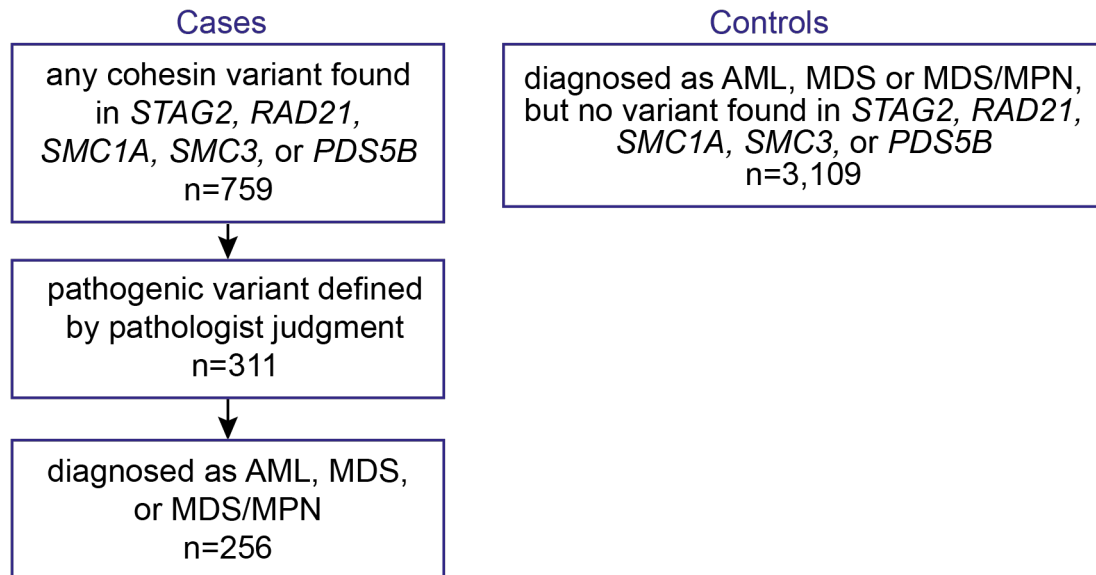
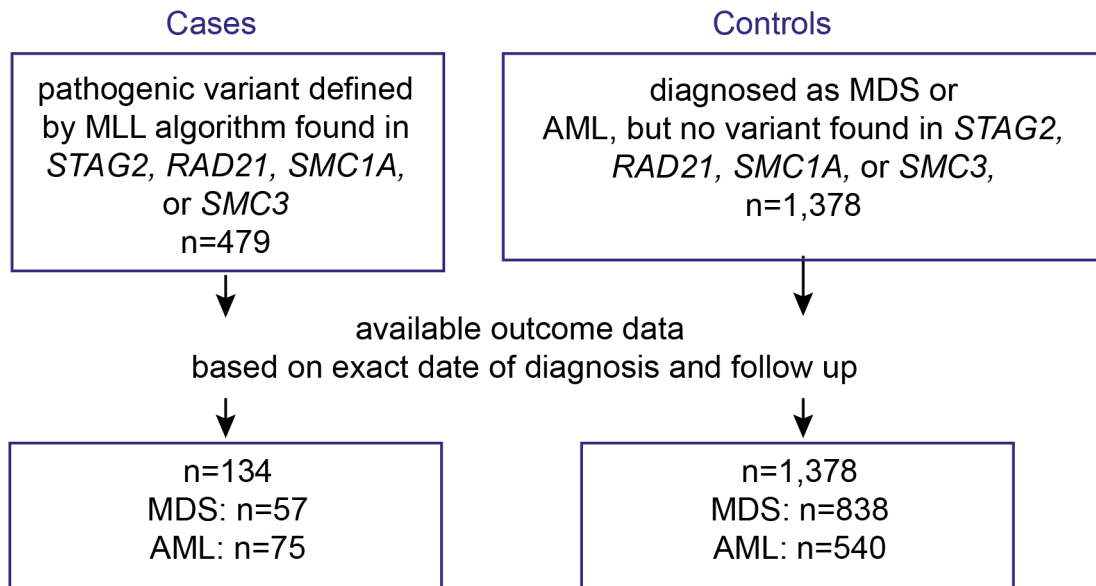


