



UNIVERSITÀ
DEGLI STUDI
DI MILANO

ELK1 ChIP-Seq data analysis

Giulia Di Virgilio

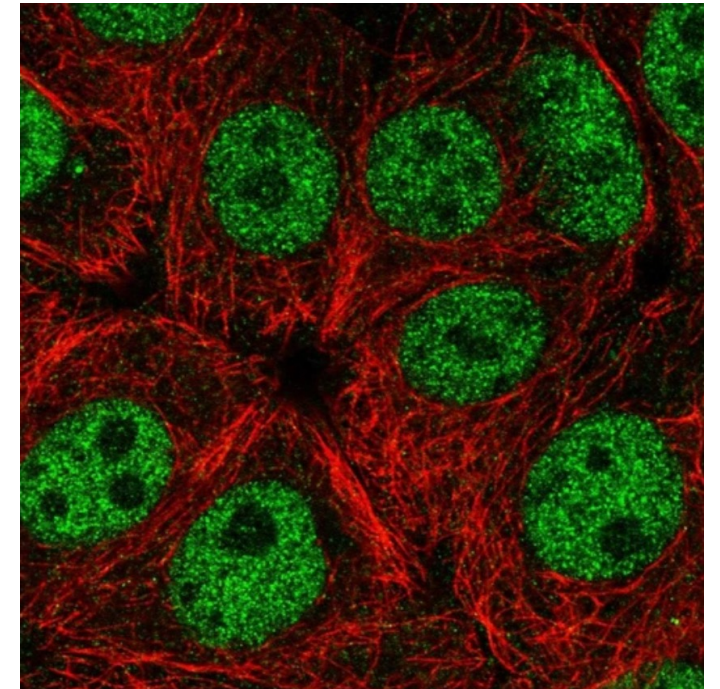
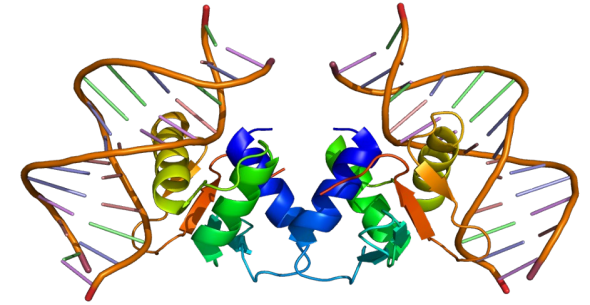
ELK1: some background info...

ETS Like-1 protein **Elk-1** is a protein encoded by the **ELK1** gene. It is a member of the **ETS family**.

What is ELK1's role?

ELK1 has several biological roles, let's take a look at them:

- **Gene expression regulation:** key regulator of immediate early genes (binding to SRF)¹
- **Cellular processes:** essential for controlling cell proliferation, apoptosis and development
- **Dual function:** can act as a repressor and an activator²
- **Signal transduction:** downstream target of signaling cascades

