

sSNAPPY: an R/Bioconductor package for single-sample directional pathway perturbation analysis

Wenjun Liu^{1,2,3}, Ville-Petteri Mäkinen^{4,5}, Wayne D. Tilley¹, and Stephen M. Pederson^{1,6,7}

¹Dame Roma Mitchell Cancer Research Laboratories, Adelaide Medical School, Faculty of Health and Medical Sciences, University of Adelaide, Adelaide, Australia

²Adelaide Centre for Epigenetics, School of Biomedicine, Faculty of Health and Medical Sciences, University of Adelaide, Adelaide, Australia

³The South Australian Immunogenomics Cancer Institute, Faculty of Health and Medical Sciences, University of Adelaide, Adelaide, Australia

⁴Computational Medicine, Faculty of Medicine, University of Oulu, Oulu, Finland

⁵Center for Life Course Health Research, Faculty of Medicine, University of Oulu, Oulu, Finland

⁶Black Ochre Data Laboratories, Indigenous Genomics, Telethon Kids Institute, Adelaide, Australia

⁷John Curtin School of Medical Research, Australian National University, Canberra, Australia

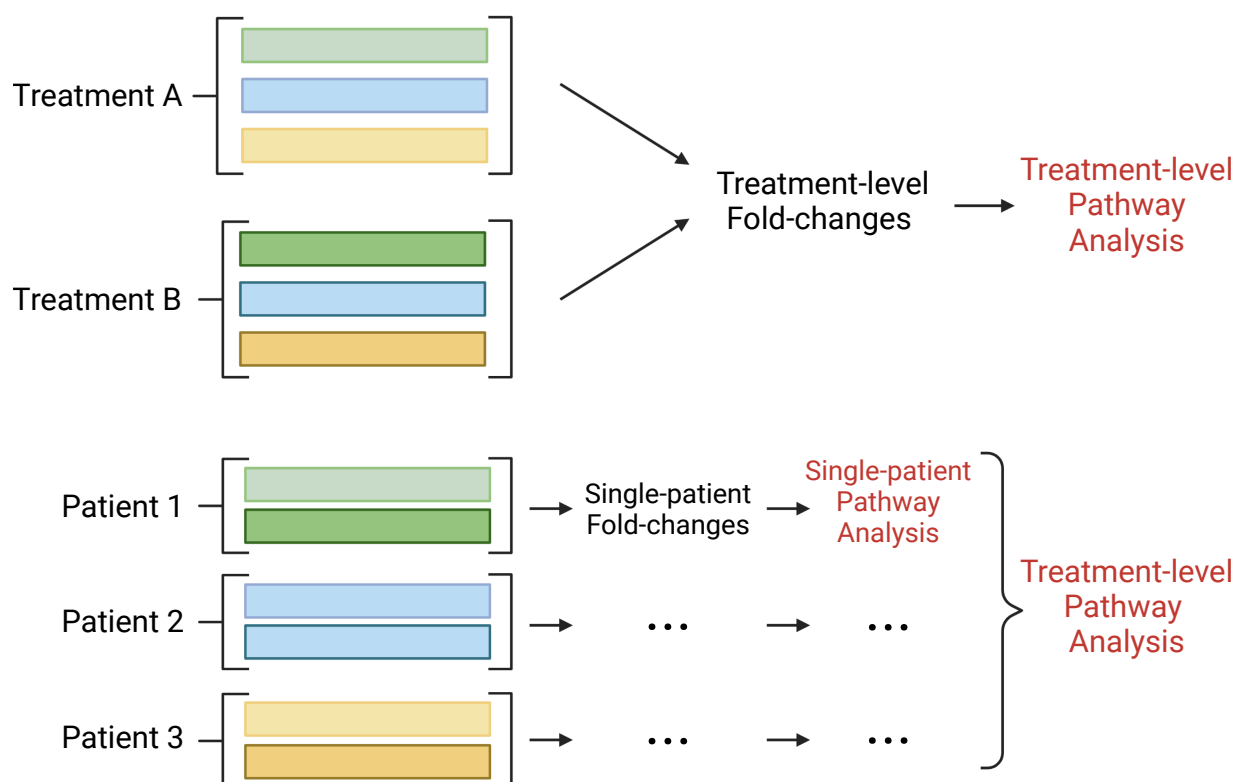


Figure 1. Schematic illustration of the differences between conventional pathway analysis methods and sSNAPPY. Instead of being limited to treatment-level analyses, sSNAPPY allows the detection of pathway perturbation in individual samples by using sample-specific estimates of fold-change instead of experiment-wide estimates. (Created with BioRender.com).

