Phosphoprotein Keratin 23 accumulates in MSS but not MSI colon cancers *in vivo* and impacts viability and proliferation *in vitro*

Supplementary Figure 1

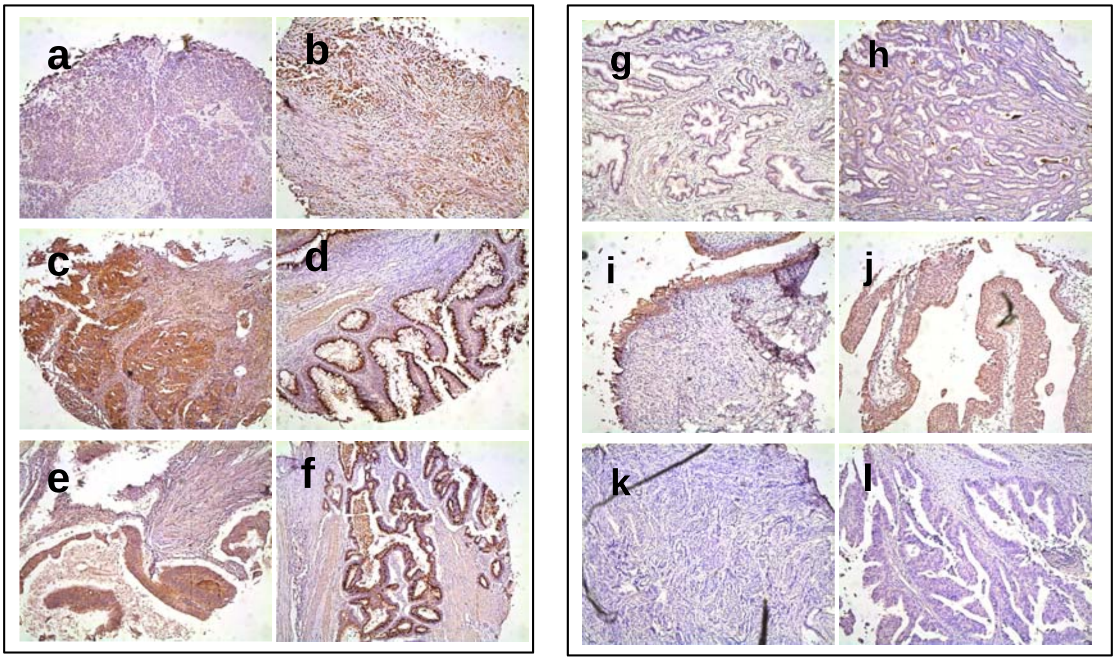
Immunohistochemical analyses were performed applying the rabbit anti-K23 antibody in a 1:600 dilution to a tissue microarray with various human tumors from different organ sites and matching or non-matching normal tissues (100 x magnification).

a)-f) K23 is upregulated in adenocarcinomas compared to normal mucosa from the same organ. a) normal pancreas (exocrine only); b) pancreas adenocarcinoma; c)-f) gastrointestinal adenocarcinomas of the c) stomach d) small intestine e) colon and f) rectum

g)-l) K23 is not differentially expressed in epithelial cells of the prostate, bladder or uterus.