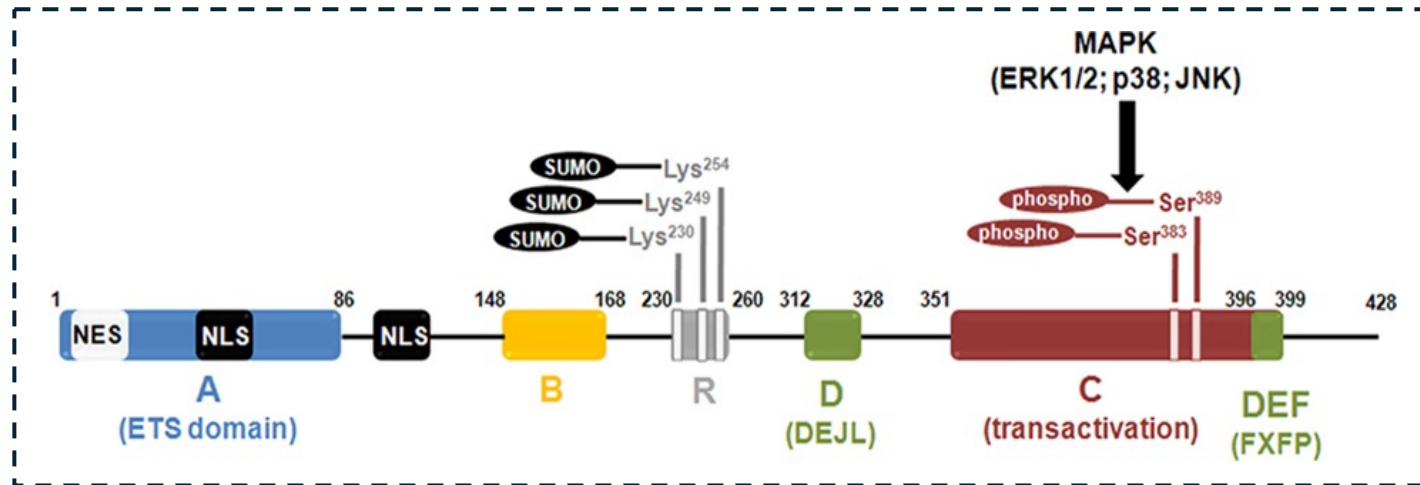


ELK1's localization in the nucleoplasm with immunofluorescence


ELK1: some background info...

What about its DNA binding domain?

The **ETS domain (A)** is required for the binding to DNA: it is localized in the **N-terminal region**



... and its binding motif?

 The picture can't be displayed.

Binding motif image taken from JASPAR dataset

Do they reach the quality control thresholds?

Mapping reads and quality control

Replicate number	Total reads	Mapped reads	Unique reads	Multi-mapping reads	Non-mapping reads
Rep1	36874705	31270739	28926976 (78.5%)	2343763 (6.4%)	5603966 (15.2%)
Rep2	34913246	29442431	27368968 (78.4%)	2073463 (5.9%)	5470815 (15.7%)
Input	46498956	39915516	37465668 (80.6%)	2449848 (5.3%)	6583440 (14.2%)



Peak calling with MACS2

MACS2 was used to perform peak-calling with $q=0.05$.

Thresholds	
• Redundant rate:	< 0.2
• Estimated fragment size:	< 200bp

Replicate	Number of peaks	Redundant rate	Estimated fragment size (d)
Rep1	98	0.09	121
Rep2	139	0.08	137
Merge	55	0.09	128



peak counts are very **low**: they can be a cause of concern

it is important to mention that ENCODE reports **687 peaks**

How many peaks in common?

40.8% of peaks in common
between the two replicates

Intersection between replicates	Peaks in common (basic overlap)	Peaks in common (within 100bp of each other)
Rep1 vs. Rep2	40	49

And a comparison with the merge...

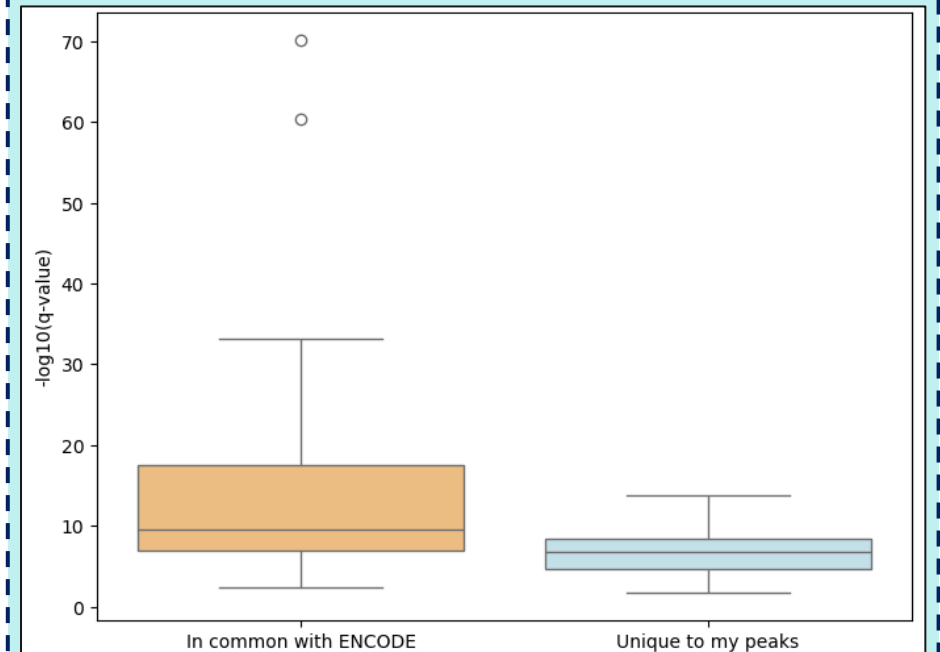
Intersection with merge	Peaks in common (basic overlap)	% peaks in common
Rep1 vs. Merge	35	35.7%
Rep2 vs. Merge	51	36.7%

