

Figure S1

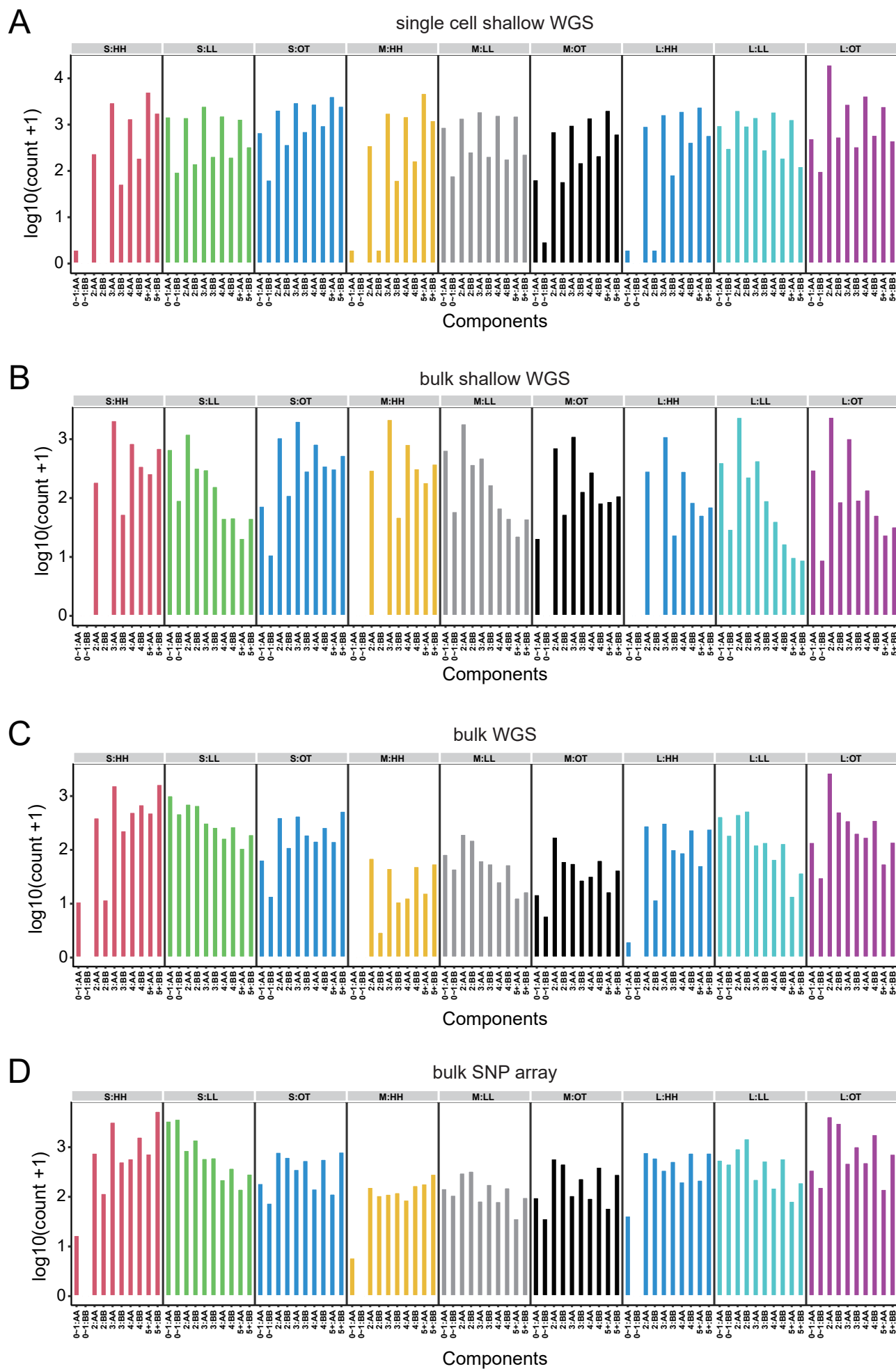


Figure S1. The distribution of the CNA features in different datasets.

A, CNA feature matrix distribution of single-cell sWGS data.

B, CNA feature matrix distribution of bulk sWGS data.

C, CNA feature matrix distribution of bulk WGS data.

D, CNA feature matrix distribution of bulk SNP array data.