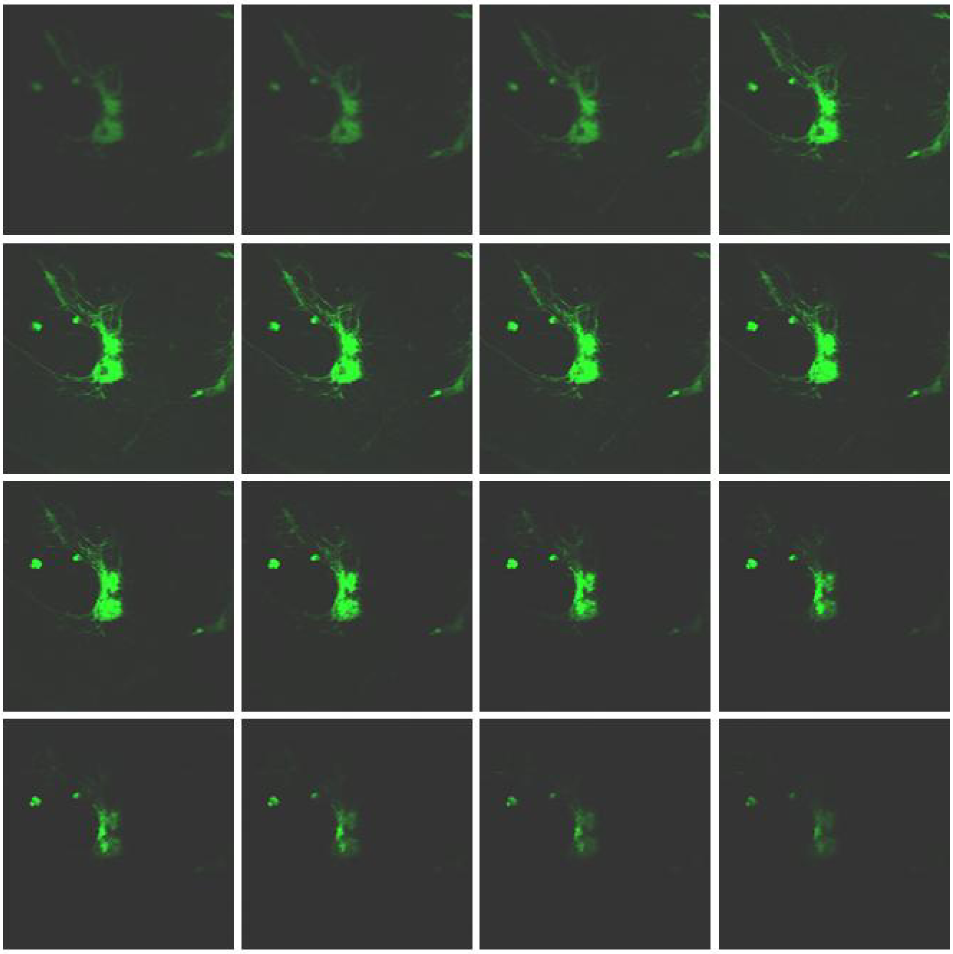
g) normal prostate ; h) prostate adenocarcinoma; i) normal urothelium j) bladder transitional cell carcinoma grade II; k) normal endometrium; l) adenocarcinoma of the endometrium.

g)-l) K23 is not differentially expressed in epithelial cells of the prostate, bladder or uterus. g) normal prostate ; h) prostate adenocarcinoma; i) normal urothelium j) bladder transitional cell carcinoma grade II; k) normal endometrium; l) adenocarcinoma of the endometrium.



