Supplemental Material for

Design principles of concentration dependent transcriptome deviations in drug-exposed differentiating stem cells

By: Tanja Waldmann, Eugen Rempel, Nina V. Balmer, André König, Raivo Kolde, John Antonydas Gaspar, Margit Henry, Jürgen Hescheler, Agapios Sachinidis, Jörg Rahnenführer, Jan G. Hengstler, Marcel Leist

This material is available free of charge via the Internet at http://pubs.acs.org.

Table of contents

Figure S1

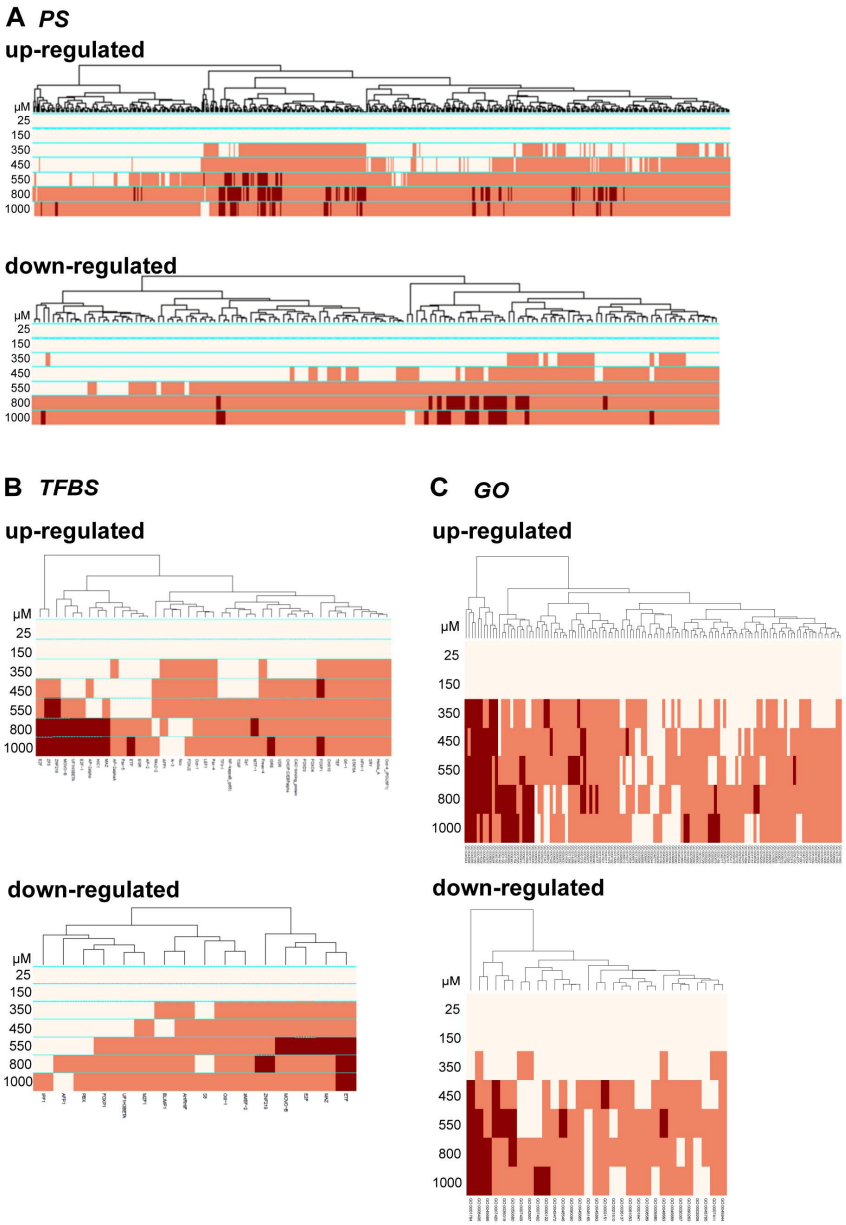

Figure S1: Heatmaps and clustering of concentration dependent altered gene expression and expression motives