Figure S2

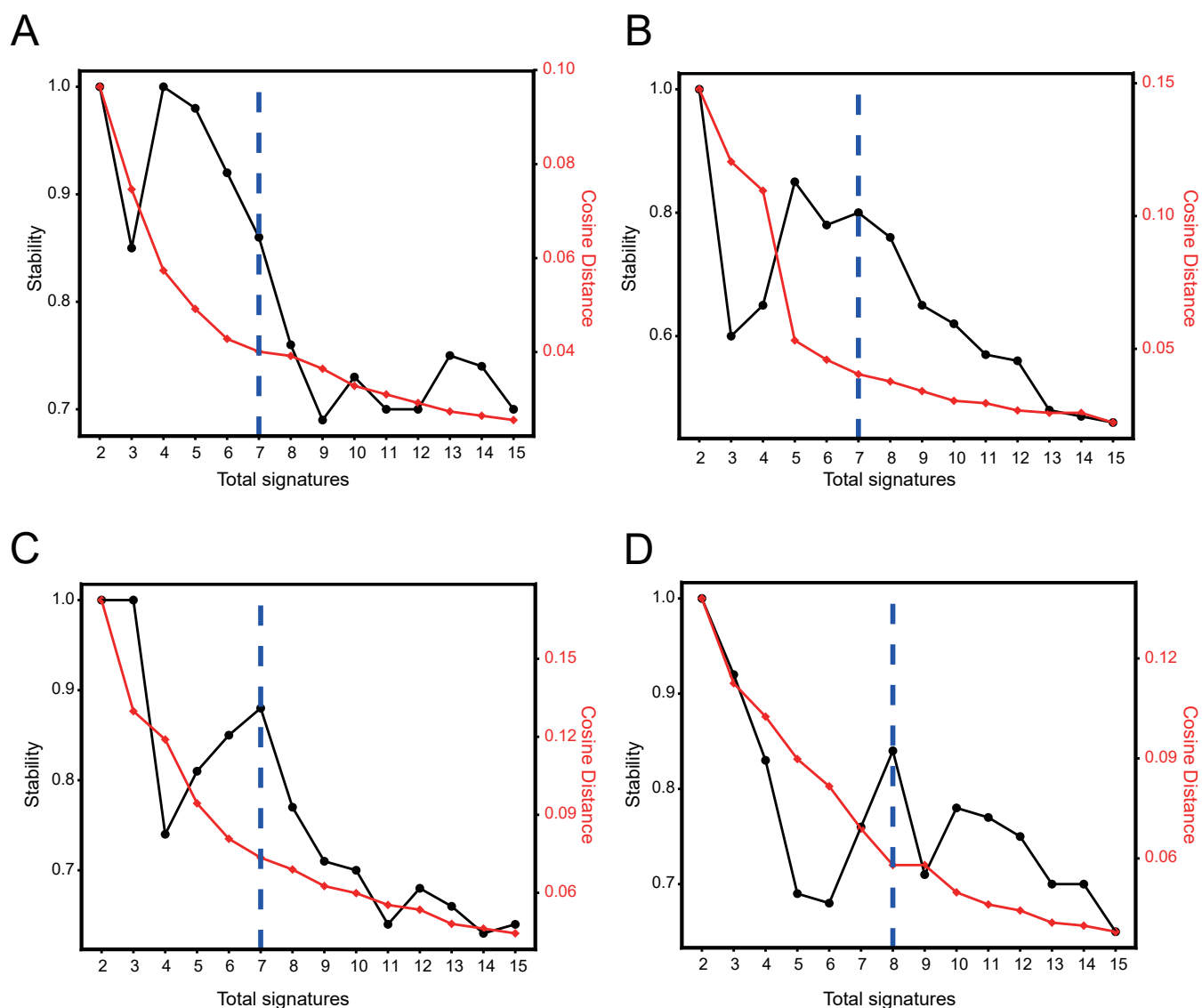


Figure S2. Determining the number of CNA signatures.

The abscissa is the number of signatures, the y-axis on the left is the stability of each signature, and the right is the similarity between signatures. Signatures have stability and low similarity. A-D are single-cell sWGS data (A), bulk sWGS (B), bulk WGS (C), and bulk SNP array (D) data.

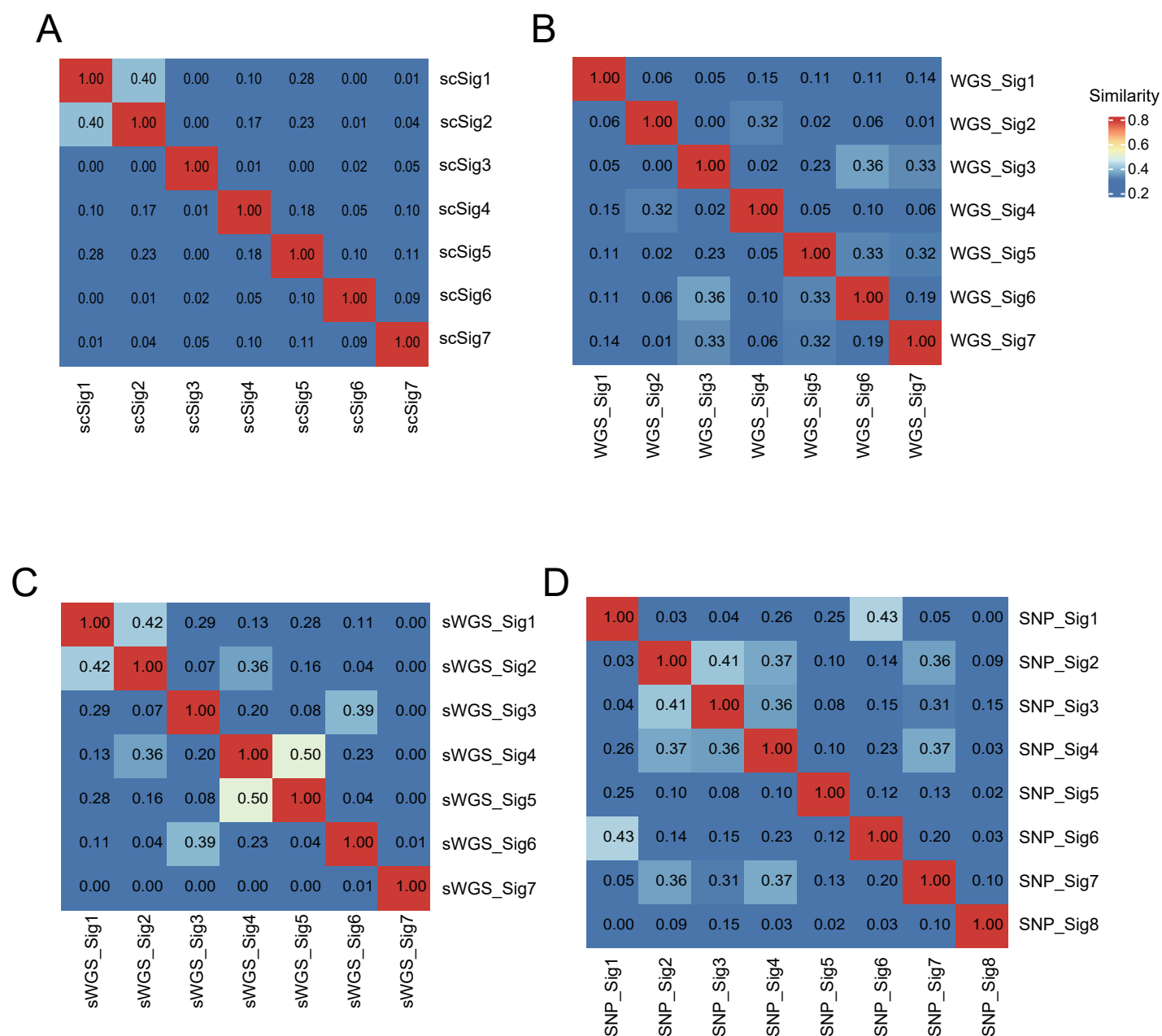


Figure S3. Correlations between copy number signatures.

Inter-correlation analysis of the profiles of different dataset CNA signatures. The numbers are cosine similarity values comparing each pair of CNA signatures. A-D are the signatures of scSig (A), sWGS_Sig (B), WGS_Sig (C), and SNP_Sig (D) datasets.