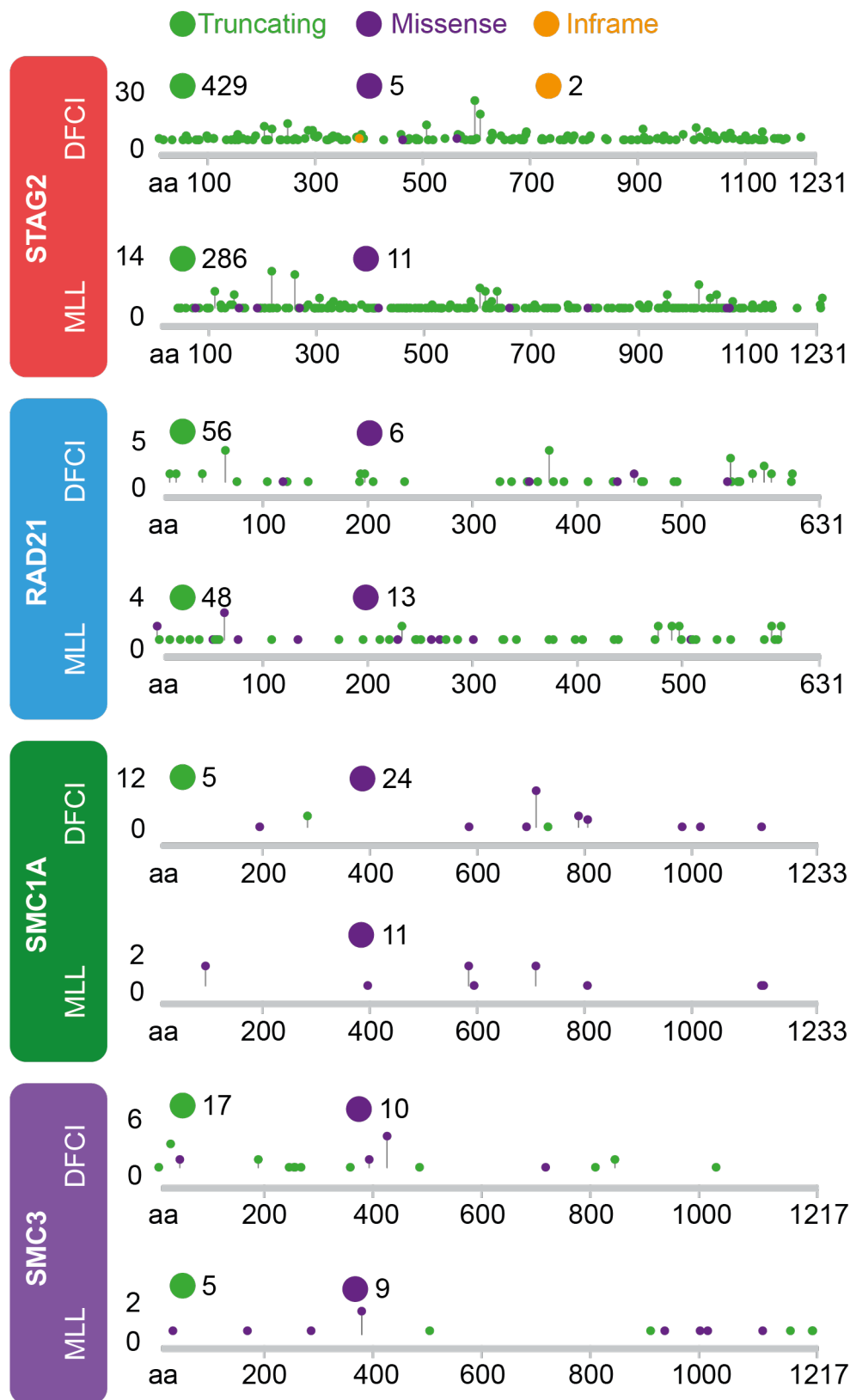
DFCI n=3,868



MLL n=1,857

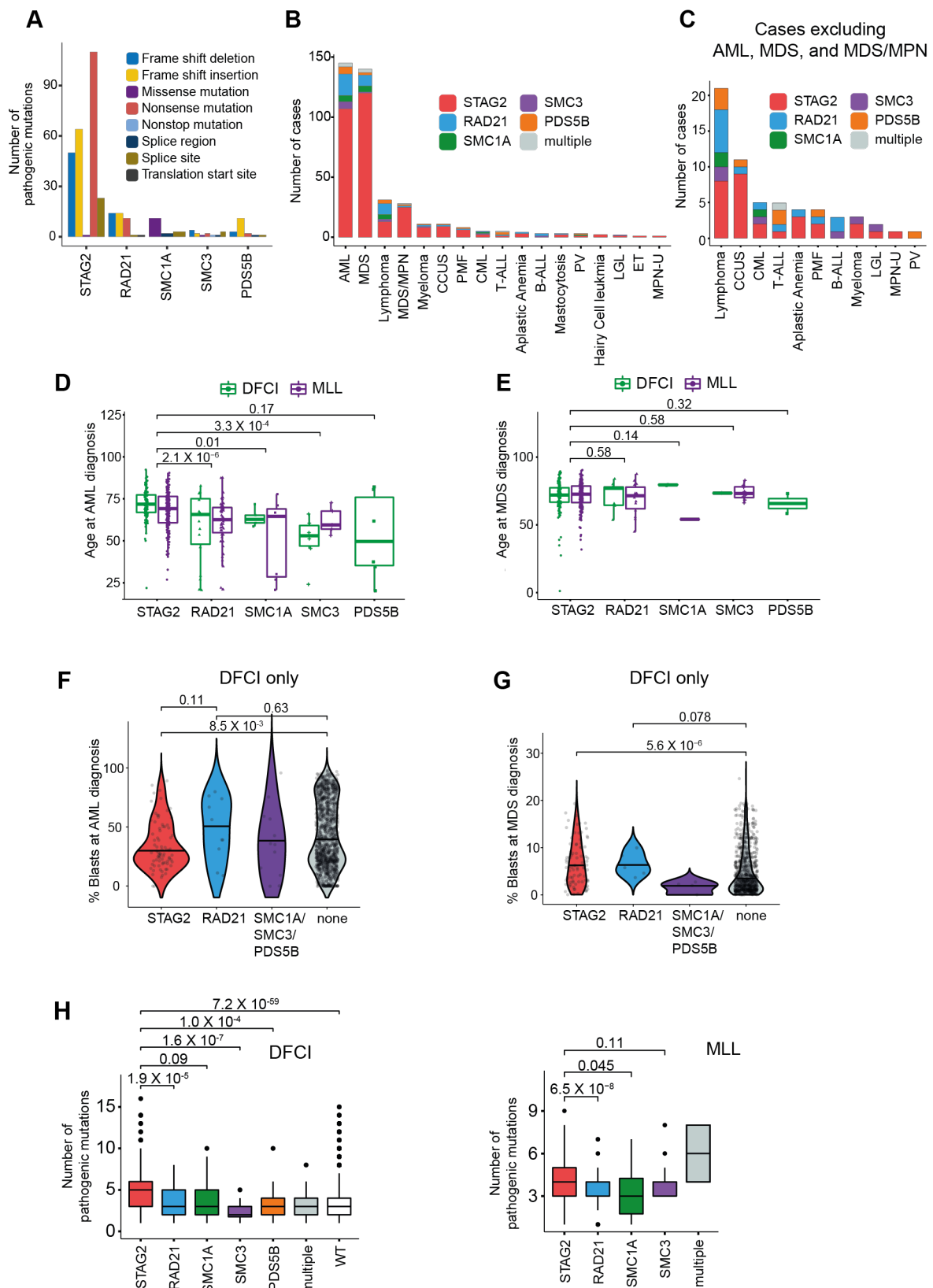


Supplementary Figure 1: CONSORT diagram of patient flow in this study



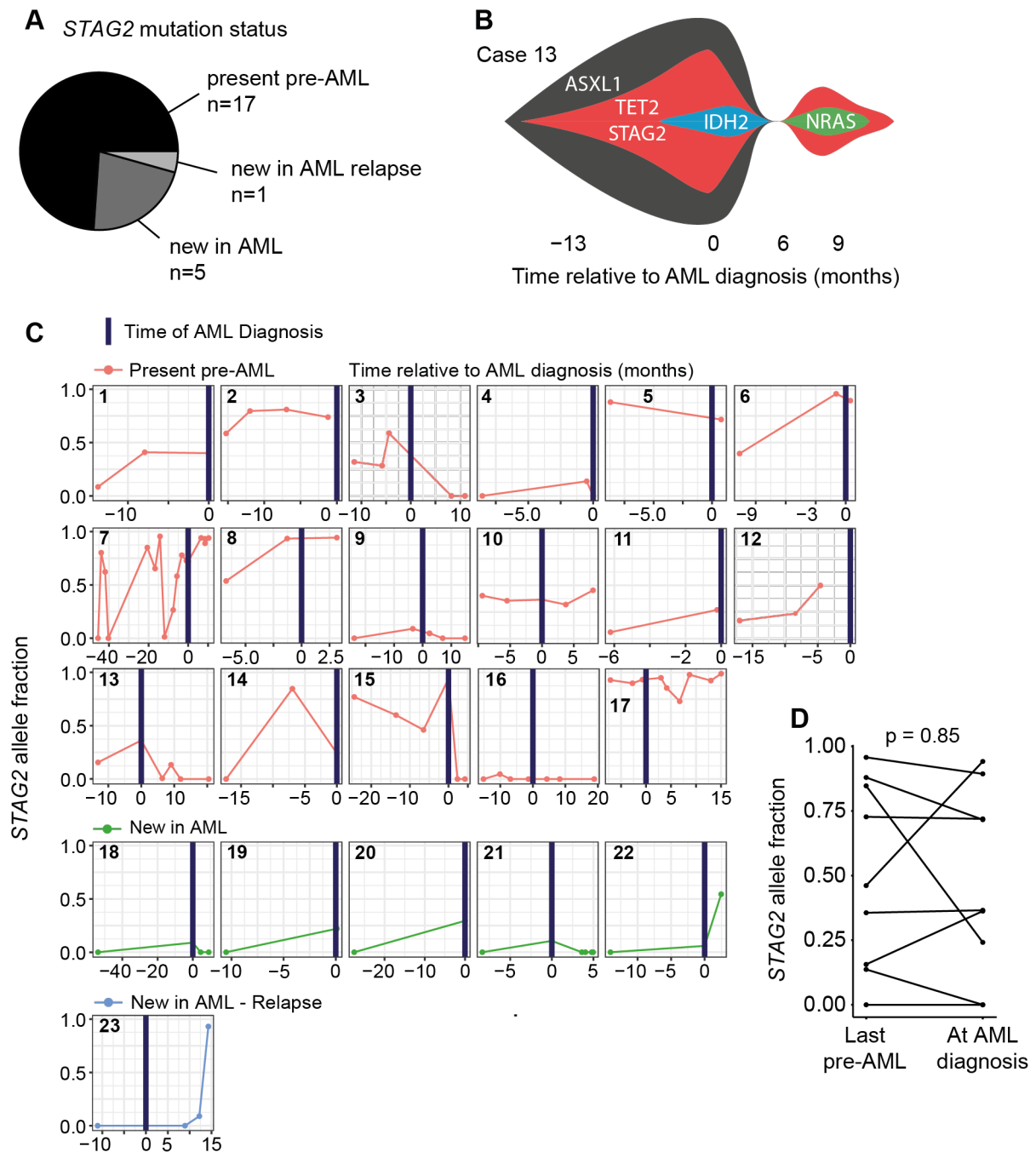
Supplementary Figure 2: Molecular characterization of cohesin mutations in hematologic malignancies separated by cohort.

Lollipop plot panel of cohesin mutations separated by cohort and cohesin gene. DFCI = Dana-Farber Cancer Institute cohort, MLL = Munich Leukemia Laboratory cohort.



