

Results CSCP3

2015-10-15

European Bioinformatics Institute (EBI-EMBL)

Cambridge, UK

Christof Angermueller

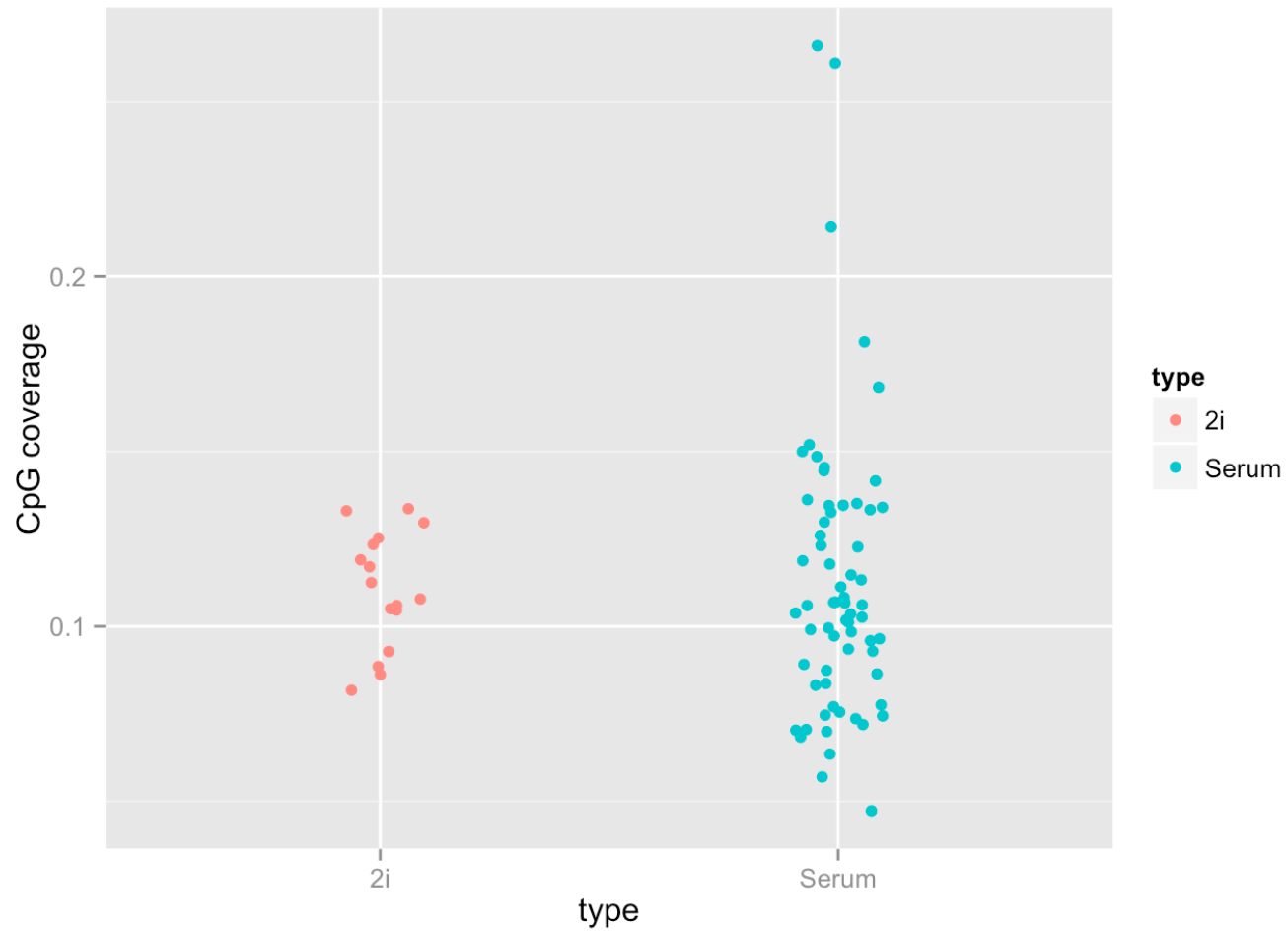
<https://cangermueller.com>

cangermueller@ebi.ac.uk