Figure S4

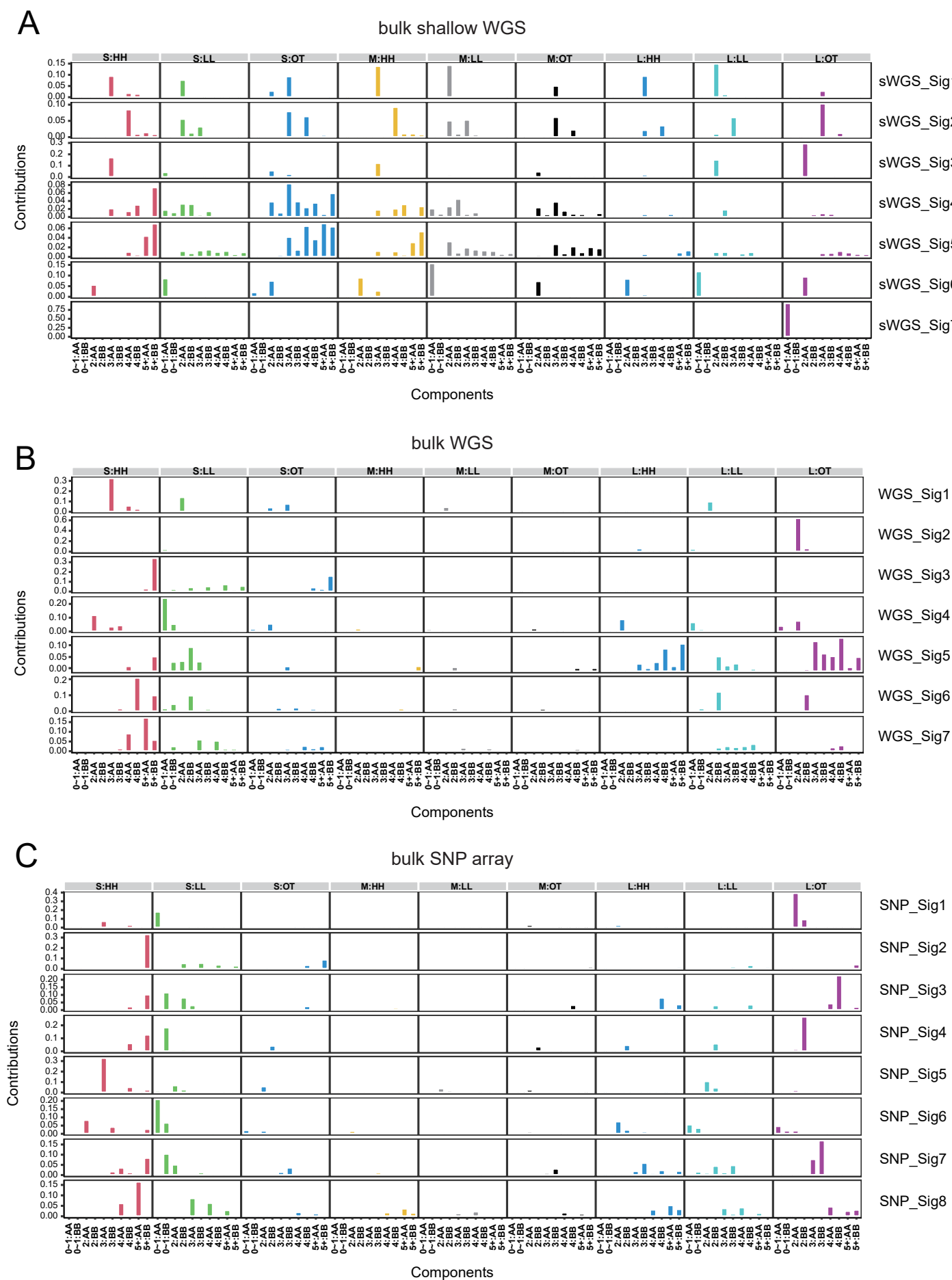
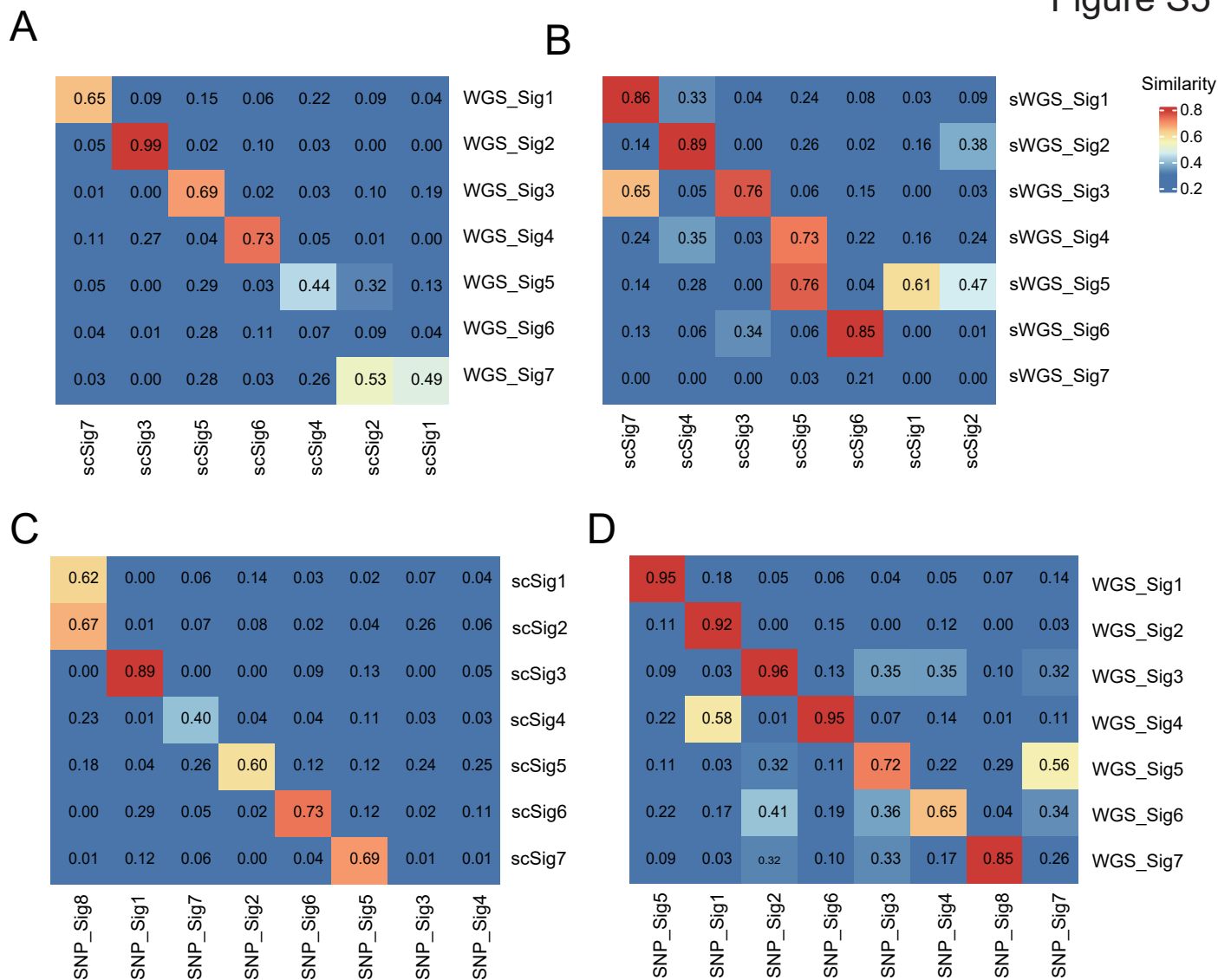


Figure S4. Profiles of CNA signatures extracted from different types of datasets.