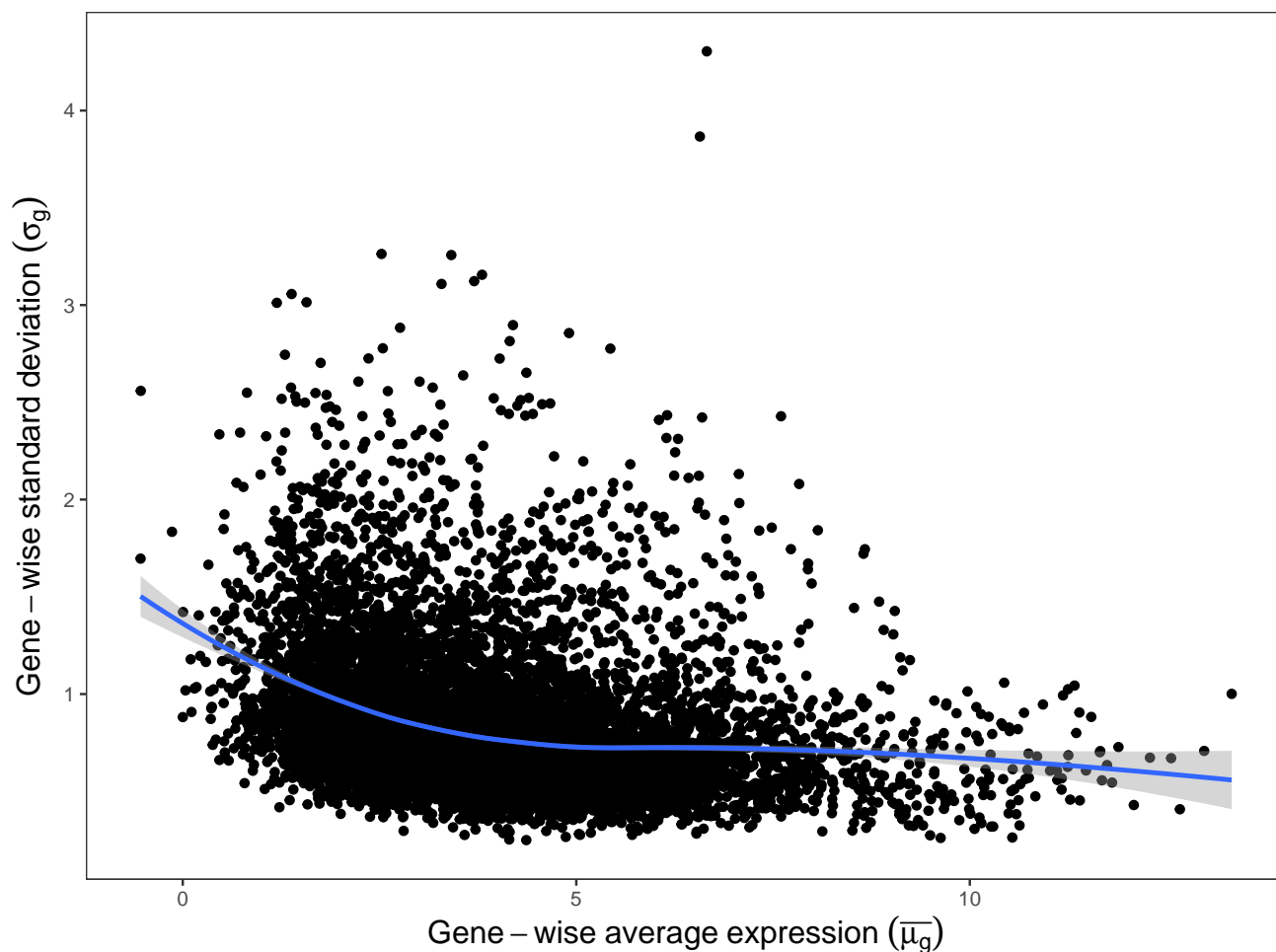


Figure 2. Gene-wise standard deviations are plotted against the mean logCPM values with mean-variance trend modelled by a loess fit. Whilst standard deviations are shown here for the purposes of visualisation, gene-level weights are calculated using variances at this stage of the sSNAPPY algorithm.

