Supplementary figure legends

Figure S1: Preliminary immunophenotypic analysis prior to Cytobank export and patient cohort details

A. Table illustrating the sample volume and percentage purity obtained after the first passage (Run1) and second passage (Run2) through the magnet during magnetic bead enrichment of CD34+ HSPCs. All samples had a purity of >95% prior to culture.

B. Gating strategy on D0 prior to Cytobank import. Identification of singlets on forward scatter-area (FSC-A) versus (vs) FSC-height (H) plots was followed by the identification of cells on FSC-A vs side scatter-area (SSC-A), viable cells on SSC-A versus Viability and then the population of interest (POI) on FSC-A vs SSC-A. Zombie Violet (ZV) was used as the viability dye. Cells are coloured using banded colouring and plots are in a linear scale for all plots including C and D below.

C. Gating strategy on D7 following the same method as mentioned in B above.

D. Table displaying the characteristics of the HEU cohort. All mothers had a viral load (VL) <100 copies/mL within 3 months of delivery (D) (LTDL-Lower than detectable limit) and all except one were taking TLD as antiretrovirals (ARVs). One mother was on TEE (TDF, Emtricitabine and Efavirenz); Polymerase chain reaction (PCR) results for all infants were negative as was the HIV Gene Xpert on the umbilical cord blood (UCB).

Figure S2: Concatenated viSNE plots showing differences between HEU and HUU CD34+Lin- subsets

A. Concatenated viSNE plot of D0 HUU samples coloured according to marker expression on the z-axis and scale of expression on the right of each channel. Only the expression of the Lin-CD34+ population is shown for the following markers: CD38 ECD, CD133 PE-Vio770, CD45RA APC, CD117 PE, CD90 BV510, and CD49f SB780 (N=4).

B. Concatenated viSNE plot of D0 HEU samples with plot characteristics as described in A above (N=5).

C. D7 viSNE concatenated samples for the HUU VC condition portraying the difference in immunophenotypic populations compared to the D0 samples. Cells cluster completely differently than for D0, and HSPC populations are predominantly CD38- on D7, compared to CD38+ on D0.

D. Equivalent viSNE concatenated samples for the HEU VC portraying the differences compared to HUU. The HSPCs cluster differently. A large bright CD90+ population present in HUU samples is absent in HEU samples; however, HEU groups have widespread dim expression of CD90 in the majority of cells. In contrast, HUU groups appear to have more dim expression of CD49f than HEU.

E. FlowSOM clustering maps coloured according to the metaclusters (MCs), illustrated on the left of the image, for HUU and HEU samples (D0, D7 VC, D7 SR1). Only one sample is shown for ease of visualisation. Overt differences are the prominence of MC3 and 13 in HUU and MC 9 and 12 in HEU samples.
F. Concatenated image of viSNE plot of D0 HUU samples showing only two channels (CD38-ECD and CD133 PE-Vio-770). The manual gates drawn around the cell clusters are shown with the names which were assigned to each population. (N=5) See also Table 1 for the phenotypic description of each population. Gates were guided by the FlowSOM MC in Figure S2E.
G. Complimentary viSNE plot of D7 VC also only showing the two channels mentioned above in F. The plot clearly shows the marked difference in clustering between D0 and D7 samples. (N=5)

Fig. S3 Statistically significant metacluster differences in Lin-CD34+ sub-population between HEU and HUU populations.

A. Box plot of the 17 FlowSOM MCs (x-axis) and frequency of each MC in HUU (orange) and HEU (blue) on D0. MC13, 15, and 17 were statistically different, with MC13 being higher in HUU and MC15 and 17 being higher in HEU. N=4 (HEU) and N=5 (HUU) for all experiments. *(p=<0,05)
B. Heatmap of the statistically significant MCs identified in A, portraying the phenotype of each MC. Scale of the plot is provided at the bottom of the image. MC13, is in keeping with ExtCMP.
C. Box plot of the D7 VC for HEU and HUU populations illustrates the statistically significant MC3, 11 and 15. *(p=<0,05)
D. Corresponding heatmaps for C showing the phenotype of the MCs. MC3 corresponds to ExtLMPP2.
E. Box plot of D7 SR1 for HEU and HUU populations. The MCs of significance are 2, 3, 4, 6, 11, and 12. *(p=<0,05)
F. The complementary heatmap for E shows the following correlation with phenotype. MC3 and 4 correspond to ExtLMPP2, MC2 to HSPC1 and 2 and MC12 to ExtLMPP1 and 3.

Figure S4: CITRUS cluster analysis showing differences between D0 and D7 HUU and HEU CD34+Lin- subsets

A. Citrus analysis for D0 HUU vs HEU samples. Each CITRUS algorithm was repeated five times, using Significance Analysis of Microarrays (SAM) and abundance for each conditon to obtain the common significant clusters. Equal events were used for each sample. Figure shows two of the five runs. Each circle depicts one cluster, and the red clusters are those found to be significantly different. The red clusters which give rise to other clusters are seen as the dominant clusters and those are used for export and phenotypic description. Significant MCs are 17935 (Lin-CD34+ CD38+CD133dimCD45RA-CD117+CD90-CD49f-: correlates to ExtEMP2), 17936 (Lin-CD34+CD38+CD133+CD45RA-CD117+CD90-CD49f-: correlates to ExtEMP), 17953 (Lin-CD34+CD38dimCD133+CD45RA-CD117+CD90-CD49fdim-: correlates to ExtMPP), 17959 (Lin-CD34+CD38-CD133vCD45RA+CD117-CD90-CD49fv: correlates to ExtLMPP), and 17962 (Lin-CD34+CD38dimCD133+CD45RAvCD117+CD90-CD49f-: correlates to ExtCMP)

B. Citrus analysis showing two sample runs for HUU D7 VC vs HEU D7 VC. Clusters 62995 (Lin-CD34+CD38-CD133+CD45RA+CD117vCD90-CD49f+: correlates to ExtLMPP1), and 62997 (Lin-CD34+CD38-CD133+CD45RA+CD117vCD90-CD49fv: correlates to ExtLMPP3)

C. Similar analysis for HUU D7 SR1 vs HEU D7 SR1 samples is shown. Significant clusters are 269995 (Lin-CD34+CD38CD133v CD4RA+CD117dim/-CD90-CD49f+: correlates to ExtLMPP3) and 269996 (Lin-CD34+CD38-CD133+CD45RA+CD117vCD90dimCD49f-: correlates to ExtLMPP2).

Fig. S5 D0 CITRUS cluster immunophenotype analysis of Lin-CD34+ sub-population between HEU and HUU subgroups.

A. Histogram plots of significant clusters identified during CITRUS analysis of D0, HUU (N=7) vs HEU (N=5). The blue shaded area depicts all the cells that were included in the algorithm and the red shaded area shows the histogram image of the specific cluster, and thus the phenotype. The immunophenotypic channel is shown above the figure and the cluster identification is on the right of each histogram.

B. Box plot of the identified clusters of significance which illustrated the group within which the higher/lower event count (y-axis) was found (either HEU or HUU). On D0, 2 clusters (17953 and 17959, corresponding to ExtMPP and ExtLMPP respectively) were more prominent in HEU samples and 3 in HUU (17935, 17936, 17962, ExtEMP2, ExtEMP and ExtCMP, respectively).

Figure S6: D7 CITRUS cluster immunophenotype analysis of Lin-CD34+ sub-population between HEU and HUU subgroups.

A. Histograms of the two significant clusters identified in the D7 VC, HEU (N=4) vs HUU (N=5) samples from the CITRUS analysis. The CITRUS analysis was performed using the Significance Analysis of Microarrays (SAM) model in Abundance mode, with equal events across all subgroups. A false discovery rate (FDR) of 5% was applied to identify significant clusters, which were further visualised using cluster and box plot analyses. The blue shaded area depicts all cells included in the algorithm, while the red shaded area shows the histogram of the specific cluster, and thus the phenotype. The immunophenotypic channel is shown above the figure and the cluster identification is on the right side of each histogram.
B. Box plot of exported significant CITRUS clusters on D7 VC, depicting clusters 62995 and 62997, corresponding to ExtLMPP1 and 3, found in a higher proportions in HEU samples.
C. Histograms of the two significant clusters identified in the D7 SR1, HEU (N=4) vs HUU (N=5) CITRUS analysis.
D. Box plot of significant CITRUS clusters, one which was higher in HEU samples (269995, corresponding to ExtLMPP3) and one in HUU samples (269996, corresponding to ExtLMPP2).

Figure S7: Concatenated UMAP plots showing differences between D0 and D7 HEU and HUU CD34-Lin- subsets

A. UMAP image of HUU D0 Lin-CD34- population showing all the phenotypic markers above on the z-scale. Map is coloured by the scale of expression on the right. CD38-ECD, CD133 PE-Vio770, CD45RA APC, CD117 PE, CD90 BV510, CD49f SB 780 are shown (N=5). All events were included in the analysis.
B. Complimentary UMAP portraying the HEU D0 population is shown with all the characteristics and channels as described above. (N=4). Visually the differences between HUU and HEU samples on D0 are not easily perceptible.

C. Concatenated UMAP of HUU D7 VC with all features and channels as mentioned in A above. The D7 population analysis is not linked to the D0 analysis as the cells plated on D0 are a pure Lin-CD34+ population.
D. UMAP of the HEU D7 VC population depicting easily perceptible changes as compared to HUU D7 VC. CD90 expression is much more widespread in HEU samples.
E. The HUU D7 SR1 concatenated UMAP showing similar characteristics to the HUU D7 VC UMAP in B above.
F. Complimentary HEU D7 SR1 UMAP confirming increased expression of CD90 in HEU samples.
G. ViSNE plot showing the manual gates on an HUU samples guided by the FlowSOM MCs. Gate names are provided. Plot scale is on the right.
H. FlowSOM MC dot overlays of D0, D7 VC, and D7 SR1 of one HUU sample showing the MCs which guided manual gates.

Figure S8: Population frequencies and statistically significant differences in Lin-CD34- sub-populations

A. Table depicting the sub-population frequencies of the Lin-CD34- populations between HEU and HUU samples
B. Bar graph of statistically significant MC in D7 VC. Orange bars are HEU and blue bars HUU samples. MC3, 13, and 16 are higher in HEU samples, while MC11 and 12 are higher in HUU samples.
C. Corresponding bar graph of D7 SR1 conditions. MC13 and 16 are higher in HEU samples and MC6, 11, and 12 higher in HUU samples.

Figure S9: Differences in CFU phenotype between HEU and HUU samples.

A. FlowSOM Dot overlay map showing the 17 identified MCs on the HUU samples, coloured according to the MC map. All 4 HUU samples are shown.
B. Equivalent HEU FlowSOM map showing all 4 HEU samples. The MC confirm barely perceptible changes between sample groups
C. FlowSOM MC box plot shows no significant difference in immunophenotype between the HEU and HUU sub-groups. *(p=<0.05)

Figure S10: Volcano plot showing genes that were significantly up- (red dots, 108 genes) and downregulated (blue dots, 163 genes) between HEU and HUU samples based on their p-value (y-axis) and fold change (x-axis).

