

Table 2: Selected TF motifs extracted) by ISMARA and sorted by z-value in the inlet vs middle area analysis.

“TF motif” represents a set of transcription factors with similar target nucleotide sequence pattern identified by CAGE. “PCC” is the gene with strongest Pearson correlation coefficient for the TF motif identified. “Pattern” describes the nucleotide sequence pattern of the motif for the PCC selected genes.

TF motif	Z-value of motif activity	PCC	Pattern	Potential liver functions¹
PATZ1_KLF4 Up in the middle area	1,30	PATZ1	GGGGGAG GGGCCG	Negative regulation of neoglucogenesis Wnt, Notch, p53 signaling
WT1_MTF1_ZBTB7B Up at the inlet	1,28	MTF1	GCGGGGG CGGGGG	Bile salt transport, Wnt/ β catenin, EMT, AP1
IKZF1 Up in the middle area	1,07	IKZF1	TTGGGAG A	Lipid and cholesterol metabolism
ESRA_ESR2 Up in the middle area	1,03	ESRA	CAAAGGT CAG	Estrogen regulation
SIN3A_CHD1 Up at the inlet	1,02	SIN3A	GGAGGCG GAGGTGG GAGAG	Retinoic acid, cellular defence, insulin reactome, β catenin
GATA3 Up in the middle area	0,71	GATA3	AGATGG	Xenobiotic reactome
RCOR1_MTA3 Up in the middle area	0,67	RCOR1	GGAGGCG GGG	Glutathione, beta catenin, and lipid metabolism
SP4_PML Up at the inlet	0,66	PML	GGGGCCA GGGGGGG GGCGGGG CCG	Wnt/ β catenin, SMAD, VEGF, BMP signaling
E2F7_E2F1 Up at the inlet	0,56	E2F1	CTTTTCCC GCCA	Collagen I trimer Hepatocyte homeostasis
NFIX_NFIB Up in the middle area	0,56	NFIB	CTTGCA	Oxygen metabolism

¹ From Supp. File 4, data extracted with top target analysis in ISMARA