

HMGA1 regulates the Plasminogen activation system in the secretome of breast cancer cells

Giulia Resmini^{1,§,#}, Serena Rizzo^{1,#}, Cinzia Franchin^{2,3}, Rossella Zanin¹, Carlotta Penzo¹, Silvia Pegoraro¹, Yari Ciani⁴, Silvano Piazza^{4,£}, Giorgio Arrigoni^{2,3}, Riccardo Sgarra^{1,*} and Guidalberto Manfioletti^{1,*}

¹Department of Life Sciences, University of Trieste, Trieste, Italy

²Department of Biomedical Sciences, University of Padova, Padova, Italy.

³Proteomics Center, University of Padova and Azienda Ospedaliera di Padova, Padova, Italy.

⁴National Laboratory CIB (LNCIB), Area Science Park, Trieste, Italy

[§]Present address: I.R.C.C.S. Policlinico San Donato, Milano, Italy

[£]Present address: Bioinformatics Core Facility, Centre for Integrative Biology, CIBIO, University of Trento, Trento, Italy

[#]these authors contributed equally.

^{*}co-last and co-corresponding authors.

Correspondence:

Guidalberto Manfioletti & Riccardo Sgarra

Department of Life Sciences, University of Trieste

Via L. Giorgieri, 5 - 34127 Trieste

Italy

Phone: +39 040 558-8720 / -8721

E-mail: rsgarra@units.it / manfiole@units.it

Supplementary Figure legends

Supplementary Figure S1. Effect of serum starvation on the growth of MDA-MB-231 and MDA-MB-231 shA1_3 cells and the evaluation by SDS-PAGE of a protocol for the enrichment of glycosylated proteins. a) Metabolic activity of MDA-MB-231 and MDA-MB-231_shA1_3 not induced (NI) and induced (I – 10 days). Cells were seeded at a density of 5000 cells per well in 96-multiwell. After 72 hours, complete medium has been removed and replaced with serum-free medium for the serum starvation (SS). As a control, cells were also maintained in complete medium. MTS assay was performed at 0, 24, 48, 72, and 96 h after serum starvation. Values shown are averages $\pm$ SD (n=3). The metabolic activity of cells at time 0 was set to 100% arbitrarily. The dashed vertical line indicates the 30h time chosen for the collection of secreted proteins. b) Silver staining of proteins separated by SDS-PAGE (T=15%) to verify glycoproteins purification protocol. CM Input: protein concentrate obtained from the culture medium (CM). Samples purified with a control resin (Empty Resin, i.e. without lectins) or with a mixture of ConA- and WGA-derivatized resins are shown (ConA+WGA). Unbound: unbound proteins (flow through), lanes 2 and 5. Competitive elution: proteins retained by the resin and eluted by competitive elution with methyl- α -D-mannopyranoside and N-acetyl-glucosamine, lanes 3 and 6. SDS elution: protein retained by the resin and eluted by SDS, lanes 4 and 7. Molecular weights markers (kDa) are shown on the left. c) The conA+WGA affinity purified proteins (lane 6, panel b) have been identified by LC-MS/MS analyses and subjected to bioinformatic evaluation (gene-annotation enrichment analysis).

Supplementary Figure S2. Checking (i) the efficiency of HMGA1 silencing in MDA-MB-231_shA1_3 cells, (ii) the normalization of protein samples subjected to iTRAQ screening, and (iii) the enrichment of secreted/glycosylated proteins in the ConA+WGA affinity purification. a) 12 out of 60 independent 175 cm² cell culture flask were randomly selected for WB analyses to check for HMGA1 expression levels in MDA-MB-231 shA1_3 I and NI cells. After collection of culture media, total protein lysates obtained from cells were analyzed by SDS PAGE (T=15%) and western blot. A blue coomassie (BC) stained gel is shown to check protein quantity normalization. The expression of HMGA1 was detected by anti-HMGA1 antibody (α -A1). Molecular weight markers (kDa) are shown on the right. b) Blue Coomassie staining of proteins separated by SDS-PAGE (T=15%) to verify glycoproteins quantification and normalization by the

