

Application of High-Throughput Transcriptomics for Mechanism-Based Biological Read-Across of Short-Chain Carboxylic Acid Analogues of Valproic Acid

Supplementary Data

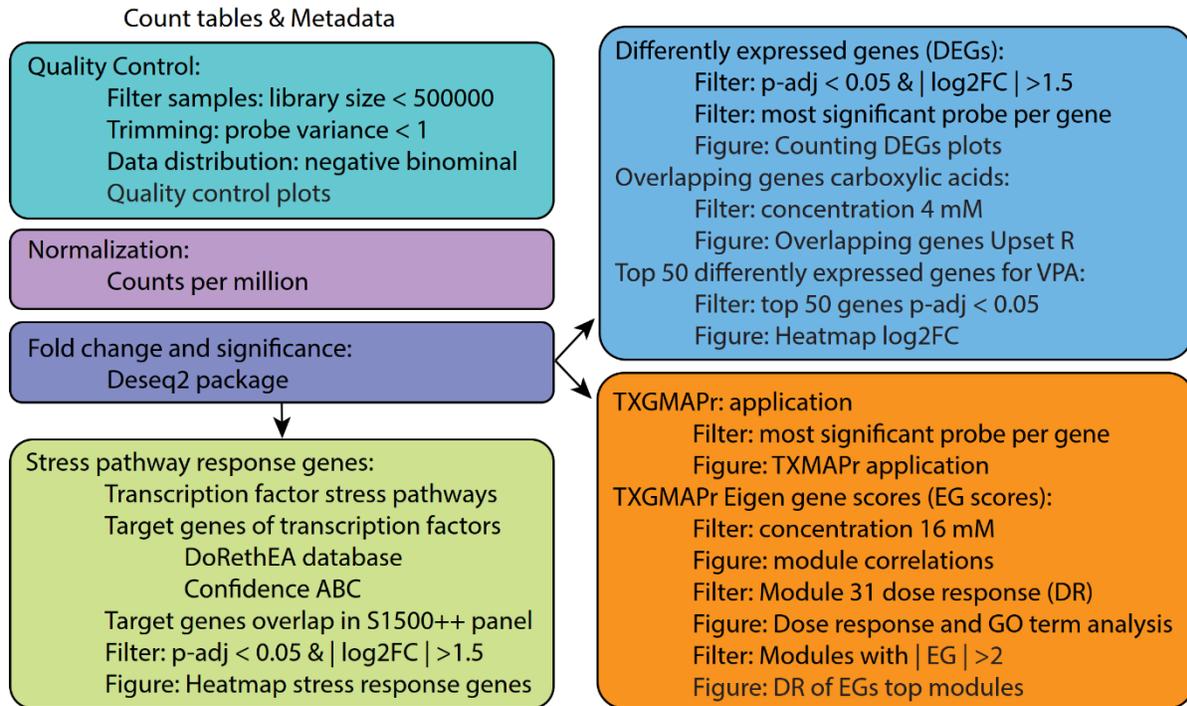


Fig. S1: Analysis flow of transcriptomics analysis

TempO-Seq count tables were prepared by BioClavis. After quality control checks and normalization, DEGs, fold changes and p-values were calculated. Fold change and significance data was used in the TXG-MAPr.

