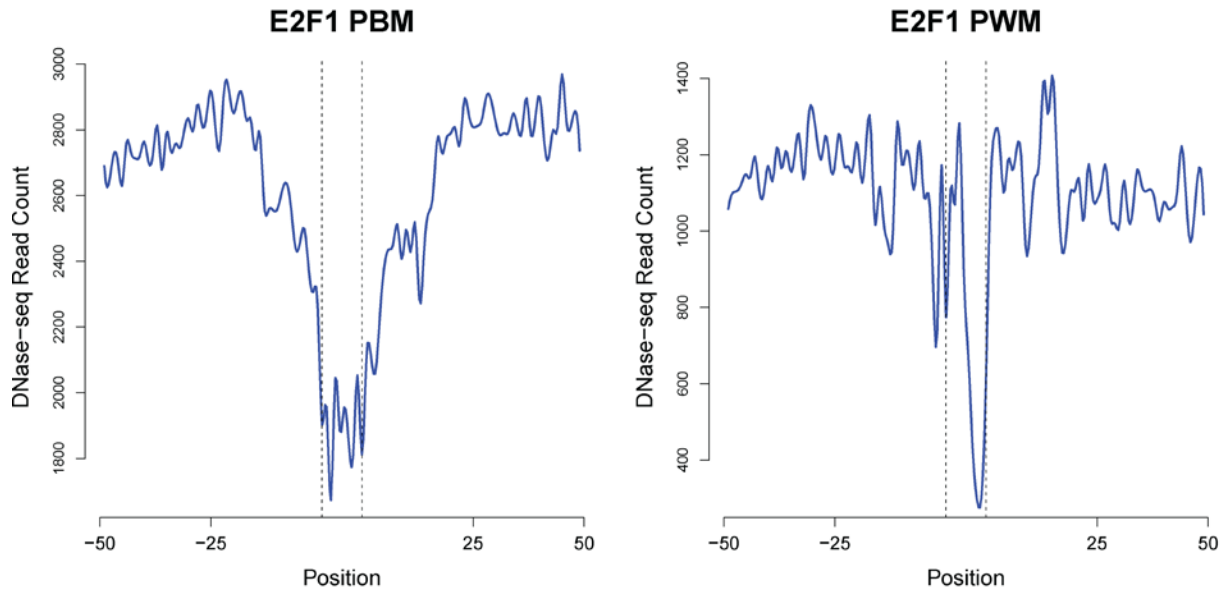Supplemental Figure 1. auROC (A) and sensitivity (B) values for D-s and FLR values in for classification of all sequence motifs detected across the genome
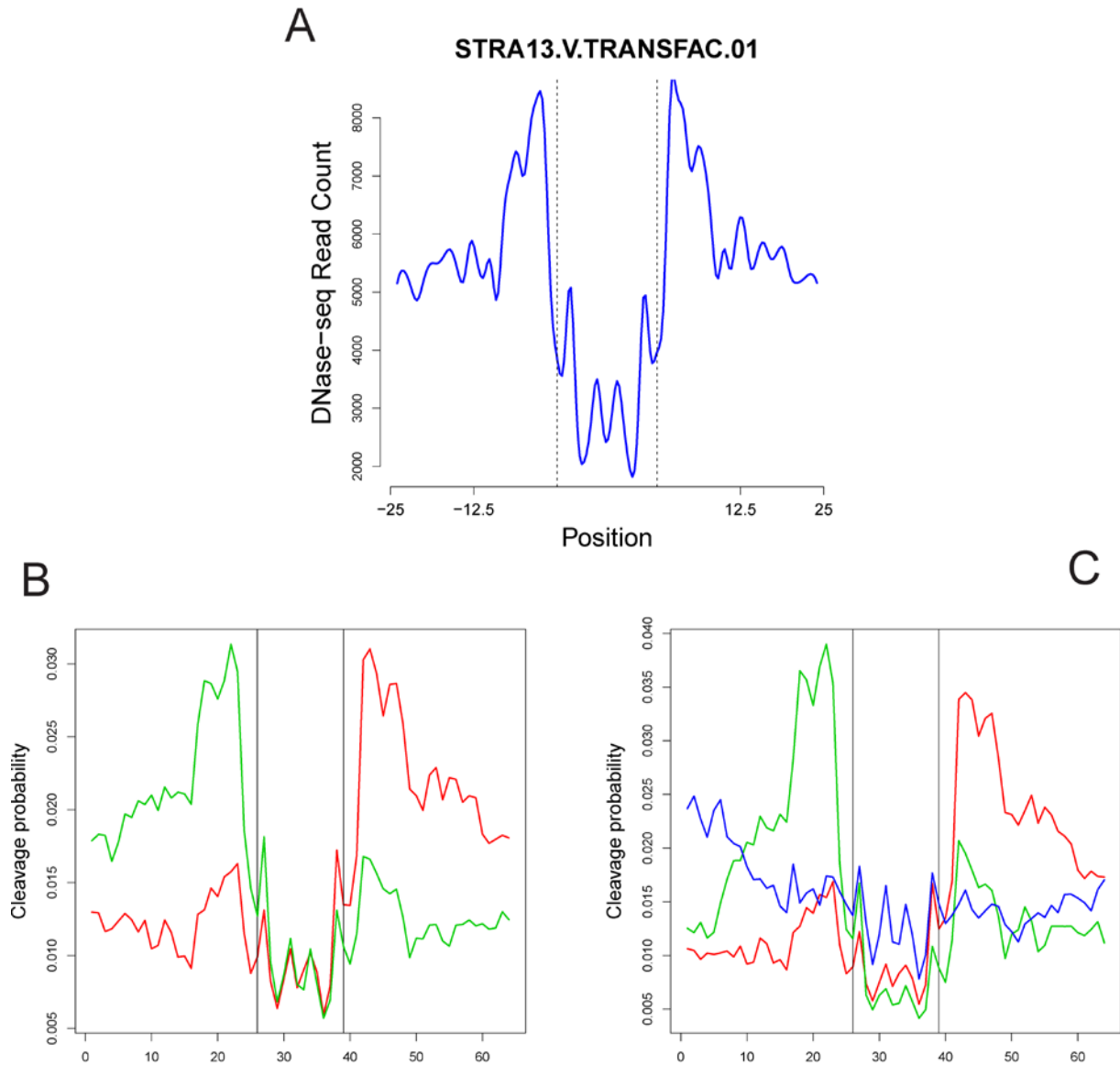
Supplemental Figure 2. Aggregate DNase-seq plots for E2F1 sequence motif matches in E2F1 ChIP-seq peaks from MCF7 cell line, obtained by either PBM scan (left) and PWM scan (right). Aggregate plot obtained using PBM scan show a substantially cleaner footprint shape compared to PWM