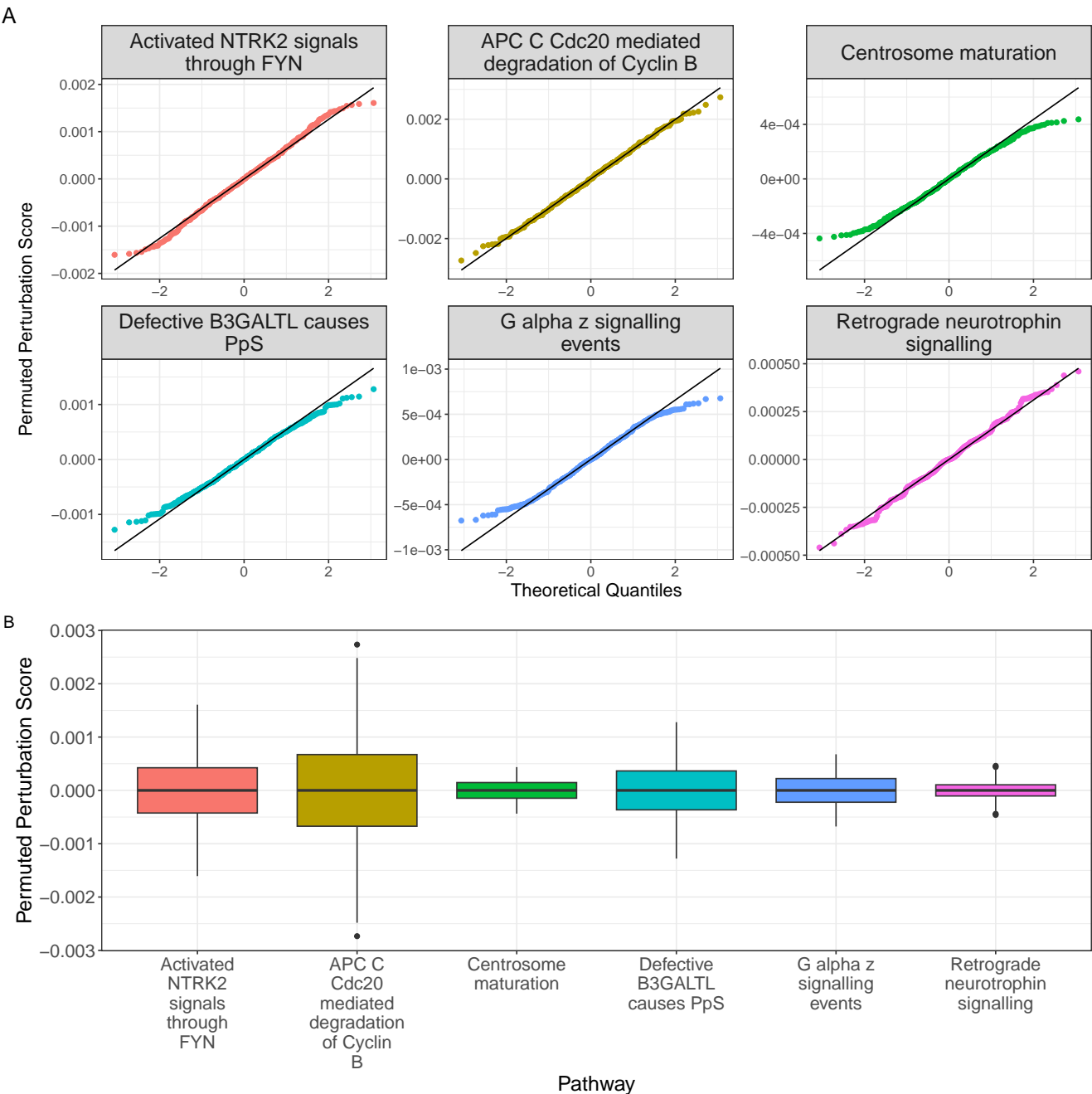


Figure 3. (A) Q-Q plot and (B) distributions of permuted perturbation scores of six randomly selected pathways. All sampled empirical distributions are approximately normally distributed with a mean of zero.