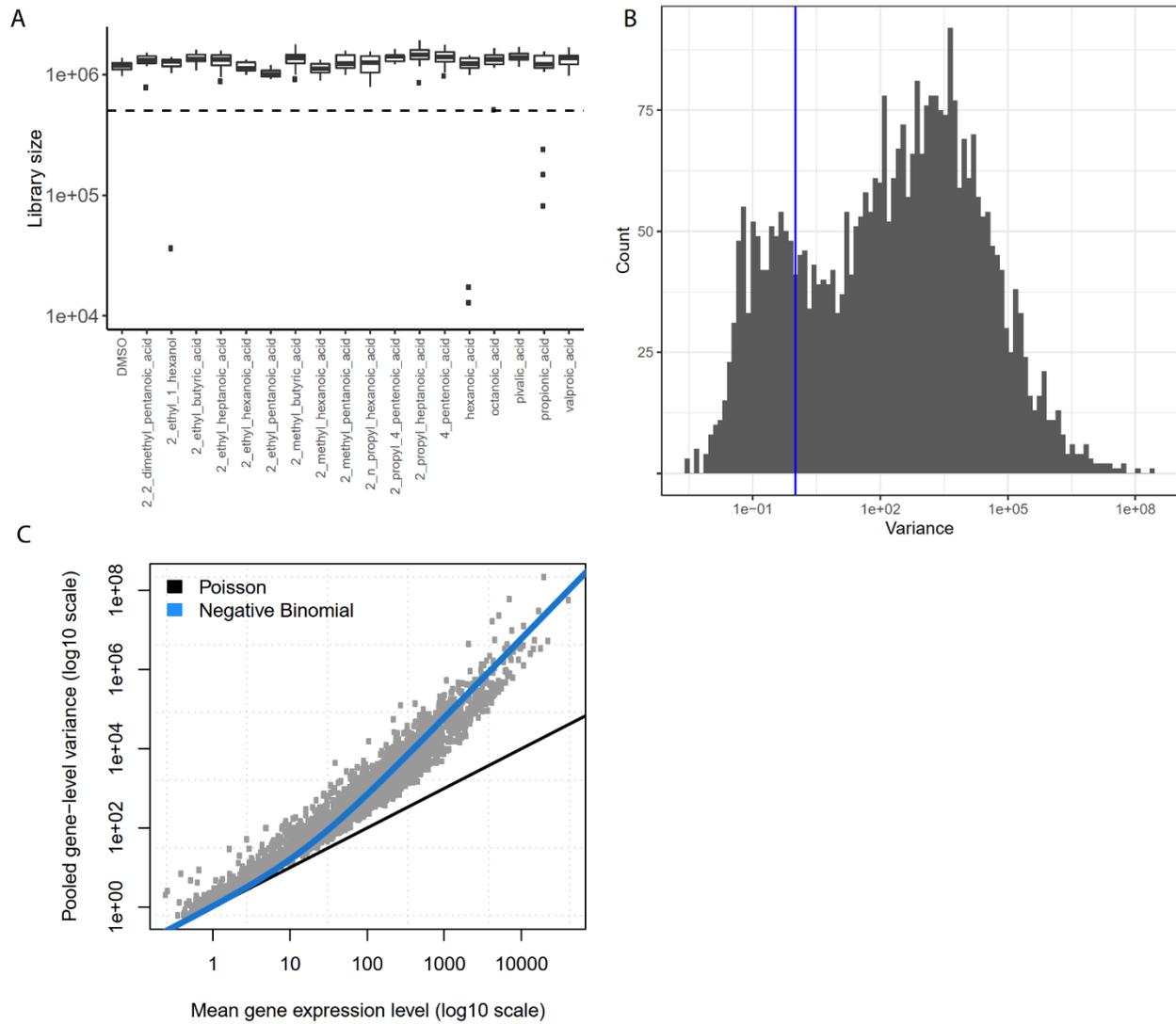


Fig. S2: Quality control plots for primary human hepatocyte samples

A) Library size for all samples. Six samples below the threshold of 500,000 counts were excluded from analysis. All replicates from the highest concentration of propionic acid were below the threshold, so this condition was not taken along. B) Histogram of gene counts. Probes with a variance below 1 were excluded from analysis, ~2700 genes were included (right side of blue cutoff line). C) Transcriptomic data follows a negative binomial distribution. Therefore, data was analyzed with the Deseq2 package in R.