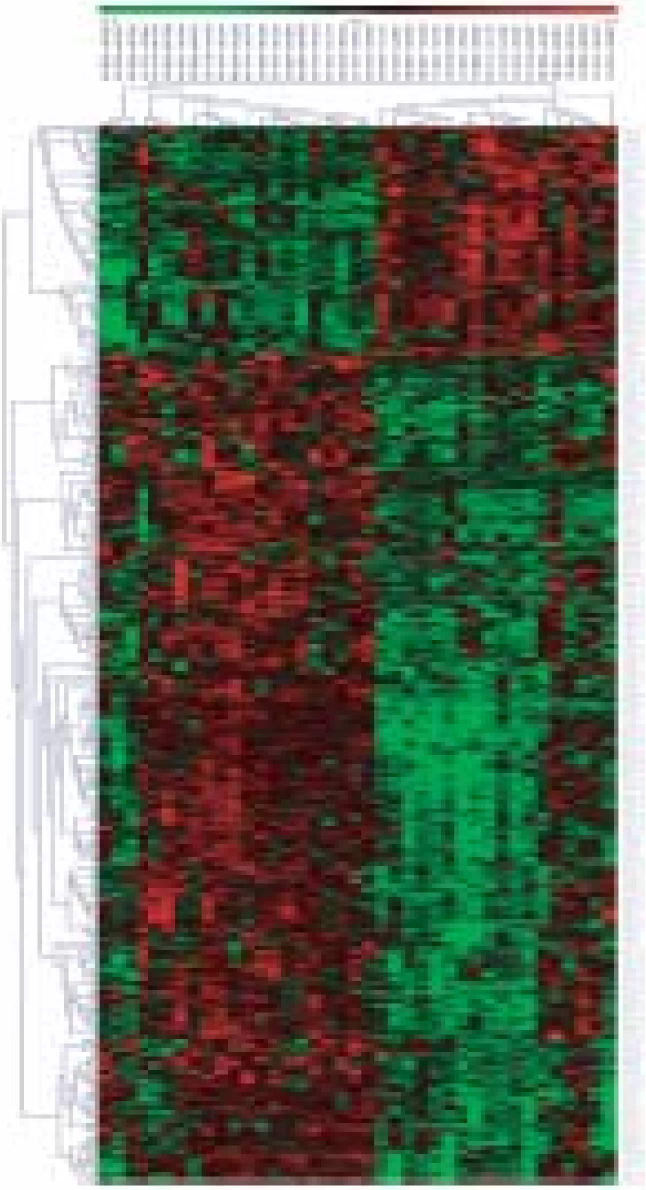
Supplementary Figure 2

Confocal microscopy of COS7 cells overexpressing *KRT23*. Cells were labeled with the polyclonal rabbit anti-K23 antibody, applied in a 1:600 dilution. K23 expression (green) was visualized applying a 1:2000 dilution of the secondary antibody AlexaFlour 488 goat anti-rabbit IgG highly cross-adsorbed. (Leica DMRS confocal microscope, 63x, 2x zoom; Z-stack, 0.5m interval)