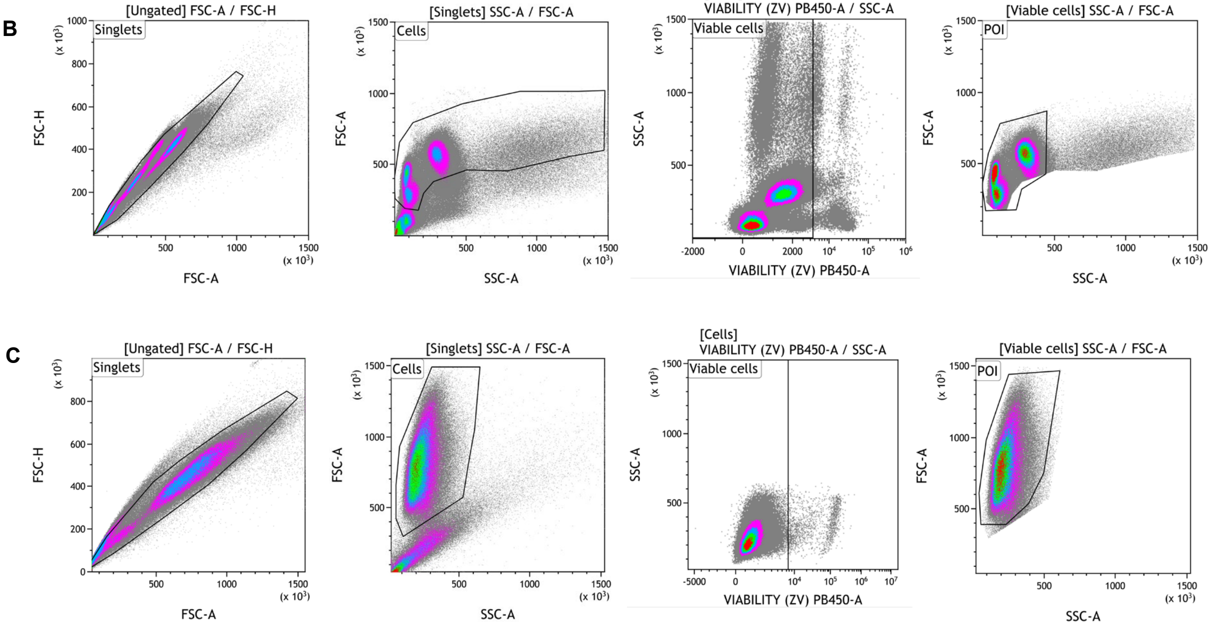
Figure S11: Full gene list displaying all 271 differentially expressed genes, fold change, p-value, gene symbol, and description.

Figure S12: Functional analysis of DEGs by Gene Ontology biological processes (GOBP) depicting 11 biological processes which are involved. Only upregulated but not downregulated genes were involved in these processes

Figure S1: Preliminary immunophenotypic analysis prior to Cytobank export and patient cohort details

A

	Sample volume (ml)	Purity after Run 1 (%)	Purity after Run 2 (%)
HUU Sample 1	98	79,37	97,04
HUU Sample 2	75	72,08	98,29
HUU Sample 3	57	77,3	95,87
HUU Sample 4	108	21,98	98,65
HUU Sample 5	74	80,38	99,38
HEU Sample 1	64	66,95	96,64
HEU Sample 2	72	71,46	98,64
HEU Sample 3	67	60,89	95,02
HEU Sample 4	63	57,78	97,54



D

Sample	Maternal VL timepoint relative to delivery	VL result (copies/ml)	ARVs	HIV PCR in infant (timepoint relative to delivery)	HIV PCR result	Gene Xpert result on UCB
1	D+1 day	39	TLD	D	Negative	Negative
2	D-6 weeks	LTDL	TLD	D+72 hours	Negative	Negative
3	D	LTDL	TLD	D+96 hours	Negative	Negative
4	D-3 months	LTDL	TEE	D+1 day	Negative	Negative
5	D+1 day	79	TLD	D+1 day	Negative	Negative

VL, Viral load; ARV-Antiretrovirals; D-Day of delivery; LTDL-Lower than detectable limit; PCR Polymerase chain reaction; TLD- TDF, Lamivudine, Dolutegravir; TEE- ;Tenofovir disoproxil fumarate(TDF), Emtricitabine and Efavirenz

Figure S2: Concatenated viSNE plots showing differences between HEU and HUU CD34+Lin- subsets

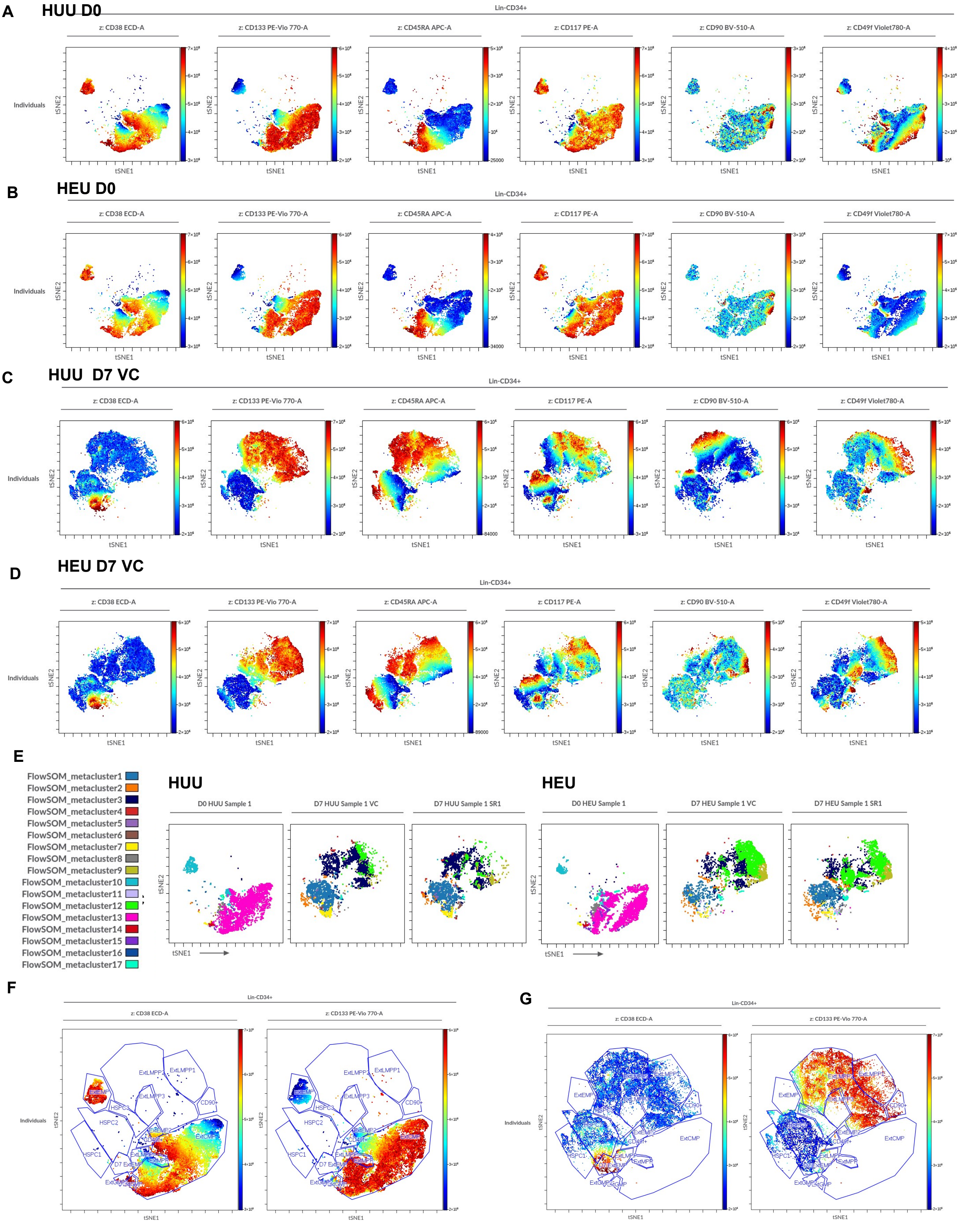


Fig. S3 Statistically significant metacluster differences in Lin-CD34+ sub-population between HEU and HUU populations.

