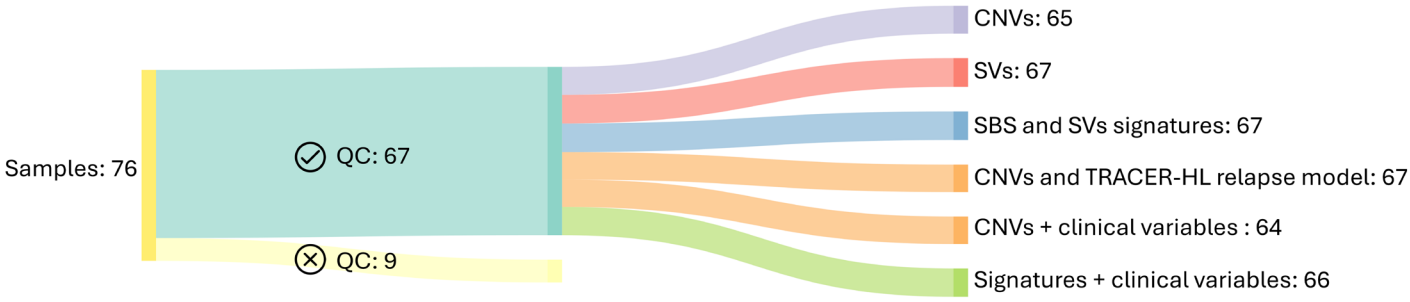
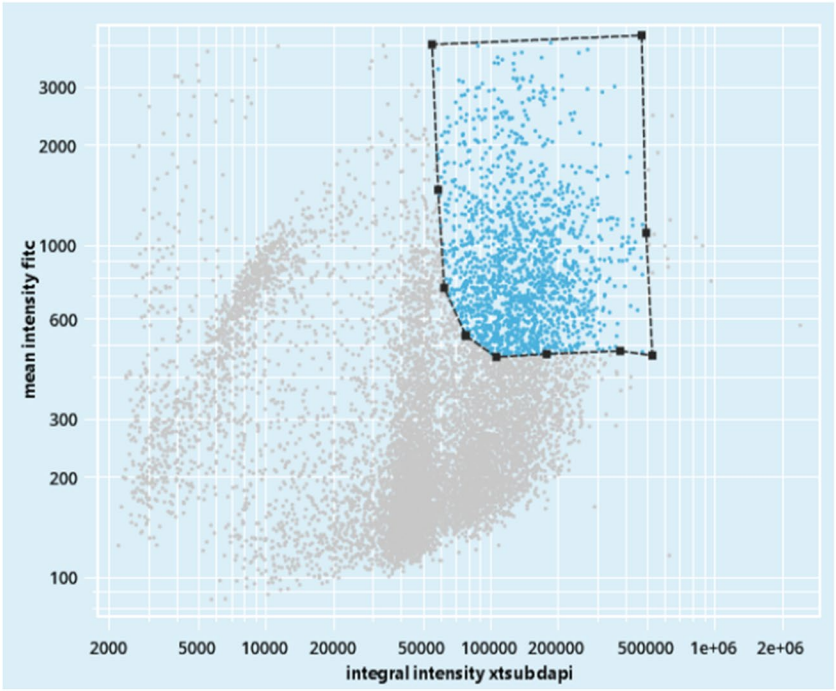
Supplementary Fig. S1



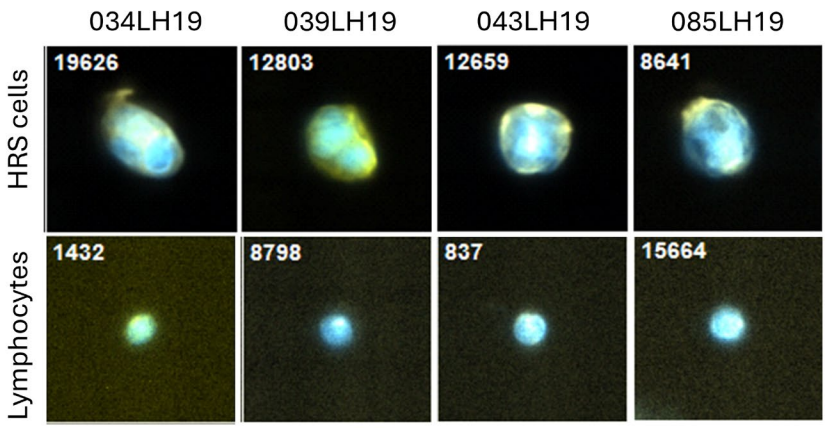
Supplementary Fig. S1. Sample flow diagram illustrating patient inclusion across analytical layers. Sankey plot showing the number of patients from the initial cohort (n = 76) that were successfully included in each genomic analysis after sequential quality control steps. Samples were excluded at each stage based on bioinformatic criteria specific to each analytical pipeline.