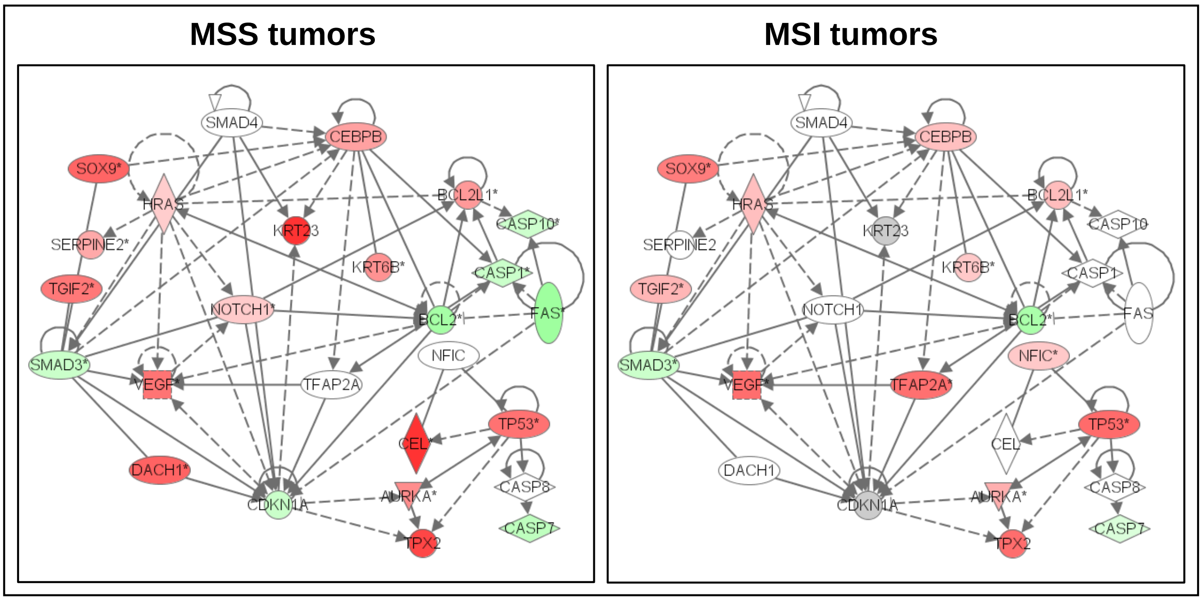
Supplementary Figure 3

Supervised Cluster analyses. Heatmap of the top 371 transcripts separating tumor with status “MSS-high Keratin23” and “MSI-low/no Keratin 23” as identified by immunohistochemistry followed by supervised clustering of transcript profiling data. Genes with a p-value < 0.05 in a SAM analysis and with a variance across all samples >0.5 were clustered using Cluster and Treeview (Eisen). Red are upregulated and green are downregulated genes.



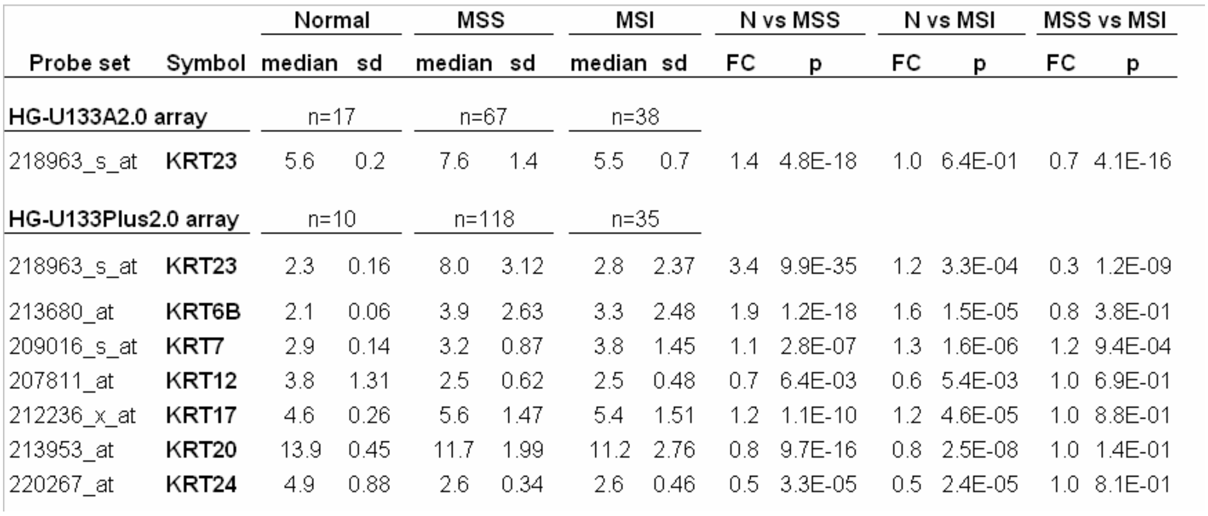
Supplementary Figure 4

Ingenuity pathway analyses (IPA5.0). Transcript profiling data (U133plus2.0) from 10 normal colon mucosas were compared to those from 118 MSS and 35 MSI colon adenocarcinomas as well as to colon cancer cell lines SW480, CaCo2 (MSS) and HCT15, HCT116 (MSI). *KRT23*, *CEL*, *DACH1*, *CEBPB*, *TGIF2*, *SERPINE/2 and NOTCH1* were upregulated in MSS tumors, while *TFAP2a* and *NFIC* were upregulated in MSI tumors. Expression patterns of *KRT23* and *CEBPB* (MSS) and *TFAP2a* and *NFIC* (MSI) are reproducible in the cell lines analyzed.