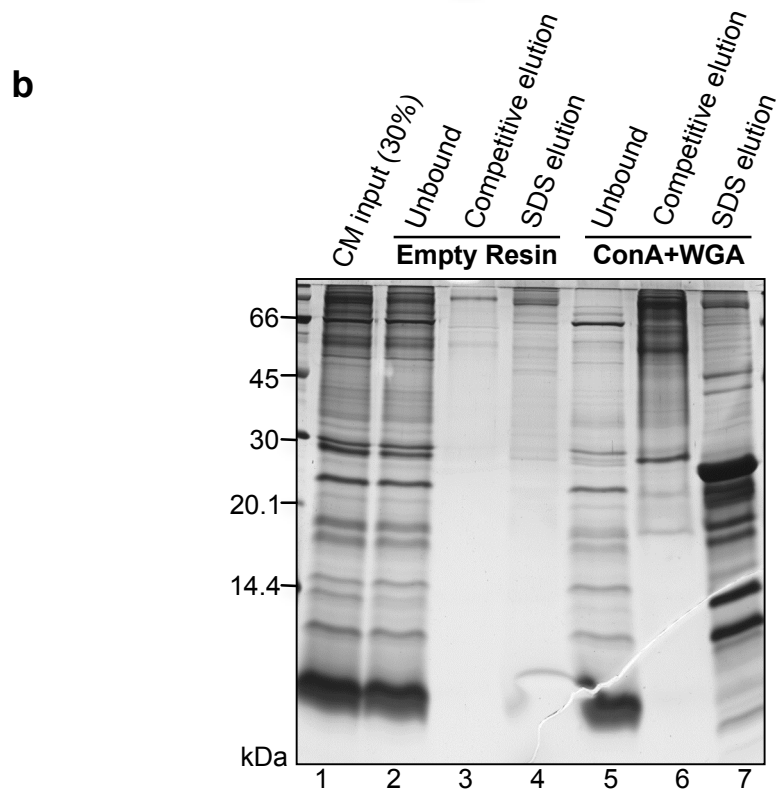
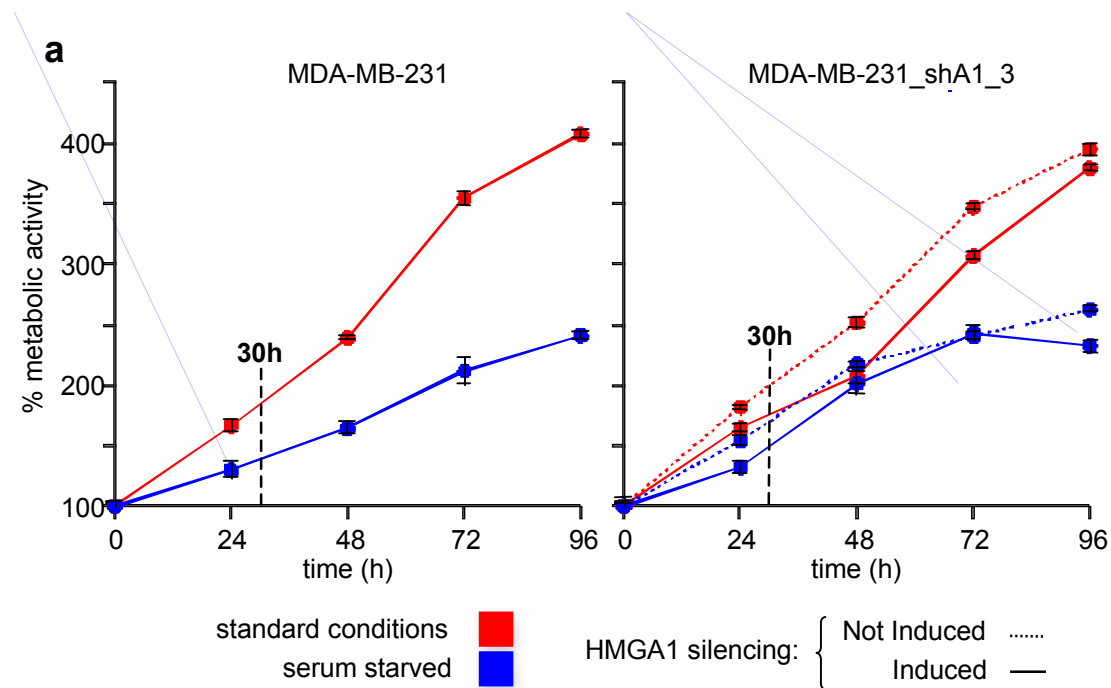
Lowry method. Equivalent amounts (4.5 µg) of the glycoproteins purified in the three biological replicates (NI and I) were loaded. c) The conA+WGA affinity purified proteins have been identified by the iTRAQ LC–MS/MS screening. All the identified proteins have been subjected to bioinformatic evaluation (gene-annotation enrichment analysis).