The x-axis represents 90 types of CNA features and the y-axis indicates the contribution of each feature. A-C are the CNA feature matrices of bulk sWGS (A), bulk WGS (B), and bulk SNP array (C) data.

Figure S5

**Figure S5. Correlations of CNA signatures in different datasets**

A, Inter-correlations between the CNA signatures extracted in single-cell sWGS dataset and bulk WGS dataset. Cosine similarity values are reported for each comparison.

B, Inter-correlations between the CNA signatures extracted in single-cell sWGS dataset and bulk sWGS data dataset.

C, Inter-correlations between the CNA signatures extracted in single-cell sWGS dataset and bulk SNP array dataset.

D, Inter-correlations between the CNA signatures extracted in bulk WGS dataset and bulk SNP array dataset.

Figure S6

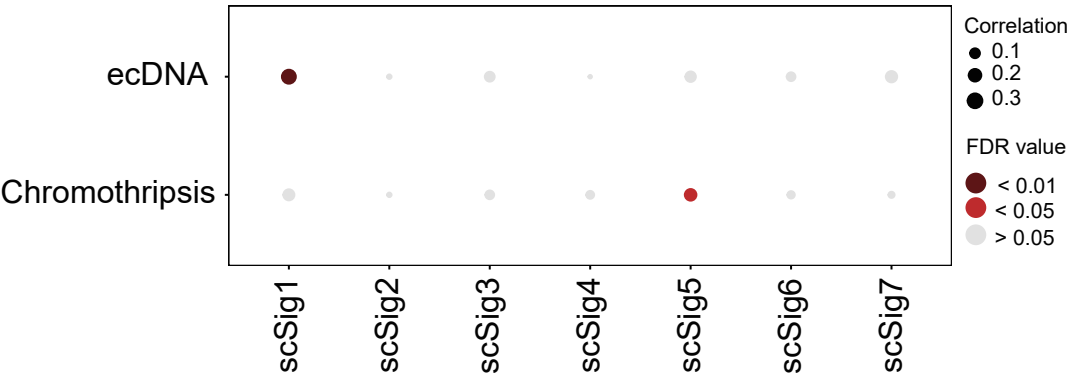


Figure S6. Associations between ecDNA, chromothripsis and CNA signatures

The Pearson correlation coefficient is shown, the circle's size indicates the correlation level, and the color indicates significance. Significant associations with FDR adjusted $P < 0.05$ are showed. scSig1 show the strongest correlation with ecDNA event, and scSig5 show the strongest correlation with chromothripsis event.

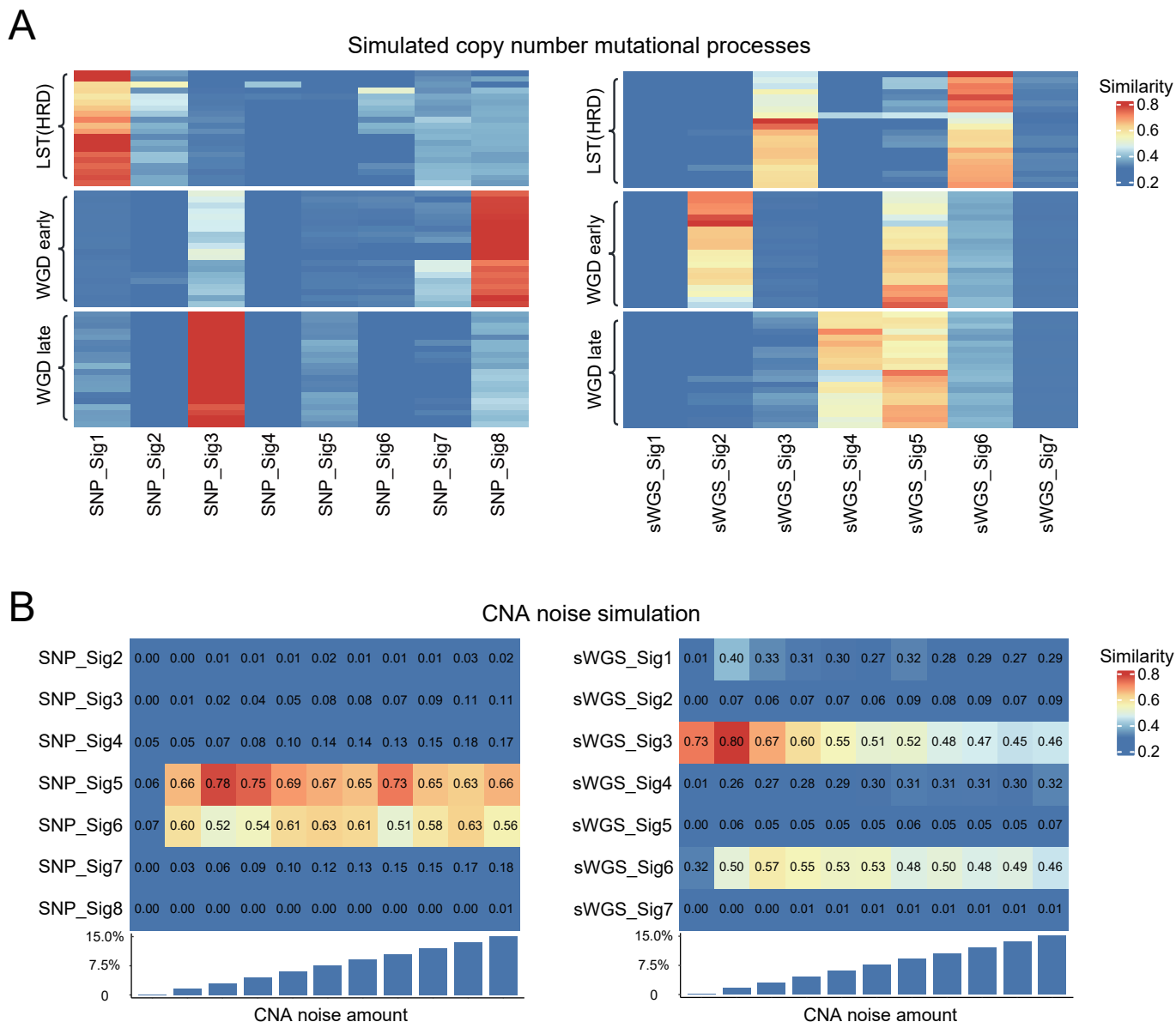


Figure S7. CNA mutational processes and noise simulation analysis.