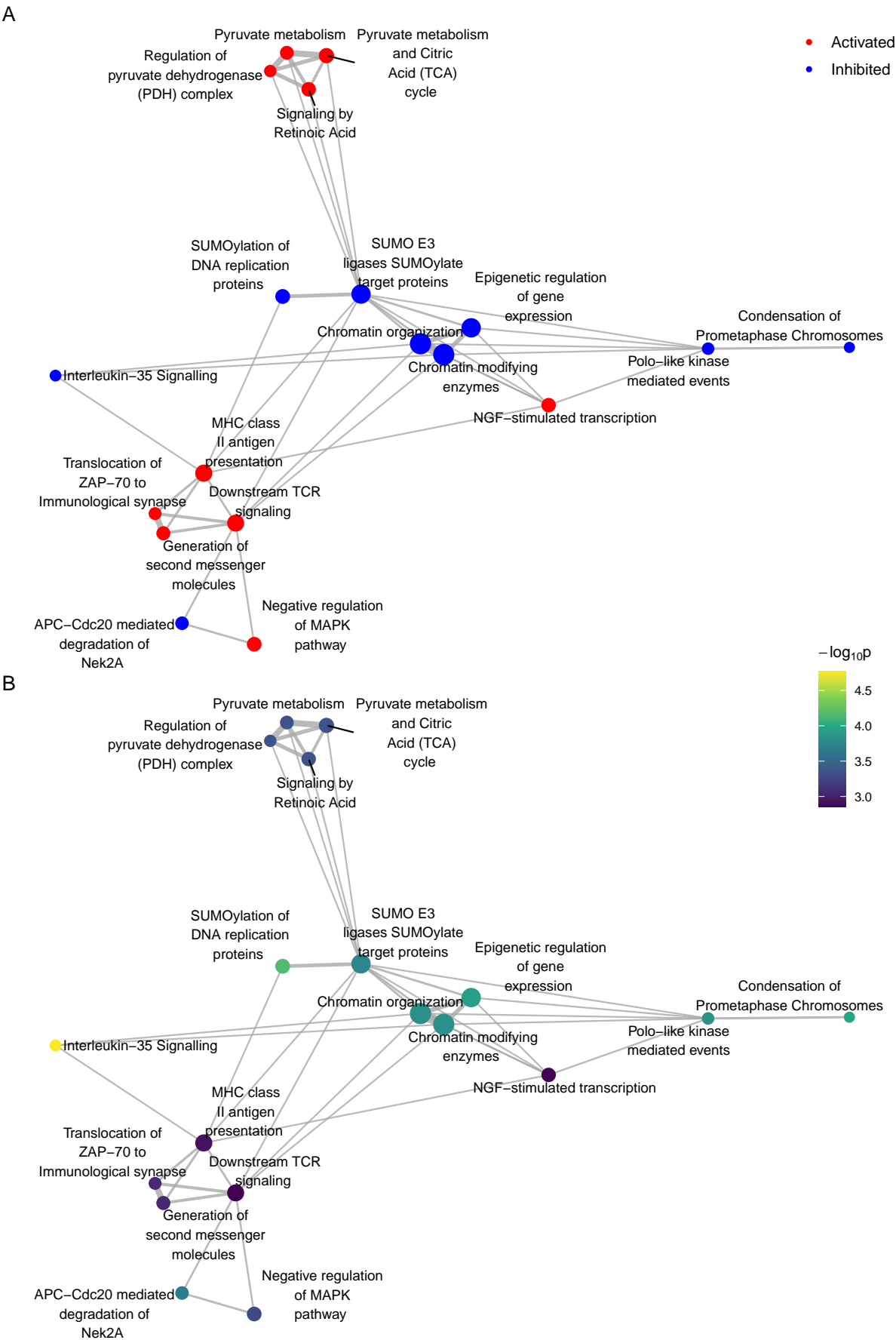


Figure 4. Significantly perturbed Reactome pathways identified among post-chemotherapy samples using sSNAPPY, colored by (A) predicted direction of change and (B) $-\log_{10}(p)$. Only the 10 most significantly inhibited and 10 most significantly activated pathways are shown.