scan, showing that PBM scan accurately pinpoints exact locations of TF-DNA interaction.
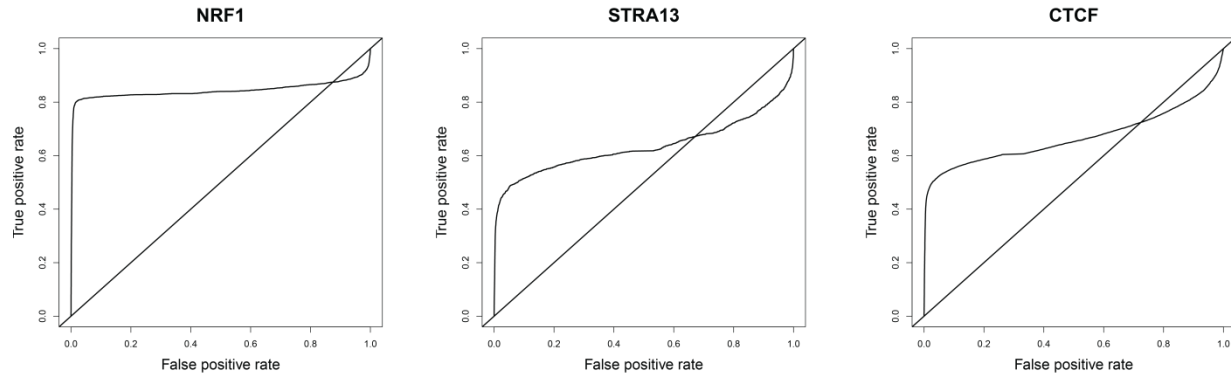
Supplemental Figure 3. Receiver operating characteristic (ROC) curves for 3 transcription factors using FLR as a predictor for all genome wide sequence motif matches. A rapid increase in true positive rate at high (>0.9) false positive rates is a common trend for many factors, indicating that a portion of sequence motifs that are in ChIP-seq peaks get low FLR scores.
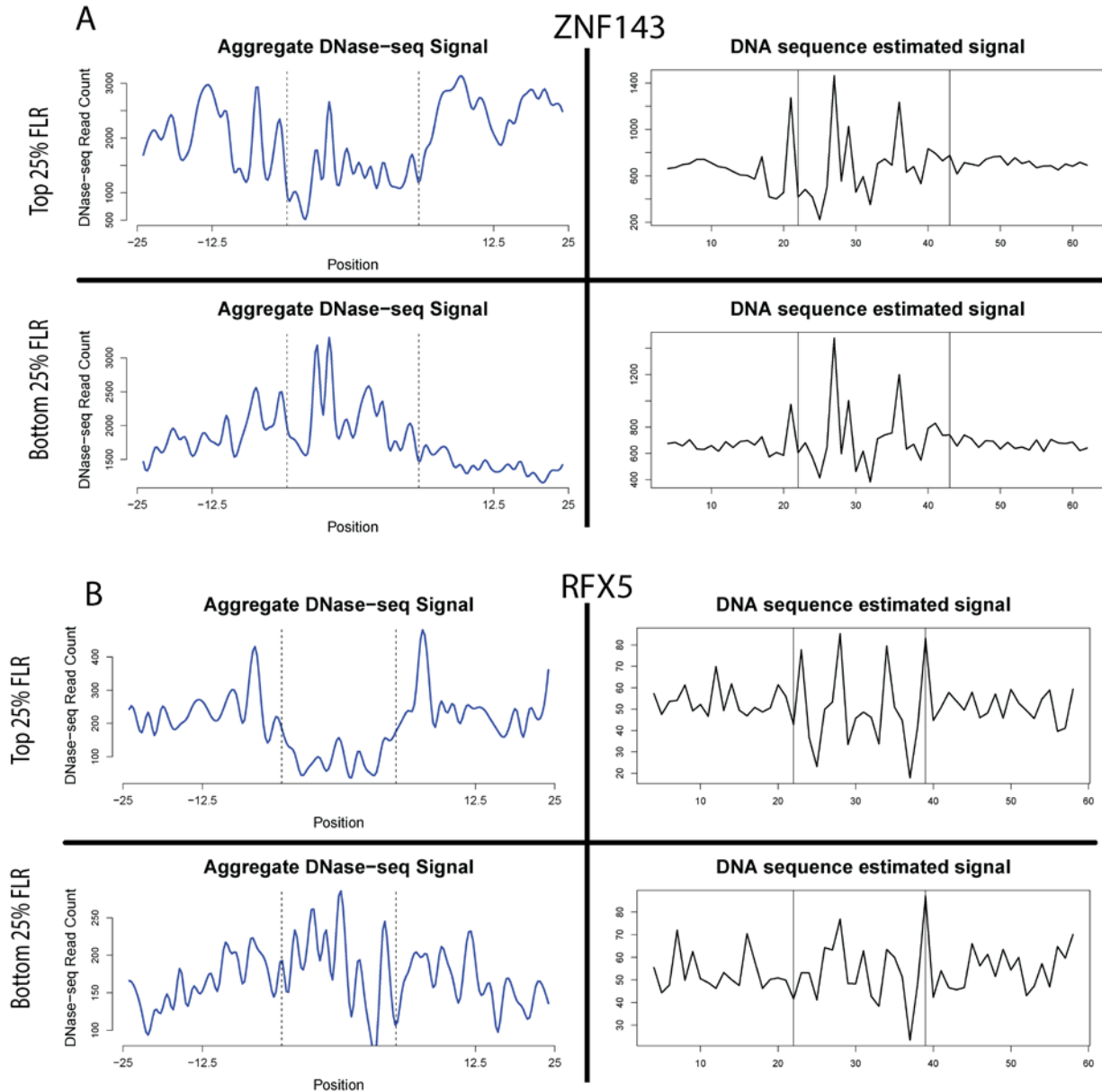
Supplemental Figure 4. A) Aggregate DNase-seq signal around top 25% and bottom 25% of sequence motifs for ZNF 143 - sorted by FLR score - are shown on left side. Top 25% profile has a footprint shape as expected, but no footprint shape in the lower 25% set. On right side we see the DNA sequence estimated signal for the same two quartiles, and neither show a footprint shape profile. This indicates that sequence bias is not a main contributor to footprint shape. B) Same plot for the RFX transcription factor.