Comparing results with ENCODE

Intersection	Peaks in common	Jaccard index
Rep1 vs. ENCODE	25	0.016
Rep2 vs. ENCODE	69	0.051
(Rep1 \cap Rep2) vs. ENCODE	23	0.015
Merge vs. Encode	71	0.027

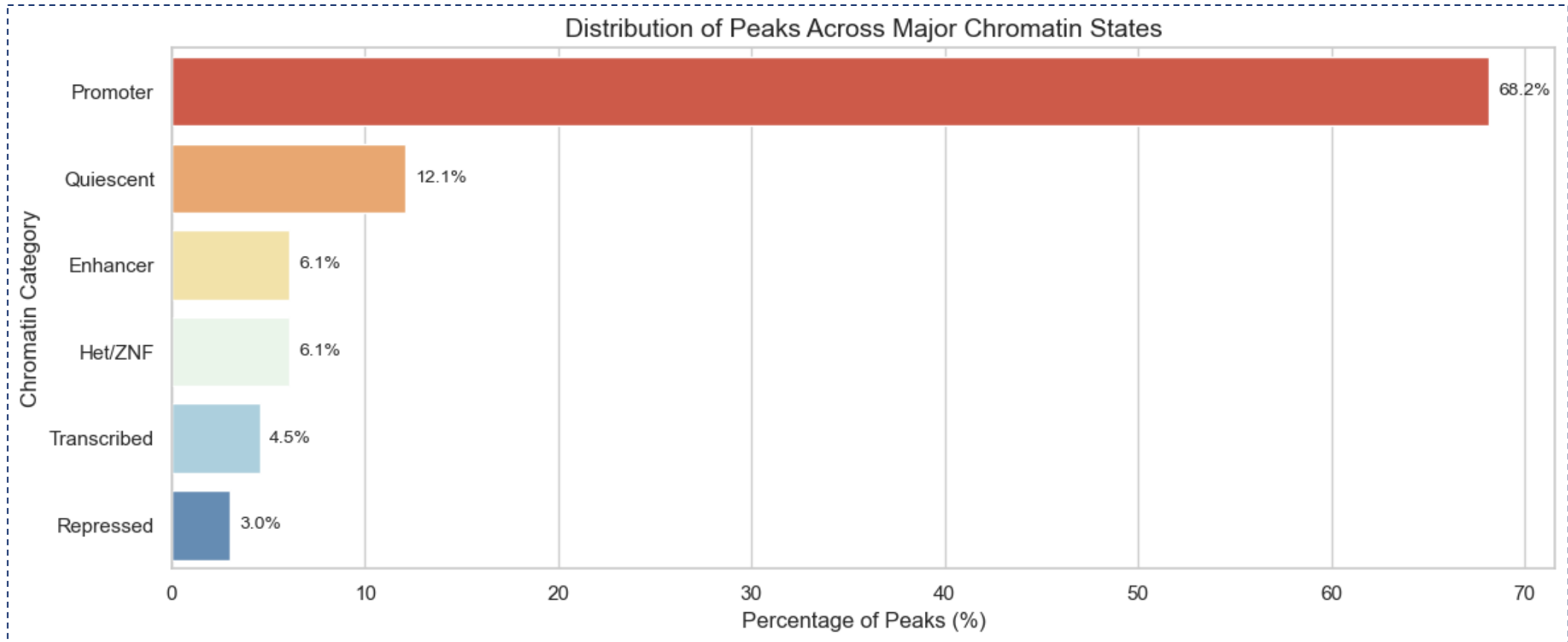
No peaks were present in the intersection with the blacklisted regions.