Supplementary Fig. S2

A

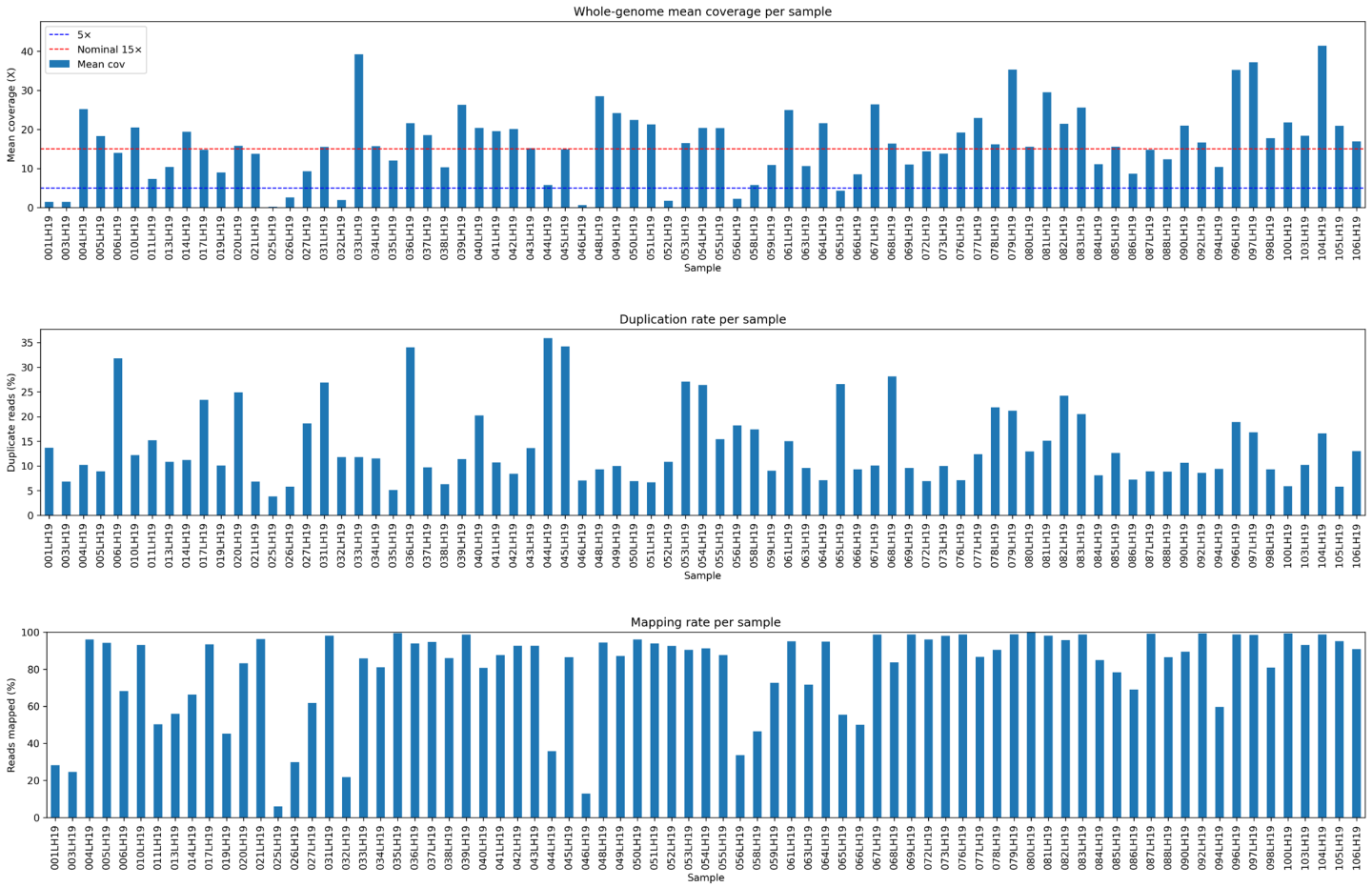


B



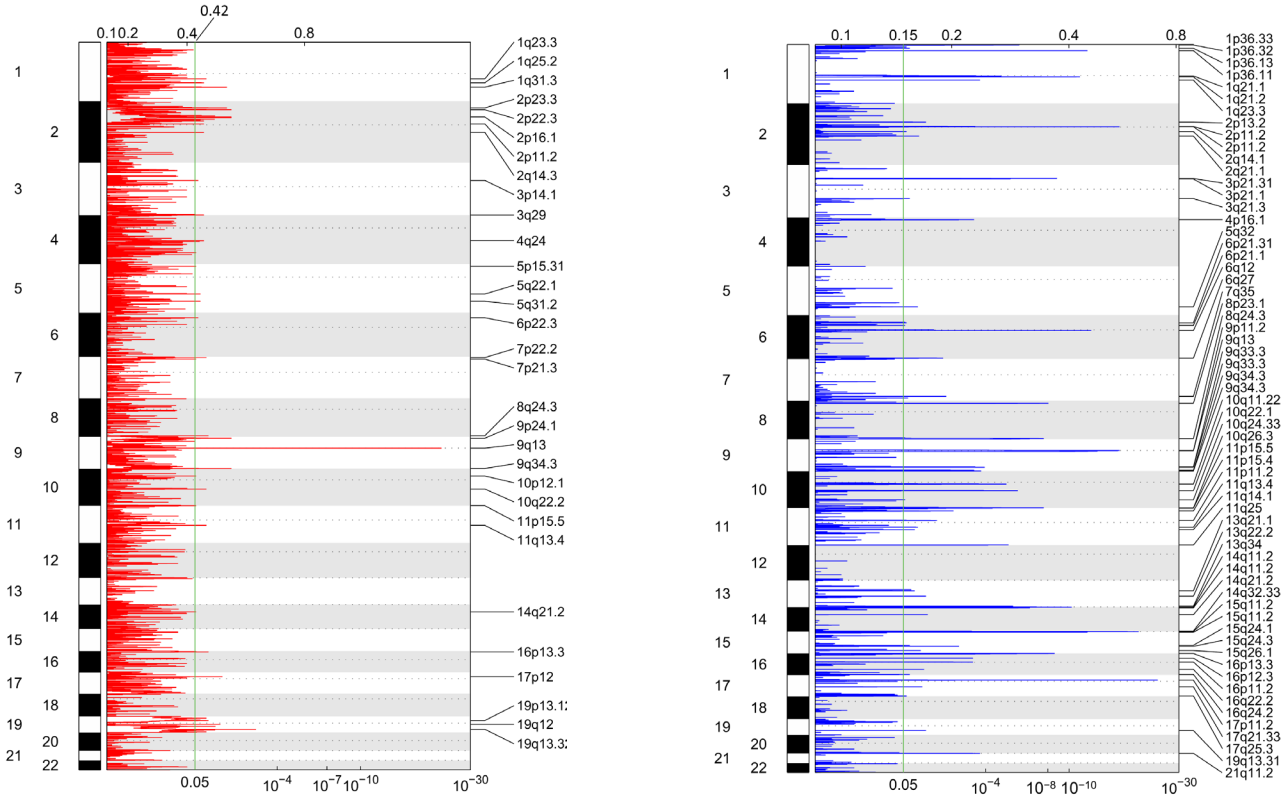
Supplementary Fig. S2. Selection of HRS cells from FFPE samples using DEPArray™ digital sorting. (A) Two-dimensional scatterplot representing single-cell events based on DAPI (X-axis) and CD30 (Y-axis) fluorescence intensity. HRS cells were selected based on high expression of CD30 and strong nuclear staining, consistent with their large size. **(B)** Representative immunofluorescence images from four independent FFPE samples showing the morphology and marker expression of individually selected HRS cells and background lymphocytes. HRS cells were positive for CD30 and DAPI, while non-malignant lymphocytes were identified based DAPI staining and size.