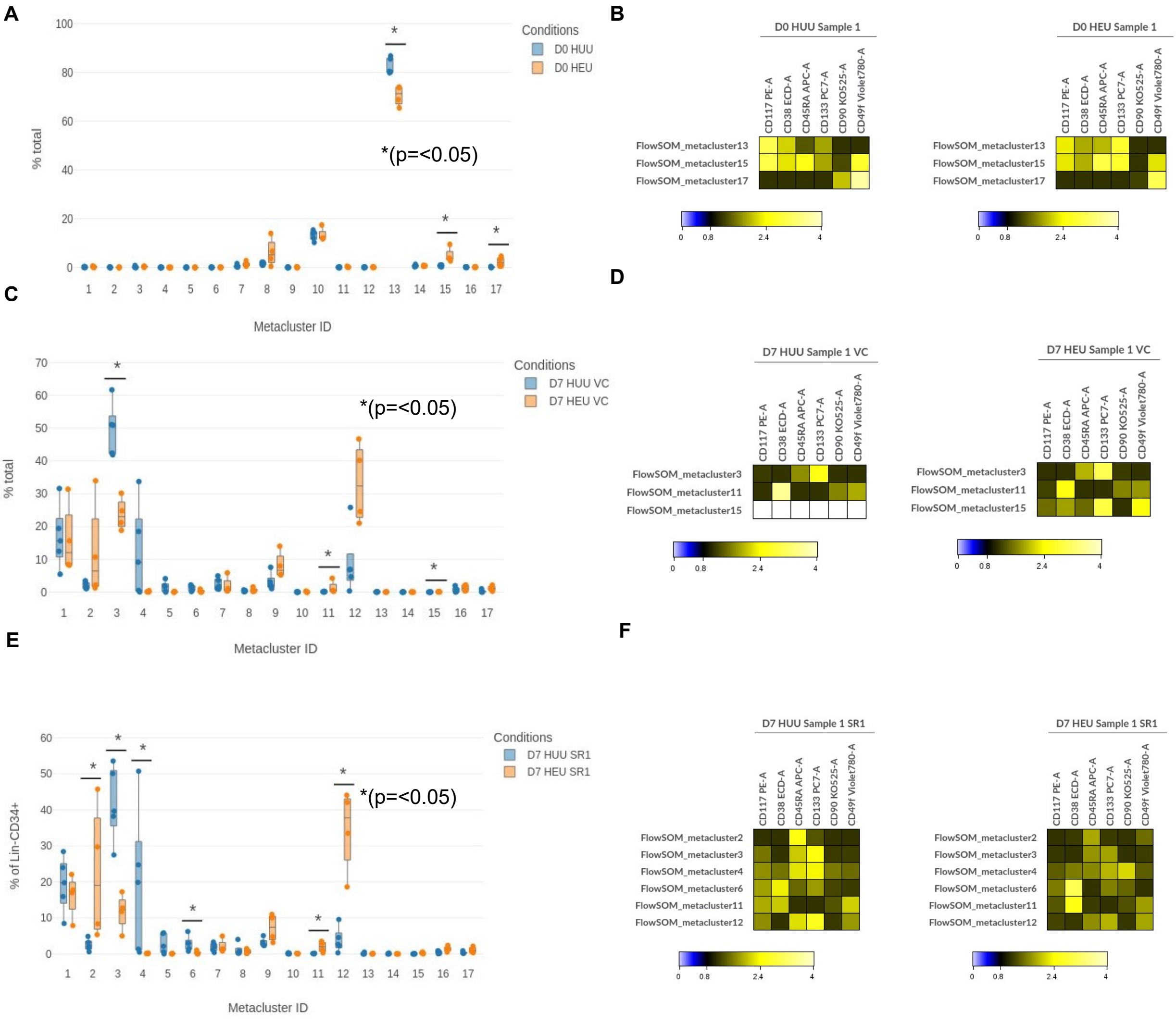
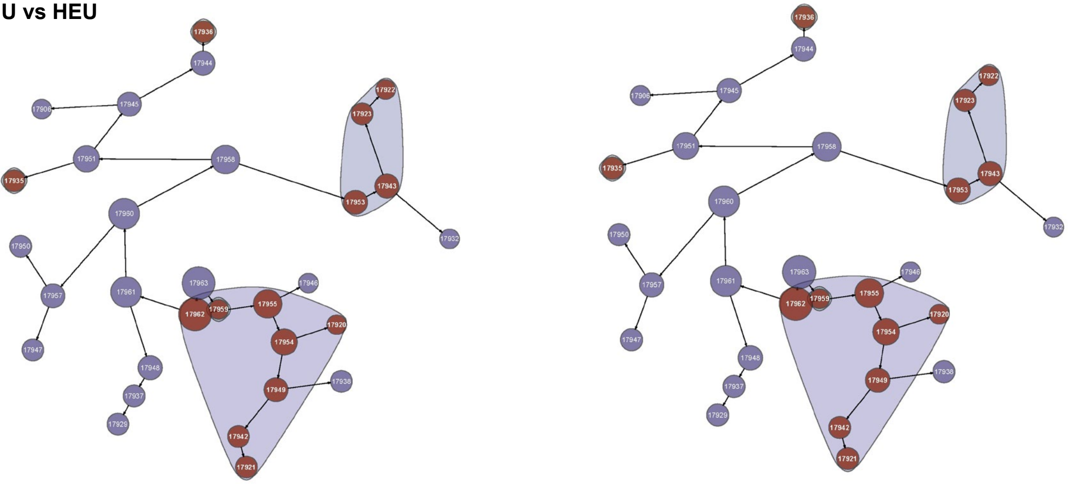
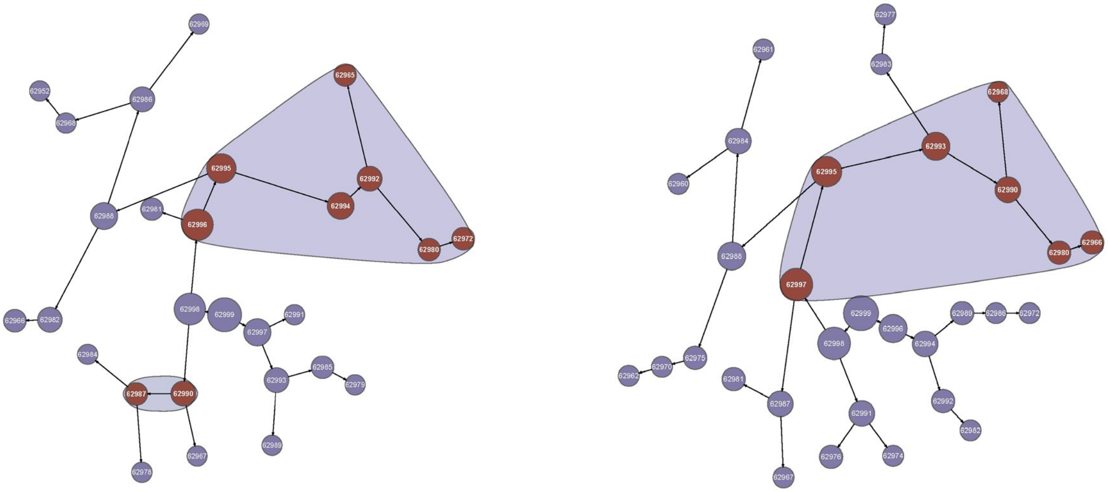


Figure S4: CITRUS cluster analysis showing differences between D0 and D7 HUU and HEU CD34+Lin- subsets

A D0 HUU vs HEU



B D7 VC HUU vs HEU



C D7 SR1 HUU vs HEU

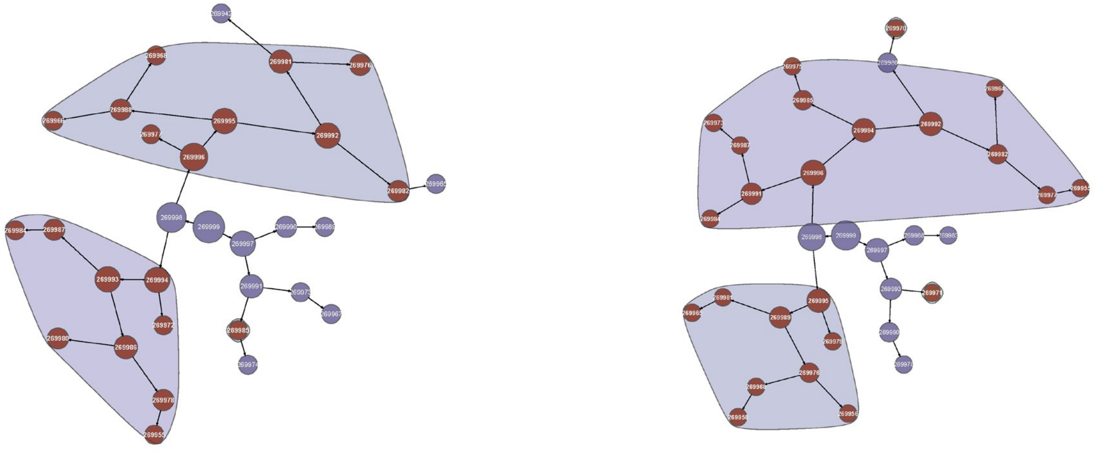


Fig. S5 D0 CITRUS cluster immunophenotype analysis of Lin-CD34+ sub-population between HEU and HUU subgroups.

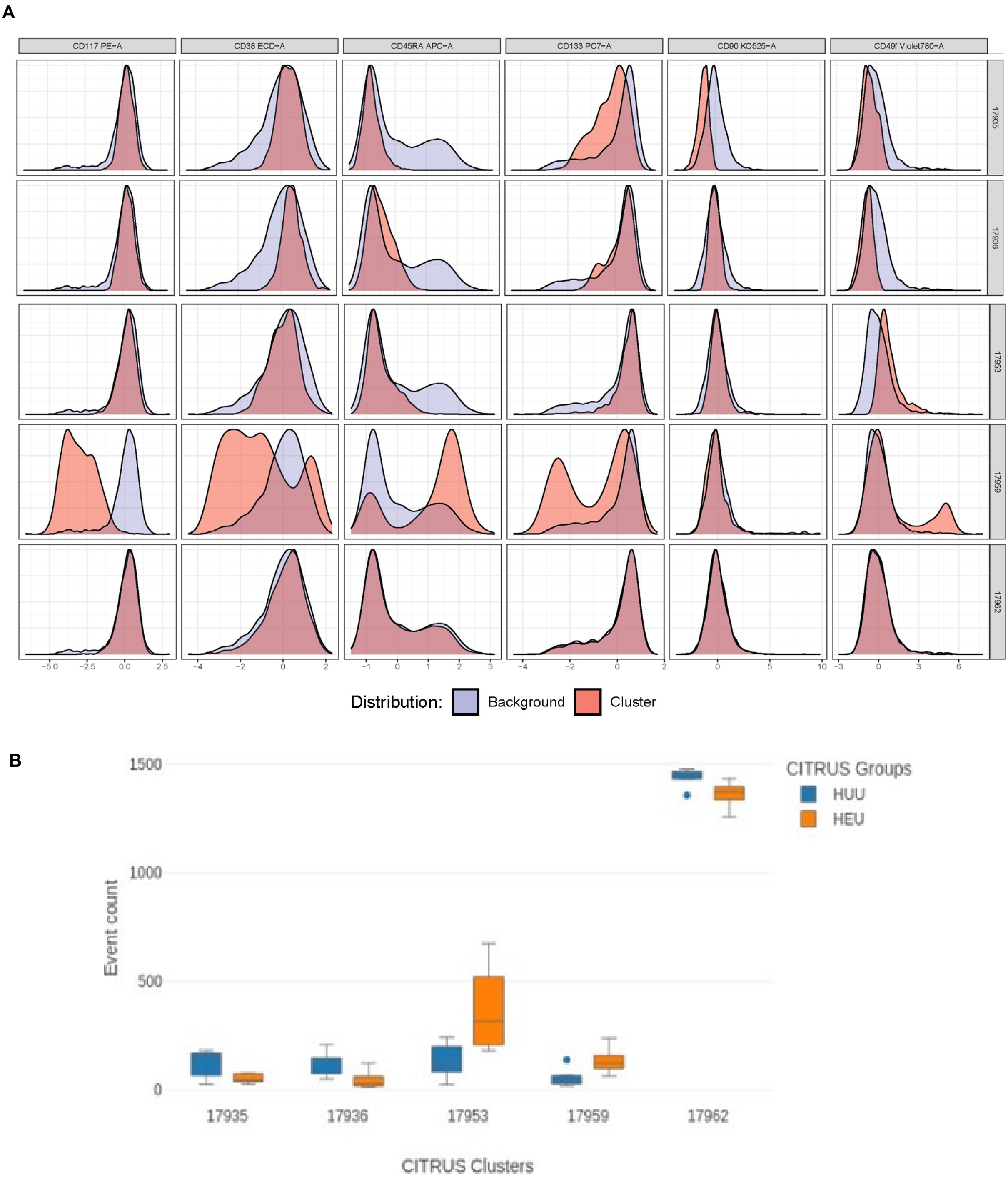


Fig. S6 D7 CITRUS cluster immunophenotype analysis of Lin-CD34+ sub-population between HEU and HUU subgroups.