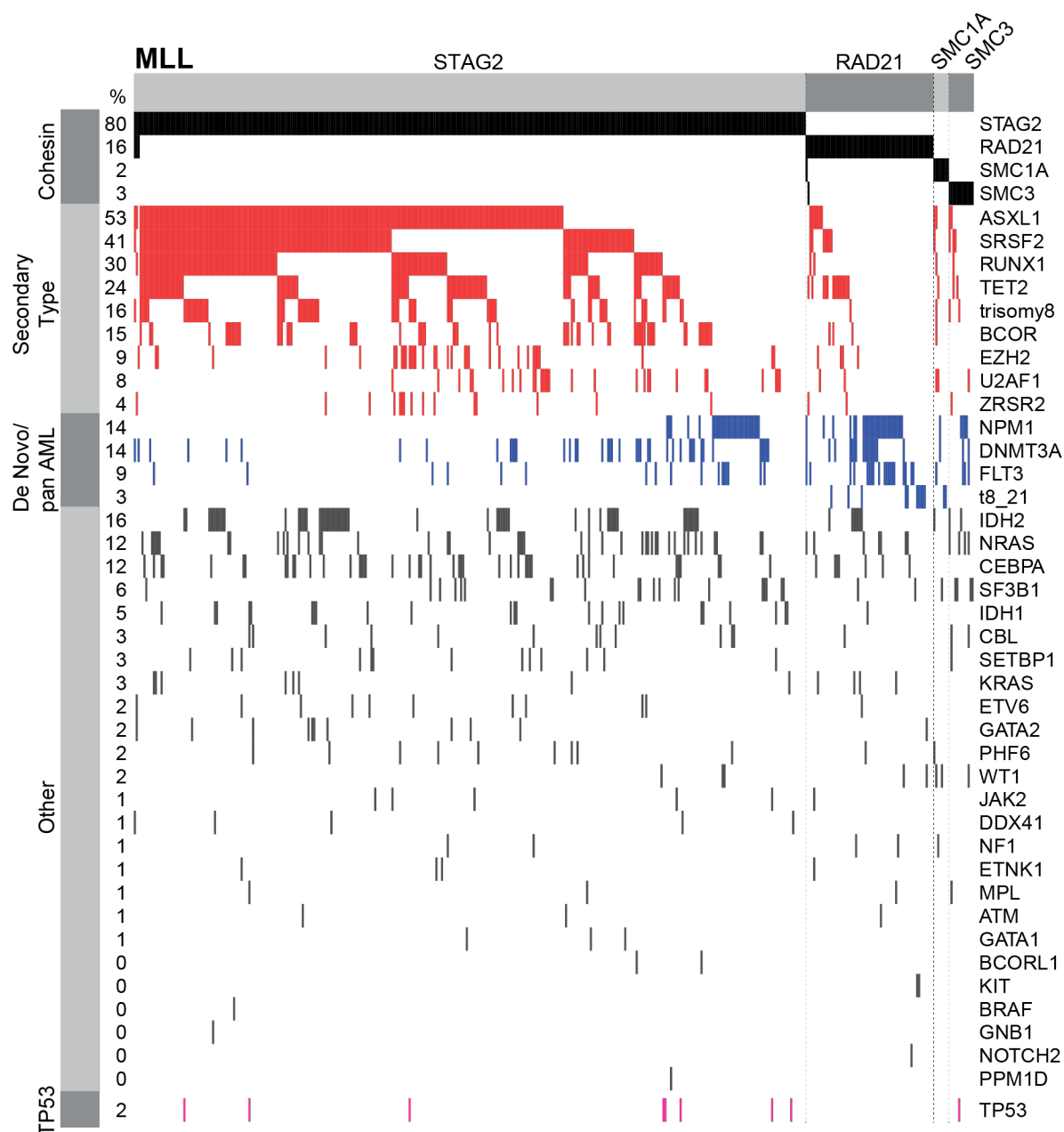
Supplementary Figure 3: Clinical characteristics of cohesin-mutant patients separated by cohort

(A) Bar plot of the total number of pathogenic mutations in the DFCI cohort. (B) Bar plot of the total number of cases per diagnosis for patients with a pathogenic cohesin mutation in the DFCI cohort. (C) Bar plot of the total number of patients with a pathogenic cohesin mutation, in whom no diagnosis of AML, MDS, or MDS/MPN was made. Multiple diagnoses can be counted per single patient, n=55. (D-E) Box plots of age at diagnosis for MDS or AML stratified by cohesin mutation. Benjamini Hochberg adjusted p values for specific mutations are shown. (F-G) Violin plots of %blasts in diagnostic bone marrow specimens for patients diagnosed with AML or MDS in the DFCI cohort, stratified by cohesin mutation and compared to the cohesin WT cohort. Benjamini Hochberg adjusted p values are shown. (H) Box plot of total number of pathogenic mutations determined by myeloid target capture sequencing for MDS and AML patients stratified by cohesin mutation status, separated by DFCI and MLL cohorts.