Supplementary Figure S3. Transcription factor motif analysis of the *PLAU* promoter/enhancer region (about 10 kbp upstream the transcription start site). The analysis of the promoters sequences was performed through Biobase Transfac using the Match analysis method. Promoter sequences were obtained from the Transfac internal database (version 2016.3) selecting a distance of 10kb upstream and 1kb downstream the TSS. The sequence of the *PLAU* promoter/enhancer region is shown. Putative transcription factor binding sites are evidenced with arrows indicating the binding orientation.

Supplementary Figure S4. Transcription factor motif analysis of the *SERPINE1* promoter/enhancer region (about 10 kbp upstream the transcription start site). The analysis of the promoters sequences was performed through Biobase Transfac using the Match analysis method. Promoter sequences were obtained from the Transfac internal database (version 2016.3) selecting a distance of 10kb upstream and 1kb downstream the TSS. The sequence of the *SERPINE1* promoter/enhancer region is shown. Putative transcription factor binding sites are evidenced with arrows indicating the binding orientation.

Supplementary figure S5. Focal Adhesion Kinase (FAK) expression levels are linked to HMGA1. MDA–MB–231 cells have been treated with siRNA targeting HMGA1 (siA1) or with control siRNA (siCtrl). Proteins have been extracted with SDS lysis buffer and analysed by SDS–PAGE and western blot. FAK has been detected using a α-FAK antibody. The membrane after western blot has been stained with red ponceau and it is shown as a loading control. Molecular markers (kDa) are shown on the right.

Supplementary Figure 1



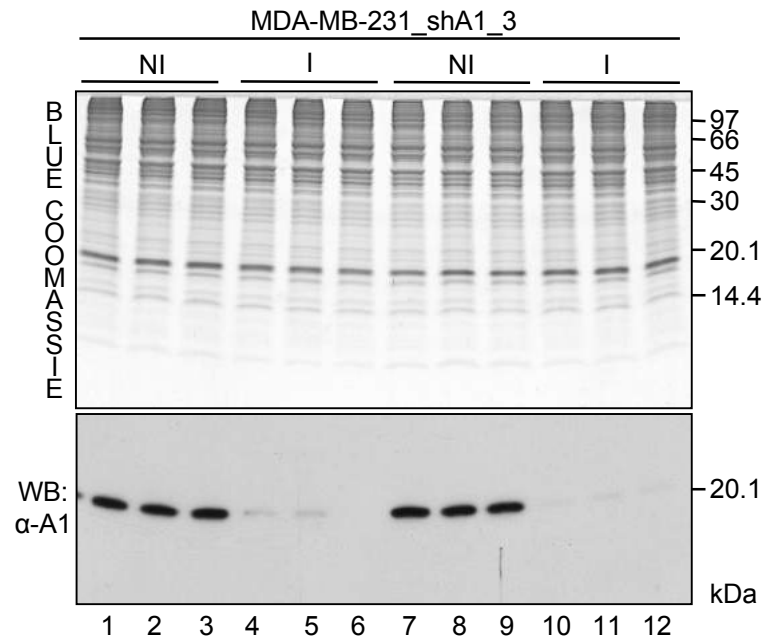
c

MS

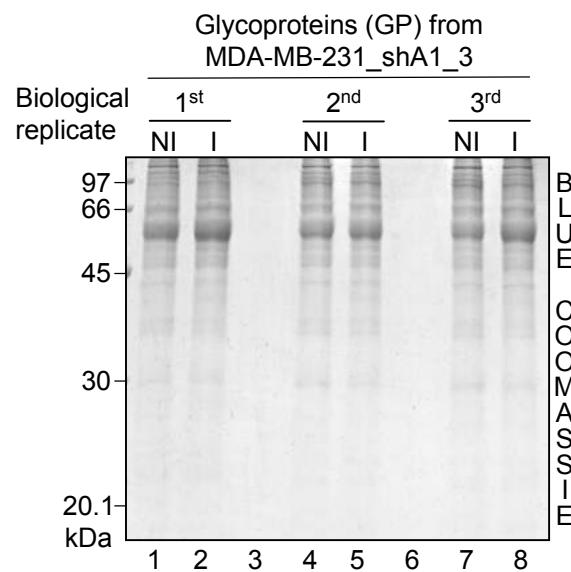
Term	Count	%	PValue	Fold Enrichment	FDR
Glycoprotein	186	85,71	7,06E-89	3,87	9,22E-86
glycosylation site:N-linked (GlcNAc...)	179	82,49	3,18E-84	3,93	5,04E-81
Signal	174	80,18	1,09E-80	3,98	1,43E-77
extracellular exosome	150	69,12	3,96E-71	4,50	5,25E-68
extracellular space	108	49,77	1,29E-62	6,66	1,71E-59
Secreted	94	43,32	1,36E-39	4,65	1,78E-36