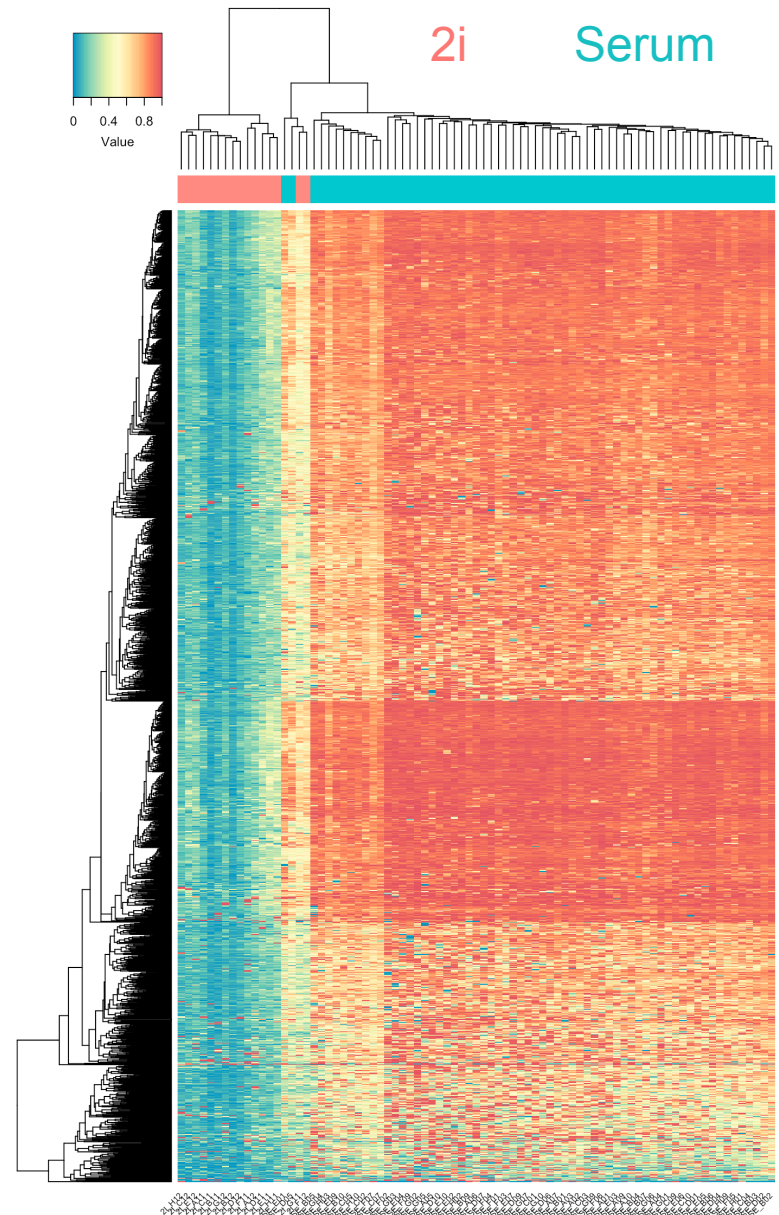
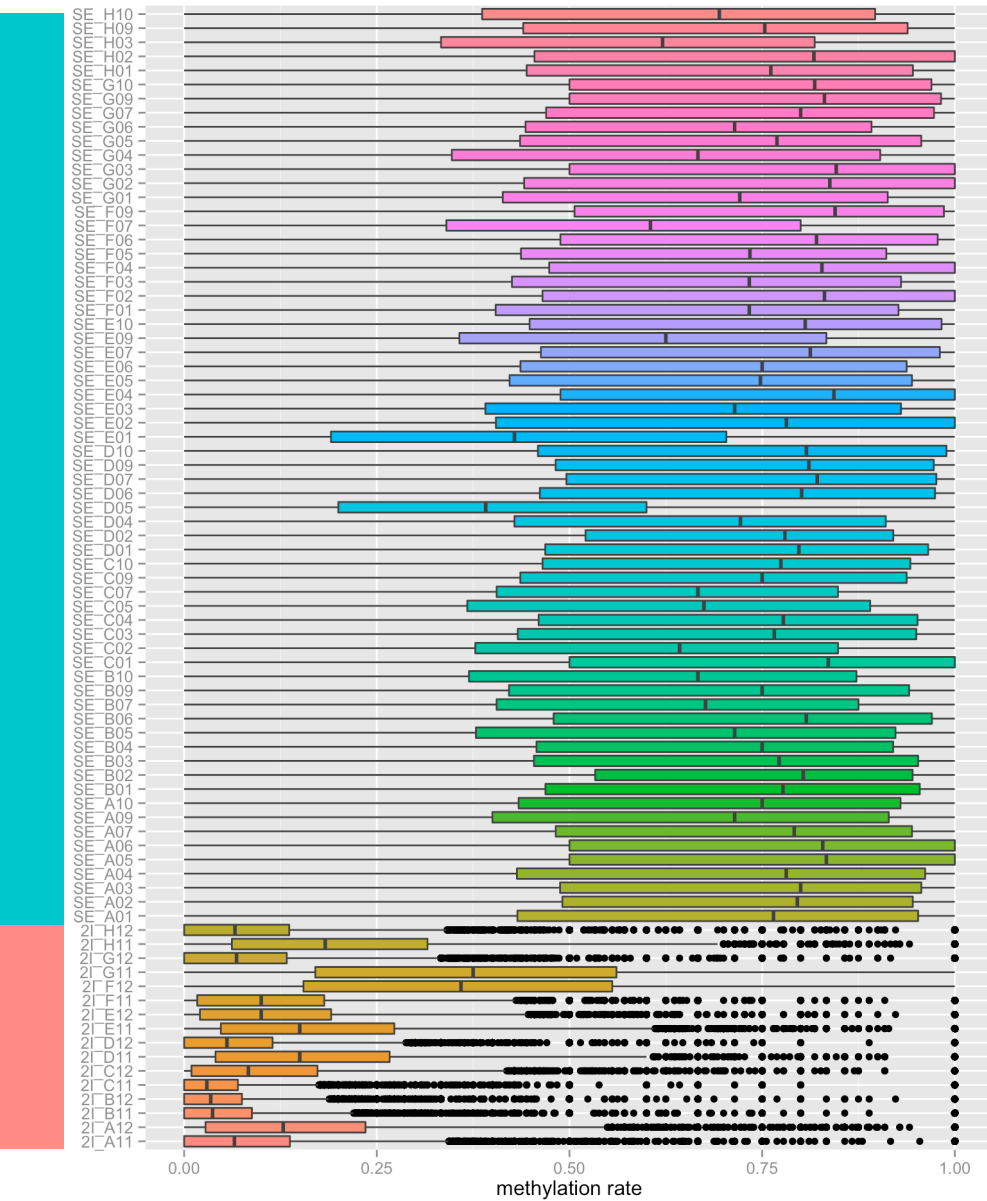
Quality Control

- ▶ 81 cell passed QC
- ▶ CpG coverage good (10% - 20%)
- ▶ 2i and serum cluster separately
- ▶ Two 2i cells seem to be “serum like”