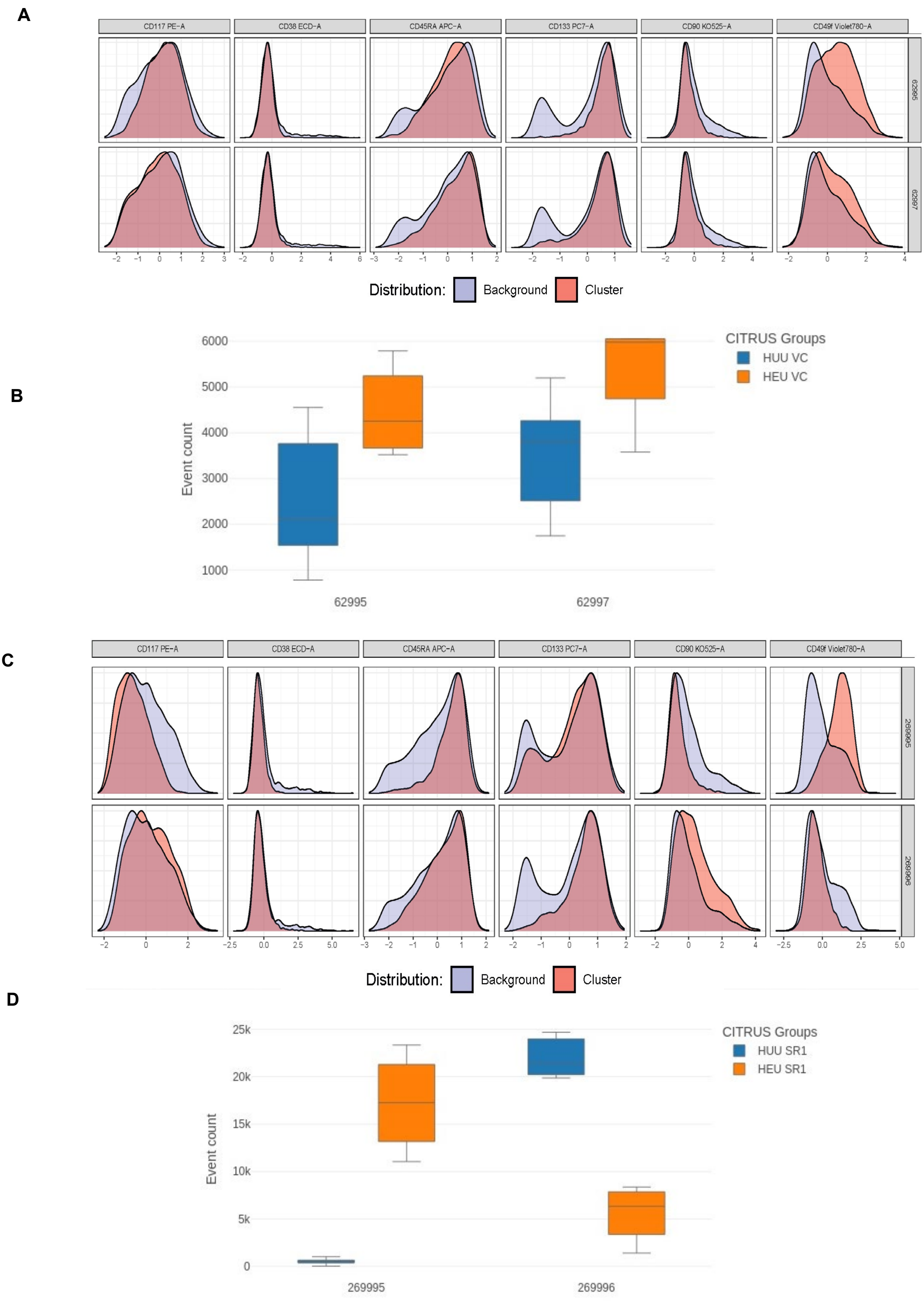


Figure S7: Concatenated UMAP plots showing differences between D0 and D7 HEU and HUU CD34-Lin- subsets

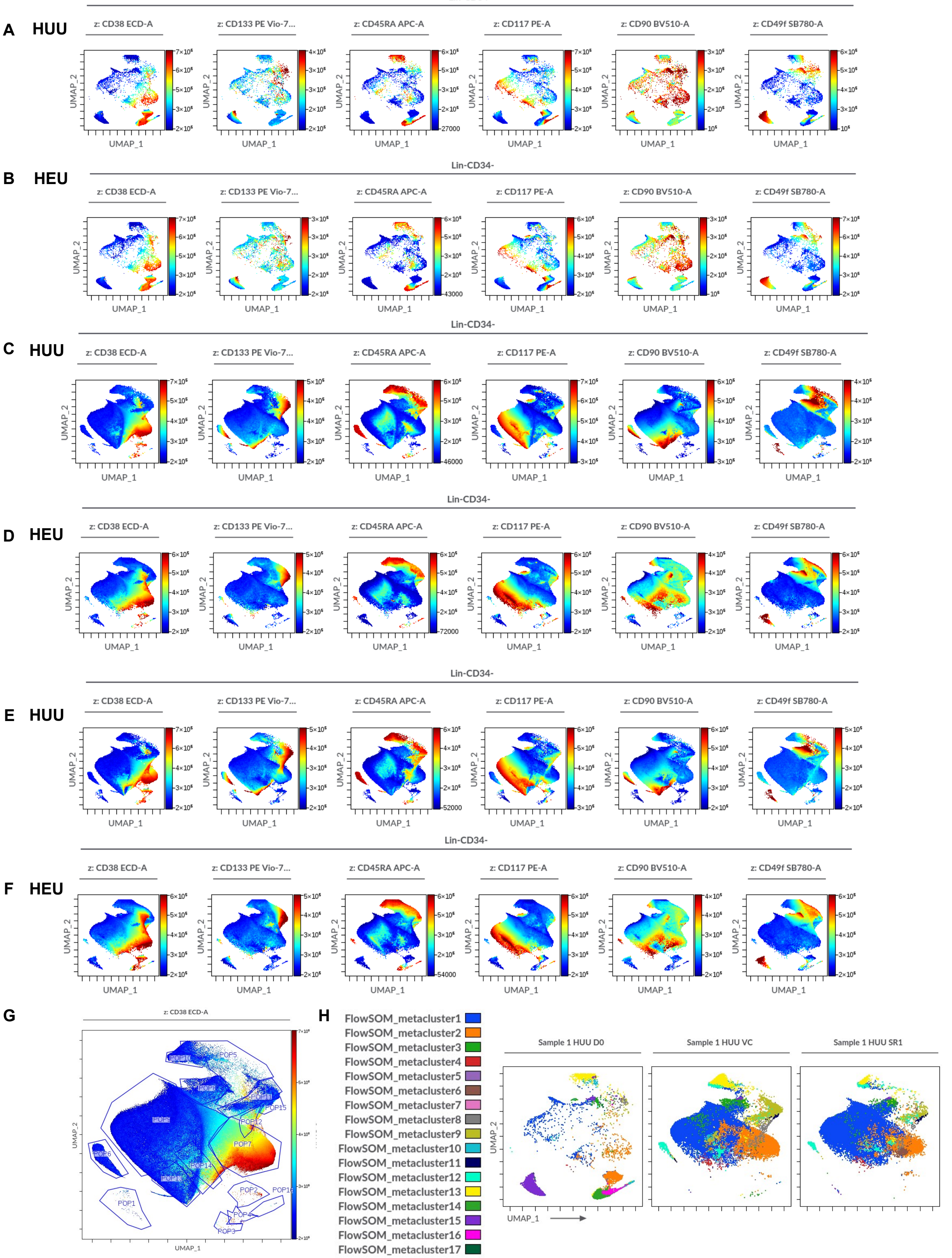


Figure S8: Population frequencies and statistically significant differences in Lin-CD34- sub-populations

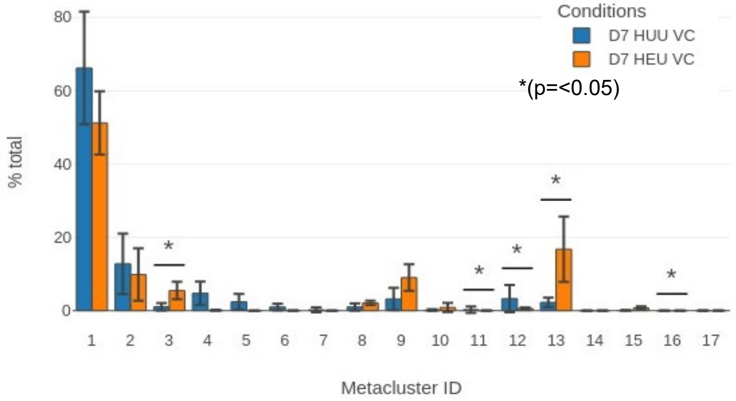
A

Average percentage of immunophenotypic sub-populations present in Lin-CD34- on D0 and D7 between HUU and HEU samples

Population name	Phenotype	% D0 HUU	% D0 HEU	% D7 VC HUU	% D7 VC HEU	% D7 SR1 HUU	% D7 SR1 HEU
POP1	Lin-CD34-CD38-CD133vCD45RA-CD117-CD90dim CD49f+	30,49	41,4	0,02	0,24	0,1	0,62
POP2	Lin-CD34-CD38+CD133dimCD45RA-CD117vCD90dimCD49f-	26,73	16,03	0,04	0,05	0,15	0,29
POP3	Lin-CD34-CD38vCD133-CD45RA+CD117-CD90-CD49f-	15,89	14,79	0,01	0,01	0,01	0,06
POP4	Lin-CD34-CD38+CD133-CD45RAdimCD117-CD90dimCD49f-	11,04	11,35	0,01	0,01	0,05	0,04
POP5	Lin-CD34-CD38-CD133vCD45RA+CD117-CD90vCD49f+	3,41	5,47	2,53	20,29	2,12	38,41
POP6	Lin-CD34-CD38-CD133vCD45RA+CD117-CD90+ CD49f-	0	0	3,02	0,03	2,91	0,03
POP7	Lin-CD34-CD38+CD133-CD45RAvCD117vCD90vCD49f-	1,14	1,07	19,35	11,99	21,28	7,69
POP8	Lin-CD34-CD38-CD133-CD45RA-CD117vCD90dimCD49f+	0,65	0,74	4,14	9,94	3,24	7,45
POP9	Lin-CD34-CD38-CD133-CD45RAvCD117vCD90vCD49f-	4,37	4,65	56,03	41,68	53,64	31,49
POP10	Lin-CD34-CD38-CD133dimCD45RA+CD117-CD90+CD49f-	0,02	0,03	6,4	2,09	1,7	1,1
POP11	Lin-CD34-CD38vCD133vCD45RAvCD117vCD90vCD49fdim	0,23	0,15	2,61	7,32	1,75	7,06
POP12	Lin-CD34-CD38vCD133-CD45RA-CD117-CD90dimCD49fdim	0,12	0,08	1,24	1,71	1,6	2,02
POP13	Lin-CD34-CD38-CD133vCD45RA-CD117+CD90+CD49f-	0,01	0	7,26	0,28	11,89	0,49
POP14	Lin-CD34-CD38vCD133vCD45RA-CD117dimCD90+CD49f-	0,77	0,56	3,69	1,03	5,92	1,14
POP15	Lin-CD34-CD38-CD133+CD45RA+CD117+CD90dimCD49fdim	0	0	0,56	0,33	0,3	0,09
POP16	Lin-CD34-CD38dimCD133-CD45RA+CD117+CD90-CD49fv	3,27	2,28	0,01	0,03	0	0,07