A, Simulation of real CNA events, LST, early WGD, and late WGD were simulated in twenty samples. The similarity between the CNA profile of each sample and CNA signature is calculated through cosine similarity.

B, Simulate different levels of CNA noise (0~15%), and calculate the cosine similarity between the CNA profile and CNA signature under each noise condition. Each noise level is simulated 100 times, and the mean similarity is used for plotting.

Figure S8

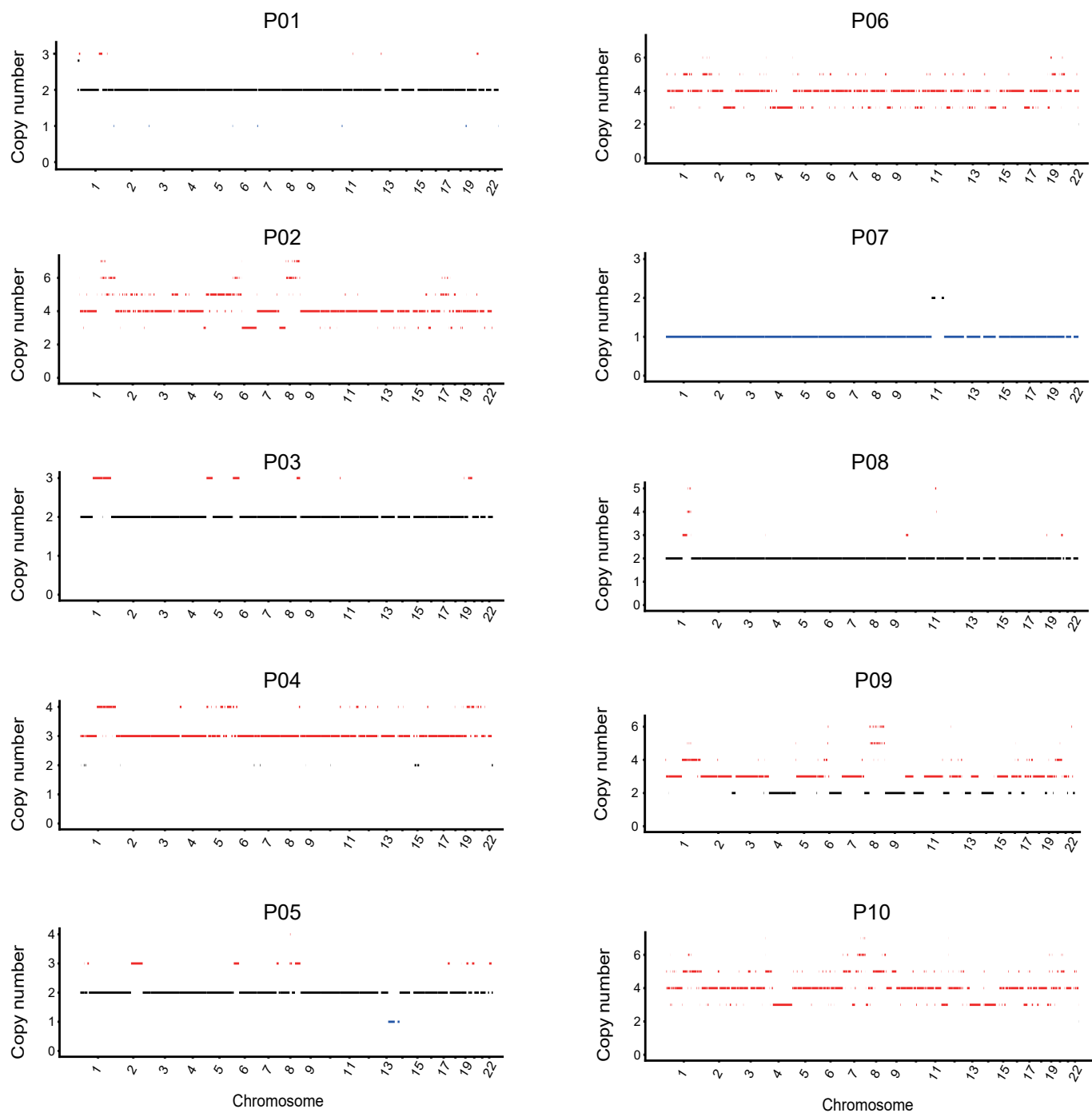


Figure S8. CNA profiles of bulk tissue samples for each corresponding single-cell samples.

Based on the raw data of single cells of each sample, the corresponding bulk tissue data is generated, and then the copy number is extracted. Black lines represent normal genome, blue lines represent deletions, and red lines represent amplifications.