Supplemental Figure 5. Aggregate DNase-seq signal around known binding sites (per ChIP-seq) of STRA13 (A). The shape of the footprint profile is obvious from this plot. Parameters learned from de novo EM training when number of components is set to two (B) and three (C). Two component mixture learns breaks down the main footprint profile into two symmetric two footprints models, shown in red and green. Adding a third component to the mixture learns the background signal reliably (shown in blue).

Transcription Factor Name	PWM ID	Lowest PWM score in top 50K	10^{-4} P-value PWM threshold
E2F4	M00738	7.97	7.87367
NFYA	MA0060.2	8.36	7.08848
NFYB	MA0060.2	8.36	7.08848
NRF1	M00652	3.9	1.58485
RFX1	M00280	8.63	7.21743
ZNF143	M00262	7.96	5.56315
BHLHE40	M00985	7.98	3.99022
USF1	MA0093.2	10.35	9.24034
USF2	M00726	9.35	7.65913
YY1	M00793	8.74	7.14483
CTCF	MA0139.1	8.09	5.86003
MEF2A	MA0052.2	9.29	4.81889
REST/NRSF	MA0138.2	5.89	3.83553
Pax5	MA0014.2	8.29	6.95939
PU1	MA0080.3	8.69	6.50746
SP1	MA0079.3	9.17	7.64199
SRF	MA0083.2	7.25	4.88501
TCF3	MA0091.1	9.35	8.35423
ZEB1	MA0103.2	7.91	6.68559
C-Myc	MA0059.1	9.39	8.28472
MAX	MA0058.2	10.22	6.76693