CD, cluster of differentiation; CMP, common myeloid progenitor; D0, day 0; D7, day 7; EMP, erythro-myeloid progenitor; Ext, extended; GMP, granulocyte-macrophage progenitor; HSC, haematopoietic stem cell; LMPP, lympho-myeloid primed progenitor; MPP, multipotent progenitor; SR1, stemregenin1; v, variable; VC, vehicle control

B



C

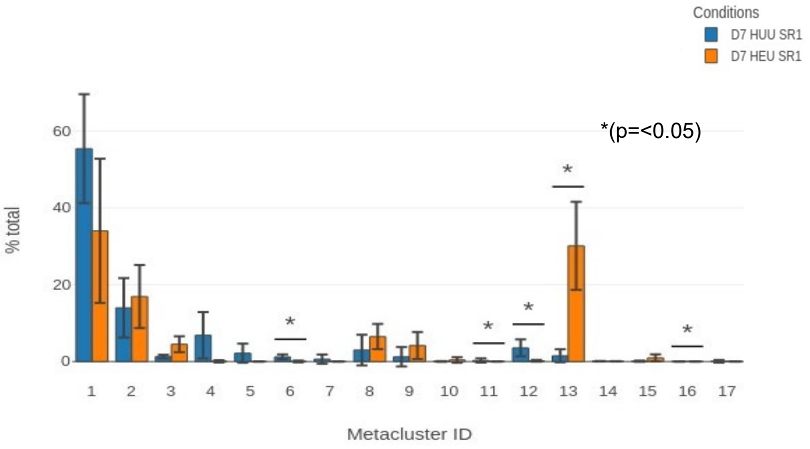


Figure S9: Differences in CFU phenotype between HEU and HUU samples.

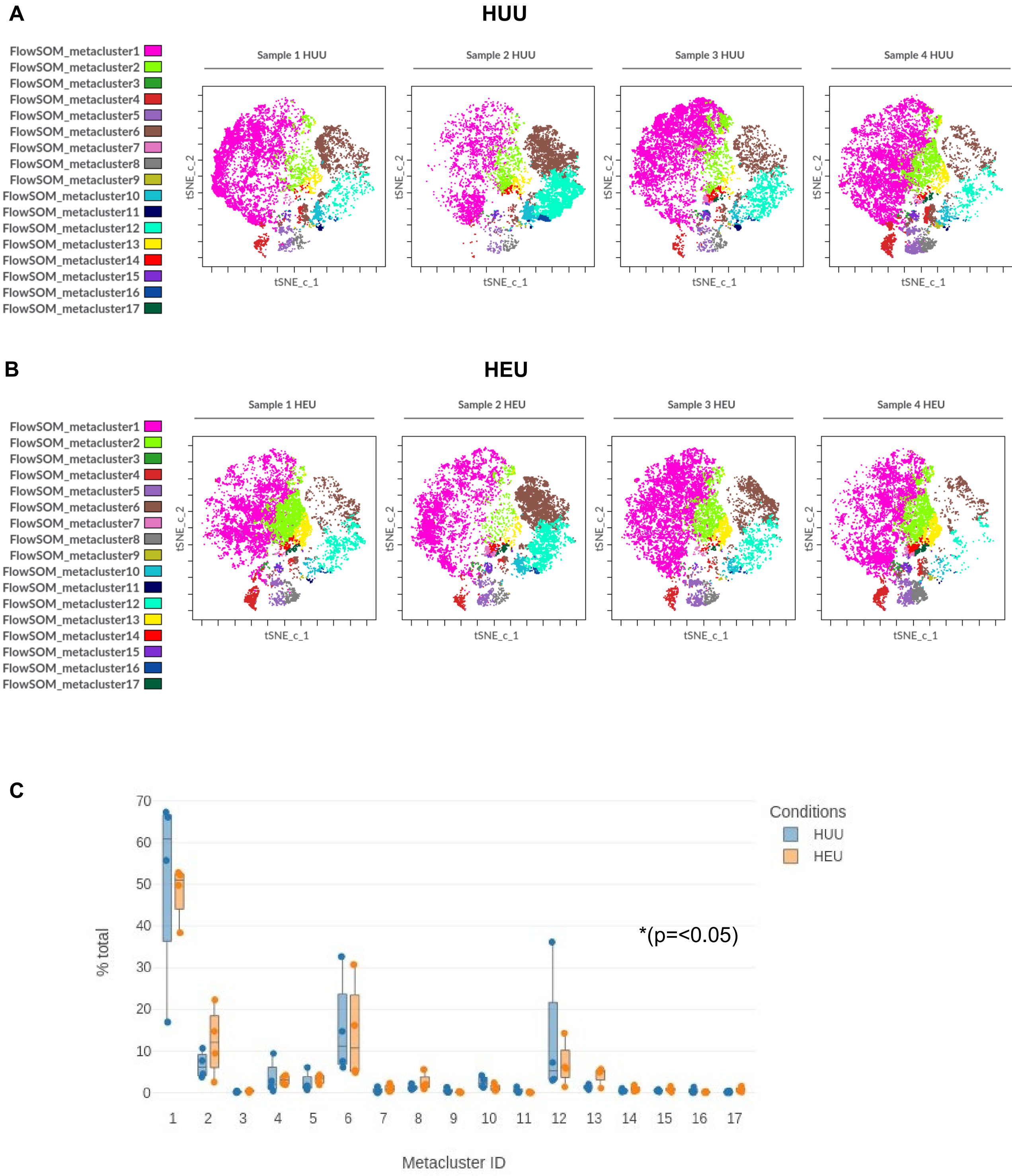


Figure S10: Volcano plot showing genes that were significantly up- (red dots, 108 genes) and downregulated (blue dots, 163 genes) between HEU and HUU samples based on their p-value (y-axis) and fold change (x-axis).

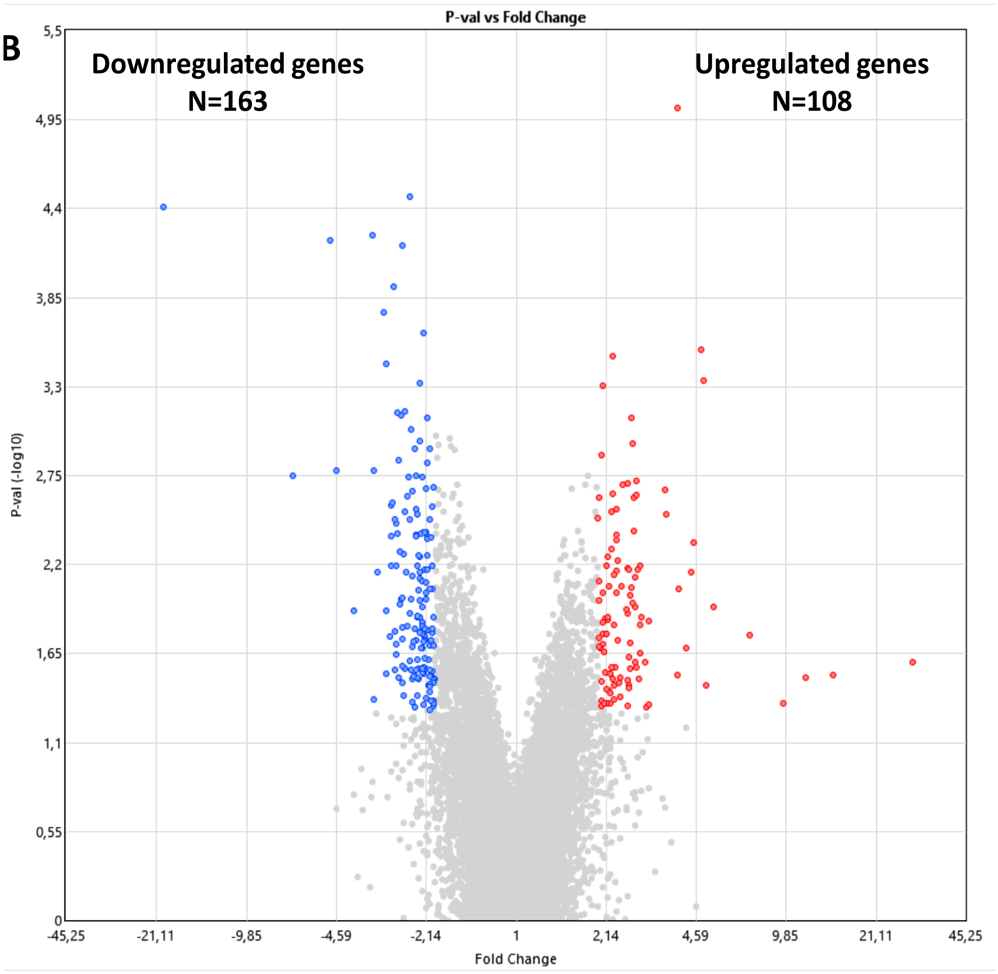


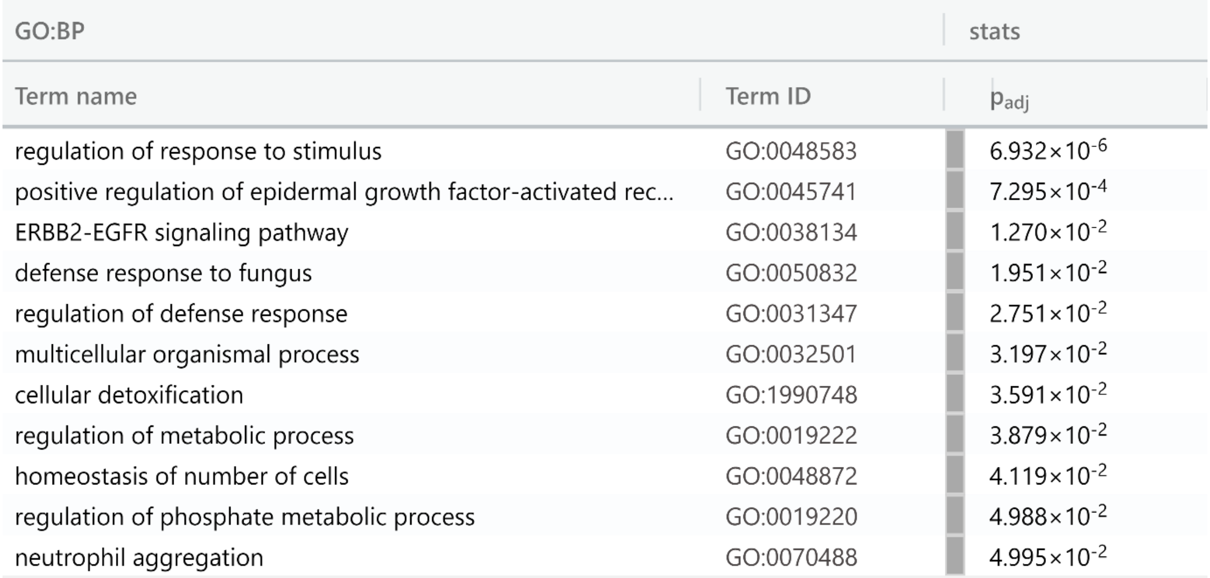
Figure S11: Full gene list displaying all 271 differentially expressed genes, fold change, p-value, gene symbol, and description.