Supplementary Fig. S3



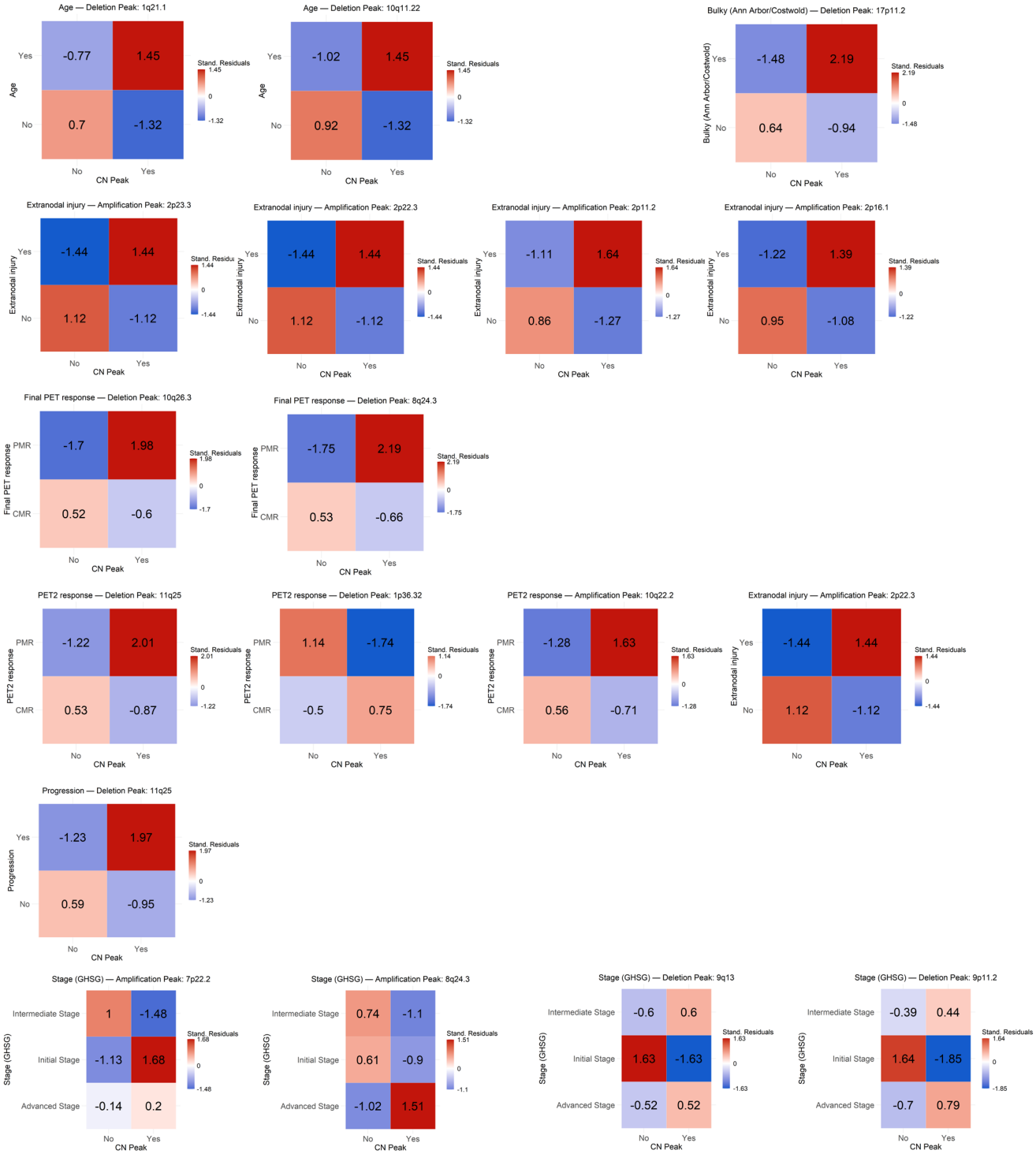
Supplementary Fig. S3. Whole-genome sequencing quality control metrics for purified HRS-cell samples. **Top panel:** mean whole-genome coverage per sample; blue bars indicate the average depth (X), the blue dashed line marks the minimum threshold of 5× and the red dashed line the nominal target coverage of 15×. **Middle panel:** duplication rate per sample, expressed as the percentage of reads marked as PCR/optical duplicates. **Bottom panel:** mapping rate per sample, showing the percentage of reads aligned to the reference genome.