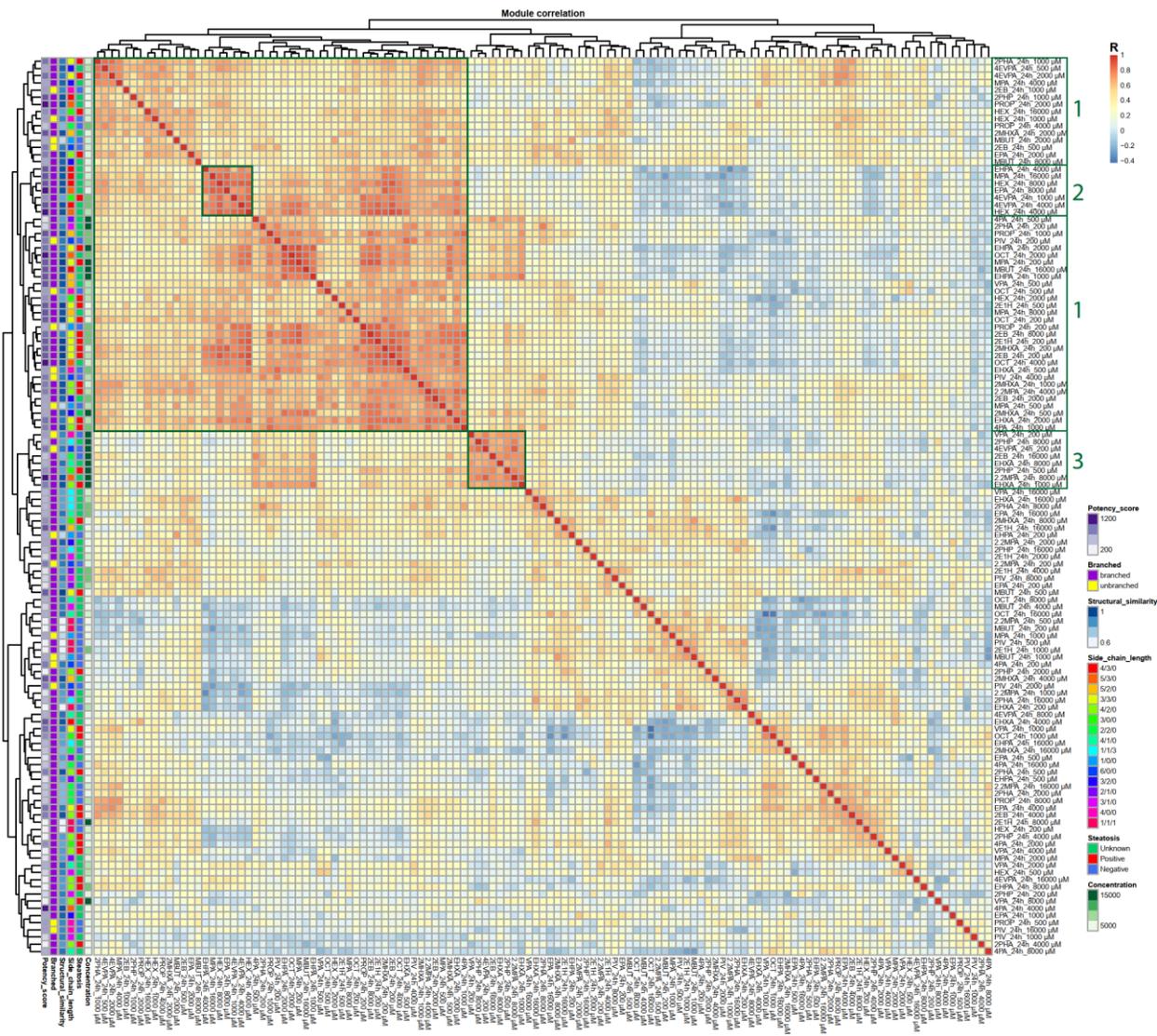


Fig. S3: TXG-MAPr correlation of all carboxylic acids for all concentrations
 Heatmap of correlation score PearsonR for all modules in TXG-MAPr. Row and column names are indicated by abbreviation of compound_exposure time_concentration. Clusters of highly correlating compound concentrations are enclosed in green boxes, in which box 1 covers all potent carboxylic acids, box 2 is a cluster of 2-propylhexanoic acid, and box 3 is a cluster of high concentration potent and less potent compounds.