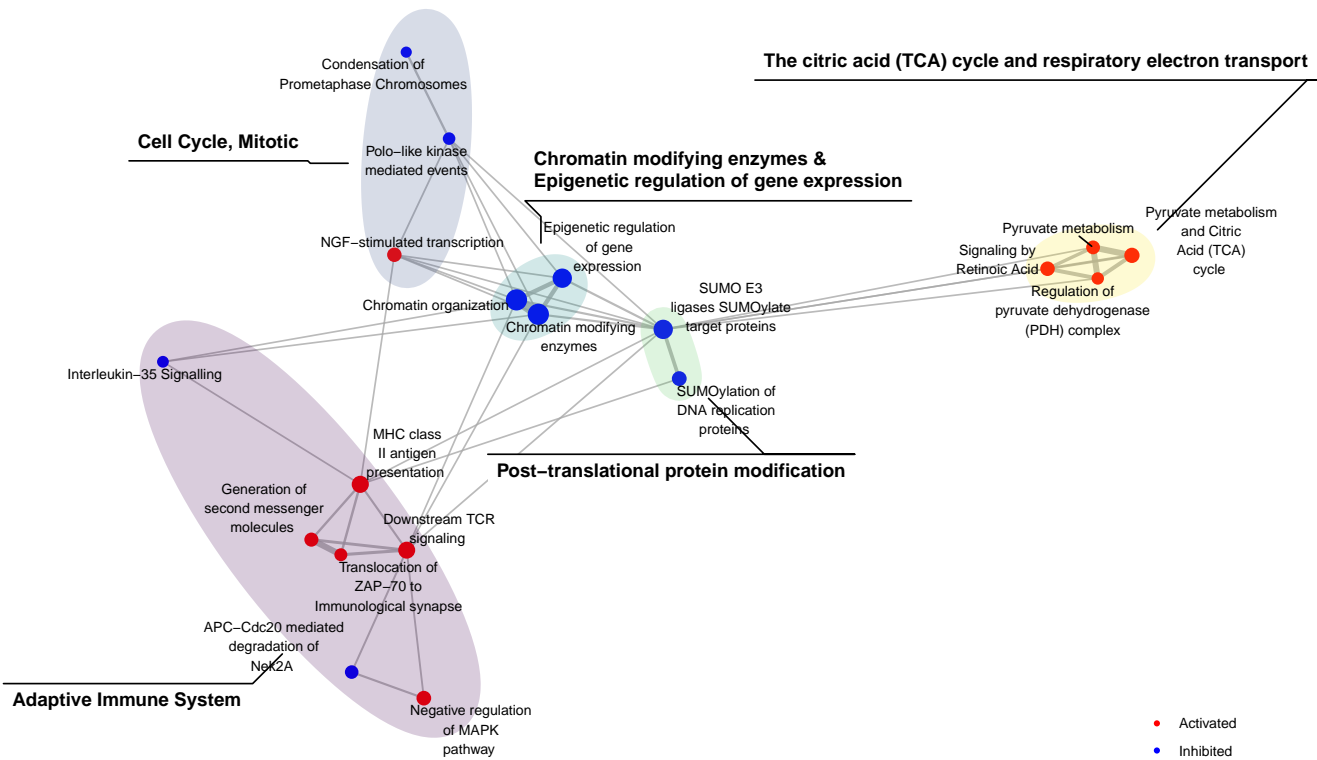


Figure 5. Significantly perturbed Reactome pathways identified among post-chemotherapy samples using sSNAPPY, colored by community structures detected through the louvain algorithm. The main biological processes associated with the top 20 pathways that wer most perturbed by the chemotherapy were shown.