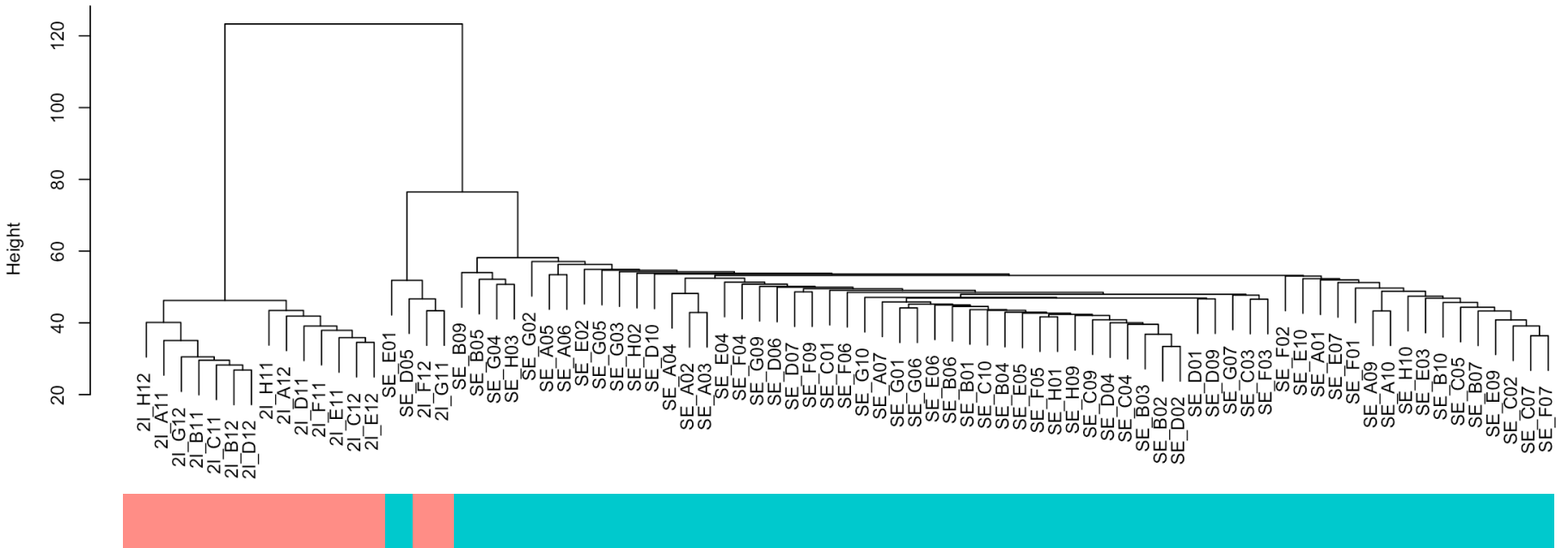
Genome-wide CpG coverage

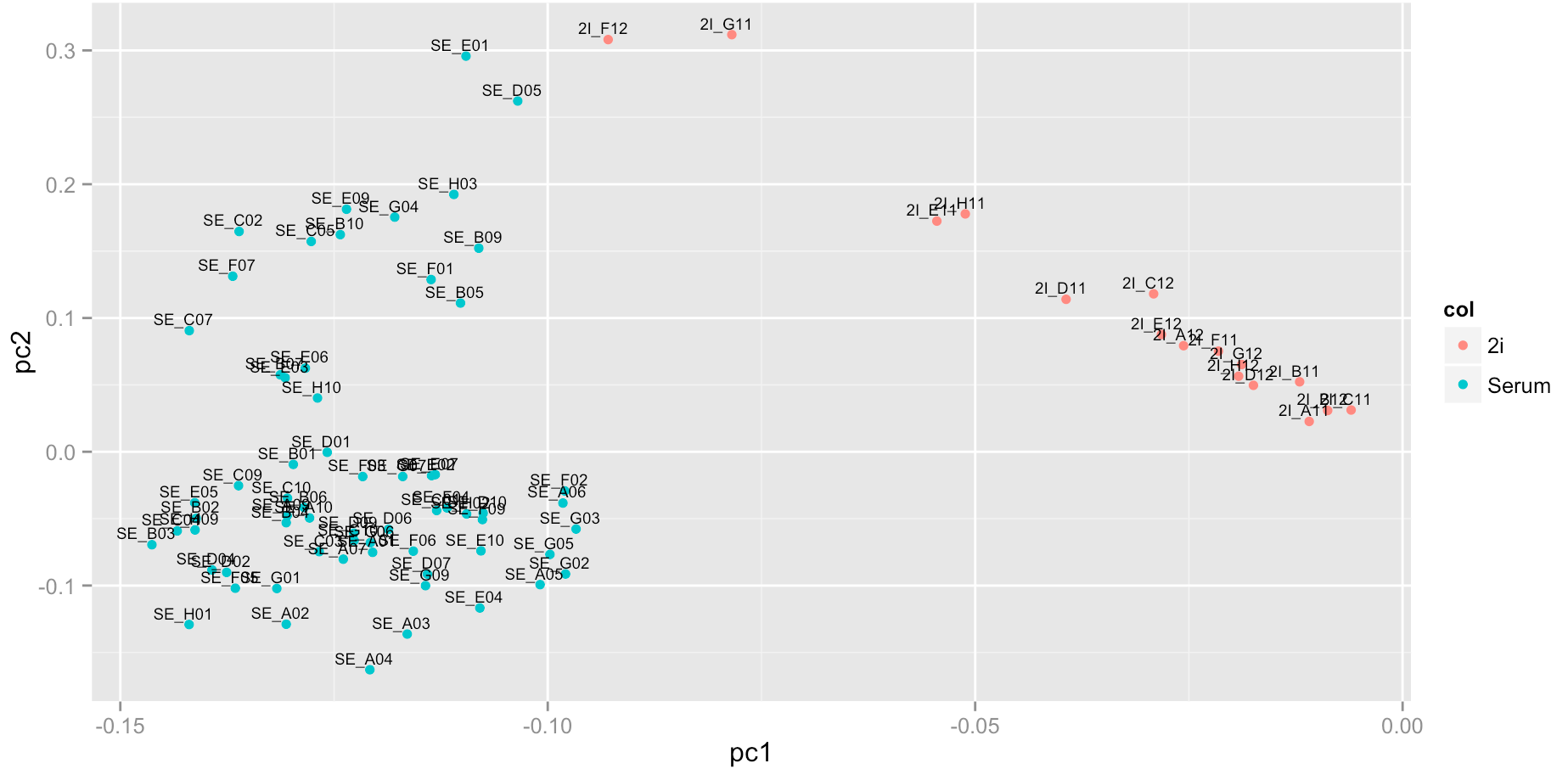


Serum and 2i cluster



Two serum-like 2i cells

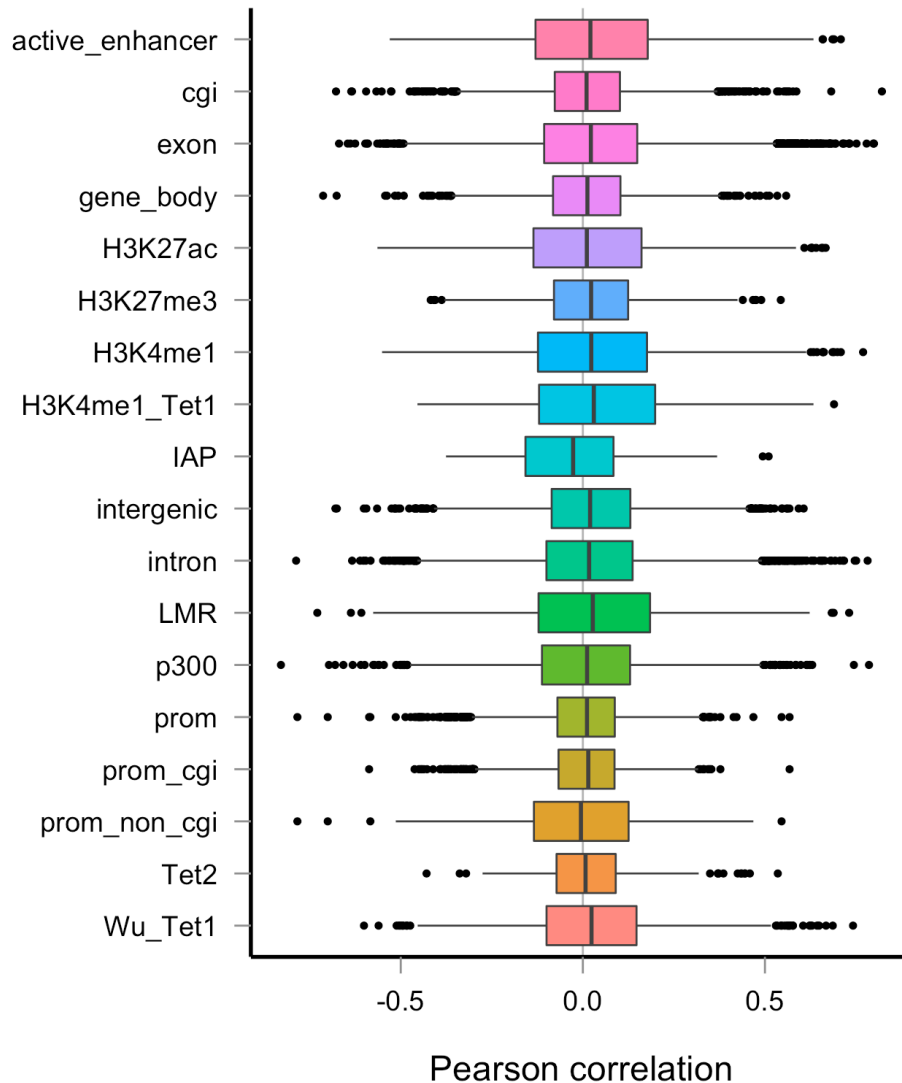




Gene-specific correlation

- ▶ 61 cells
 - ▶ 81 passed scBS-Seq QC
 - ▶ 61 passed scBS-Seq QC and RNA-Seq QC
- ▶ I can correlate considerably more sites than before
- ▶ I can decrease the significance threshold from 0.3 to 0.1
- ▶ I find considerably more significant genes
- ▶ Non-CGI promoter slightly anti-correlated
- ▶ More positive than negative LMR and p300 correlations
 - ▶ In contrast to before!
- ▶ Many positive H3K4me1 and H3K27ac correlations