Supplemental Table 1. PWM score thresholds that correspond to top 50000 highest scoring candidate binding sites (column 2) and 10^{-4} P-value (column 3). For all PWMs, we note that the threshold that yields top 50000 candidate binding sites is higher than 10^{-4} P-value threshold.

Factor Name	Matrix ID
E2F4	M00738
NFYA	MA0060.2
NFYB	MA0060.2
NRF1	M00652
RFX5	M00280
ZNF143	M00262
BHLHE40	M00985
USF1	MA0093.2
USF2	M00726
YY1	M00793
CTCF	MA0139.1
MEF2A	MA0052.2
REST/NRSF	MA0138.2
Pax5	MA0014.2
PU1	MA0080.3
SP1	MA0079.3
SRF	MA0083.2
TCF3	MA0091.1
ZEB1	MA0103.2
C-Myc	MA0059.1
MAX	MA0058.2
OCT4::Sox2	MA0142.1
KLF4	MA0039.2

Supplemental Table 2. ENCODE antibody names used for ChIP experiments transcription factors and the IDs of PWMs used to look for sequence motif matches in these peaks. TRANSFAC and JASPAR databases were used to find corresponding PWMs.

E2F4, Direct Associations	P-value
Irf3	0.000283
Cfos	0.000314
NFYA, Direct Associations	P-value
Irf3	8.61E-06
NFYB, Direct Associations	P-value
Cfos	3.79E-41
Nfya	1.80E-39
Irf3	3.26E-28
Rfx5	1.10E-16
Chd2	1.76E-14
Pbx3	3.19E-13
E2f4	3.87E-11
Sp1	2.83E-10
Gabp	2.03E-08
Sin3	1.72E-07
NFYB, Indirect Associations	P-value
Batf	4.38E-08
Ikzf1	1.34E-05
P300	2.33E-05
Bcl11a	6.50E-05
NRF1, Direct Associations	P-value
Pol2	6.21E-05
Atf3	0.000368
NRF1, Indirect Associations	P-value
Ikzf1	2.42E-06
P300	2.01E-05
Irf4	0.000136
Ebf1	0.000185
ZNF143, Direct Associations	P-value
Ets1	1.08E-11
Six5	4.76E-09
Chd2	2.25E-07
Gabp	7.11E-07
E2f4	1.18E-05
Zbtb33	0.000134
Atf3	0.000182
Bclaf1	0.000225

Sin3	0.000355
Sp1	0.000481
ZNF143, Indirect Associations	P-value
Rad21	4.44E-12
Smc3	1.05E-10
Ctcf	1.27E-06
P300	7.30E-05
BHLHE40, Direct Associations	P-value
Nfe2	8.23E-37
Atf3	6.39E-34
Cfos	8.87E-32
Taf1	4.87E-18
Pol2	8.83E-17
Nrf1	1.16E-15
Nfyb	2.89E-15
Six5	9.18E-15
Egr1	3.55E-13
Gabp	4.74E-13
BHLHE40, Indirect Associations	P-value
Ebf1	4.33E-06
P300	8.82E-05
Batf	0.000113
Ikzf1	0.000164
YY1, Direct Associations	P-value
Six5	5.19E-25
Znf143	1.59E-16
Gabp	4.73E-16
Sp1	9.53E-15
Elk1	3.82E-14
Ets1	6.26E-14
Tbp	1.88E-13
Srf	6.03E-13
Rfx5	2.84E-12
Cfos	2.44E-11
CTCF, Direct Associations	P-value
Yy1	6.94E-171
Elf1	2.55E-139
Maz	4.50E-123
Foxm1	9.05E-119
Znf143	7.07E-117
Pax5	1.08E-115
Runx3	1.34E-82
Pou2f2	1.88E-82

Whip	4.14E-64
Irf4	3.67E-62
CTCF, Indirect Associations	P-value
P300	4.14E-10
Nfic	1.69E-09
Bcl11a	8.99E-09
Pol2	1.45E-07
Taf1	1.52E-07
Ikzf1	1.97E-06
Mta3	5.47E-06
Batf	2.25E-05
Tbp	0.000113
MEF2A, Direct Associations	P-value
Cfos	1.43E-05
Nfya	0.000241
NRSF/REST, Direct Associations	P-value
Znf143	5.45E-71
Corest	1.43E-12
Sin3	0.000137
Yy1	0.00024
Mxi1	0.000262
NRSF/REST, Indirect Associations	P-value
Pol2	1.41E-20
Taf1	2.89E-15
Tcf3	3.66E-15
Pml	2.99E-13
Tcf12	4.29E-13
Elf1	1.03E-12
Runx3	1.29E-12
Pou2f2	1.47E-12
Sp1	4.83E-12
Mxi1	8.30E-12
PU1, Direct Associations	P-value
Six5	2.03E-11
Gabp	7.91E-11
Sp1	4.24E-10
Pml	7.21E-10
Srf	7.99E-10
Elk1	1.32E-08
Taf1	1.37E-08
Tbp	1.60E-08
Egr1	2.59E-08
Ets1	4.01E-08