ID	Fold Chang	P-val	Gene Symbol	Description
TC0100015858.hg.1	28,72	0,0255	S100A8	S100 calcium binding protein A8
TC0100015856.hg.1	14,66	0,0306	S100A12	S100 calcium binding protein A12
TC0400007857.hg.1	11,66	0,0316	AREG	amphiregulin
TC1200008275.hg.1	9,64	0,0454	SYT1	synaptotagmin I
TC0X00010473.hg.1	7,24	0,0172	TSC22D3	TSC22 domain family, member 3
TC0200014414.hg.1	5,29	0,0114	CXCR4	chemokine (C-X-C motif) receptor 4
TC0200012429.hg.1	4,98	0,0353	PPM1B	Memczak2013 ANTISENSE, coding, INTERNAL, intronic best transcript NM_001033556
TC1200009800.hg.1	4,88	0,0005	SLC2A3	solute carrier family 2 (facilitated glucose transporter), member 3
TC1200009796.hg.1	4,78	0,0003	SLC2A14	solute carrier family 2 (facilitated glucose transporter), member 14
TC0100010032.hg.1	4,49	0,0046	S100A9	S100 calcium binding protein A9
TC1500007355.hg.1	4,42	0,007	AQP9	aquaporin 9
TC1200007652.hg.1	4,23	0,0205	GRASP	GRP1 (general receptor for phosphoinositides 1)-associated scaffold protein
TC1100006749.hg.1	3,95	0,009	ZNF215	zinc finger protein 215
TC0100015598.hg.1	3,91	9,55E-06	TXNIP	thioredoxin interacting protein
TSUnmapped00000354.hg.1	3,9	0,0302	EIF3F	Eukaryotic translation initiation factor 3 subunit F [Source:UniProtKB/Swiss-Prot;Acc:O00303]
TC0200015885.hg.1	3,56	0,0031	FAM124B	family with sequence similarity 124 member B
TC1200010006.hg.1	3,53	0,0022	PLBD1	phospholipase B domain containing 1
TC2000009023.hg.1	3,08	0,0141	SAMHD1	SAM domain and HD domain 1
TC1900008900.hg.1	3,07	0,046	PPP6R1	Memczak2013 ANTISENSE, CDS, coding, INTERNAL best transcript NM_014931
TC1300008706.hg.1	3,01	0,0476	ELF1	Transcript Identified by AceView, Entrez Gene ID(s) 1997; 100128628
TC0800011150.hg.1	2,97	0,0253	TP53INP1	tumor protein p53 inducible nuclear protein 1
TC1700007777.hg.1	2,9	0,0132	THRA	thyroid hormone receptor, alpha
TC1700011815.hg.1	2,86	0,0148	MXRA7	matrix-remodelling associated 7
TC2200008661.hg.1	2,86	0,0222	LGALS2	lectin, galactoside-binding, soluble, 2
TC0100010241.hg.1	2,85	0,0063	MNDA	myeloid cell nuclear differentiation antigen
TC1700012262.hg.1	2,81	0,0323	PSME3	proteasome activator subunit 3
TC2200008475.hg.1	2,79	0,0067	PIK3IP1	phosphoinositide-3-kinase interacting protein 1
TC0700013442.hg.1	2,78	0,0019	LSMEM1	leucine-rich single-pass membrane protein 1
TC0700011584.hg.1	2,78	0,0273	FGL2	fibrinogen-like 2
TC1900008881.hg.1	2,76	0,0023	FCAR	Fc fragment of IgA receptor
TC0800009819.hg.1	2,75	0,0115	DOK2	docking protein 2
TC0300010770.hg.1	2,75	0,025	CSRNP1	cysteine-serine-rich nuclear protein 1
TC1200010559.hg.1	2,73	0,0076	VDR	vitamin D (1,25- dihydroxyvitamin D3) receptor
TC0600013231.hg.1	2,7	0,0024	SGK1	serum/glucocorticoid regulated kinase 1
TC2000007117.hg.1	2,7	0,0039	ASXL1	additional sex combs like transcriptional regulator 1
TC1900008141.hg.1	2,68	0,0108	CYP2S1	cytochrome P450, family 2, subfamily S, polypeptide 1
TC0100016944.hg.1	2,68	0,0011	ARL8A	ADP-ribosylation factor like GTPase 8A
TC0400007851.hg.1	2,65	0,0088	EPGN	epithelial mitogen
TC1700011910.hg.1	2,65	0,0008	CYTH1	cytohesin 1
TC0500011852.hg.1	2,64	0,0276	LOX	lysyl oxidase
TC0700010604.hg.1	2,64	0,0279	CPVL	carboxypeptidase, vitellogenic-like
TC0200010795.hg.1	2,63	0,0097	SLC11A1	solute carrier family 11 (proton-coupled divalent metal ion transporter), member 1
TC2200007356.hg.1	2,61	0,0193	KDELR3	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 3
TC0500012238.hg.1	2,61	0,0234	HBEGF	heparin-binding EGF-like growth factor
TC0300007016.hg.1	2,6	0,0351	ARPP21	Transcript Identified by AceView, Entrez Gene ID(s) 10777
TC1500010160.hg.1	2,59	0,0068	CTSH	cathepsin H
TC0400007853.hg.1	2,59	0,0365	EREG	epiregulin
TC0100008625.hg.1	2,58	0,0127	PDE4B	phosphodiesterase 4B, cAMP-specific
TC1100013012.hg.1	2,58	0,002	STX3	syntaxin 3
TC1500010731.hg.1	2,58	0,047	TRIM69	tripartite motif containing 69
TC0500007138.hg.1	2,58	0,0326	IL7R	interleukin 7 receptor
TC0600012009.hg.1	2,58	0,0067	DEFB114	defensin, beta 114
TC0500012519.hg.1	2,55	0,0118	SPARC	secreted protein, acidic, cysteine-rich (osteonectin)
TC0100015789.hg.1	2,47	0,002	POGZ	Transcript Identified by AceView, Entrez Gene ID(s) 23126
TC0700013581.hg.1	2,43	0,0086	STEAP4	STEAP family member 4
TC0M00006437.hg.1	2,42	0,0313	COX2	cytochrome c oxidase subunit II
TC0600008109.hg.1	2,4	0,0415	VEGFA	vascular endothelial growth factor A
TC1800008743.hg.1	2,4	0,0336	CCDC68	coiled-coil domain containing 68
TC0600011843.hg.1	2,36	0,0185	TBCC	tubulin folding cofactor C
TC1700009528.hg.1	2,36	0,006	CXCL16	chemokine (C-X-C motif) ligand 16
TC0200009662.hg.1	2,34	0,0029	TNFAIP6	tumor necrosis factor, alpha-induced protein 6
TC0300008550.hg.1	2,34	0,0068	CSTA	cystatin A (stefin A)
TC0300011516.hg.1	2,34	0,0045	EIF4E3	eukaryotic translation initiation factor 4E family member 3
TC0800007284.hg.1	2,33	0,0094	KCNU1	potassium channel, subfamily U, member 1
TC0200007261.hg.1	2,32	0,0041	QPCT	glutaminyl-peptide cyclotransferase
TC1000009865.hg.1	2,31	0,027	FAM107B	Memczak2013 ALT_ACCEPTOR, ALT_DONOR, coding, INTERNAL, intronic best transcript NM_031453
TC0700011571.hg.1	2,29	0,0073	YWHAG	Memczak2013 ALT_ACCEPTOR, ALT_DONOR, coding, INTERNAL, intronic best transcript NM_012479
TC0300009692.hg.1	2,28	0,0148	DVL3	dishevelled segment polarity protein 3
TC1000011807.hg.1	2,28	0,0433	SMNDC1	survival motor neuron domain containing 1
TC1200011961.hg.1	2,28	0,0349	TMEM116	transmembrane protein 116
TC0100007832.hg.1	2,26	0,0003	ZC3H12A; MIR6732	zinc finger CCCH-type containing 12A; microRNA 6732
TC1700011523.hg.1	2,26	0,0023	PSMD12	proteasome 26S subunit, non-ATPase 12
TC0500010265.hg.1	2,26	0,0318	CDH18	cadherin 18, type 2
TC1100010809.hg.1	2,24	0,003	OR4C12	olfactory receptor, family 4, subfamily C, member 12
TC0600013273.hg.1	2,23	0,0051	BCLAF1	Memczak2013 ALT_DONOR, coding, INTERNAL, intronic best transcript NM_014739
TC1400010085.hg.1	2,23	0,0273	TCL1A	T-cell leukemia/lymphoma 1A
TC0300012816.hg.1	2,23	0,0455	P2RY12	purinergic receptor P2Y, G-protein coupled, 12
TC0400012178.hg.1	2,23	0,0297	DCHS2	dachsous cadherin-related 2
TC0600012562.hg.1	2,21	0,0395	BACH2	Memczak2013 ALT_ACCEPTOR, ALT_DONOR, coding, INTERNAL, intronic best transcript NM_001170794
TC0600014360.hg.1	2,2	0,0085	IPCEF1	interaction protein for cytohesin exchange factors 1
TC0200007200.hg.1	2,18	0,0375	LTBP1	latent transforming growth factor beta binding protein 1
TC1100006831.hg.1	2,18	0,0132	ADM	adrenomedullin
TC0600013570.hg.1	2,18	0,0056	ZBTB2	zinc finger and BTB domain containing 2
TC0900008148.hg.1	2,17	0,0138	TDRD7	tudor domain containing 7
TC1600007936.hg.1	2,14	0,0065	SLC6A2	solute carrier family 6 (neurotransmitter transporter), member 2
TC1500006811.hg.1	2,14	0,037	CHRM5	cholinergic receptor, muscarinic 5
TC1100010733.hg.1	2,14	0,0458	SPI1	Spi-1 proto-oncogene
TC1200007050.hg.1	2,14	0,0168	IAPP	islet amyloid polypeptide
TC0700012467.hg.1	2,13	0,0136	TMEM229A	transmembrane protein 229A
TC0700008190.hg.1	2,12	0,0292	CD36	CD36 molecule (thrombospondin receptor)
TC1600011426.hg.1	2,11	0,0452	ATXN1L	ataxin 1-like
TC0500010553.hg.1	2,09	0,0217	FYB	FYN binding protein
TC0100008026.hg.1	2,09	0,0169	YBX1	Zhang2013 ALT_ACCEPTOR, ALT_DONOR, coding, INTERNAL, intronic best transcript NM_004559
TC1000007883.hg.1	2,09	0,0195	SRGN	serglycin
TC0100006805.hg.1	2,08	0,0094	RBP7	retinol binding protein 7, cellular