Distribution correlation coefficients



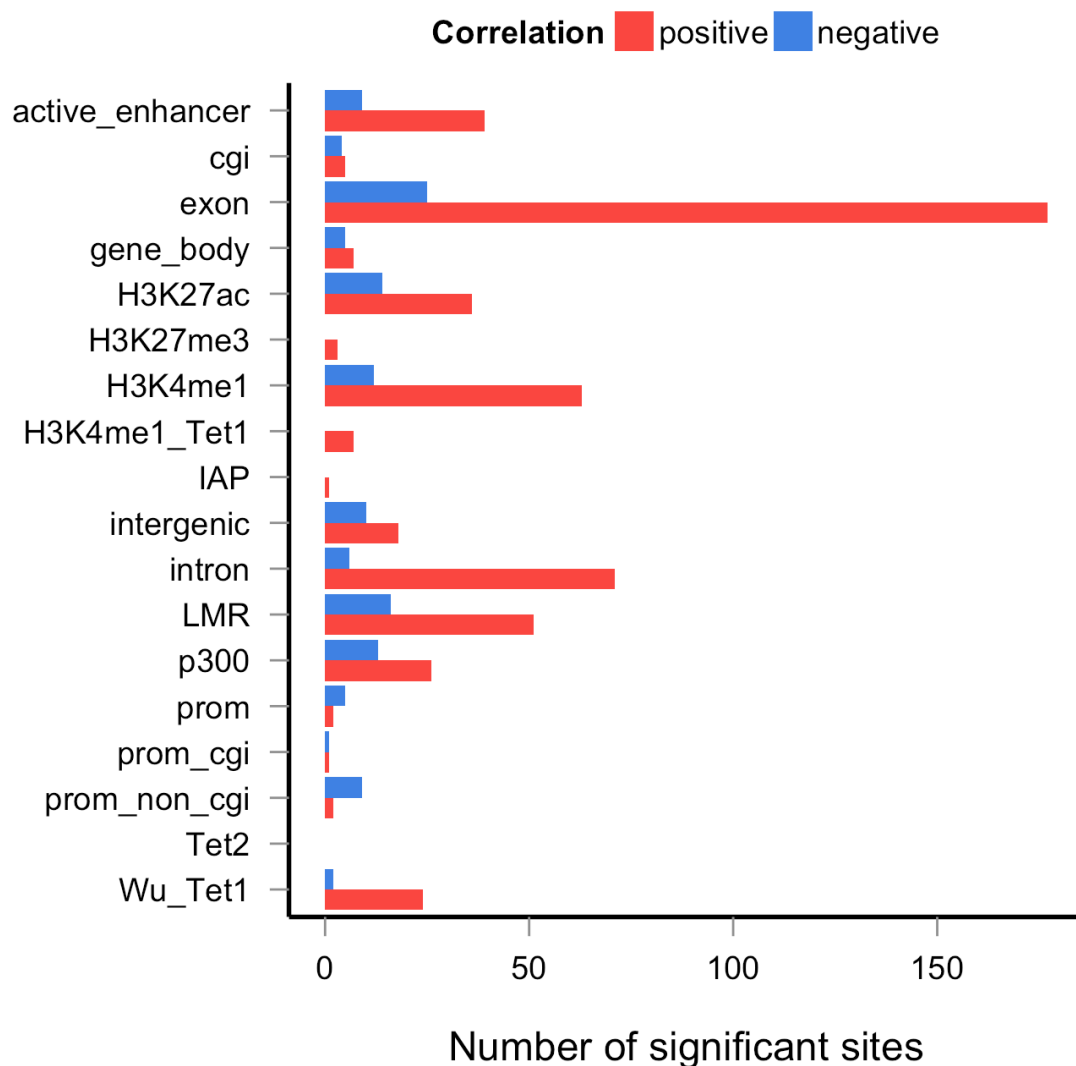
	name	r_mean	n	sig	sig_pos	sig_neg
1	active_enhancer	0.03	1397	48	39	9
2	cgi	0.01	3598	9	5	4
3	exon	0.03	13900	202	177	25
4	gene_body	0.01	4445	12	7	5
5	H3K27ac	0.02	1877	50	36	14
6	H3K27me3	0.03	591	3	3	0
7	H3K4me1	0.03	2317	75	63	12
8	H3K4me1_Tet1	0.04	315	7	7	0
9	IAP	-0.02	98	1	1	0
10	intergenic	0.02	6690	28	18	10
11	intron	0.02	23583	77	71	6
12	LMR	0.03	2255	67	51	16
13	p300	0.01	2906	39	26	13
14	prom	0.00	3883	7	2	5
15	prom_cgi	0.01	3325	2	1	1
16	prom_non_cgi	-0.02	263	11	2	9
17	Tet2	0.02	328	0	0	0
18	Wu_Tet1	0.03	3660	26	24	2