Peak $-\log_{10}(\text{q-value})$ distribution: peaks in common with ENCODE vs. my unique peaks



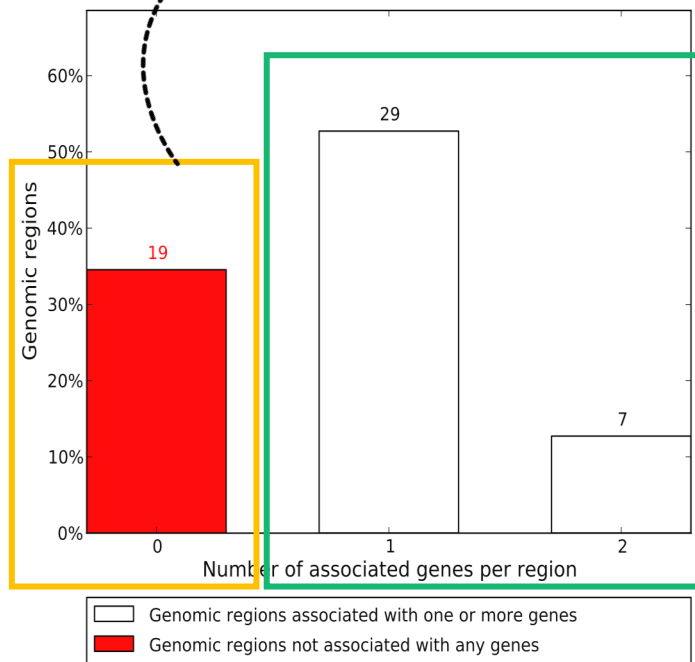
Chromatin states

Most peaks are located at or around promoters!



Target genes: peak-gene association using GREAT

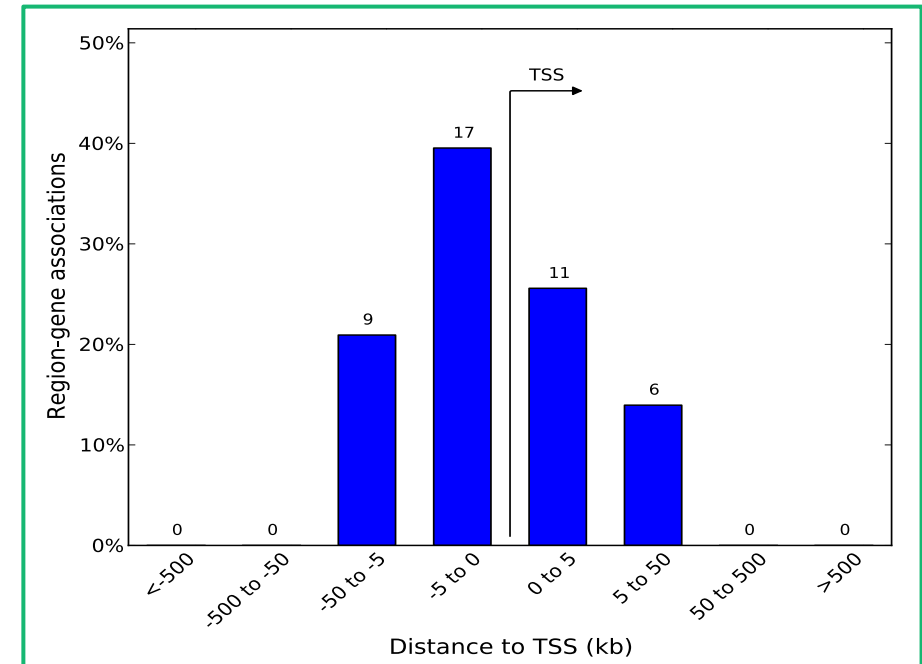
Peaks left unassigned to any gene: **19** (35%)



- Target genes bound at their promoter: **26**
- Target genes bound at distal regulatory regions: **15**