Supplementary Table 1

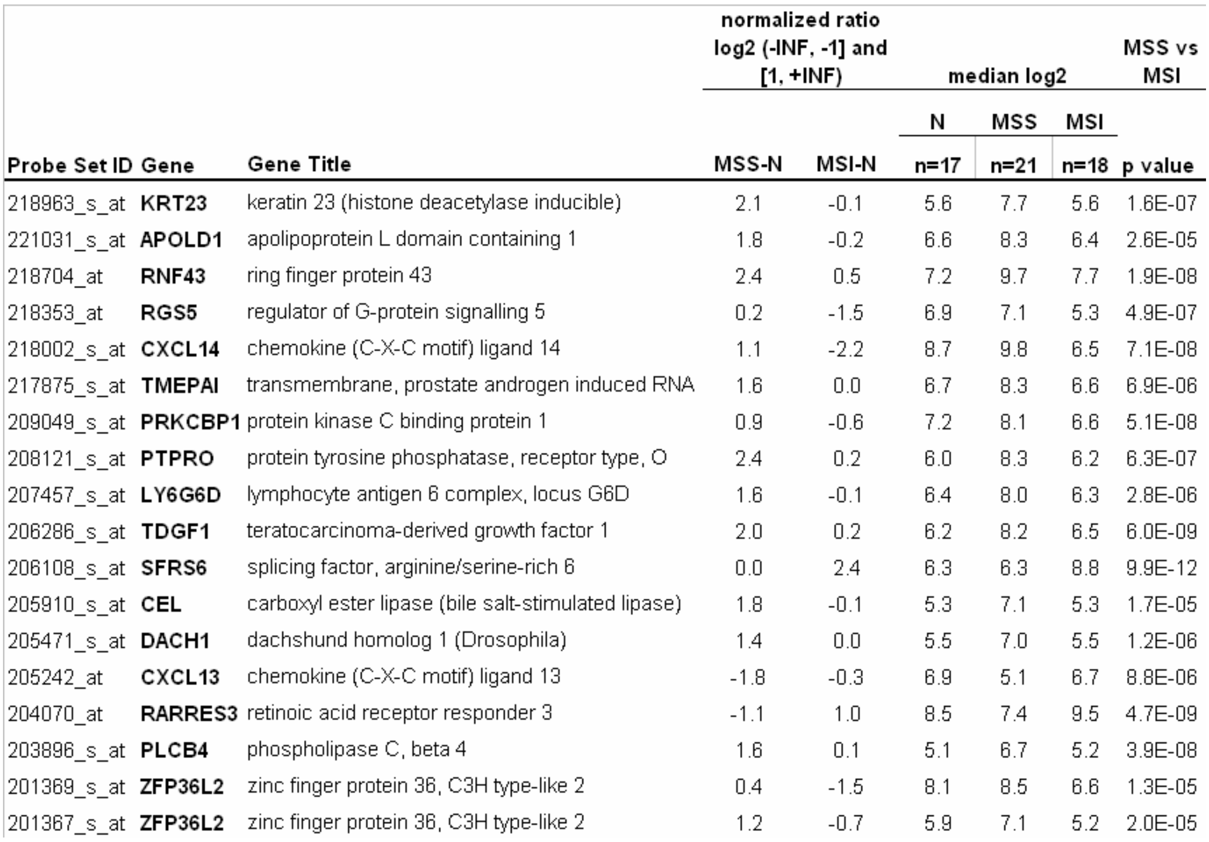
Transcript profiling of seven keratins. Expression profiling of 122 colon samples using HG-U133A2.0 arrays and of 163 colon samples using HG-U133plus2.0 arrays shows strong significant upregulation of *KRT23* in MSS tumors compared to low transcript levels in MSI tumors and normal colon mucosa. Genome wide profiling identified only six other keratins to be differentially expressed in colon adenocarcinomas compared to normal mucosas. Extremely high expression of a keratin restricted to MSS tumors is thus a unique feature for *KRT23*. Median transcript signal, SD and fold change (FC) are given as log2 values.