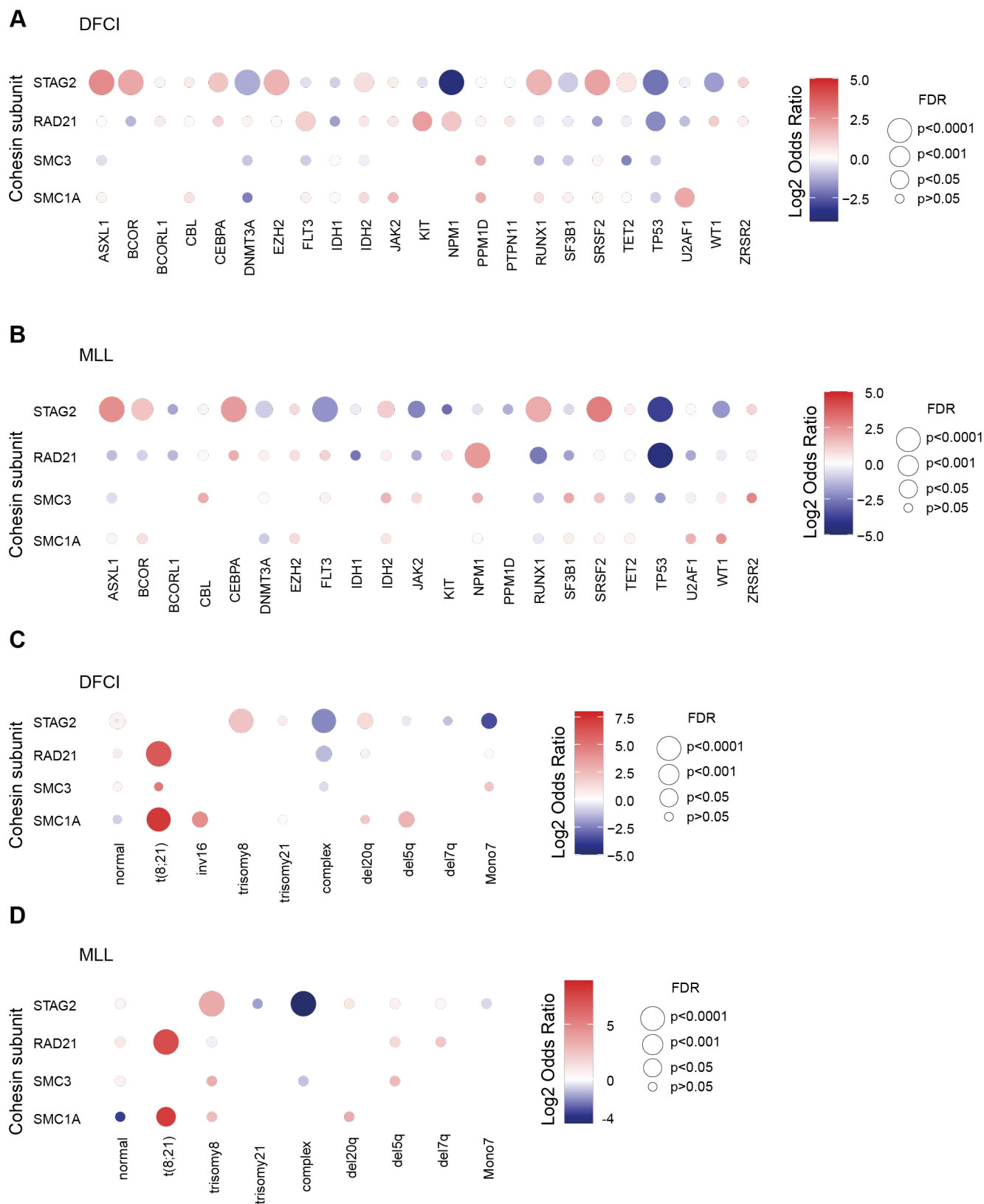
Supplementary Figure 4: Acquisition of STAG2 mutations relative to the onset of AML

(A) Pie chart of STAG2 mutation detection relative to AML diagnosis. (B) Fish plot for clonal evolution of a representative STAG2-mutant case, with STAG2 mutation preceding the AML diagnosis. Clone sizes are extrapolated from variant allele fraction (VAF) as previously reported⁴². (C) VAF quantification for STAG2-mutant patients with at least 1 available mutational panel testing prior to AML diagnosis. Each box represents a single patient. Patients were grouped according to the presence or absence of the STAG2 mutation prior to AML diagnosis using a STAG2 VAF cutoff of 0.01. Thick black line indicates timepoint of AML diagnosis. (D) VAF quantification for paired mutational assessment pre-AML and at the time of AML diagnosis. Two sided paired t-test was used to determine statistical significance.