Supplementary Fig. S4



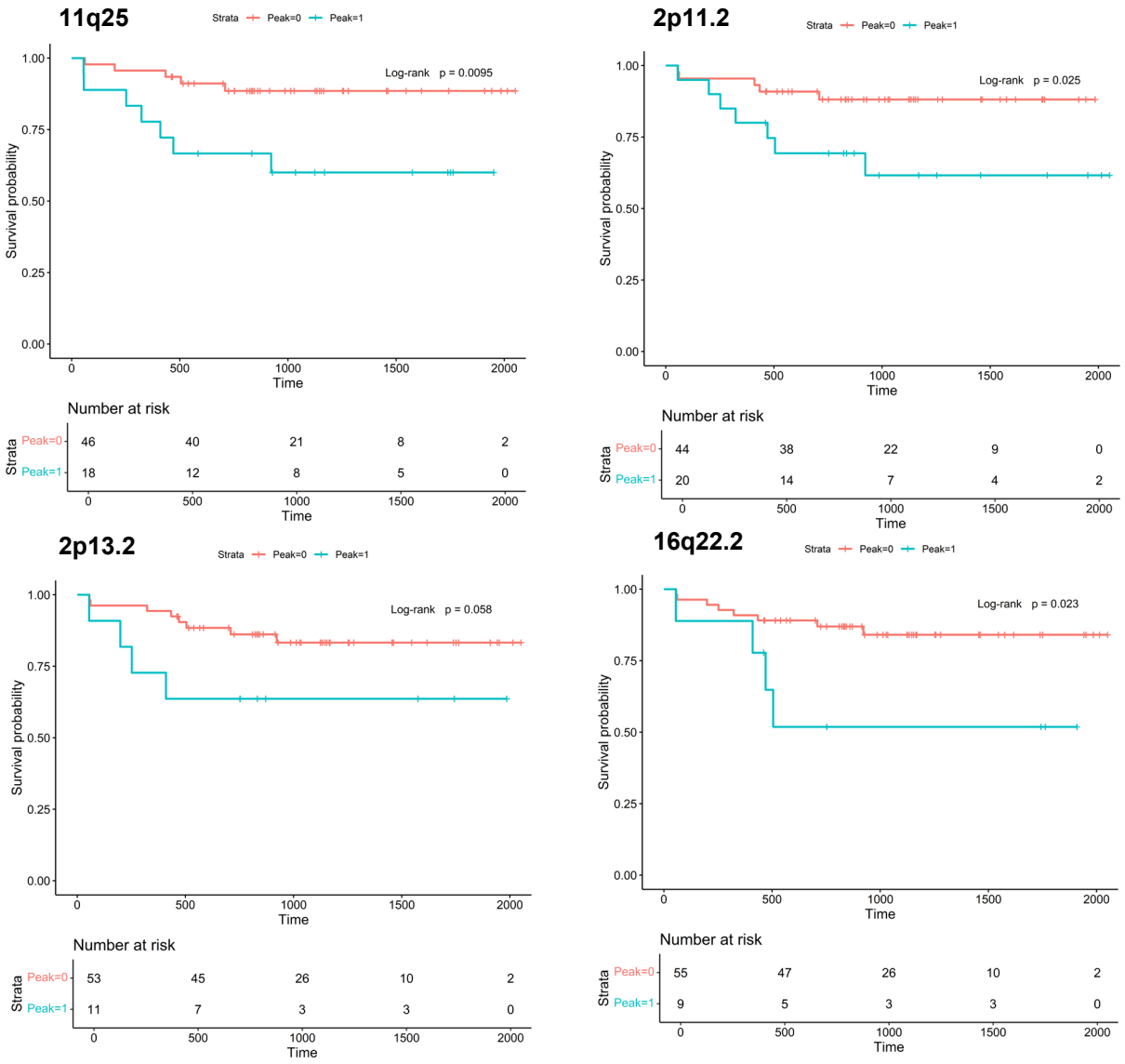
Supplementary Fig. S4. Genomic landscape of CNVs in HRS cells. Recurrent focal CNVs peaks identified by GISTIC2 (v2.0.23) from GATK-derived segments; statistically significant peaks ($q < 0.05$) are shown, with amplifications in the left panel (red) and deletions in the right panel (blue).

Supplementary Fig. S5



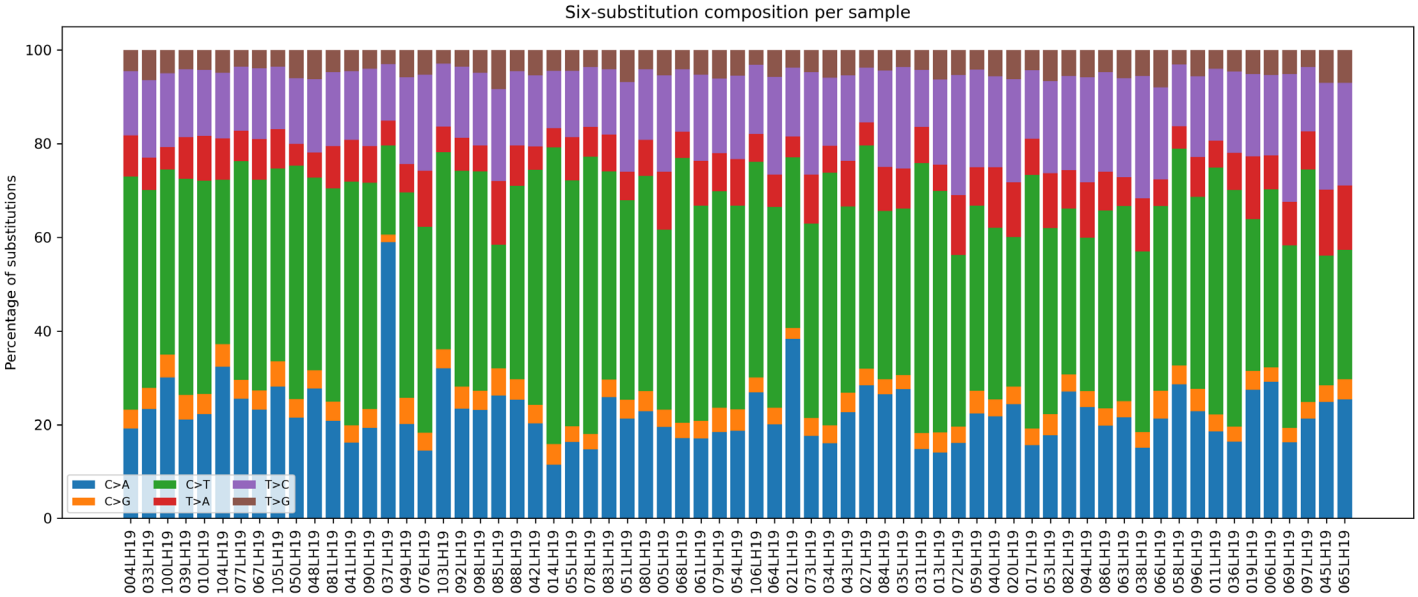
Supplementary Fig. S5. Mosaic plots displaying standardized residuals from contingency table analyses between recurrent CNVs events and relevant clinical variables. Residuals reflect the deviation between observed and expected frequencies, highlighting overrepresented (red) or underrepresented (blue) combinations.