More sites

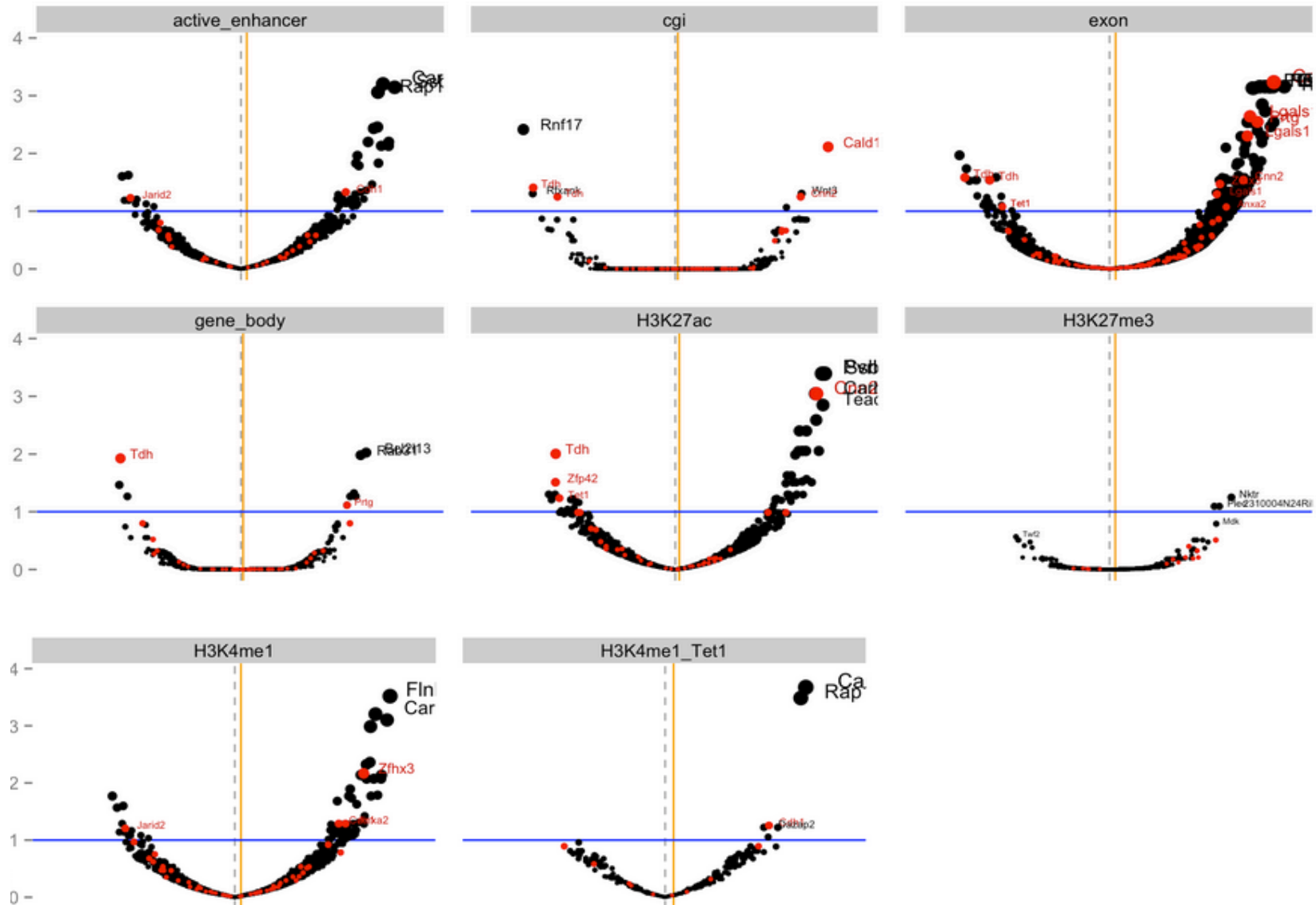
More significant

More significant correlations

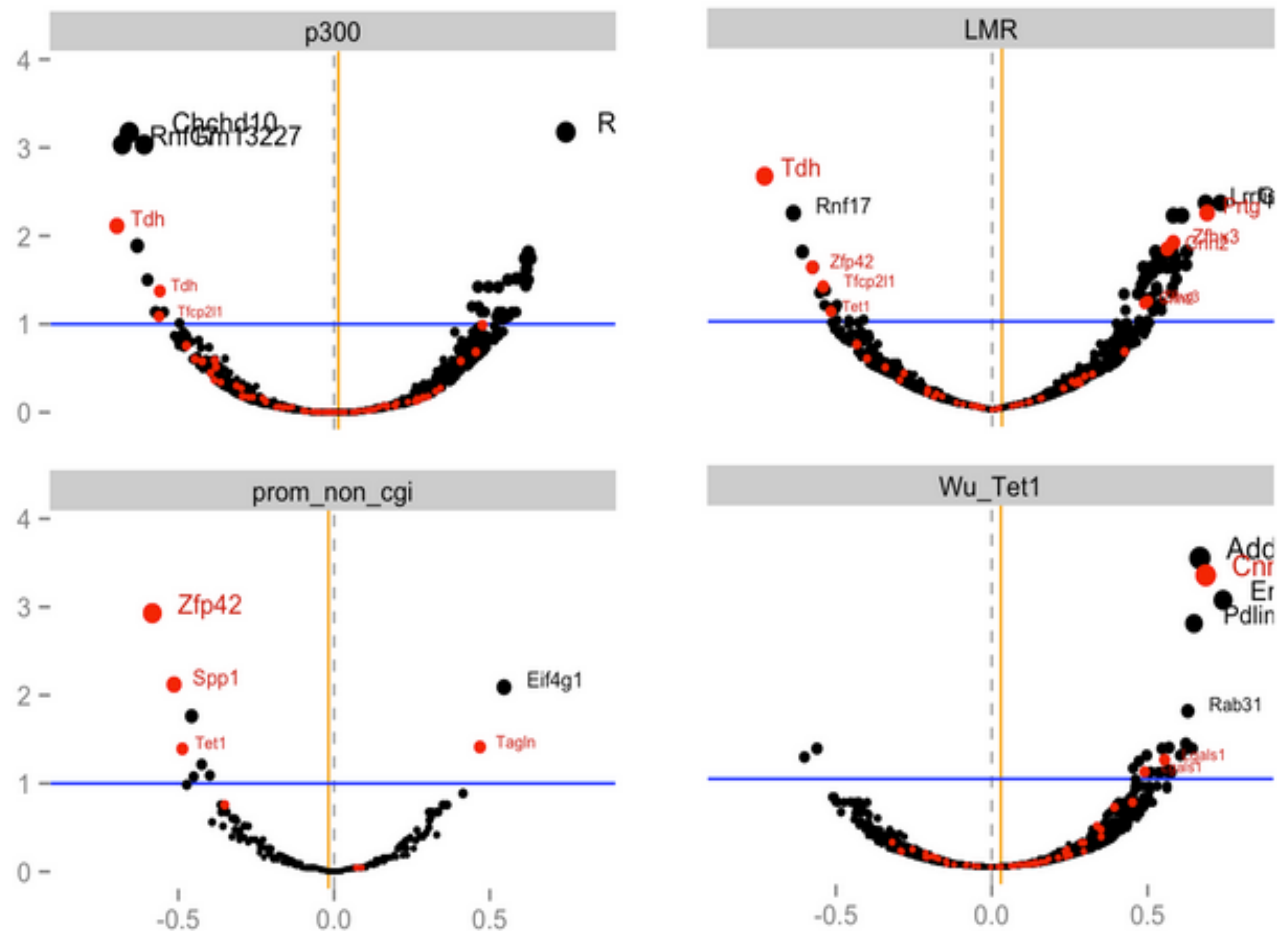
alpha = 0.1



Enrichment pluripotency genes (red)