Supplementary Figure 2

a



b



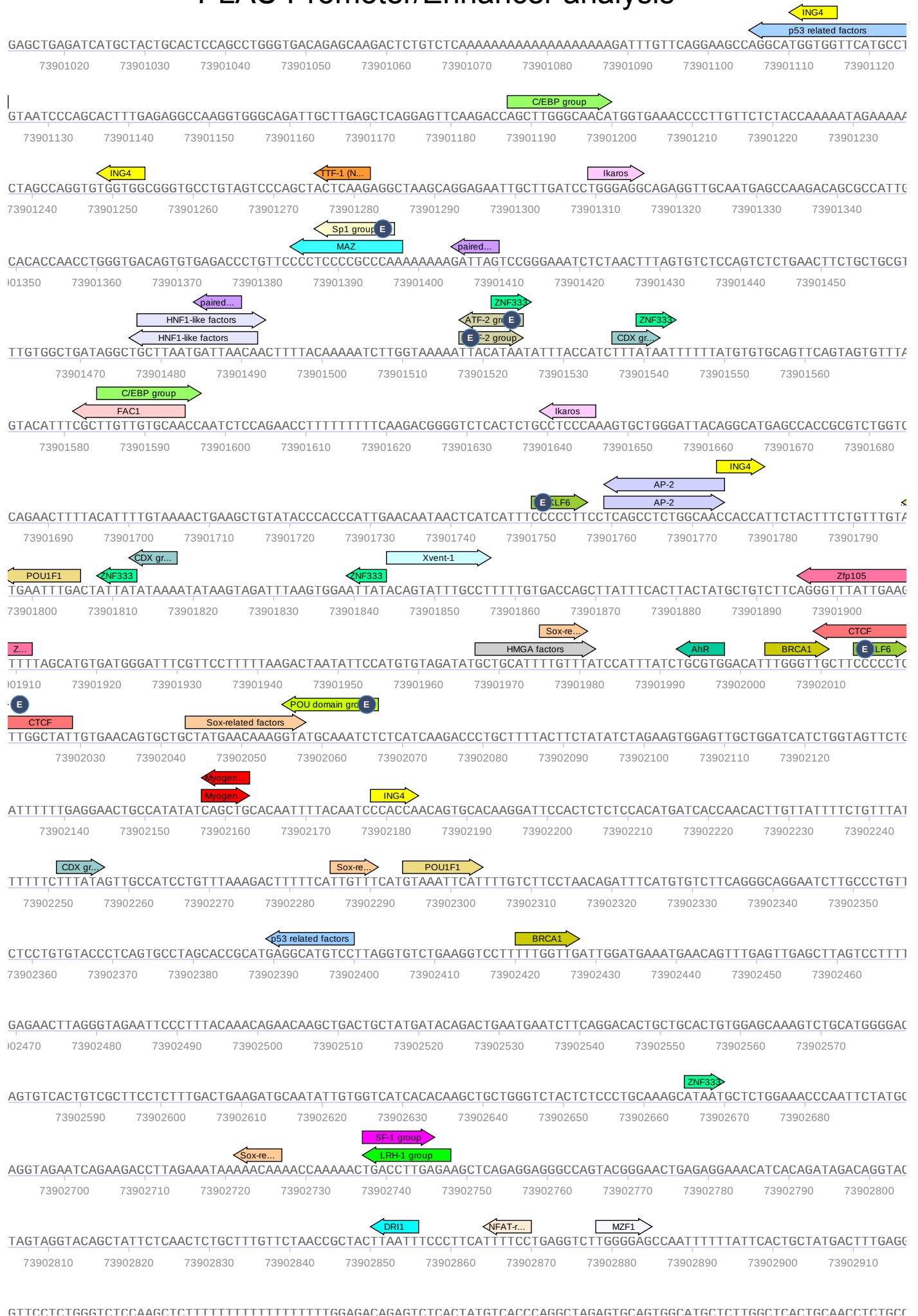