Figure S9

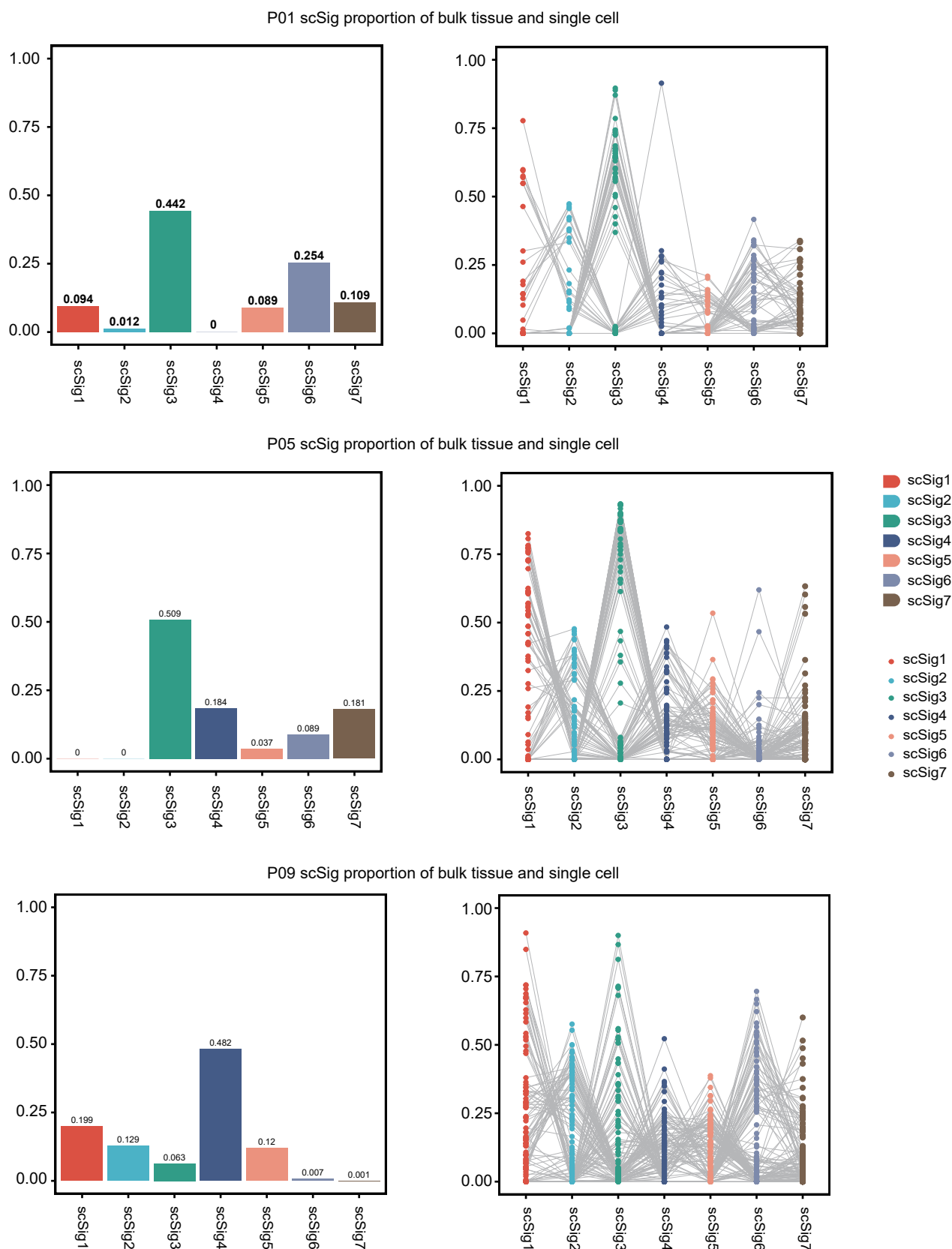


Figure S9. Comparison of scSig proportions between bulk samples and single-cell samples

Distribution of scSig proportions in bulk tissue and single cells in samples P01, P05 and P09. There is a significant difference between scSig in bulk samples and scSig in single cells.

Figure S10

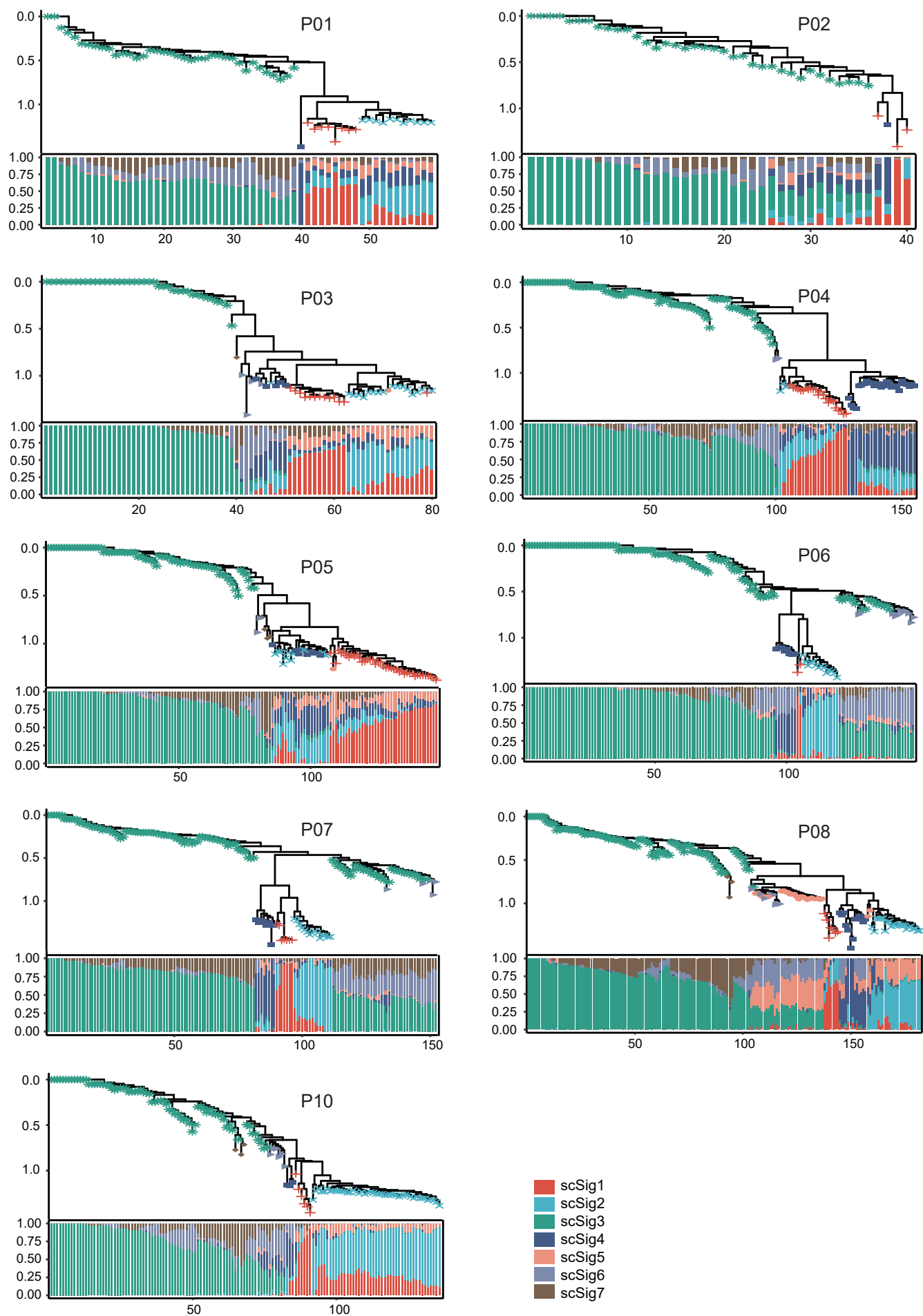


Figure S10. The scSig evolutionary tree of single-cell samples.