(A) probesets, (B) overrepresented transcription factor binding sites, and (C) overrepresented GO groups. Similarity scores, indicated by the dendrogram at the top, were calculated based upon p-values obtained by limma t-test in (A) or by enrichment analysis in (B) and (C). Inclusion criterium for the analysis was a p-value < 0.001, taken over all concentrations. The *colours* of the heatmap describe the p-value of the according entity at the given concentration: *dark brown* means less than or equal to 10^{-4}, *light brown* less than or equal to 0.05.

Figure S2

[VPA]		25 µM	150 µM	350 µM	450 µM	550 µM	650 µM	800 µM	1000 µM
raw	UP	0	62	482	775	1038	1170	1785	2201
	DOWN	0	3	78	208	421	297	1091	1682
Concentration dependent	UP	0	0	273	554	955	656	1724	2070
	DOWN	0	0	33	110	344	118	910	1273
Overall adjustment	UP	0	34	379	596	919	699	1585	1785
	DOWN	0	1	49	124	320	132	749	994

Figure S2: Comparison of the number of differentially expressed genes using different statistical approaches.

Numbers of differentially expressed PS identified per concentration, are indicated for three types of false discovery rate (FDR) adjustment. Analysis was performed separately for up-regulation and down-regulation. UP refers to PS with fold change of at least 2, DOWN to PS with FC of at most 0.5. “Raw”: PS with p-value $p \leq 0.05$ without FDR adjustment. “Concentration dependent”: PS with p-value $p \leq 0.05$ after adjustment separately for each concentration according to FDR adjustment using the Benjamini Yekutieli method (BY-FDR). “Overall adjustment”: Same as above, but all p-values for all concentrations were adjusted according to BY-FDR simultaneously across all concentrations.

Figure S3

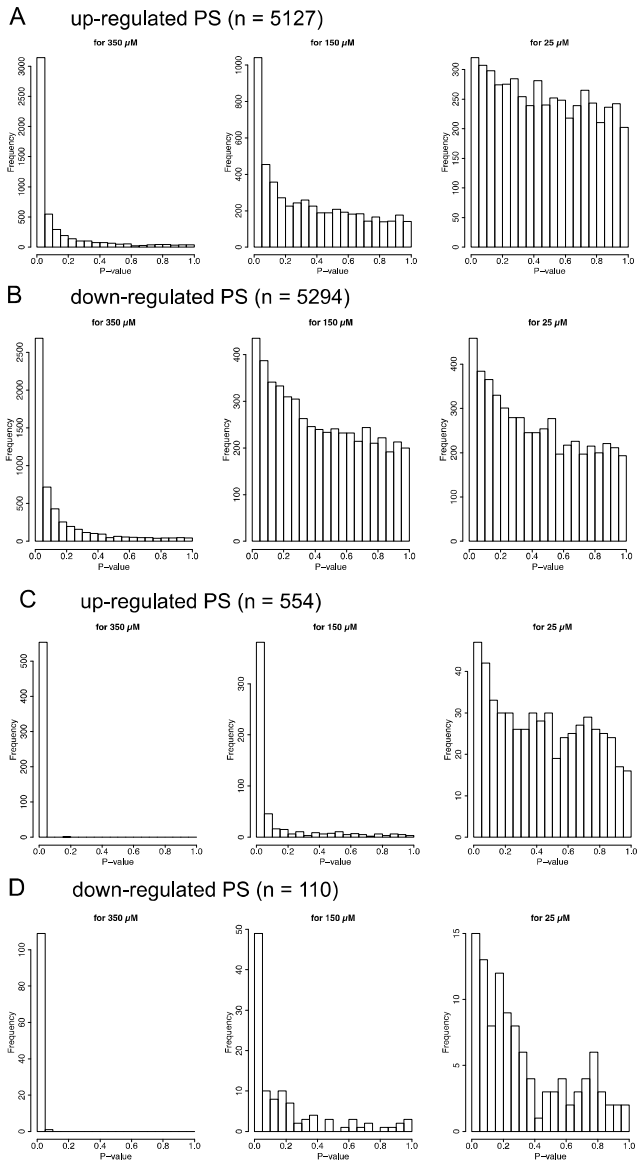

Figure S3 Distribution of p-values of a defined set of probesets (PS) at lower drug concentrations