Supplementary Fig. S6



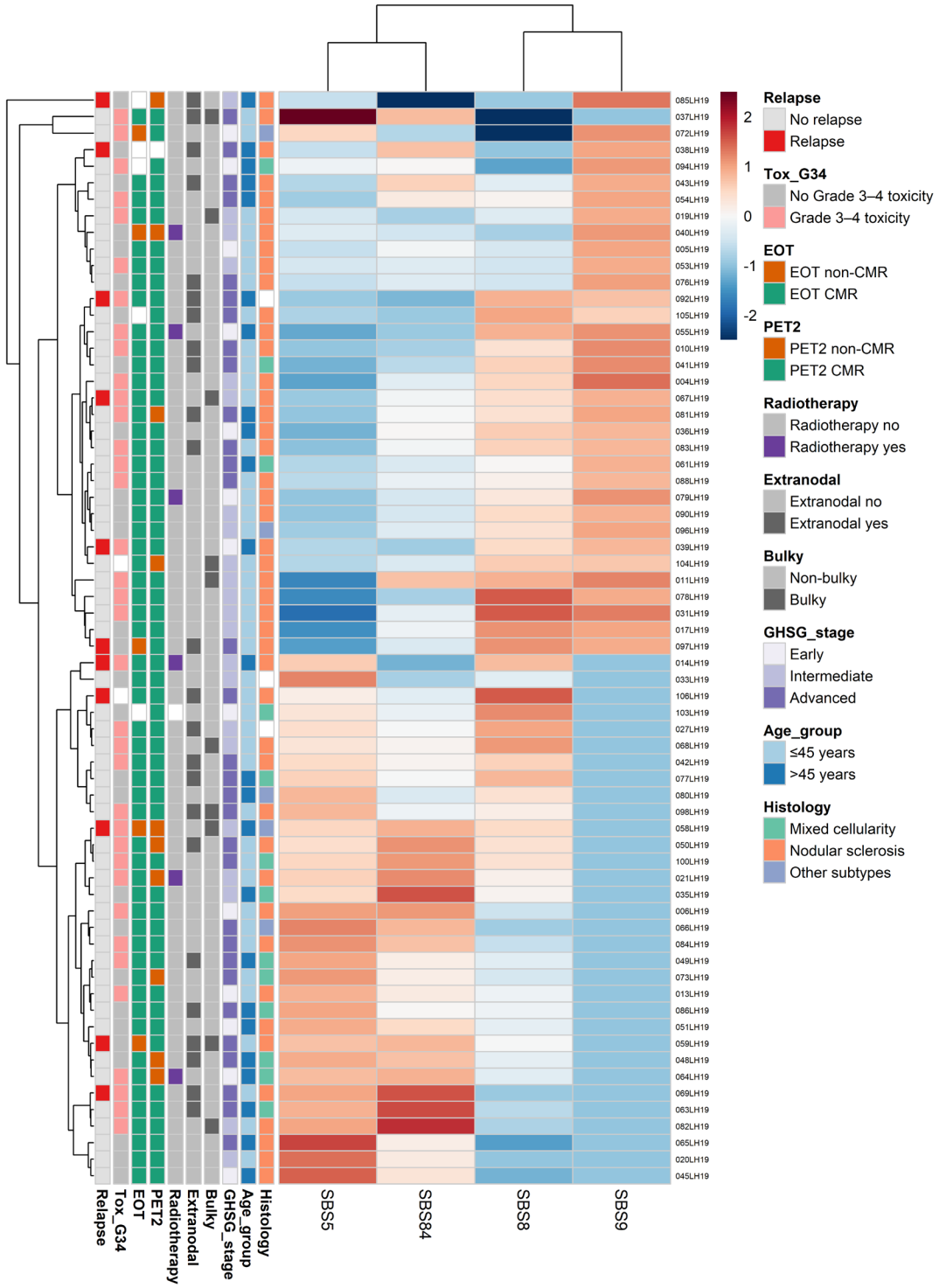
Supplementary Fig. S6. Kaplan–Meier progression-free survival curves stratified by the presence or absence of focal CNVs. All shown CNAs were statistically significant (log-rank $p < 0.05$), with the deletion at 2p13.2 displaying a trend towards significance ($p = 0.058$).

Supplementary Fig. S7



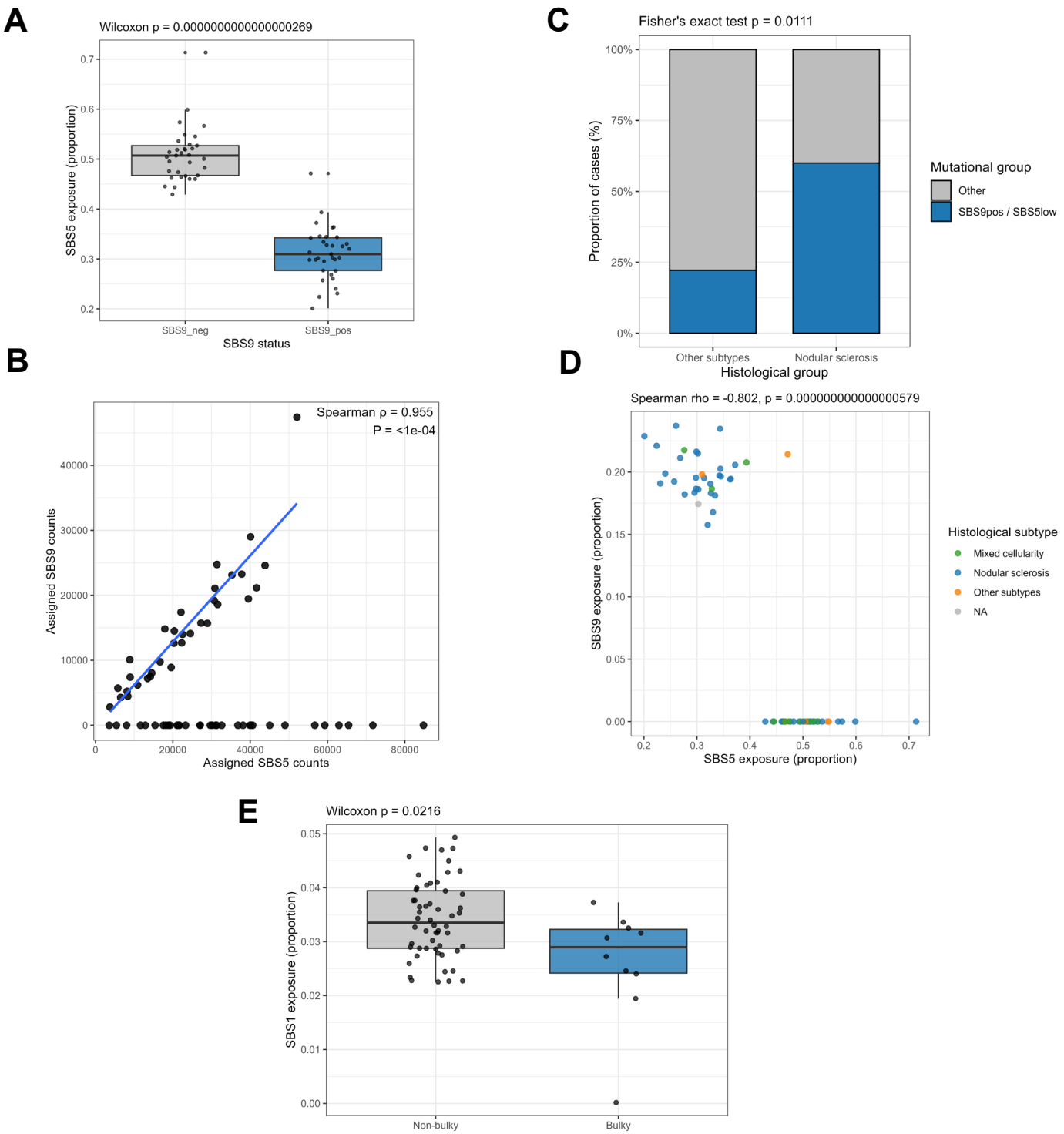
Supplementary Fig. S7. Six-substitution composition of 96-channel SBS mutational profiles across cHL cases. Stacked bar plot showing the relative contribution of the six base-substitution classes in each tumor. Although C>T substitutions were frequent across the cohort, the mutational spectra were not dominated exclusively by this class and retained substantial contributions from other substitution types across cases.