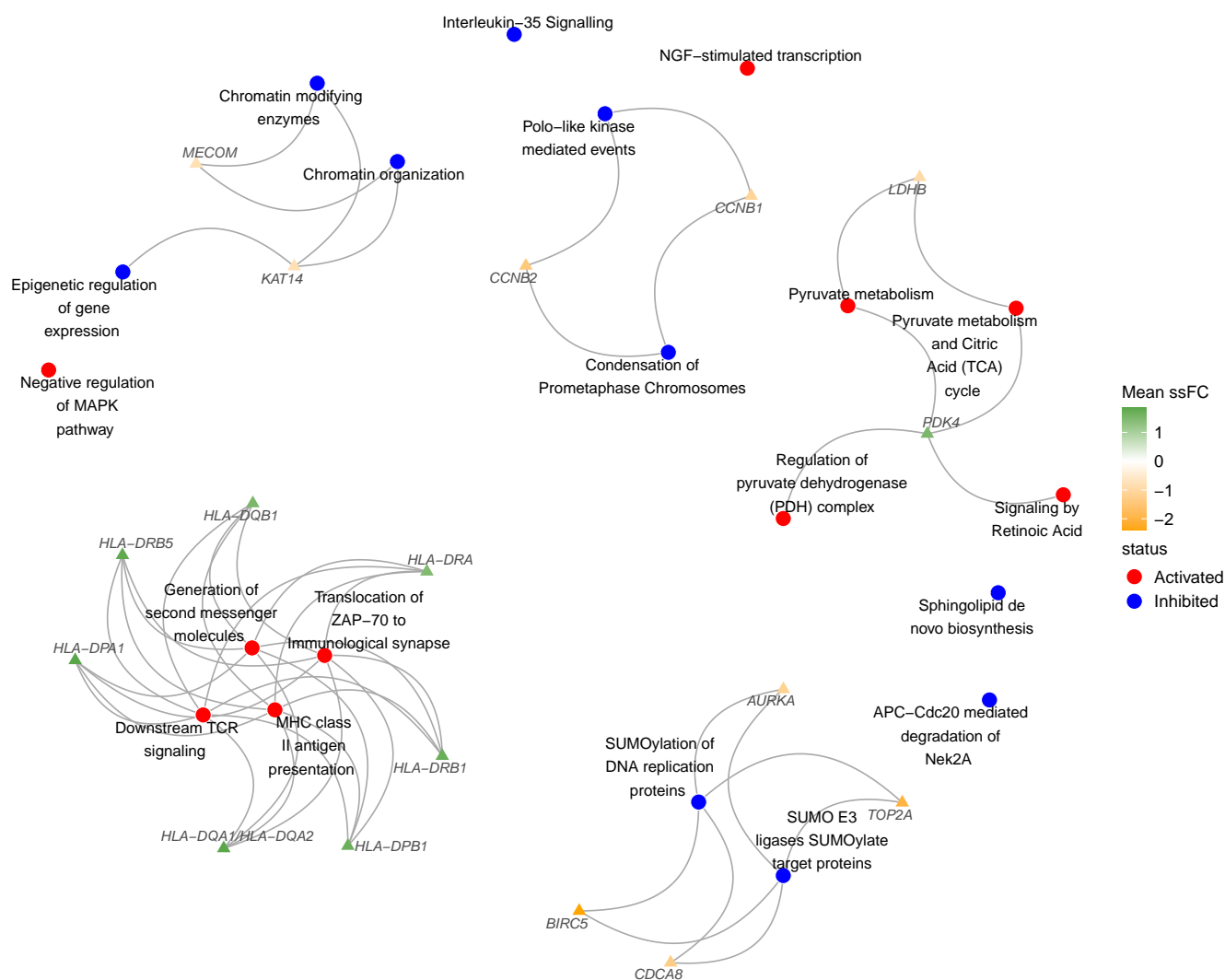


Figure 6. Significantly perturbed Reactome pathways identified among post-chemotherapy samples using sSNAPPY, showing any genes in the top 500 ranked by magnitude of change in expression, and which pathways they are likely contributing to. Only the 10 most significantly inhibited and 10 most significantly activated pathways are shown.