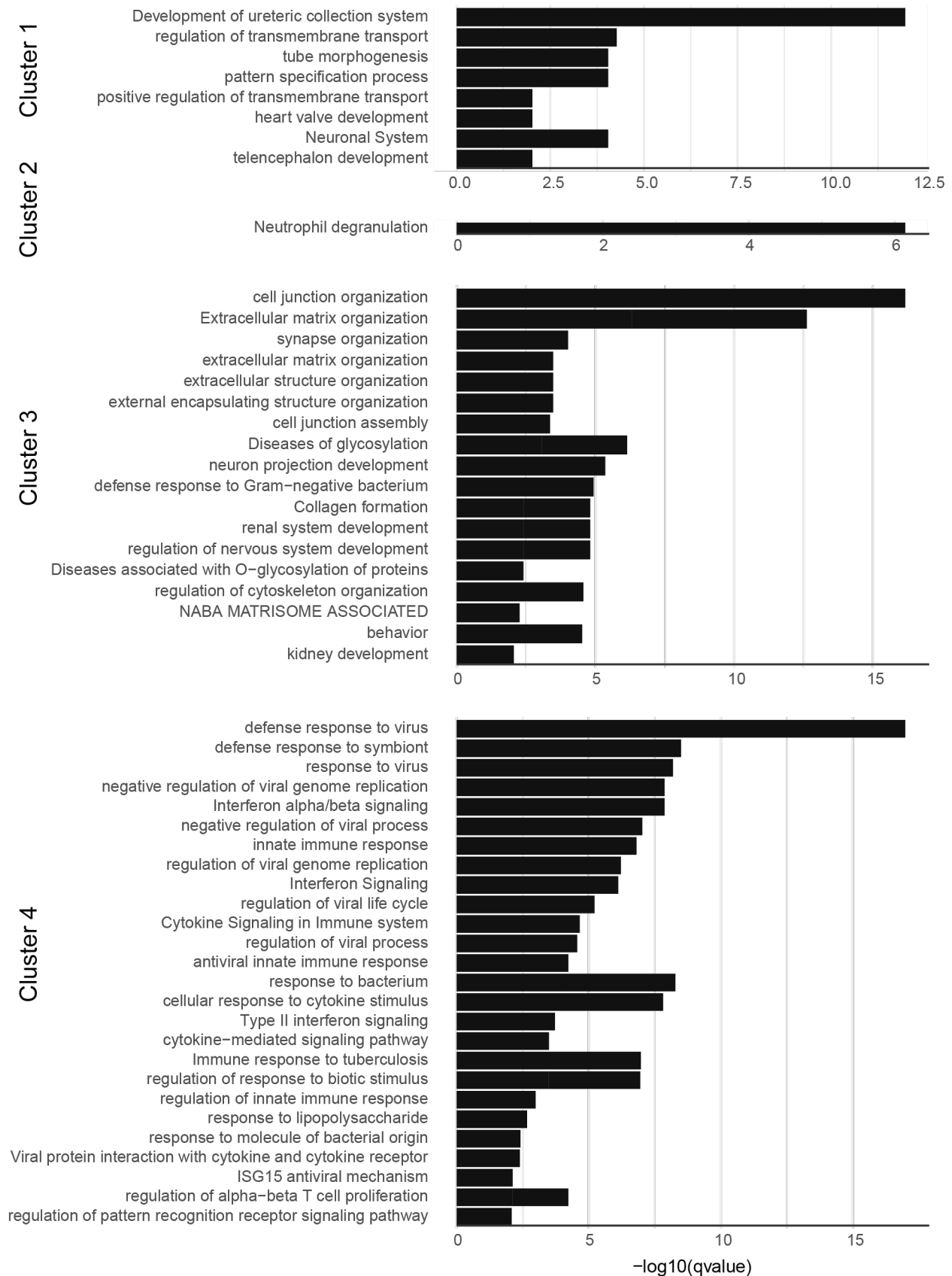
Supplementary Figure 5: Co-mutational plot for different cohesin mutations in the MLL cohort

Oncoprint for MDS, AML, and MDS/MPN patients with *STAG2*, *RAD21*, *SMC3*, and *SMC1A* mutations sorted by co-mutational pattern.



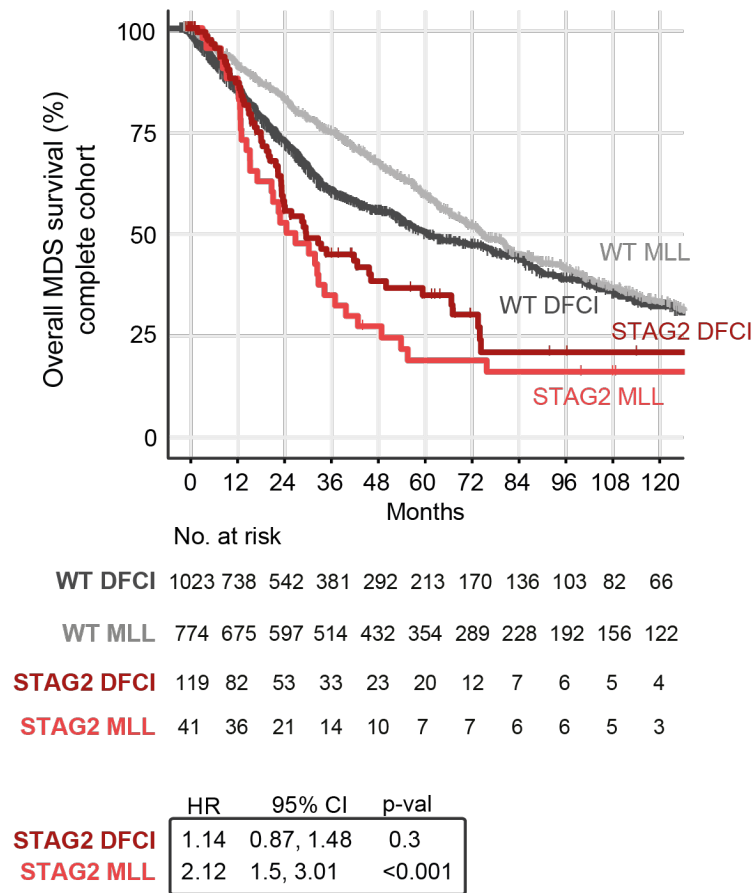
Supplementary Figure 7: Mutations in different cohesin subunits display unique mutational and chromosomal abnormality characteristics separated by cohort

(A-B) Balloon plot for relative enrichment of co-occurrence of cohesin subunit mutations with other myeloid driver mutations and **(C-D)** chromosomal aberrations, separated by cohort. Cohesin WT cohort was used as a reference cohort to calculate enrichment, which is indicated as log2 odds ratio (OR). Combinations with $q < 0.05$ or 5% mutational frequency in the total cohort are shown.