Supplementary Fig. S8



Supplementary Fig. S8. SBS COSMIC mutational signatures. Heatmap showing unsupervised hierarchical clustering of cHL samples based on the relative contribution of the four most prevalent COSMIC SBS signatures (SBS5, SBS8, SBS84 and SBS9). For each tumour, signature exposures are expressed as z-scores across the cohort (blue = lower, red = higher than the cohort mean). Rows represent individual patients, and columns represent SBS signatures. Dendrograms depict hierarchical clustering of samples and signatures using Euclidean distance and complete linkage. The annotation bar on the right summarises clinical and pathological features for each case.

Supplementary Fig. S9

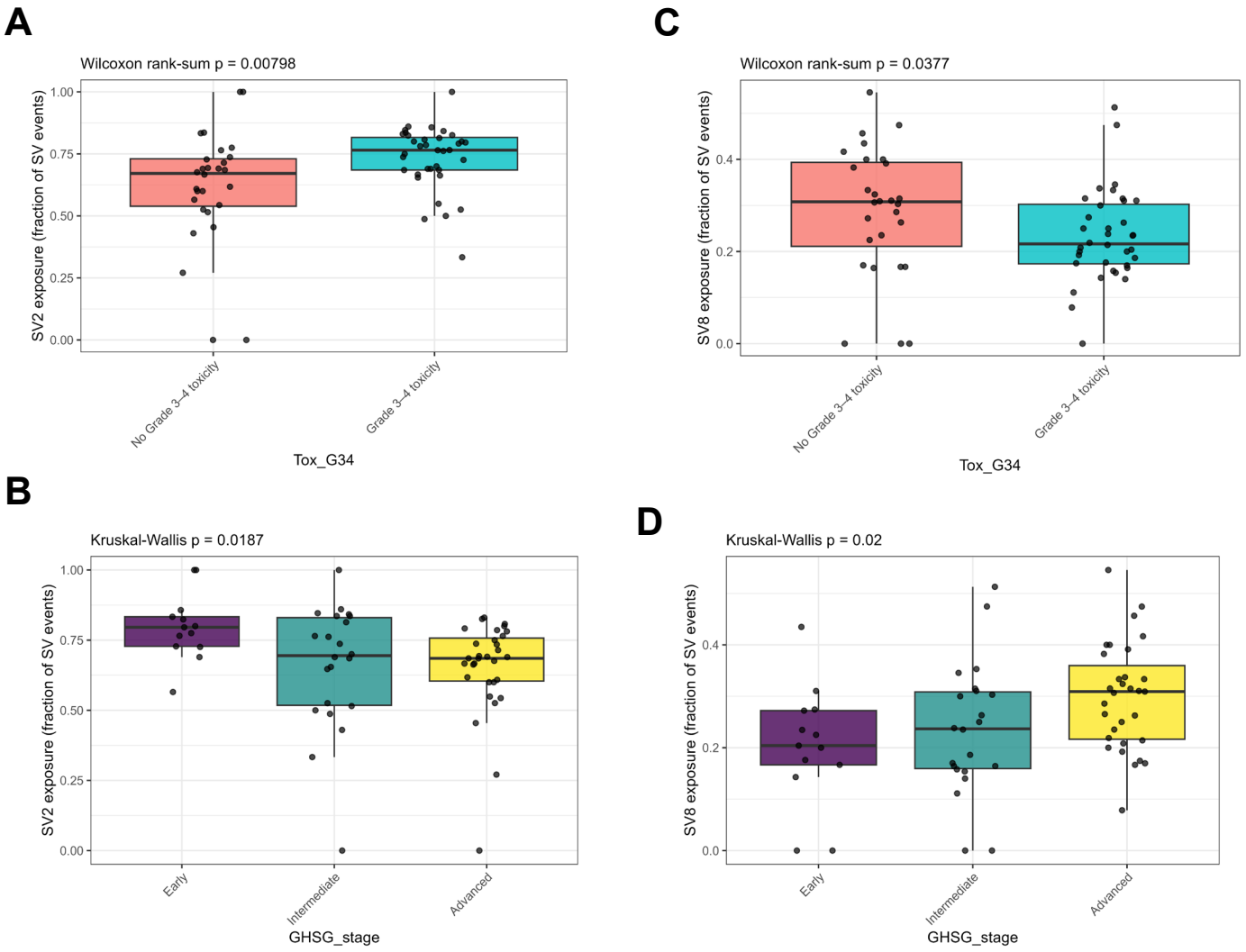


Supplementary Fig. S9. Clinical correlates of SBS1 and AID-like SBS9/SBS5 mutational patterns. **(A)** Boxplots showing SBS5 exposure (proportion of total SBS burden) in tumours classified as SBS9-negative (SBS9_neg) or SBS9-positive (SBS9_pos). Each dot represents an individual case; boxes indicate median and interquartile range. **(B)** Per-sample scatter plot of SBS5 and SBS9. SBS9 was present in only a subset of tumours, whereas SBS5 was broadly distributed across the cohort. **(C)** Proportion of patients classified within the AID-like mutational subgroup (SBS9-positive/SBS5-low; blue) versus all other cases (grey) across histological subtypes. **(D)** Scatterplot illustrating the inverse relationship between SBS5 and SBS9 exposure, coloured by histological subtype. Each point represents one tumour, with signature exposures expressed as proportions of total SBS mutations. **(E)** Boxplot showing SBS1 exposure (proportion of SBS mutations) according to presence of bulky disease at diagnosis. Each dot represents an individual case; boxes indicate median and interquartile range.