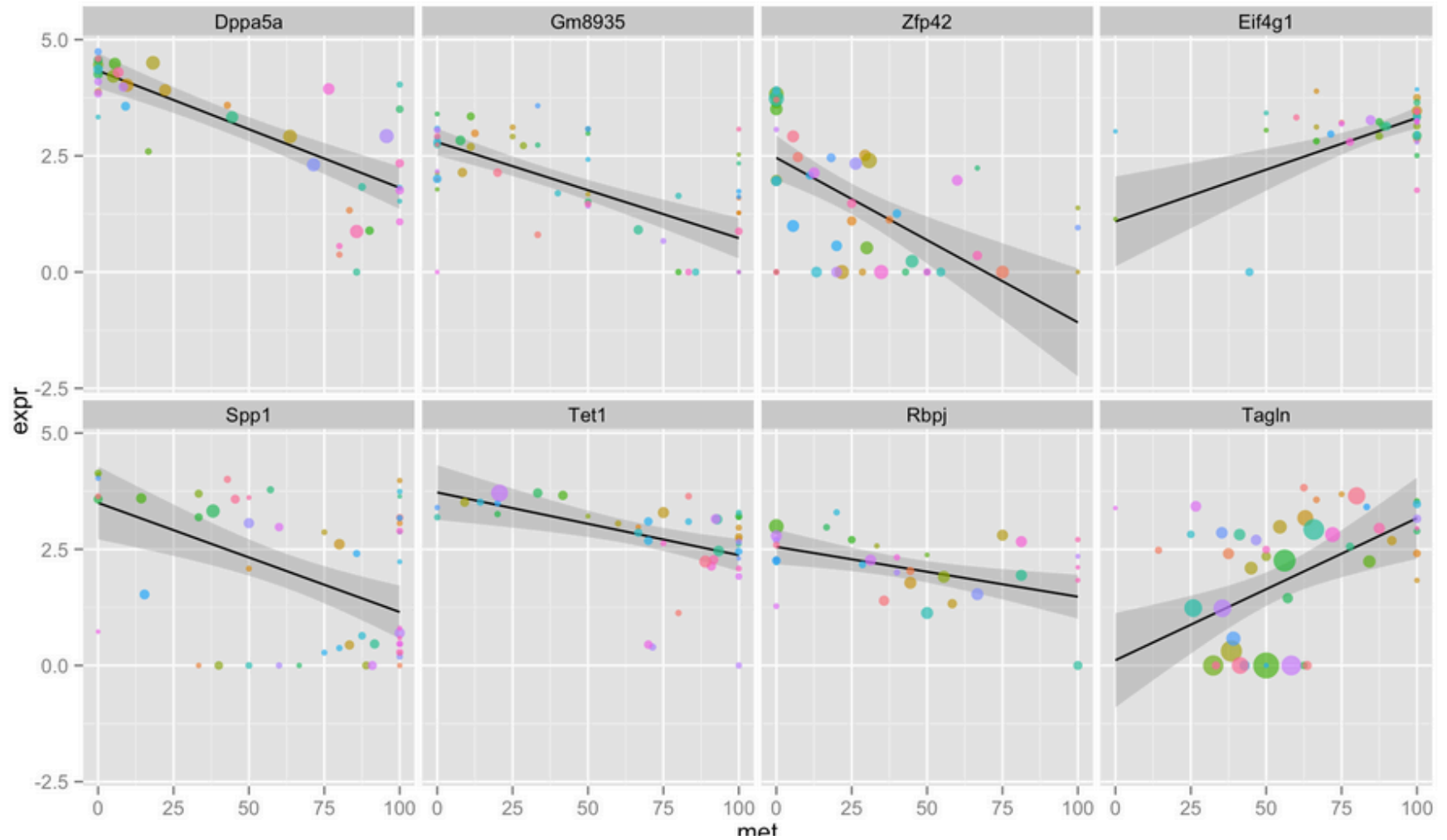


Enrichment pluripotency genes (red)