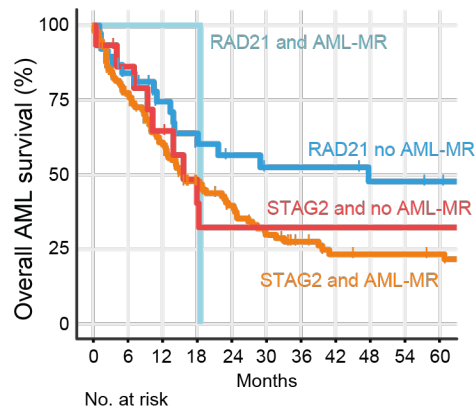
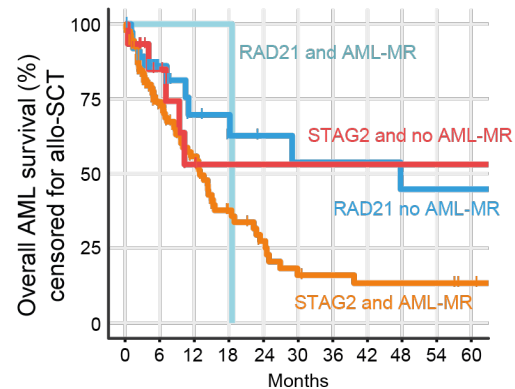
Supplementary Figure 8: Gene set enrichment analysis of cohesin-mutant AML RNA-Seq data in the BeatAML³⁶ cohort.

Gene sets with an FDR<0.01 were plotted for each cluster derived from unsupervised k-means clustering from Fig. 3C. Clusters 1-4 include 194, 401, 463, and 270 genes, respectively.



Supplementary Figure 9: MDS-related outcomes separated by cohort

Survival analysis using the Kaplan-Meier method and log rank test for overall MDS survival stratified for STAG2-mutant and cohesin-WT status in the DFCI and MLL cohorts.

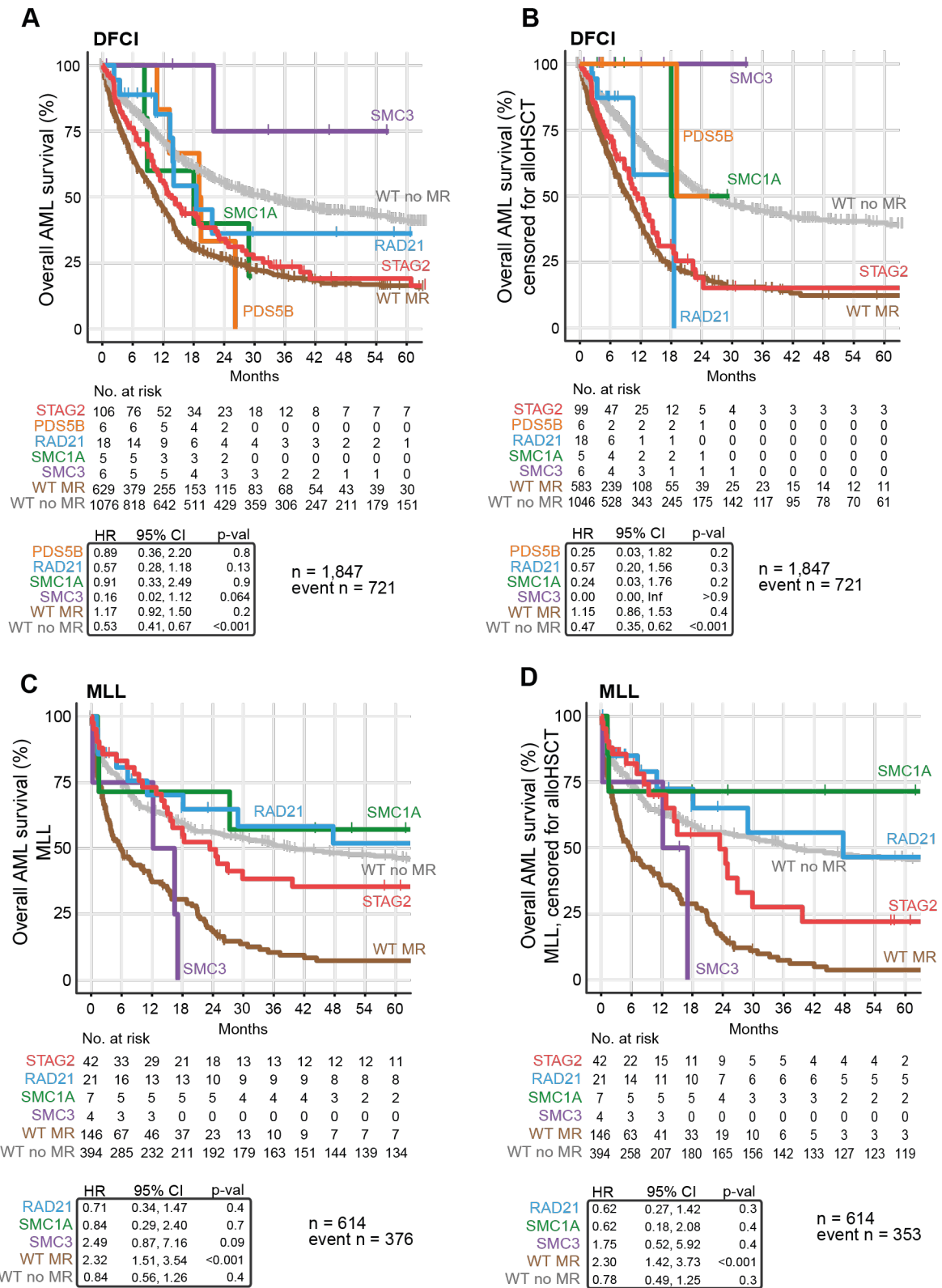
A**B**

RAD21 and AML-MR	1	1	1	1	0	0	0	0	0	0	0
RAD21 no AML-MR	38	29	21	18	14	13	12	12	10	10	9
STAG2 and AML-MR	133	97	72	50	37	27	21	16	15	15	14
STAG2 and no AML-MR	15	12	9	5	4	4	4	4	4	4	4