Evolutionary relationship of scSig. Calculate the distance between cells based on scSig, infer the evolutionary relationship based on the distance, and then draw an evolutionary tree. Symbols on each branch represent the most prevalent scSig within that lineage.

Figure S11

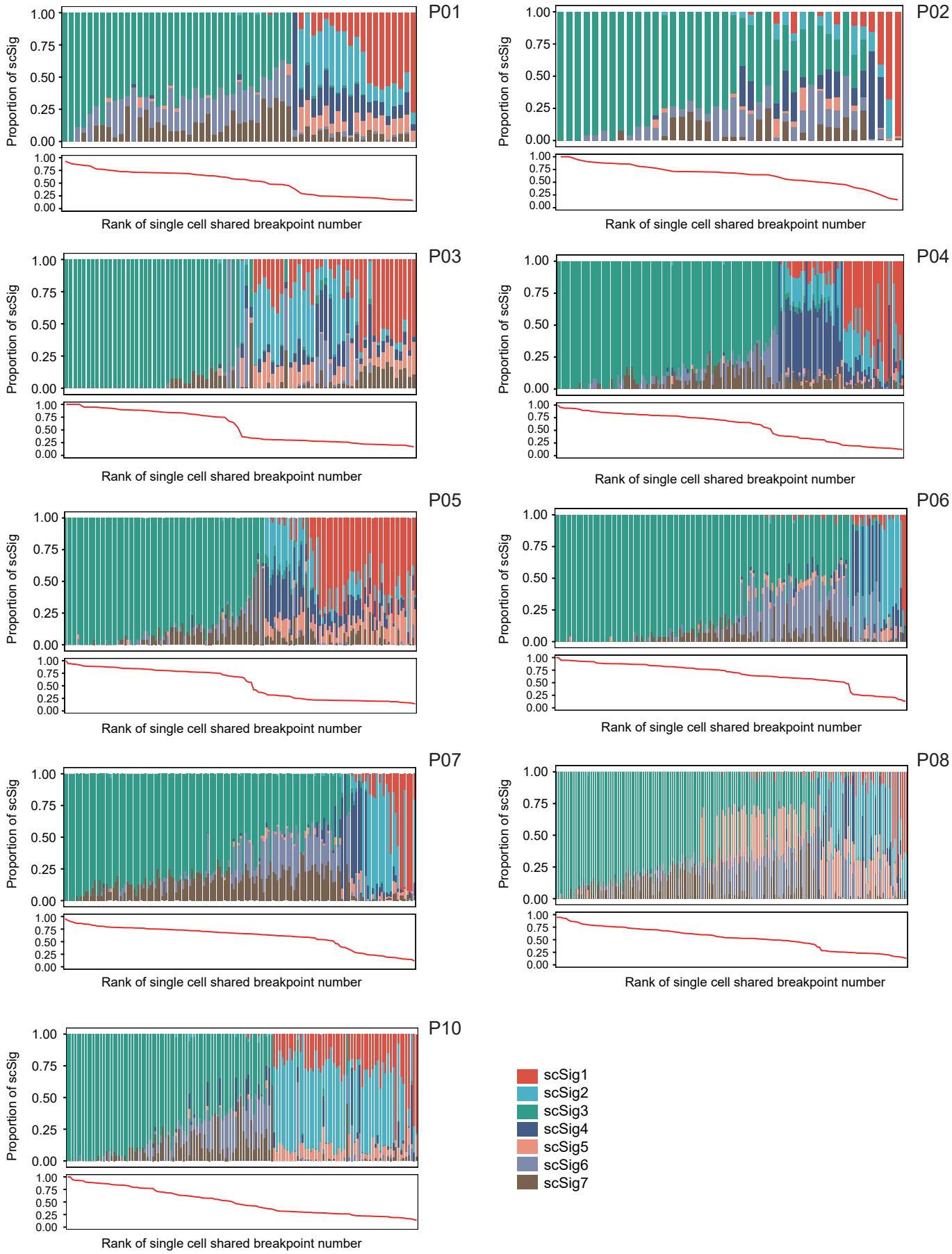


Figure S11. scSig evolutionary trajectory analysis based on shared breakpoints.

The average number of shared breakpoints per cell within the sample is calculated, and cells are ranked according to the ratio of shared breakpoints. Cells with a higher ratio of shared breakpoints are positioned early in the clonal evolution, as subsequent subclones retain the breakpoints of earlier cells.