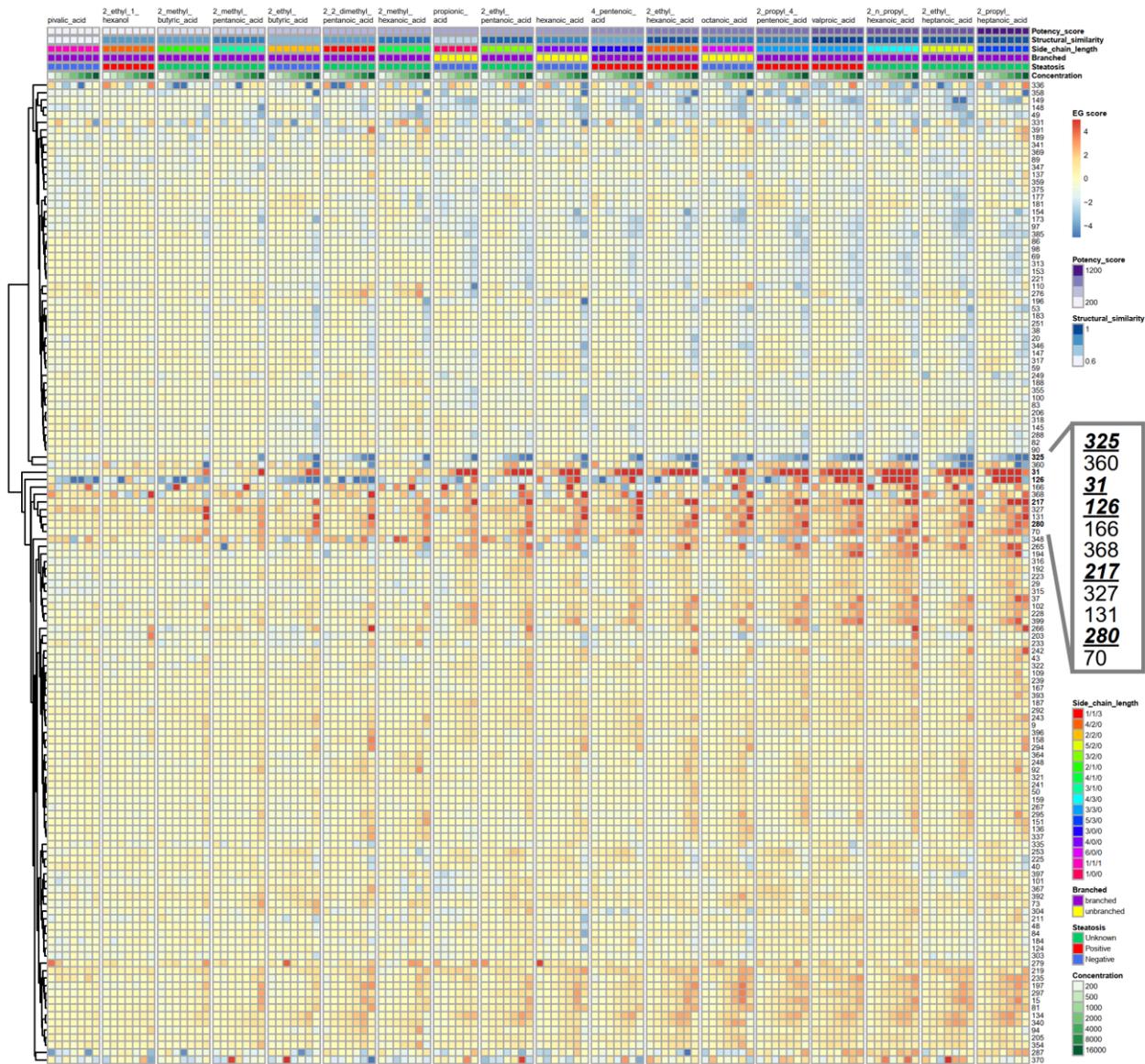


Fig. S4: Dose response of the EG scores of top modules

Heatmap of the EG scores of the modules (with $-2 > EG > 2$) from TXG-MAPr for all carboxylic acids in a dose response manner. The bold and underlined modules on the y-axis are the hit modules annotated in Figure 2A. These modules cluster together and mostly have a higher EG score for VPA and other potent carboxylic acids in a dose dependent manner.