Supplementary Fig. S10

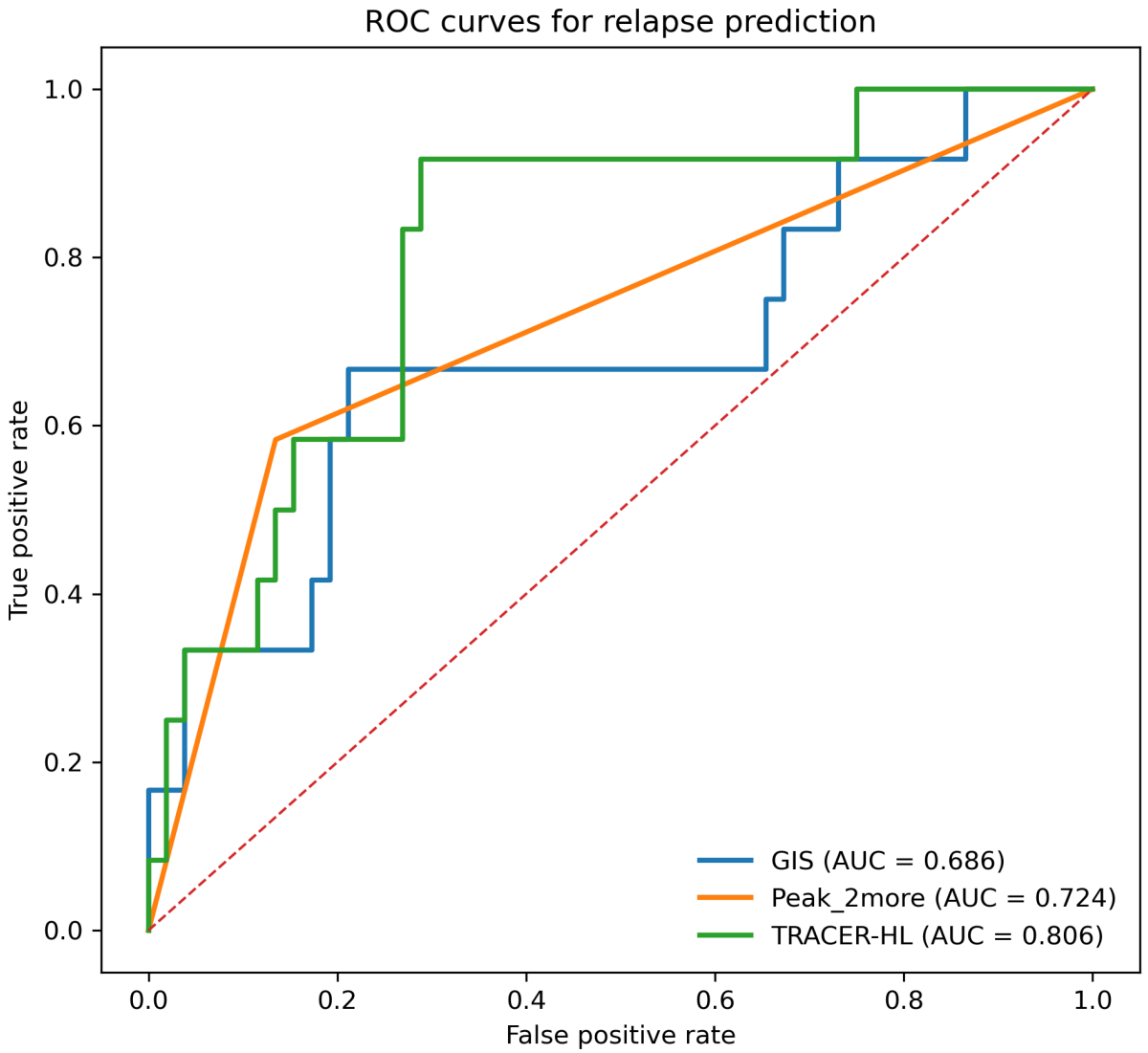
Supplementary Fig. S10. Unsupervised hierarchical clustering of SVs mutational signature exposures across cHL samples. Rows represent individual cases and columns show signatures with non-zero contribution in at least one sample are displayed. The annotation bars depict clinical and treatment variables, including relapse status, grade 3–4 toxicity (Tox_G34), end-of-treatment PET response (EOT), interim PET2 response, use of radiotherapy, extranodal involvement, bulky disease, GHSG risk stage, age group and histological subtype. Unavailable clinical data are shown in white in the plot.