		Normal		MSS		MSI		N vs MSS		N vs MSI		MSS vs MSI	
Probe set	Symbol	median	sd	median	sd	median	sd	FC	p	FC	p	FC	p
HG-U133A2.0 array		n=17		n=67		n=38							
218963_s_at	**KRT23**	5.6	0.2	7.6	1.4	5.5	0.7	1.4	4.8E-18	1.0	6.4E-01	0.7	4.1E-16
HG-U133Plus2.0 array		n=10		n=118		n=35							
218963_s_at	**KRT23**	2.3	0.16	8.0	3.12	2.8	2.37	3.4	9.9E-35	1.2	3.3E-04	0.3	1.2E-09
213680_at	**KRT6B**	2.1	0.06	3.9	2.63	3.3	2.48	1.9	1.2E-18	1.6	1.5E-05	0.8	3.8E-01
209016_s_at	**KRT7**	2.9	0.14	3.2	0.87	3.8	1.45	1.1	2.8E-07	1.3	1.6E-06	1.2	9.4E-04
207811_at	**KRT12**	3.8	1.31	2.5	0.62	2.5	0.48	0.7	6.4E-03	0.6	5.4E-03	1.0	6.9E-01
212236_x_at	**KRT17**	4.6	0.26	5.6	1.47	5.4	1.51	1.2	1.1E-10	1.2	4.6E-05	1.0	8.8E-01
213953_at	**KRT20**	13.9	0.45	11.7	1.99	11.2	2.76	0.8	9.7E-16	0.8	2.5E-08	1.0	1.4E-01
220267_at	**KRT24**	4.9	0.88	2.6	0.34	2.6	0.46	0.5	3.3E-05	0.5	2.4E-05	1.0	8.1E-01

Supplementary Table 2

Top 17 genes (top 18 probe sets) differentially expressed between MSS tumors with high Keratin23 and MSI tumors with low Keratin 23 as identified by immunohistochemistry followed by supervised clustering of transcript profiling data. Ratios are given as normalized ratio log2 (-INF, -1] and [1, +INF).



			normalized ratio log2 (-INF, -1] and [1, +INF)		median log2			MSS vs MSI
					N	MSS	MSI	
Probe Set ID	**Gene**	**Gene Title**	**MSS-N**	**MSI-N**	**n=17**	**n=21**	**n=18**	**p value**
218963_s_at	**KRT23**	keratin 23 (histone deacetylase inducible)	2.1	-0.1	5.6	7.7	5.6	1.6E-07
221031_s_at	**APOLD1**	apolipoprotein L domain containing 1	1.8	-0.2	6.6	8.3	6.4	2.6E-05
218704_at	**RNF43**	ring finger protein 43	2.4	0.5	7.2	9.7	7.7	1.9E-08
218353_at	**RGS5**	regulator of G-protein signalling 5	0.2	-1.5	6.9	7.1	5.3	4.9E-07
218002_s_at	**CXCL14**	chemokine (C-X-C motif) ligand 14	1.1	-2.2	8.7	9.8	6.5	7.1E-08
217875_s_at	**TMEPAI**	transmembrane, prostate androgen induced RNA	1.6	0.0	6.7	8.3	6.6	6.9E-06
209049_s_at	**PRKCBP1**	protein kinase C binding protein 1	0.9	-0.6	7.2	8.1	6.6	5.1E-08
208121_s_at	**PTPRO**	protein tyrosine phosphatase, receptor type, O	2.4	0.2	6.0	8.3	6.2	6.3E-07
207457_s_at	**LY6G6D**	lymphocyte antigen 6 complex, locus G6D	1.6	-0.1	6.4	8.0	6.3	2.8E-06
206286_s_at	**TDGF1**	teratocarcinoma-derived growth factor 1	2.0	0.2	6.2	8.2	6.5	6.0E-09
206108_s_at	**SFRS6**	splicing factor, arginine/serine-rich 6	0.0	2.4	6.3	6.3	8.8	9.9E-12
205910_s_at	**CEL**	carboxyl ester lipase (bile salt-stimulated lipase)	1.8	-0.1	5.3	7.1	5.3	1.7E-05
205471_s_at	**DACH1**	dachshund homolog 1 (Drosophila)	1.4	0.0	5.5	7.0	5.5	1.2E-06
205242_at	**CXCL13**	chemokine (C-X-C motif) ligand 13	-1.8	-0.3	6.9	5.1	6.7	8.8E-06
204070_at	**RARRES3**	retinoic acid receptor responder 3	-1.1	1.0	8.5	7.4	9.5	4.7E-09
203896_s_at	**PLCB4**	phospholipase C, beta 4	1.6	0.1	5.1	6.7	5.2	3.9E-08
201369_s_at	**ZFP36L2**	zinc finger protein 36, C3H type-like 2	0.4	-1.5	8.1	8.5	6.6	1.3E-05
201367_s_at	**ZFP36L2**	zinc finger protein 36, C3H type-like 2	1.2	-0.7	5.9	7.1	5.2	2.0E-05

