non-aplastic patients in blue (*different color-coding).