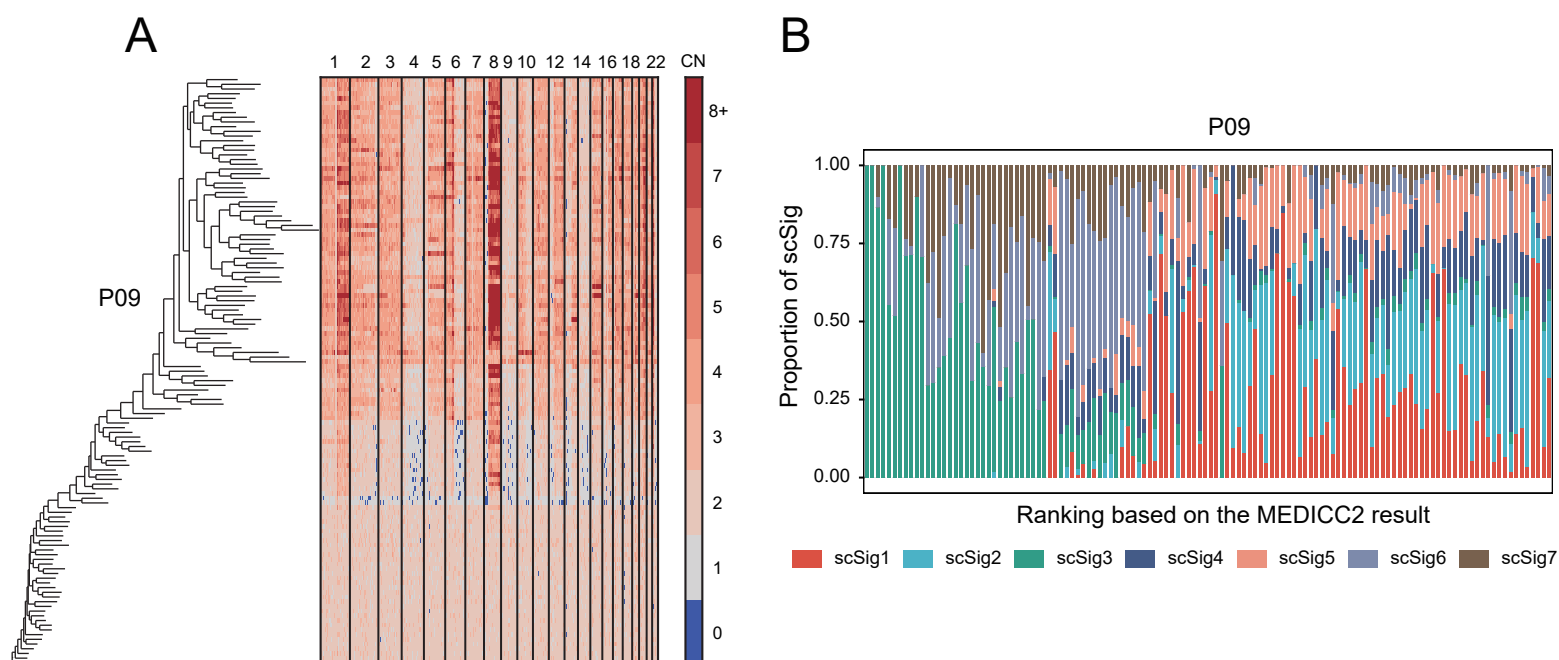


Figure S12. scSig evolutionary trajectory analysis based on the MEDICC2 algorithm.

A, Single-cell CNA evolutionary trajectory of patient P09 calculated using the MEDICC2, with each branch representing a single-cell sample.

B, scSig evolution. Sorting of cells in the sample based on the pseudotemporal sequence calculated by MEDICC2, showing the proportion of each scSig in each single cell.

Figure S13

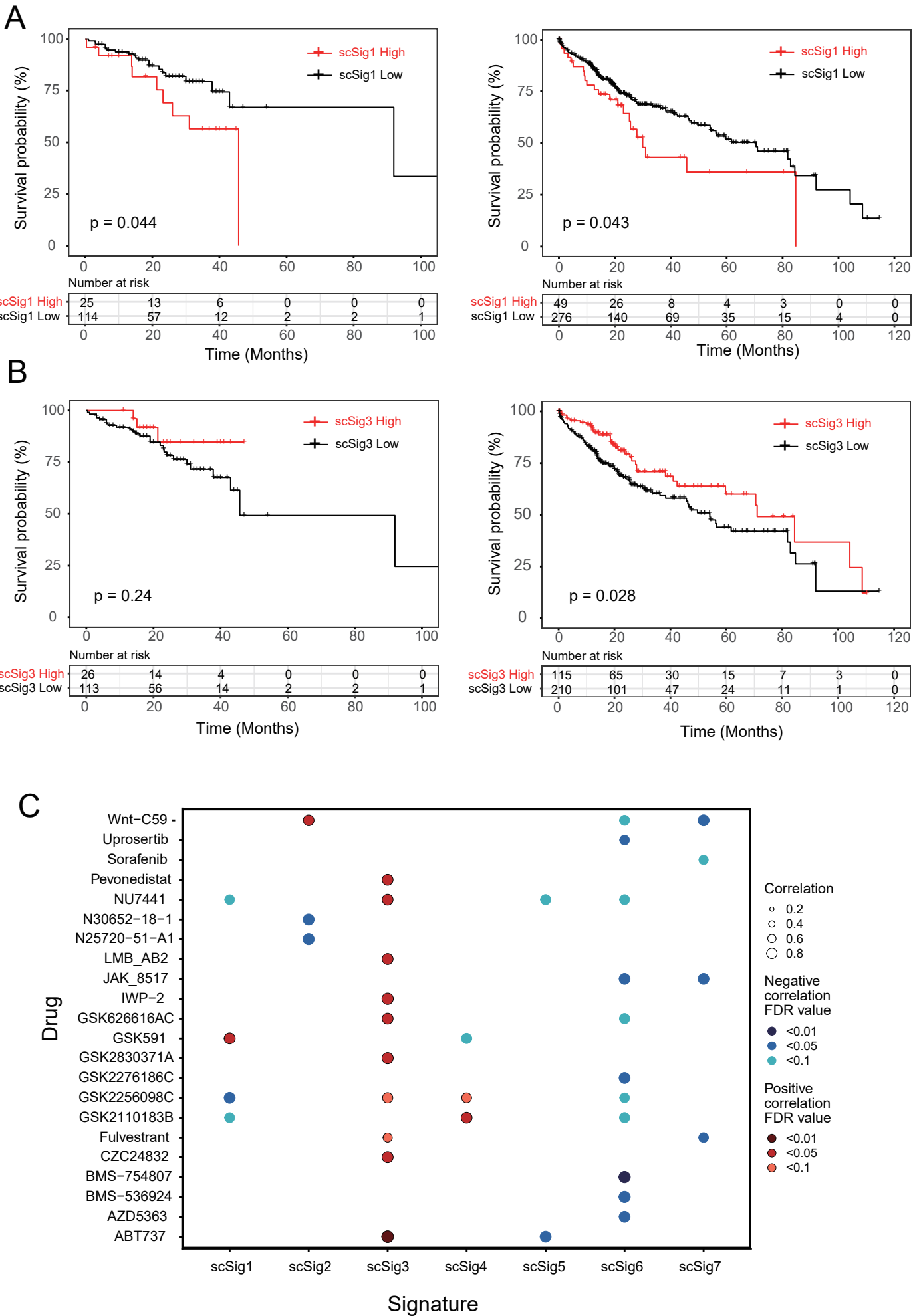


Figure S13. Application of scSig signatures in HCC prognosis and drug sensitivity prediction.

A, The prognosis performance of the scSig1 in both WGS and SNP datasets. Consistent prognostic results in two different datasets.

B, The prognosis performance of the scSig3 in both WGS and SNP datasets. Consistent prognostic results in two different datasets.

C, The relationship between scSig and drug sensitivity. Spearman correlation analysis was utilized to assess the relationship between the exposure values of scSig for each sample and IC50. The significance P-values obtained were corrected using the FDR method.