Supplementary Table 3

Transcript profiling of 16 TFs potentially targeting *KRT23* monitored on 163 tissue samples (U133plus arrays). Signal values are given as median log 2. Seven TFs with significant differences between MSS and MSI tumors with regard to the p value and the expression level are indicated with an asterisk. Three TFs, CEBPA, CEBPB and FOXQ1, shown in bold letters were regarded as highly relevant.

					median log2 signal			p value		
ID	Gene			Chr	N (n=10)	MSS (n=118)	MSI (n=35)	N vs MSS	N vs MSI	MSS vs MSI
204039_at	**CEBPA**	*	**CCAAT/enhancer binding protein (C/EBP), alpha**	**19q13.1**	**8.03**	**8.46**	**7.61**	**8.00E-02**	**1.90E-01**	**2.64E-05**
212501_at	**CEBPB**	*	**CCAAT/enhancer binding protein (C/EBP), beta**	**20q13.1**	**9.26**	**11.11**	**10.66**	**5.18E-06**	**6.86E-05**	**4.87E-05**
227475_at	**FOXQ1**	*	**forkhead box Q1**	**6p25**	**2.66**	**10.17**	**7.98**	**3.33E-33**	**2.20E-14**	**6.55E-06**
202221_s_	EP300	*	E1A binding protein p300	22q13.2	7.76	7.39	7.66	5.28E-01	6.10E-01	3.21E-02
210446_at	GATA1	*	GATA binding protein 1	Xp11.23	2.16	2.15	2.15	5.77E-01	1.56E-02	4.17E-02
209636_at	NFKB2	*	nuclear factor of kappa light polypeptide gene enhancer in B-cells 2 (p49/p100)	10q24	3.32	4.07	4.10	1.13E-04	6.71E-07	1.08E-02
40569_at	ZNF42/MZF1	*	Zinc finger protein 42 (myeloid-specific retinoic acid-responsive)	19q13.2-q13.4	6.89	7.07	7.68	8.85E-02	3.54E-03	5.50E-03
241612_at	FOXD3		forkhead box D3	1p32-p31	2.86	2.76	2.75	1.20E-01	8.21E-02	9.95E-02
209710_at	GATA2		GATA binding protein 2	3q21.3	5.29	4.86	4.88	1.58E-02	3.18E-02	7.34E-01
205038_at	IKZF1		IKAROS family zinc finger 1	7p13-p11.1	2.77	2.49	2.49	5.72E-02	3.13E-01	1.05E-01
209239_at	NFKB1		nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 (p105)	4q24	8.21	7.48	7.71	2.24E-07	6.53E-04	7.89E-02
206035_at	REL		v-rel reticuloendotheliosis viral oncogene homolog (avian)	2p13-p12	2.34	2.33	2.32	2.31E-01	1.07E-01	5.06E-01
1569638_at	SOX5		SRY (sex determining region Y)-box 5	12p12.1	2.35	2.34	2.34	3.96E-01	2.06E-01	3.52E-01
207893_at	SRY		sex determining region Y	Yp11.3	2.09	2.08	2.08	5.62E-02	3.49E-02	4.93E-01
205688_at	TFAP4		transcription factor AP-4 (activating enhancer binding protein 4)	16p13	4.26	4.50	4.37	6.44E-04	1.82E-02	6.21E-02
209338_at	TFCP2		transcription factor CP2	12q13	6.37	7.19	7.11	4.82E-05	6.16E-05	3.86E-01