SP1, Direct Associations	P-value
Nfya	6.83E-07
Nrf1	6.85E-07
Gabp	5.17E-06
E2f4	3.24E-05
Cfos	4.79E-05
Elk1	5.21E-05
Elf1	8.33E-05
Cmyc	9.49E-05
Irf3	0.000174
Pol2	0.000332
SP1, Indirect Associations	P-value
Ikzf1	0.000132
Bcl11a	0.000463
SRF, Direct Associations	P-value
Pol2	3.28E-05
Egr1	5.95E-05
Elk1	0.000264
Taf1	0.000319
E2f4	0.000497
MAX, Direct Associations	P-value
Atf3	1.15E-14
Cfos	1.19E-13
Nfe2	5.88E-13
Nrf1	2.03E-08
Six5	9.75E-08
Rfx5	3.87E-06
Pol2	5.24E-06
Gabp	6.45E-06
Elf1	3.29E-05
Egr1	4.75E-05
MAX, Indirect Associations	P-value
P300	0.000332

Supplemental Table 3. Top 10 direct and indirect associations for each transcription factor we analyzed by footprint analysis. Note that in some cases, there were less than ten associations. For some transcription factors such as C-myc, we did not detect any significant associations.

Factor, Cell Types Compared	FLR(P-value)	D-s(P-value)
NRSF, GM12878 vs. Medulo	1.18e-10	0.4
OCT4::SOX2, iPS vs. Skin	2.8e-12	2.5e-10
KLF4, iPS vs. Skin	1.05e-7	0.009
C-myc, iPS vs. Skin	0.0015	0.096

Supplemental Table 4. P-values for Kolmogorov-Smirnoff test (one-sided) for comparing the distributions of FLR and D-s metrics assigned to sequence motifs from two different cell types. The expression of transcription factors are induced or abolished in one cell type vs. the other, FLR metric is consistently more sensitive to these changes (more significant differences as measured by smaller p-values).

			AUC				Sensitivity at 1% FPR			
			Whole Genome		DHS		Whole Genome		DHS	
TF	Neg	Pos	D-s	FLR	D-s	FLR	D-s	FLR	D-s	FLR
E2F4	49532	468	0.99	0.86	0.89	0.81	0.72	0.62	0.14	0.13
NFY-A	48565	1435	0.99	0.94	0.88	0.87	0.63	0.72	0.15	0.13
NFY-B	46397	3603	0.98	0.85	0.9	0.82	0.77	0.66	0.11	0.18
NRF1	46282	3718	0.99	0.86	0.93	0.87	0.8	0.8	0.16	0.28
RFX5	49811	189	0.98	0.88	0.9	0.85	0.77	0.75	0.23	0.2
ZNF143	49130	870	0.99	0.77	0.93	0.78	0.9	0.65	0.3	0.43
BHLHE40	47651	2349	0.97	0.71	0.85	0.7	0.64	0.43	0.08	0.1
USF1	49585	415	0.97	0.85	0.77	0.77	0.59	0.44	0.11	0.11
USF2	49724	276	0.96	0.93	0.79	0.85	0.59	0.53	0.05	0.14
YY1	47905	2095	0.99	0.91	0.94	0.88	0.85	0.69	0.41	0.35
CTCF	29570	20430	0.95	0.65	0.76	0.78	0.38	0.43	0	0.1
MEF2A	48823	1177	0.97	0.79	0.85	0.75	0.56	0.44	0.06	0.14
NRSF	47566	2434	0.81	0.69	0.64	0.82	0.11	0.34	0.01	0.46
PAX5	49401	599	0.94	0.87	0.8	0.8	0.46	0.45	0.03	0.07
PU1	48731	1269	0.94	0.9	0.81	0.84	0.41	0.48	0.08	0.09
SP1	49678	322	1	0.96	0.91	0.92	0.89	0.86	0.33	0.25
SRF	49134	866	0.94	0.64	0.83	0.65	0.53	0.29	0.09	0.11
TCF3	49749	251	0.99	0.76	0.82	0.65	0.67	0.39	0.06	0.02
ZEB1	49960	40	1	0.99	0.84	0.89	0.92	0.78	0.05	0.11
C-MYC	49668	332	0.99	0.98	0.85	0.89	0.69	0.74	0.06	0.05
MAX	49317	683	0.99	0.77	0.88	0.75	0.79	0.55	0.17	0.1

Supplemental Table 5. auROC and Sensitivity at 1% FPR values for D-s and FLR for each transcription factor using *de novo* background model