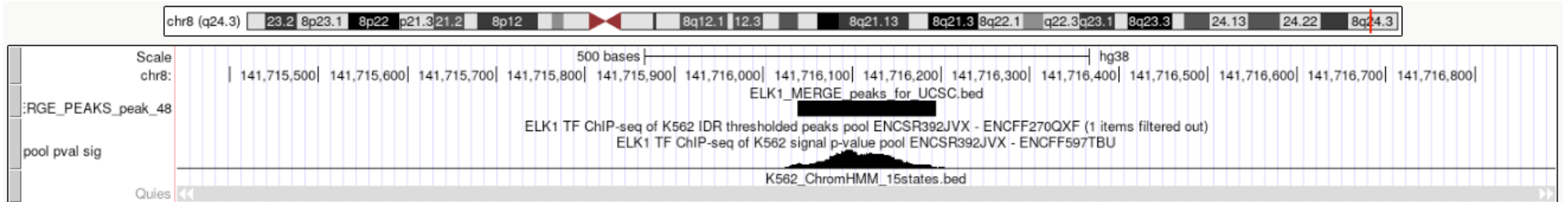
ELK1 does not regulate itself

- **Promoter: +/-1 kb**
- **Distal regulatory region: 30kb from TSS**



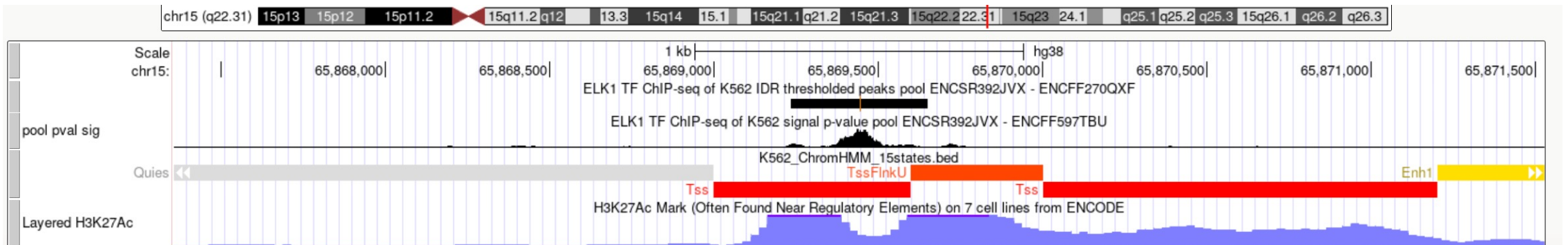
The peaks seen on on Genome Browser

A peak found by me but not by ENCODE – Chr8:141,716,040-141,716,194



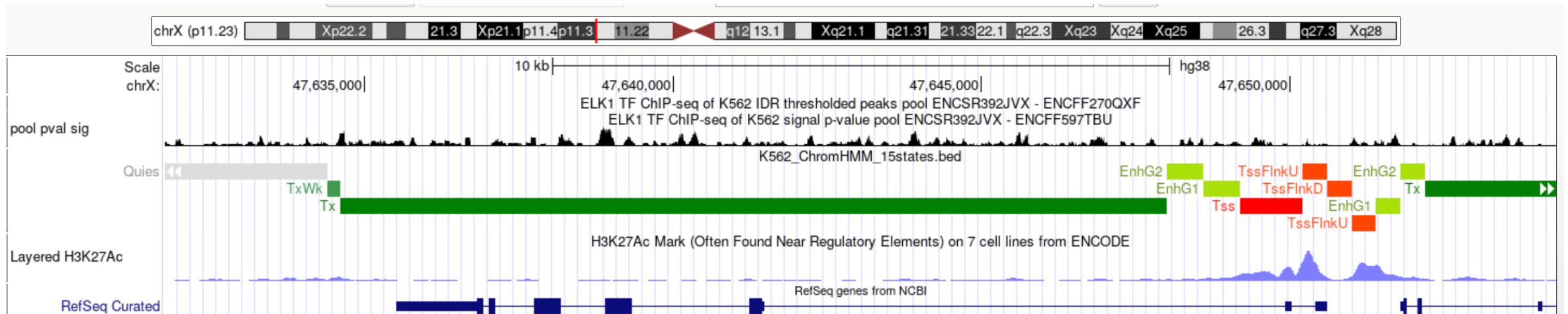
The JASPAR2024 track does **not** show any valid binding site candidates for ELK1. The chromatin is **quiescent**, which indicates that no chromatin modifications have been found in the region.