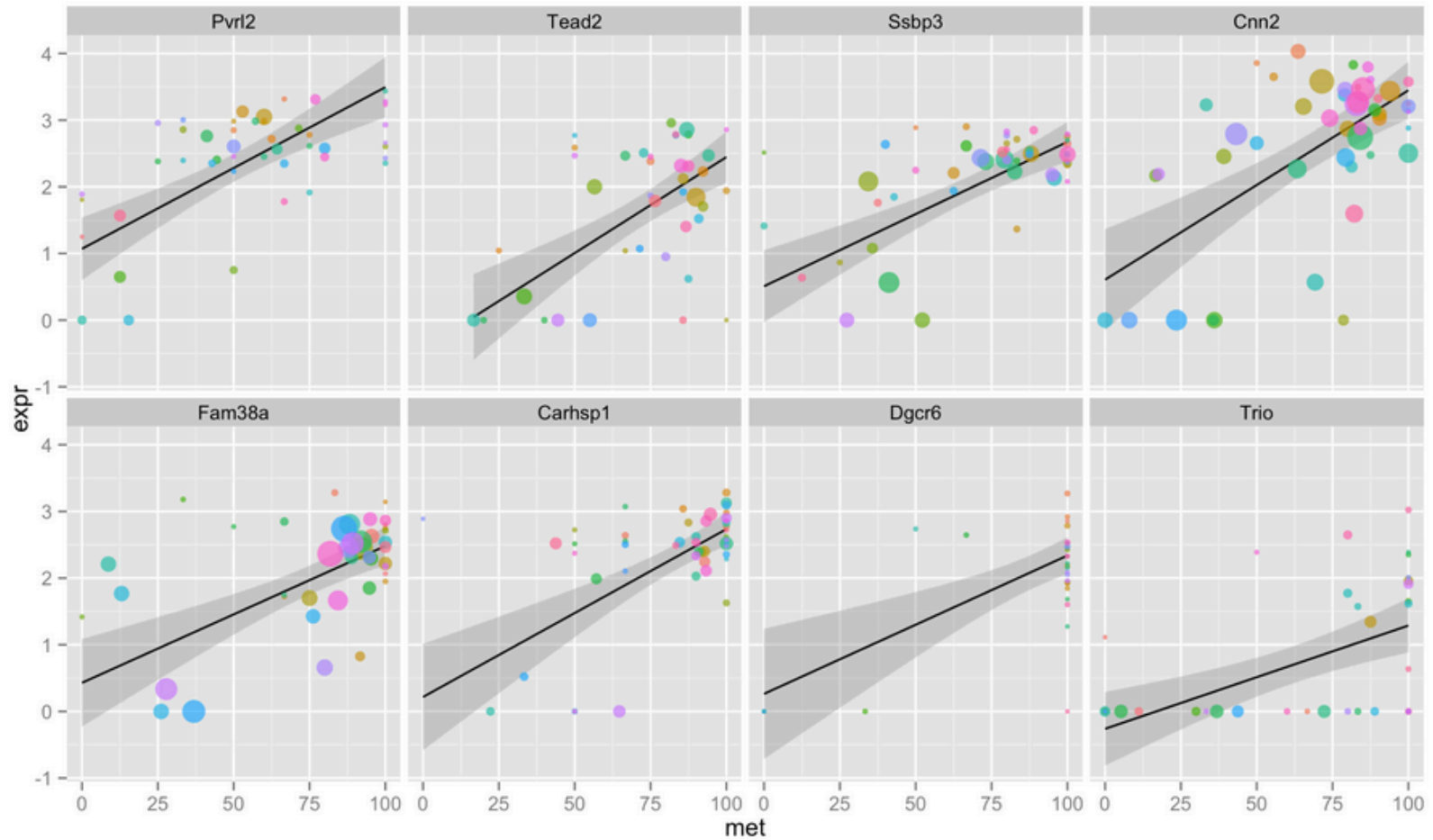
Negative correlations

Non-CGI promotor



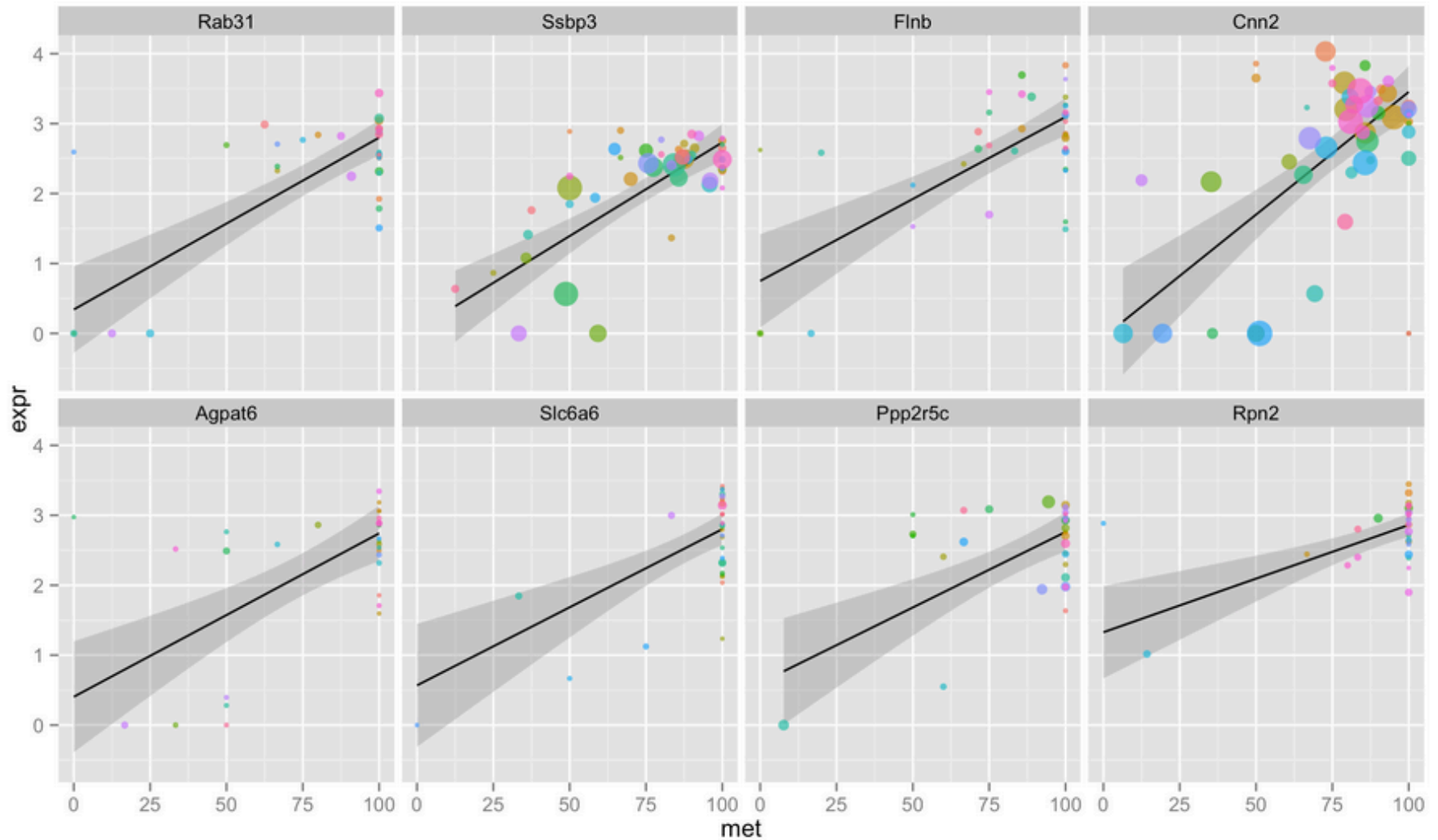
Positive correlations

H3K27ac



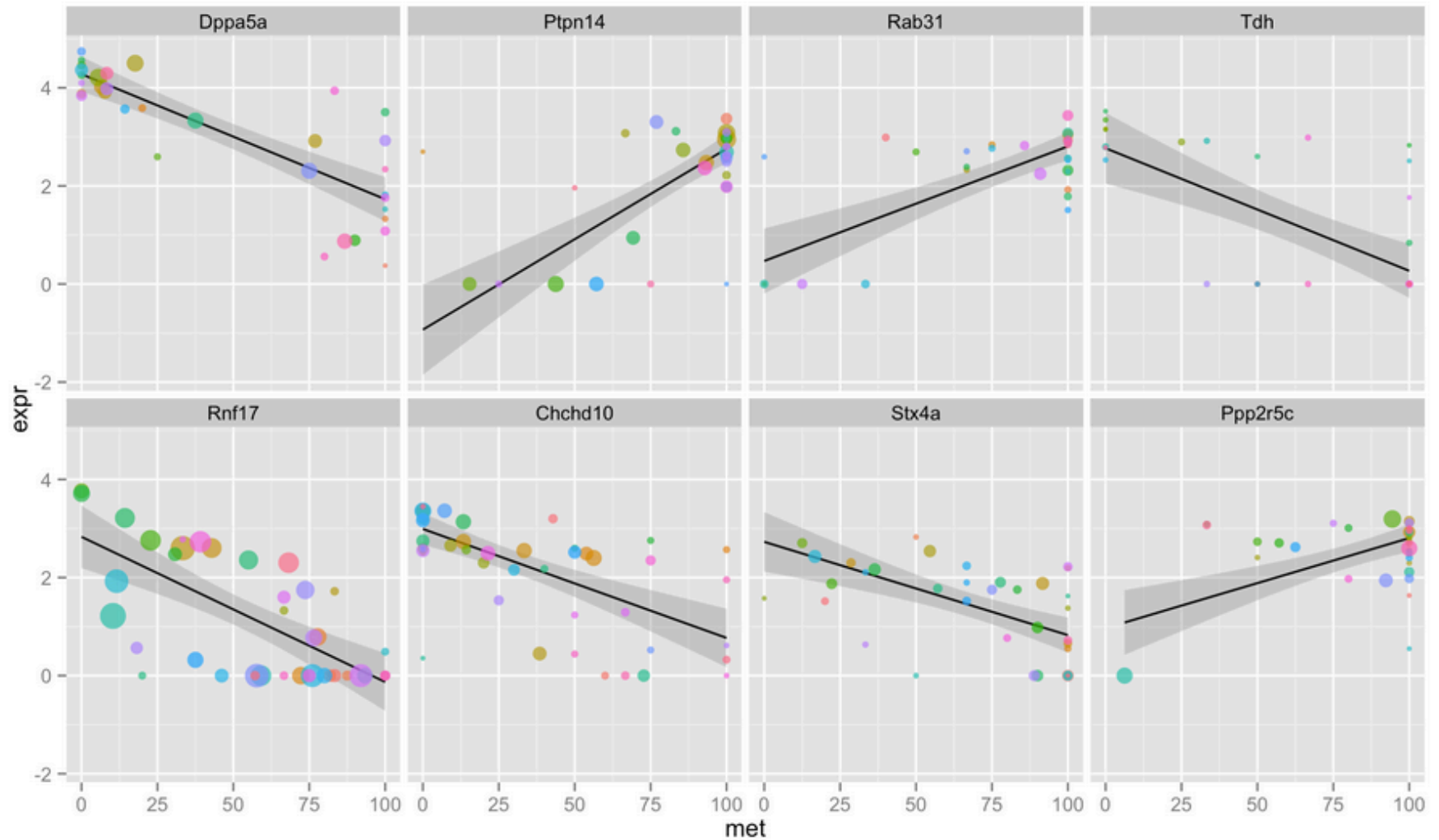
Positive correlations

H3K4me1



Positive correlations

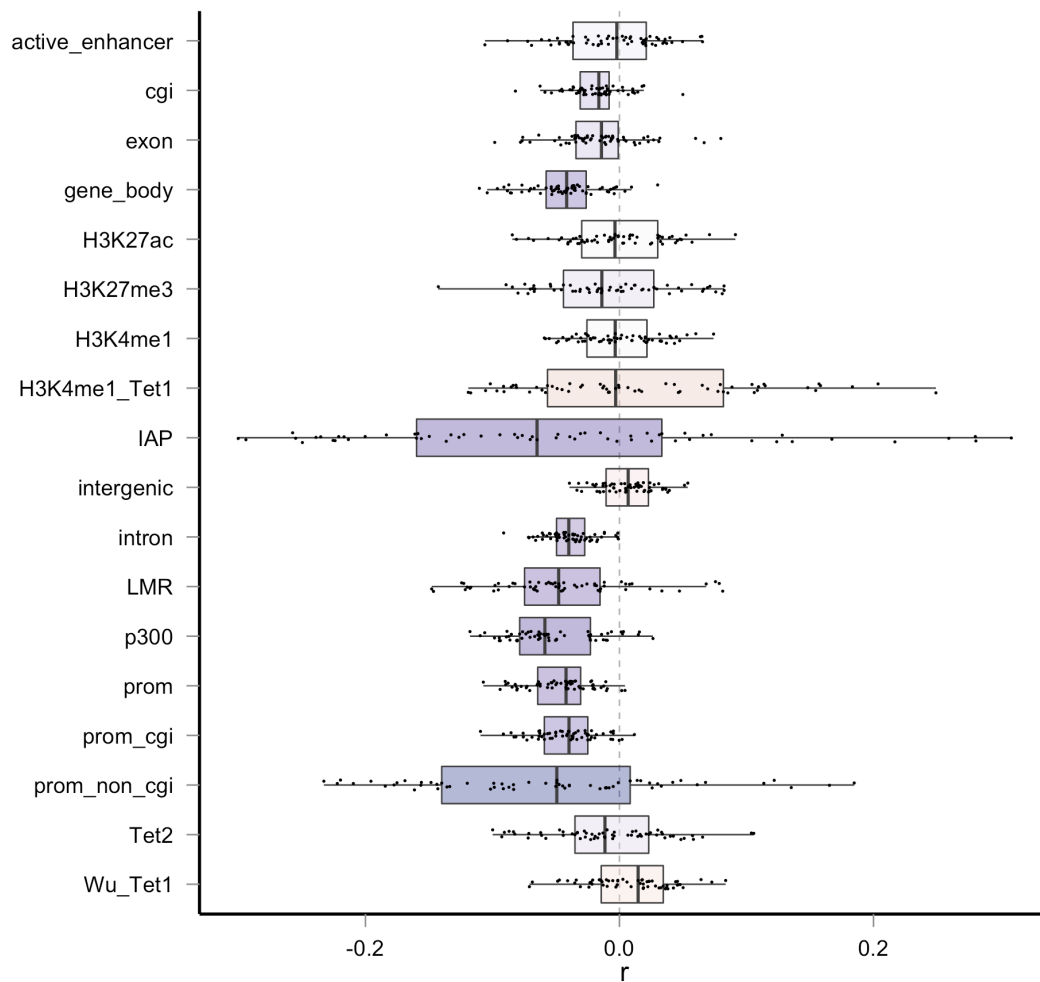
p300



Sample-specific correlation

- ▶ High variability between cells as before
- ▶ H3K4me1 and H3K27ac tend to be positively correlation
 - ▶ As in gene-specific correlation
- ▶ LMR, p300 tend to be negatively correlated
 - ▶ Unlike gene-specific correlation
- ▶ IAP negatively correlated
 - ▶ As in gene-specific correlation

Overview



	name	n	r	sig_pos	sig_neg
1	active_enhancer	932.11	-0.01	0	0
2	cgi	2533.84	-0.02	0	5
3	exon	8557.90	-0.02	0	38
4	gene_body	3981.21	-0.04	0	45
5	H3K27ac	1222.26	-0.00	0	1
6	H3K27me3	443.46	-0.01	0	0
7	H3K4me1	1503.21	-0.00	0	0
8	H3K4me1_Tet1	218.43	0.02	3	0
9	IAP	57.34	-0.05	0	0
10	intergenic	5504.02	0.01	9	0
11	intron	16049.92	-0.04	0	55
12	LMR	1355.82	-0.04	0	36
13	p300	1798.33	-0.05	0	40
14	prom	3143.70	-0.05	0	45
15	prom_cgi	2704.70	-0.04	0	31
16	prom_non_cgi	196.05	-0.06	0	9
17	Tet2	232.34	-0.01	0	0
18	Wu_Tet1	2352.30	0.01	7	0

Comparison with gene-specific correlation