Supplementary Fig. S11



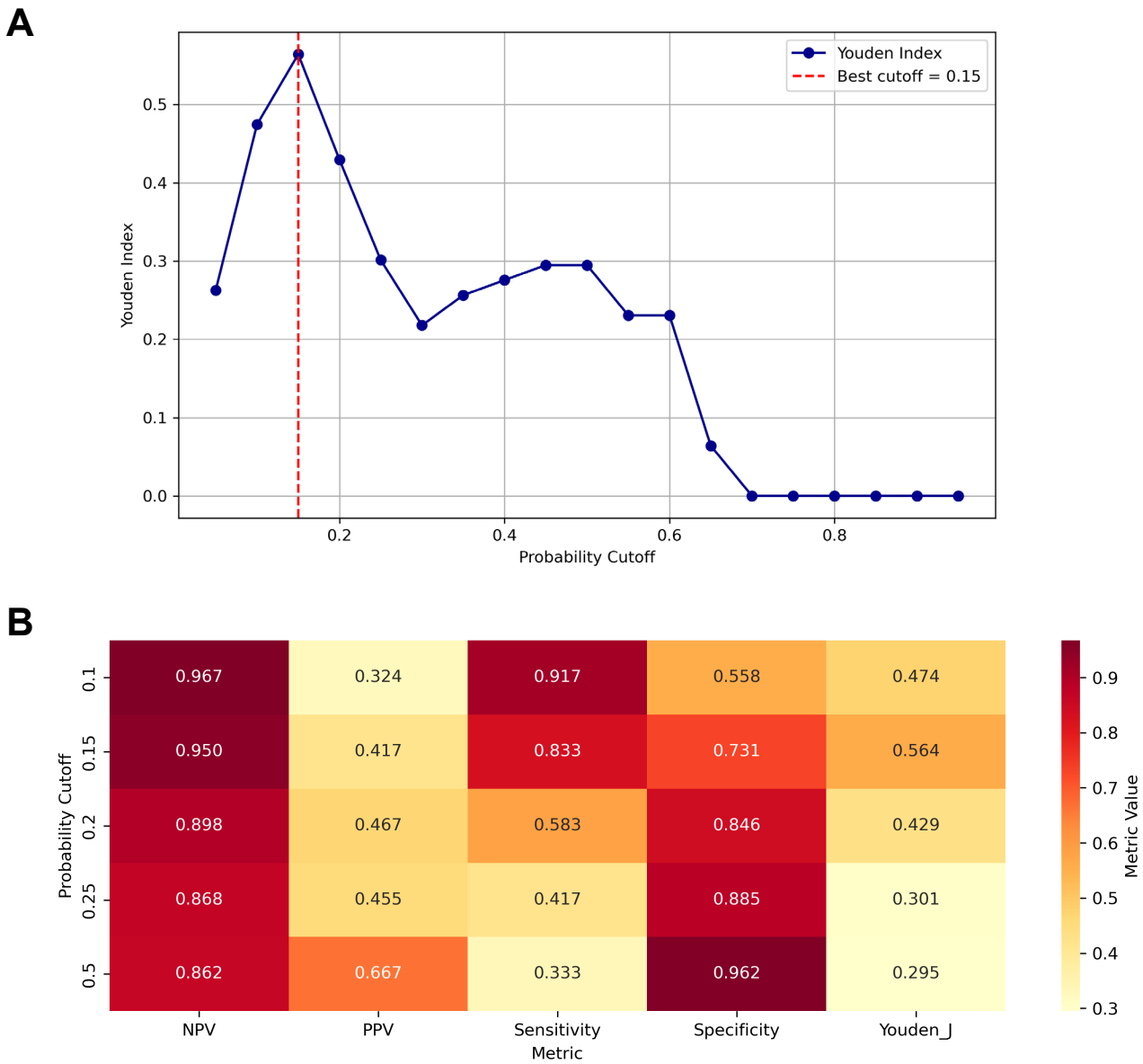
Supplementary Fig. S11. Associations between SVs mutational signatures and clinical features. **(A)** Boxplots showing SV2 exposure (fraction of SV events) according to the presence or absence of grade 3–4 treatment-related toxicity (Wilcoxon rank-sum p = 0.00798). **(B)** SV2 exposure stratified by GHSG stage (early, intermediate, advanced; Kruskal–Wallis p = 0.0187). **(C)** SV8 exposure according to grade 3–4 toxicity status (Wilcoxon rank-sum p = 0.0377). **(D)** SV8 exposure across GHSG stages (Kruskal–Wallis p = 0.02). Each dot represents one patient.

Supplementary Fig. S12



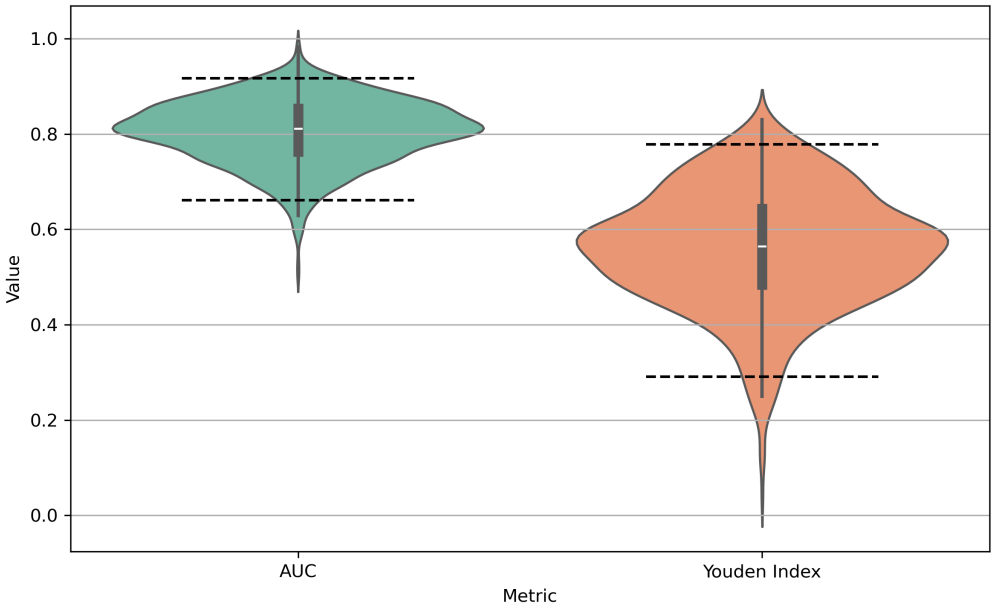
Supplementary Fig. S12. **Comparative ROC curves of individual and combined genomic predictors for relapse.** ROC curves illustrating the performance of individual genomic features and their combination in predicting relapse in cHL. The red dotted line represents random classification (AUC = 0.5).

Supplementary Fig. S13



Supplementary Fig. S13. Optimization of the TRACER-HL model cutoff using Youden index. **(A)** Line plot showing the Youden index across a range of probability thresholds (x-axis). The red dashed line marks the maximum Youden index value (J = 0.564) observed at a probability cutoff of 0.15, which was selected as the optimal threshold for classification. This value reflects the best balance between sensitivity and specificity for relapse prediction in the study cohort. **(B)** Heatmap of TRACE-HL model performance metrics across probability thresholds. Matrix plot showing the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and Youden index across selected probability cutoffs. Each cell is color-coded by the corresponding Youden index value, with annotations indicating exact metric values.

Supplementary Fig. S14



Supplementary Fig. S14. **Distribution of bootstrap-estimated AUC and Youden Index (n=1000 iterations)**. Violin plots represent the full distribution of each metric, with the white line indicating the median and dashed lines denoting the 95% confidence intervals.