A peak found by ENCODE but not by me – Chr15:65,869,235-65,869,651



The ELK1 gene on Genome Browser

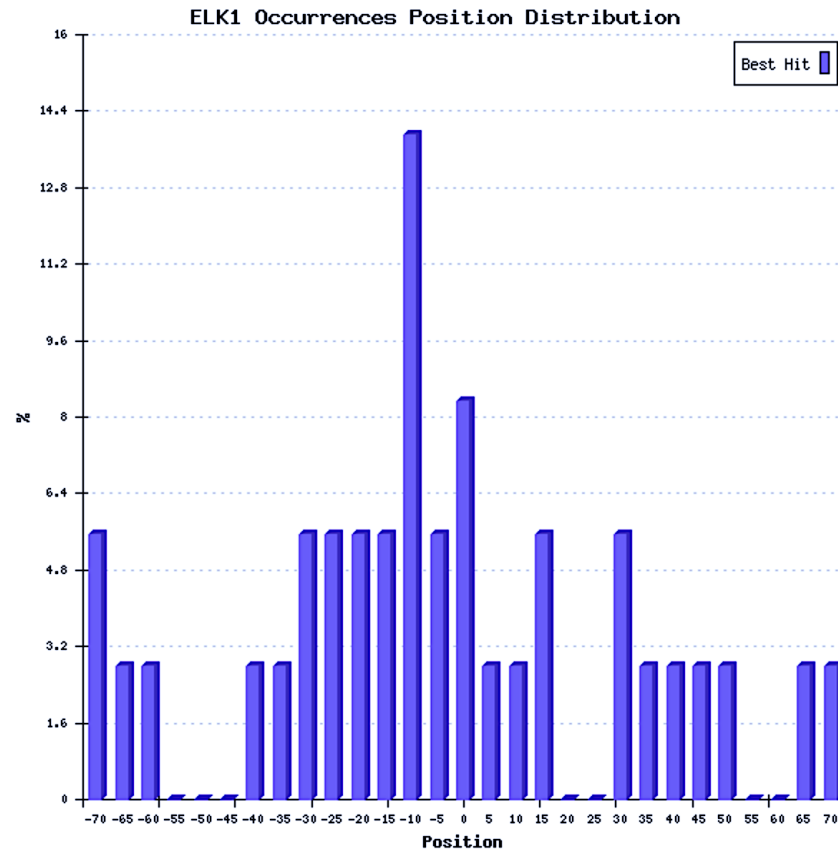
ELK1 does **not** autoregulate



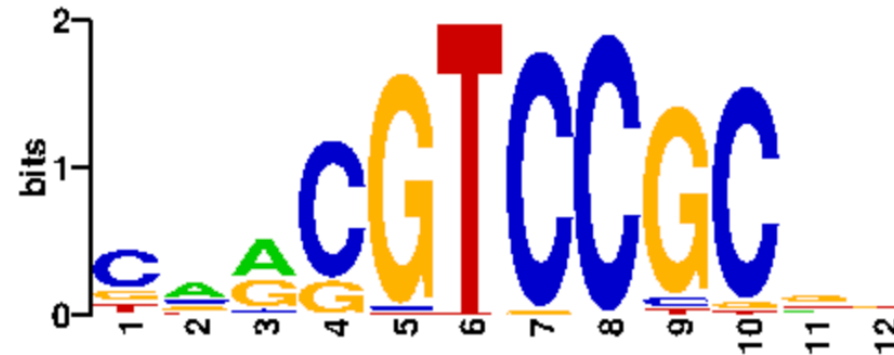
→ No peaks have been found on the gene
The chromatin signature indicates that it is active and transcribed!

Motif enrichment analysis

The most enriched motif (global p-value) is not related to the ETS family.



The most enriched motif at the global level: **HINFP**



The associated TF is **HINFP**. The literature does **not** state that there might be cooperative binding of ELK1 with HINFP.

ELK1 seems to bind mainly around the summit, but also around it.

Sources

[1] Boros J, Donaldson IJ, O'Donnell A, Odrowaz ZA, Zeef L, Lupien M, Meyer CA, Liu XS, Brown M, Sharrocks AD. Elucidation of the ELK1 target gene network reveals a role in the coordinate regulation of core components of the gene regulation machinery. *Genome Res.* 2009 Nov;19(11):1963-73. doi: 10.1101/gr.093047.109. Epub 2009 Aug 17. PMID: 19687146; PMCID: PMC2775591.