TC1900010625.hg.1	2,08	0,0143	LGALS4	lectin, galactoside-binding, soluble, 4
TC1700007415.hg.1	2,07	0,0005	CRYBA1	crystallin beta A1
TC1300006676.hg.1	2,06	0,0013	RASL11A	RAS-like, family 11, member A
TC1600006446.hg.1	2,05	0,0434	HBA1	hemoglobin, alpha 1
TC0500012486.hg.1	2,05	0,0473	DCTN4	dynactin 4 (p62)
TC1100009904.hg.1	2,05	0,0331	OR51A2	olfactory receptor, family 51, subfamily A, member 2
TC1200012753.hg.1	2,04	0,0207	KLRC2	killer cell lectin-like receptor subfamily C, member 2
TC1200006561.hg.1	2,02	0,0024	RAD51AP1	RAD51 associated protein 1
TC1400007890.hg.1	2,01	0,0105	GPR65	G protein-coupled receptor 65
TC2100007821.hg.1	2,01	0,0201	ADAMTS1	ADAM metallopeptidase with thrombospondin type 1 motif 1
TC0M00006454.hg.1	2,01	0,008	ND6	NADH dehydrogenase, subunit 6 (complex I)
TC2000008666.hg.1	2,01	0,018	NAPB	N-ethylmaleimide-sensitive factor attachment protein, beta
TC1800007063.hg.1	2	0,0033	DTNA	Transcript Identified by AceView, Entrez Gene ID(s) 1837
TC0100017722.hg.1	-2	0,0321	C1orf131	chromosome 1 open reading frame 131
TSUnmapped00000306.hg.1	-2,01	0,02	ZKSCAN7	zinc finger with KRAB and SCAN domains 7
TC0600009143.hg.1	-2,01	0,0336	SLC16A10	solute carrier family 16 (aromatic amino acid transporter), member 10
TC0500009010.hg.1	-2,01	0,0327	TCERG1	Transcript Identified by AceView, Entrez Gene ID(s) 10915
TC0100008631.hg.1	-2,01	0,0473	SGIP1	SH3-domain GRB2-like (endophilin) interacting protein 1
TC0200016659.hg.1	-2,02	0,0451	ATP6V1E2	ATPase, H+ transporting, lysosomal 31kDa, V1 subunit E2
TC1200006557.hg.1	-2,02	0,0133	TIGAR	TP53 induced glycolysis regulatory phosphatase
TC0200007195.hg.1	-2,02	0,0434	TTC27; MIR4765	tetratricopeptide repeat domain 27; microRNA 4765
TC0500013360.hg.1	-2,03	0,0021	CDKL3	cyclin-dependent kinase-like 3
TC1400010737.hg.1	-2,03	0,0028	KLHL28	kelch-like family member 28
TC0300012339.hg.1	-2,04	0,0089	GATA2	GATA binding protein 2
TC0400012961.hg.1	-2,04	0,0064	NUDT6	nudix hydrolase 6
TC0400012496.hg.1	-2,04	0,0285	AGA	aspartylglucosaminidase
TC0800010490.hg.1	-2,04	0,0165	TMEM68	transmembrane protein 68
TC1100012187.hg.1	-2,04	0,0158	KBTBD3	kelch repeat and BTB (POZ) domain containing 3
TC1700012231.hg.1	-2,05	0,0435	TMEM98	transmembrane protein 98
TC0100011421.hg.1	-2,05	0,0298	CD46	Transcript Identified by AceView, Entrez Gene ID(s) 4179
TC1900011922.hg.1	-2,06	0,0185	ZNF728	zinc finger protein 728
TC1600009514.hg.1	-2,07	0,0042	RRN3	RRN3 homolog, RNA polymerase I transcription factor
TC0100009035.hg.1	-2,07	0,0012	KIAA1107	KIAA1107
TC1200012610.hg.1	-2,07	0,0346	SSPN	sarcospan
TC0600013574.hg.1	-2,07	0,0033	RMND1	required for meiotic nuclear division 1 homolog
TC1900007267.hg.1	-2,07	0,009	AP1M1	adaptor-related protein complex 1, mu 1 subunit
TC0700009680.hg.1	-2,07	0,0494	TMEM176A	transmembrane protein 176A
TC1200012706.hg.1	-2,07	0,0387	ACAD10	acyl-CoA dehydrogenase family, member 10
TC1200007406.hg.1	-2,07	0,0281	IRAK4	interleukin 1 receptor associated kinase 4
TC1700012283.hg.1	-2,07	0,0306	LUC7L3	LUC7-like 3 pre-mRNA splicing factor
TC1200007706.hg.1	-2,07	0,0199	MFSD5	major facilitator superfamily domain containing 5
TC1900011754.hg.1	-2,08	0,0355	ZNF225	zinc finger protein 225
TC0500010628.hg.1	-2,09	0,0354	ANXA2R	annexin A2 receptor
TC0100018367.hg.1	-2,11	0,0242	CHRM3	cholinergic receptor, muscarinic 3
TC1700009402.hg.1	-2,12	0,016	SMYD4	SET and MYND domain containing 4
TC1000009894.hg.1	-2,13	0,0044	FAM188A	family with sequence similarity 188, member A
TC1700008350.hg.1	-2,13	0,0067	NOG	noggin
TC1100007266.hg.1	-2,13	0,0015	PDHX	pyruvate dehydrogenase complex, component X
TC1900011936.hg.1	-2,13	0,0055	ZNF781	zinc finger protein 781
TC0200016684.hg.1	-2,14	0,0008	TIA1	TIA1 cytotoxic granule-associated RNA binding protein
TC0200010836.hg.1	-2,14	0,042	ZFAND2B	zinc finger, AN1-type domain 2B
TC0600008515.hg.1	-2,14	0,004	KCNQ5	Jeck2013 ALT_ACCEPTOR, ALT_DONOR, coding, INTERNAL, intronic best transcript NM_001160133
TC1300009993.hg.1	-2,15	0,0081	BIVM	basic, immunoglobulin-like variable motif containing
TC1900008300.hg.1	-2,15	0,0104	RELB	v-rel avian reticuloendotheliosis viral oncogene homolog B
TC0700012159.hg.1	-2,15	0,0094	PUS7	pseudouridylate synthase 7 (putative)
TC0300009613.hg.1	-2,15	0,004	USP13	ubiquitin specific peptidase 13 (isopeptidase T-3)
TC0300006627.hg.1	-2,15	0,0021	TSEN2	TSEN2 tRNA splicing endonuclease subunit
TC0600014246.hg.1	-2,16	0,0277	C6orf229	chromosome 6 open reading frame 229
TC0300006781.hg.1	-2,17	0,0176	KCNH8	potassium channel, voltage gated eag related subfamily H, member 8
TC1900012025.hg.1	-2,17	0,0157	ZNF28	zinc finger protein 28
TC0900010792.hg.1	-2,18	0,0068	NOL8	nucleolar protein 8
TC1100013000.hg.1	-2,18	0,0185	CLP1; AP000662.4	Transcript Identified by AceView, Entrez Gene ID(s) 10978; novel transcript, sense overlapping CLP1
TC0300014077.hg.1	-2,18	0,0238	WDR49	WD repeat domain 49
TC1800008097.hg.1	-2,18	0,004	MPPE1	metallophosphoesterase 1
TC0500011173.hg.1	-2,19	0,0461	POC5	POC5 centriolar protein
TC1900009752.hg.1	-2,19	0,0272	PRDX2	peroxiredoxin 2
TC1700009254.hg.1	-2,19	0,0189	HEXDC	hexosaminidase (glycosyl hydrolase family 20, catalytic domain) containing
TC0400006756.hg.1	-2,19	0,004	TRMT44	tRNA methyltransferase 44 homolog (S. cerevisiae)
TC0700008094.hg.1	-2,2	0,0002	ZP3	zona pellucida glycoprotein 3 (sperm receptor)
TC0100011064.hg.1	-2,2	0,03	CFH	complement factor H
TC1300007731.hg.1	-2,21	0,008	GPR180	G protein-coupled receptor 180
TC1200008113.hg.1	-2,22	0,0152	NUP107	nucleoporin 107kDa
TC0600011517.hg.1	-2,22	0,0207	HLA-DPA1	major histocompatibility complex, class II, DP alpha 1
TC0900011523.hg.1	-2,22	0,0168	HSPA5	heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa)
TC0800008235.hg.1	-2,22	0,0161	FAM92A1	family with sequence similarity 92, member A1
TC2200009324.hg.1	-2,22	0,0276	DRICH1	aspartate-rich 1
TC0400011749.hg.1	-2,22	0,0293	BBS7	Bardet-Biedl syndrome 7
TC2200008257.hg.1	-2,22	0,0114	LRP5L	LDL receptor related protein 5 like
TC0100016715.hg.1	-2,23	0,0145	PTGS2	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)
TC0300013687.hg.1	-2,23	0,0018	TFRC	Transcript Identified by AceView, Entrez Gene ID(s) 7037
TC0800010553.hg.1	-2,23	0,0207	NSMAF	neutral sphingomyelinase activation associated factor
TC1400008362.hg.1	-2,23	0,004	TRAF3	TNF receptor-associated factor 3
TC0200010516.hg.1	-2,26	0,0134	CYP20A1	cytochrome P450, family 20, subfamily A, polypeptide 1
TC0700011601.hg.1	-2,26	0,0077	TMEM60	transmembrane protein 60
TC1000006863.hg.1	-2,26	0,0165	RPP38	ribonuclease P/MRP 38kDa subunit
TC0400009458.hg.1	-2,27	0,0056	CDKN2AIP	CDKN2A interacting protein