(A,B) The PS (up: 5127; down: 5294) identified to be regulated by 450 µM VPA ($p < 0.05$ BY-FDR adjusted; no fold change (FC) cutoff) were examined for their regulation at lower drug concentrations. The p-values at 350 µM, 150 µM, and 25 µM were calculated for each PS, and they were binned at steps of 0.05, and displayed as histograms. The y-axis indicates the number of PS in each bin. Note the different scaling of the axes in the panels for each concentration. (C,D) the procedure was similar as in (A,B), but the population of PS examined was the same as in Fig. 2B, i.e. the PS regulated at 450 µM with $p < 0.05$ and $FC > 2$.

Figure S4

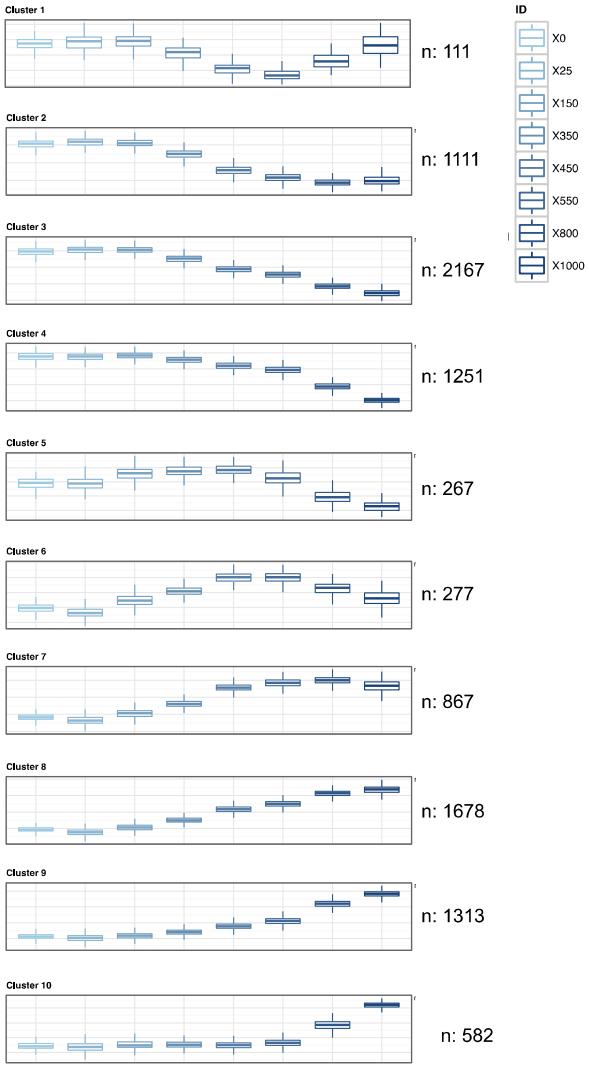

Figure S4: Box plot analysis of the concentration-dependent expression courses corresponding to k-means clusters of Fig. 5

For each concentration and each PS, averages were calculated across replicates. Then PS with a standard deviation (SD) across all concentrations of < 0.25 were excluded (non-regulated PS), and the remaining probes were scaled (to mean 0 and variance 1). Finally, k-means clustering was applied to expression vectors of the PS across concentrations. The bottom and top of each box represent the first and third quartiles, the band inside the box corresponds to the median, and the ends of the whiskers are the limits of the 1.5-fold interquartile range (IQR) added to the lower or upper quartile, respectively.