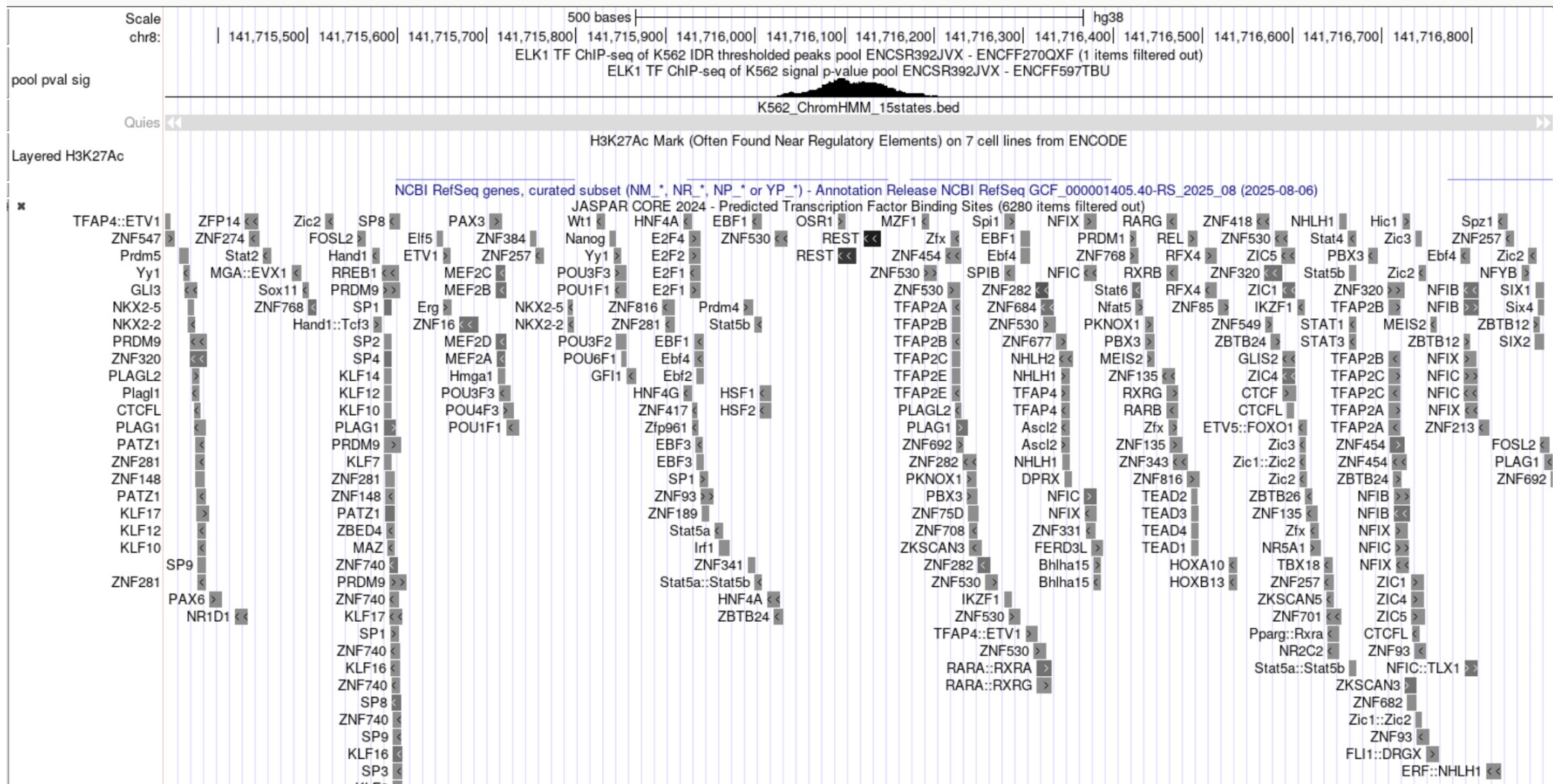
[2] Prise I and Sharrocks AD. "ELK1 has a dual activating and repressive role in human embryonic stem cells" [version 2; peer review: 2 approved]. *Wellcome Open Res* 2019, **4**:41 (<https://doi.org/10.12688/wellcomeopenres.15091.2>)

Chromatin states

Chromatin state	Number of peaks	Meaning
TssFlnkU	25	Promoter Flank, upstream of TSS
Tss	12	Promoter/TSS
Quies	8	Quiescent
TssFlnkD	4	Promoter flank, downstream of TSS
Biv		Bivalent promoter
TxWk	3	Weak transcription
ZNF/Rpts	2	ZNF genes and repeats
ReprPC		Polycomb-repressed regions (silenced chromatin)
Het		Heterochromatin (silent regions)
Enh2		Enhancer, weak or poised
EnhG2	1	Genic enhancer
Enh1		Strong enhancer

Move <<< << < > >> >>> Zoom in 1.5x 3x 10x Base Zoom out 1.5x 3x 10x 100x

chr8 (q24.3) 23.2 8p23.1 8p22 p21.3 21.2 8p12 8q12.1 12.3 8q21.13 8q21.3 8q22.1 q22.3 q23.1 8q23.3 24.13 24.22 8q24.3



ELK1 binding motif sequence logo

