

ENCODE DCC Antibody Validation Document

Date of Submission

Name:

Email:

Lab

Antibody Name:

Target:

Company/
Source:

Catalog Number, database ID, laboratory

Lot Number

Antibody Description:

Target Description:

Species Target

Species Host

Validation Method #1

Validation Method #2

Purification Method

Polyclonal/
Monoclonal

Vendor URL:

Reference (PI/
Publication
Information)

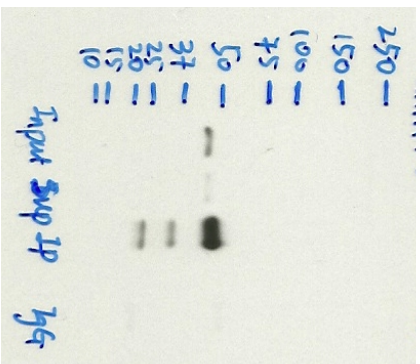
Please complete the following for antibodies to histone modifications:
if your specifications are not listed in the drop-down box,
please write-in the appropriate information

Histone Name AA modified AA Position Modification

Immunoprecipitation of K562 nuclear extract with sc-266 antibody efficiently enriches a protein of molecular weight of GATA1 (46KD).

Validation #1
Analysis

Insert Validation Image (click here)



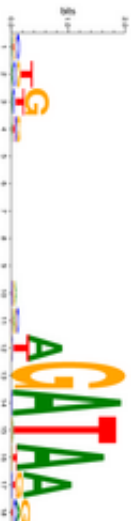
Western Blot analysis of immunoprecipitated proteins from K562 nuclear extracts detected using anti GATA1 (sc266) antibody from Santa Cruz Biotechnology. Expected protein band 47kD.

Input: Input lysate, Sup: Supernatant from IP, IP: Immunoprecipitated fraction from 07-729, IgG: Control Immunoprecipitation using IgG.

Fig. 1 GATA1 motif enrichment

Cell Line	Motif enrichment (log 2)	Fraction of Peaks containing motif
K562	4.6723	0.180272109

Fig. 2 Motif enrichment sequence and position weight matrix for the highly enriched GATA1 motif



```
>GATA_known14 GATA_14#known
Tail::Gata1_jaspar_MA0140.1 +
B 0.135962 0.451734 0.210061 0.202243
K 0.088715 0.150918 0.176751 0.583616
K 0.044497 0.061821 0.652514 0.241168
B 0.089983 0.253311 0.416978 0.239728
N 0.227165 0.215620 0.263158 0.294058
N 0.238031 0.233277 0.306621 0.222071
N 0.225390 0.252885 0.318398 0.203327
N 0.241941 0.237190 0.302341 0.218527
N 0.248898 0.224483 0.292302 0.234317
D 0.338086 0.138153 0.319077 0.204684
V 0.178947 0.397963 0.305942 0.117148
W 0.664744 0.010519 0.012555 0.312182
G 0.001018 0.002376 0.996606 0.000000
A 0.993553 0.002375 0.000679 0.003393
T 0.005433 0.008829 0.010187 0.975551
A 0.941216 0.002039 0.007815 0.048930
A 0.809378 0.017329 0.127761 0.045532
V 0.186054 0.187075 0.544558 0.082313
```

Calculations were done by Alan Boyle using a collection of known motifs. Table 1 shows the fold-enrichments and fraction of peaks which contain the motif. The motif which produced the largest value for each criterion is shown in Table 1. Note that while the maximally enriched motifs may differ from the motif with the highest enrichment b-values and the most represented motif, the motifs are highly similar (Figure 4) and thus all values are similar between motifs. Motifs were identified using a matching stringency corresponding to 4-6 (6mer) Peaks identified by DNase (1% cutoff) were used in the analysis and +/- 5000 from peak centers were considered. Enrichments are for a given motif vs. a background consisting of +/- 5000 from the centers of all DNase hypersensitive peaks. Repeat mask/simple repeats from UCSC and all genome v7 exons (including non-protein coding genes) were excluded from the analysis. Comparison to shuffle motifs were used to correct for compositional bias. Enrichment is the corrected # of motifs in ChIP peaks/corrected # of motifs in DNase1 peaks. The current ENCODE standard calls for >4-fold enrichment and >10% motif representation for this criteria to be used for validation. The K562 c-Fos dataset presented here exceed these thresholds and sc-266 is considered validated.

Validation #2
Analysis

Insert Validation Image (Click here)