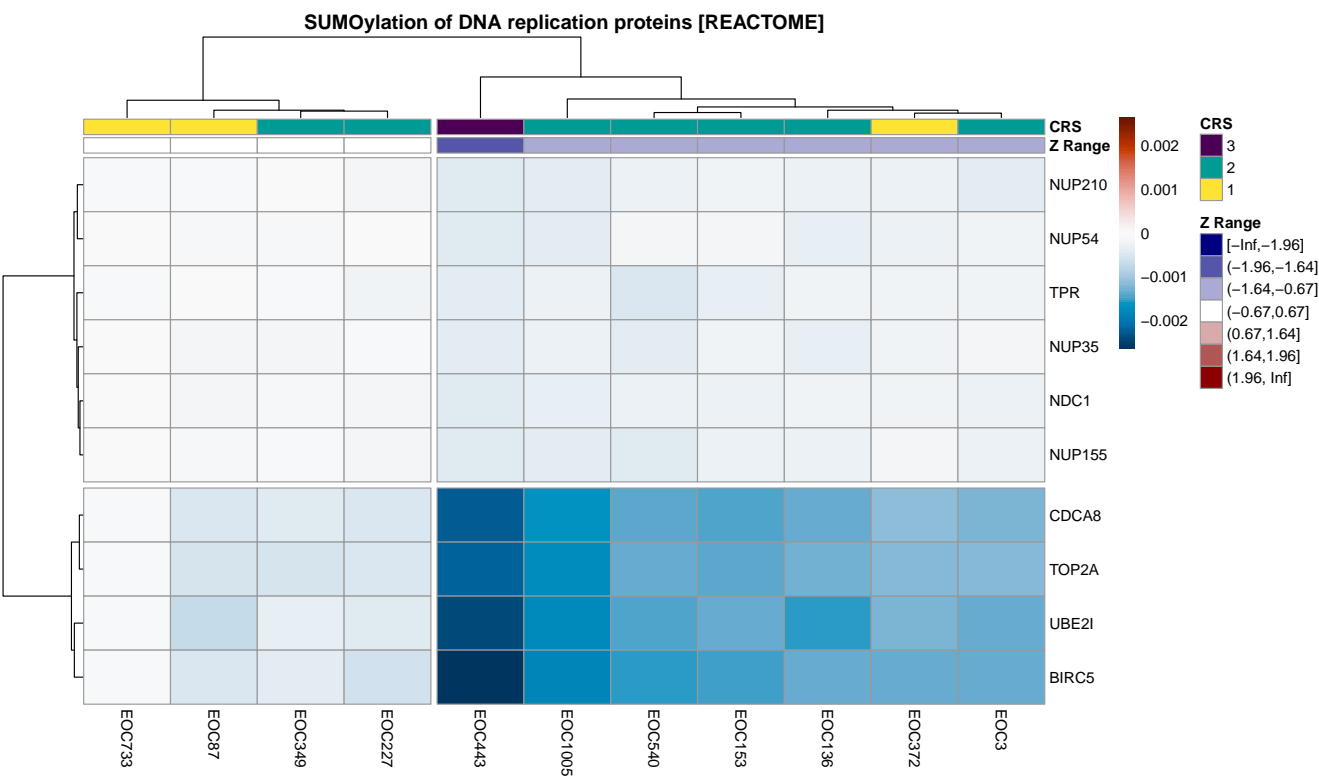


Figure 7. Gene-level perturbation scores for the top 10 genes in the "SUMOylation of DNA replication proteins" pathway ranked by average contribution to the perturbation score. Samples were annotated by patient chemotherapy response score (CRS), along with the range for sample-level Z-scores as a guide to sample-specific