TC0100011746.hg.1	-2,27	0,0165	CNIH4	cornichon family AMPA receptor auxiliary protein 4
TC0900008804.hg.1	-2,27	0,0071	ZNF79	zinc finger protein 79
TC1900010873.hg.1	-2,27	0,0005	ZNF45	zinc finger protein 45
TC1200008843.hg.1	-2,27	0,0105	FAM216A	family with sequence similarity 216, member A
TC1200010641.hg.1	-2,27	0,0166	BCDIN3D	BCDIN3 domain containing
TC0400009503.hg.1	-2,28	0,0011	PRIMPOL	primase and DNA directed polymerase
TC0700008650.hg.1	-2,28	0,0056	FAM185A	family with sequence similarity 185, member A
TC0100018446.hg.1	-2,28	0,0275	OMA1	OMA1 zinc metallopeptidase
TC0400008947.hg.1	-2,3	0,032	TMEM184C	transmembrane protein 184C
TC1100008091.hg.1	-2,3	0,009	CCS	copper chaperone for superoxide dismutase
TC1400007265.hg.1	-2,3	0,0316	TMEM260	transmembrane protein 260
TC1200012836.hg.1	-2,3	0,0064	C12orf73	chromosome 12 open reading frame 73
TC1000007427.hg.1	-2,3	0,0405	ZNF485	zinc finger protein 485
TC0400012938.hg.1	-2,31	0,0132	RASGEF1B	RasGEF domain family member 1B
TC0200008323.hg.1	-2,31	0,0133	PLGLB2	plasminogen-like B2
TC1100011398.hg.1	-2,32	0,0245	MRGPRD	MAS-related GPR, member D
TC0300012384.hg.1	-2,32	0,0031	H1FX	H1 histone family, member X
TC0600013898.hg.1	-2,33	0,0041	FAM103A1	Homo sapiens family with sequence similarity 103, member A1, mRNA (cDNA clone MGC:2560 IMAGE:2989772), complete cds.; Homo sapiens family with sequence similarity 103, member A1, mRNA (cDNA clone MGC:18029 IMAGE:3924570), complete cds.; Homo sapiens family with sequence similarity 103, member A1, mRNA (cDNA clone MGC:102778 IMAGE:5578103), complete cds.
TC0300007206.hg.1	-2,34	0,028	TCAIM	T cell activation inhibitor, mitochondrial
TC1000009922.hg.1	-2,34	0,0184	VIM	Jeck2013 ANTISENSE, coding, INTERNAL, intronic best transcript NM_003380
TC1500010725.hg.1	-2,34	0,0018	CAPN3	calpain 3
TC0700013538.hg.1	-2,35	0,0042	TRGJ2; TRGV9; TRGC1; TRGJP; TRGJP2	T cell receptor gamma joining 1; T cell receptor gamma constant 2; T cell receptor gamma joining 2; T cell receptor gamma variable 9; T cell receptor gamma constant 1; T cell receptor gamma joining P; T cell receptor gamma joining P2
TC0100008588.hg.1	-2,35	0,0028	EFCAB7; DLEU2L	EF-hand calcium binding domain 7; deleted in lymphocytic leukemia 2-like
TC0100018482.hg.1	-2,35	0,0287	HIST2H2AA4; HIST2	histone cluster 2, H2aa4; histone cluster 2, H2aa3
TC0200016516.hg.1	-2,36	0,0012	TMEM177	transmembrane protein 177
TC0100008496.hg.1	-2,36	0,0156	FGGY	FGGY carbohydrate kinase domain containing
TC1700009746.hg.1	-2,36	0,048	MYH3	myosin, heavy chain 3, skeletal muscle, embryonic
TC1900008699.hg.1	-2,38	0,0309	ZNF528	zinc finger protein 528
TC2000009392.hg.1	-2,39	0,0188	B4GALT5	UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 5
TC0100009746.hg.1	-2,4	0,0022	NUDT17	nudix hydrolase 17
TC1500009199.hg.1	-2,4	0,0074	CDAN1	codanin 1
TC1900008012.hg.1	-2,4	0,0449	CATSPERG	catsper channel auxiliary subunit gamma
TC1000011549.hg.1	-2,41	0,032	RRP12	Memczak2013 ALT_ACCEPTOR, ALT_DONOR, coding, INTERNAL, intronic best transcript NM_001145114
TC0X00008138.hg.1	-2,42	0,0202	ALG13	Transcript Identified by AceView, Entrez Gene ID(s) 79868
TC1900011650.hg.1	-2,43	0,0104	ARHGEF18	Rho/Rac guanine nucleotide exchange factor 18
TC0X00006704.hg.1	-2,43	0,028	REPS2	RALBP1 associated Eps domain containing 2
TC0300006961.hg.1	-2,45	0,0009	CMTM8	CKLF-like MARVEL transmembrane domain containing 8
TC1600007550.hg.1	-2,46	0,0125	ZNF720	zinc finger protein 720
TC1100011294.hg.1	-2,46	0,0249	CTSF	cathepsin F
TC0600007262.hg.1	-2,46	3,32E-05	HIST1H3A	histone cluster 1, H3a
TC0100016893.hg.1	-2,47	0,0033	DDX59	DEAD (Asp-Glu-Ala-Asp) box polypeptide 59
TC0500010019.hg.1	-2,49	0,0018	FASTKD3	FAST kinase domains 3
TC2000009248.hg.1	-2,51	0,0024	NEURL2	neuralized E3 ubiquitin protein ligase 2
TC0500013186.hg.1	-2,53	0,0152	PTCD2	pentatricopeptide repeat domain 2
TC2200007318.hg.1	-2,54	0,007	H1F0	H1 histone family, member 0
TC1200006444.hg.1	-2,56	0,0029	CCDC77	coiled-coil domain containing 77
TC0100012593.hg.1	-2,57	0,0007	WRAP73	WD repeat containing, antisense to TP73
TC0100009876.hg.1	-2,57	0,0277	HIST2H2AA3; HIST2	histone cluster 2, H2aa3; histone cluster 2, H2aa4
TC1700009554.hg.1	-2,59	0,0054	ZNF594	zinc finger protein 594
TC0700011050.hg.1	-2,6	0,0403	FIGNL1	fidgetin-like 1
TC1100012033.hg.1	-2,62	0,0101	MTMR2	myotubularin related protein 2
TC0X00009493.hg.1	-2,62	0,0338	EFHC2	EF-hand domain (C-terminal) containing 2
TC0100010926.hg.1	-2,62	0,0265	OCLM	Transcript Identified by AceView, Entrez Gene ID(s) 10896
TC1900011972.hg.1	-2,63	6,68E-05	ZNF404	zinc finger protein 404
TC1500008616.hg.1	-2,63	0,0182	HERC2P3	hect domain and RLD 2 pseudogene 3
TC1000008351.hg.1	-2,63	0,0153	MINPP1	multiple inositol-polyphosphate phosphatase 1
TC0700006512.hg.1	-2,66	0,0103	INTS1	Memczak2013 ANTISENSE, CDS, coding, INTERNAL best transcript NM_001080453
TC0100010553.hg.1	-2,66	0,0007	C1orf112	chromosome 1 open reading frame 112
TC2000009247.hg.1	-2,67	0,0053	SPATA25	spermatogenesis associated 25
TC1700009877.hg.1	-2,68	0,011	ZSWIM7	zinc finger, SWIM-type containing 7
TC0500009076.hg.1	-2,7	0,0315	SLC26A2	solute carrier family 26 (anion exchanger), member 2
TC1000007458.hg.1	-2,71	0,0014	TMEM72	transmembrane protein 72
TC1900006605.hg.1	-2,73	0,0041	ZNF57	zinc finger protein 57
TC1600009375.hg.1	-2,73	0,0007	TVP23A	trans-golgi network vesicle protein 23 homolog A (S. cerevisiae)
TC0600007384.hg.1	-2,75	0,0225	HIST1H2BO	histone cluster 1, H2bo
TC0200015570.hg.1	-2,76	0,0035	CRYGD	crystallin gamma D
TC1000012600.hg.1	-2,76	0,0195	FAM24B	family with sequence similarity 24, member B
TC0400012396.hg.1	-2,77	0,0064	AADAT	aminoadipate aminotransferase
TC0500011433.hg.1	-2,79	0,0162	CETN3	centrin 3
TC0900011246.hg.1	-2,81	0,028	ZNF883	zinc finger protein 883
TC1700007645.hg.1	-2,81	0,0033	CTB-75G16.3; MRM	Transcript Identified by AceView, Entrez Gene ID(s) 79922; novel transcript
TC0500008785.hg.1	-2,81	0,0001	EGR1	early growth response 1
TC0500010579.hg.1	-2,86	0,0026	TTC33	tetratricopeptide repeat domain 33
TC0300008381.hg.1	-2,87	0,0064	QTRTD1	queuine tRNA-ribosyltransferase domain containing 1
TC1200010598.hg.1	-2,88	0,0042	RND1	Rho family GTPase 1
TC0300009727.hg.1	-2,89	0,0027	VPS8	Transcript Identified by AceView, Entrez Gene ID(s) 23355
TC1100008886.hg.1	-2,9	0,0176	TMEM133	transmembrane protein 133
TC1600009953.hg.1	-3	0,0298	GDPD3	glycerophosphodiester phosphodiesterase domain containing 3
TC0100017436.hg.1	-3,01	0,0122	TAF1A	TATA box binding protein (TBP)-associated factor, RNA polymerase I, A, 48kDa
TC1800006749.hg.1	-3,01	0,0004	CEP192	centrosomal protein 192kDa
TC0600014162.hg.1	-3,09	0,0002	TSTD3	thiosulfate sulfurtransferase (rhodanese)-like domain containing 3
TC0100017617.hg.1	-3,23	0,007	HIST3H2A	histone cluster 3, H2a
TC0600007265.hg.1	-3,35	0,0017	HIST1H3C	histone cluster 1, H3c
TC0500007681.hg.1	-3,35	0,0428	OCLN	occludin
TC0600011133.hg.1	-3,39	5,79E-05	HIST1H2BE	Memczak2013 ANTISENSE, CDS, coding, upstream_start, UTR3, UTR5 best transcript NM_003523
TC0100018539.hg.1	-3,94	0,0121	LHX4-AS1; ACBD6	LHX4 antisense RNA 1; acyl-CoA binding domain containing 6
TC0800011211.hg.1	-4,59	0,0016	NIPAL2	NIPA-like domain containing 2
TC0300009531.hg.1	-4,85	6,25E-05	NLGN1	neuroligin 1
TC0600011237.hg.1	-6,63	0,0018	OR2B2	olfactory receptor, family 2, subfamily B, member 2
TC1200008667.hg.1	-19,82	3,87E-05	HSP90B1	Transcript Identified by AceView, Entrez Gene ID(s) 7184

Figure S12: Functional analysis of DEGs by Gene Ontology biological processes (GOBP) depicting 11 biological processes which are involved. Only upregulated but not downregulated genes were involved in these processes.



GO:BP		stats
Term name	Term ID	p_{adj}
regulation of response to stimulus	GO:0048583	6.932×10^{-6}
positive regulation of epidermal growth factor-activated rec...	GO:0045741	7.295×10^{-4}
ERBB2-EGFR signaling pathway	GO:0038134	1.270×10^{-2}
defense response to fungus	GO:0050832	1.951×10^{-2}
regulation of defense response	GO:0031347	2.751×10^{-2}
multicellular organismal process	GO:0032501	3.197×10^{-2}
cellular detoxification	GO:1990748	3.591×10^{-2}
regulation of metabolic process	GO:0019222	3.879×10^{-2}
homeostasis of number of cells	GO:0048872	4.119×10^{-2}
regulation of phosphate metabolic process	GO:0019220	4.988×10^{-2}
neutrophil aggregation	GO:0070488	4.995×10^{-2}