RAD21 and AML-MR	1	1	1	1	0	0	0	0	0	0	0
RAD21 no AML-MR	38	19	11	10	7	6	6	6	5	5	5
STAG2 and AML-MR	126	60	36	21	12	7	6	5	5	5	3
STAG2 and no AML-MR	15	9	4	2	2	2	2	2	2	2	2

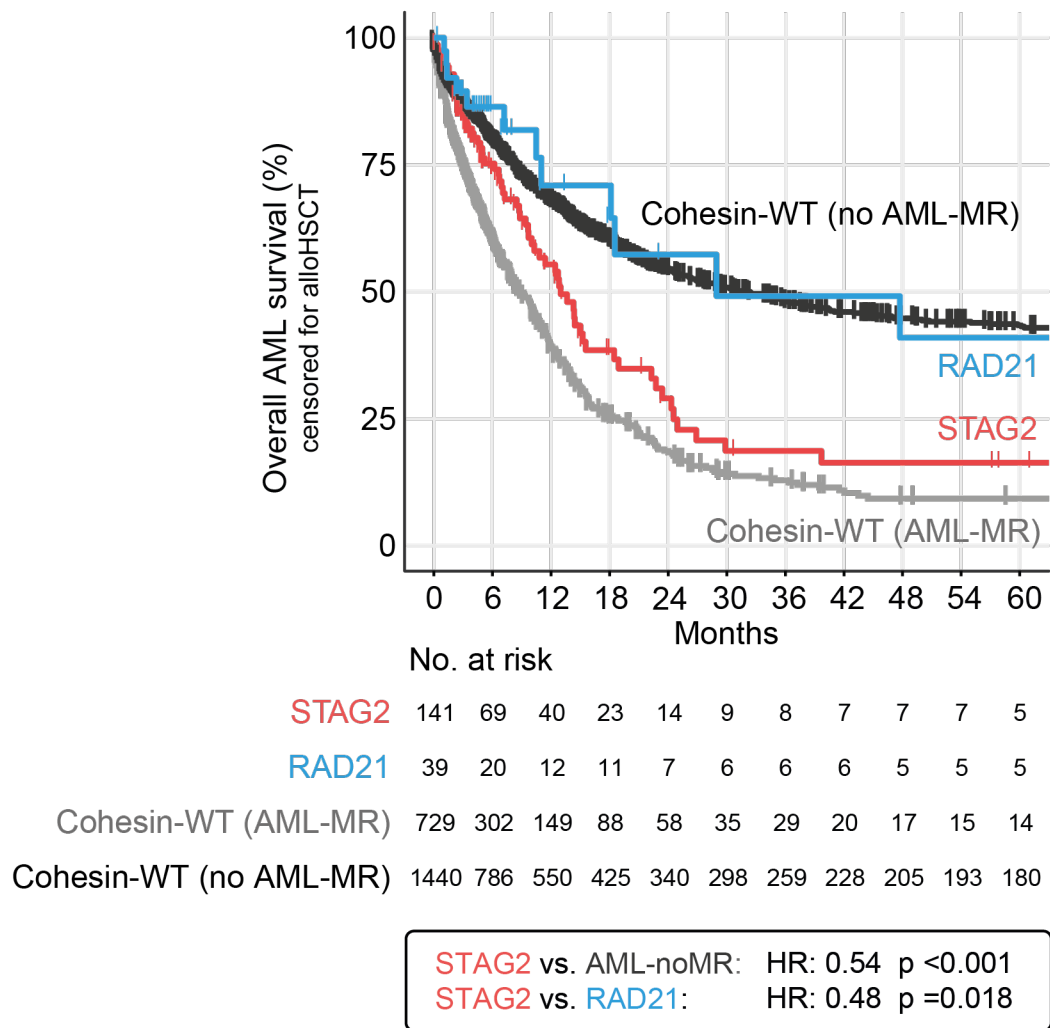
Supplementary Figure 10: AML-related outcomes in the combined dataset

(A) Survival analysis using the Kaplan-Meier method for overall AML survival stratified for cohesin mutational status and WHO2022 diagnosis. **(B)** Subgroup analysis censored at the time of allogeneic stem cell transplantation (allo-SCT).



Supplementary Figure 11: AML-related outcomes separated by cohort

(A-B) DFCI cohort: (A) Survival analysis using the Kaplan-Meier method and log-rank test for overall AML survival stratified for cohesin mutational status and cohesin-WT cases separated by AML-MR and AML-non-MR. (B) Overall AML survival for above subgroups censored at the time of allogeneic stem cell transplantation. Table indicates hazard ratio from univariate model. (C-D) Same as (A-B) above for the MLL cohort. HR=Hazard ratio, 95% CI = 95% Confidence interval.



Supplementary Figure 12: AML survival censored for allogeneic stem cell transplant

AML overall survival stratified by cohesin subunit mutation status, and cohesin-WT group by AML MR or AML-non-MR, censored at the time of first of allogeneic stem cell transplantation. HR=Hazard ratio. Statistical significance was determined using the log rank-test.