pathway perturbation. The genes CDCA8, TOP2A, UBE2I, BIRC5 were identified as possible key drivers of the inhibition of of this pathway.

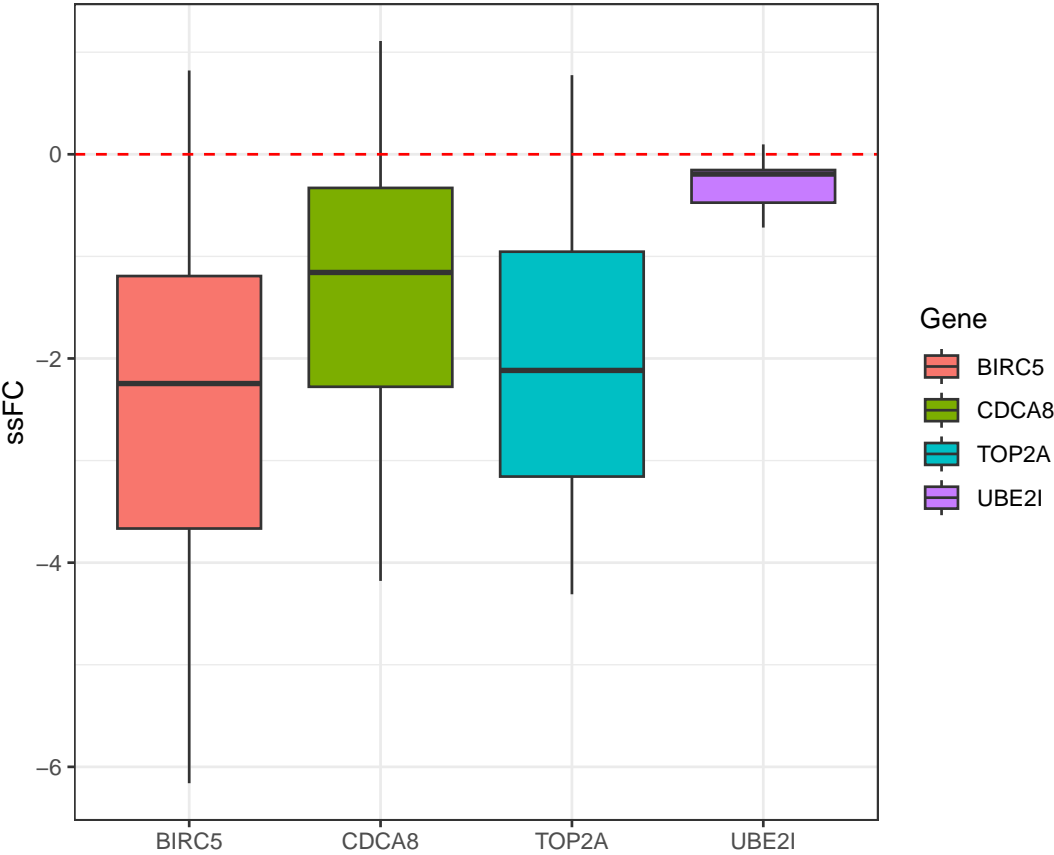


Figure 8. Single-sample logFC (ssFC) of potential key genes driving the inhibition of "SUMOylation of DNA replication proteins" pathway as a response to chemotherapy in HGSOC tumours.