c

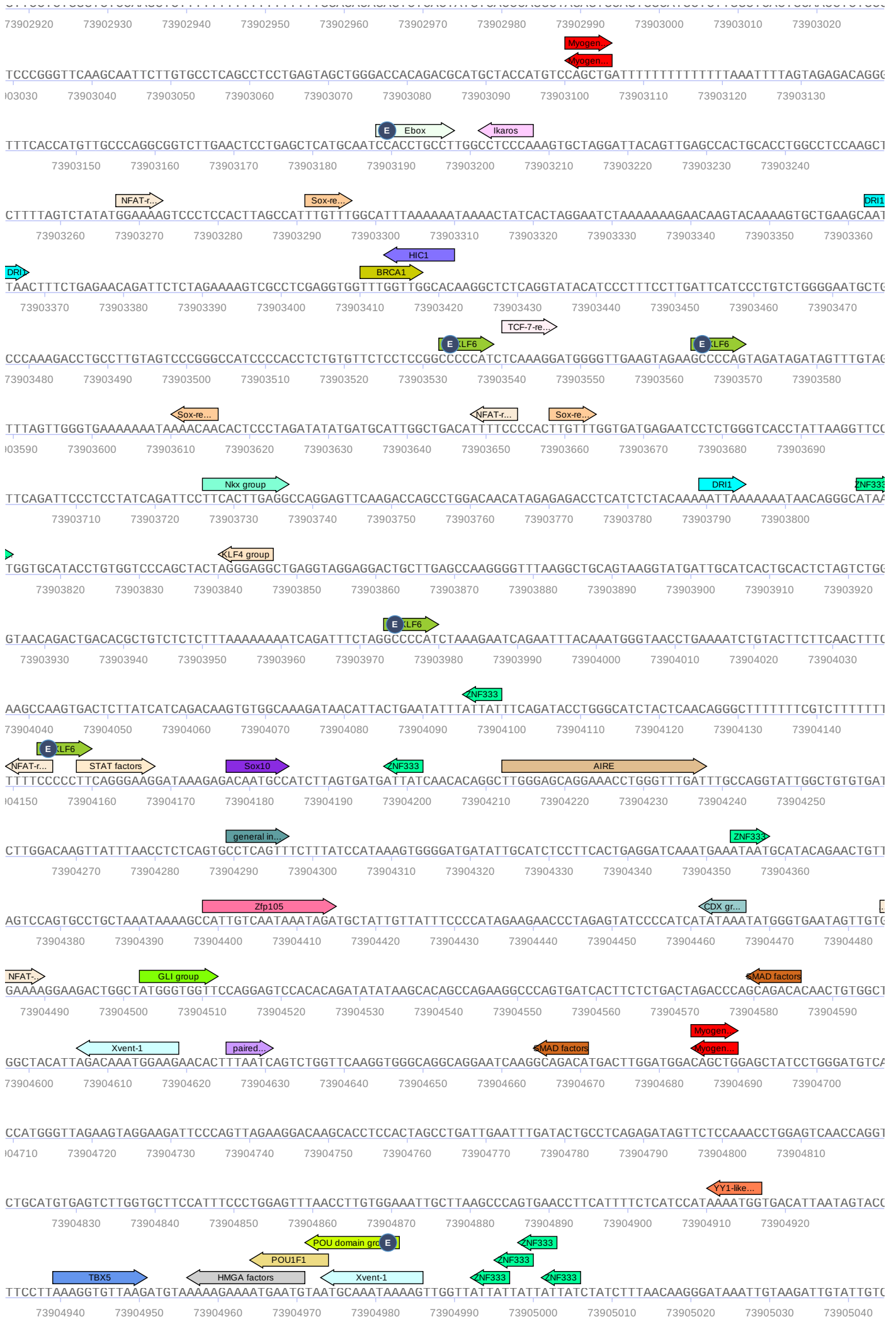
iTRAQ data: identified proteins

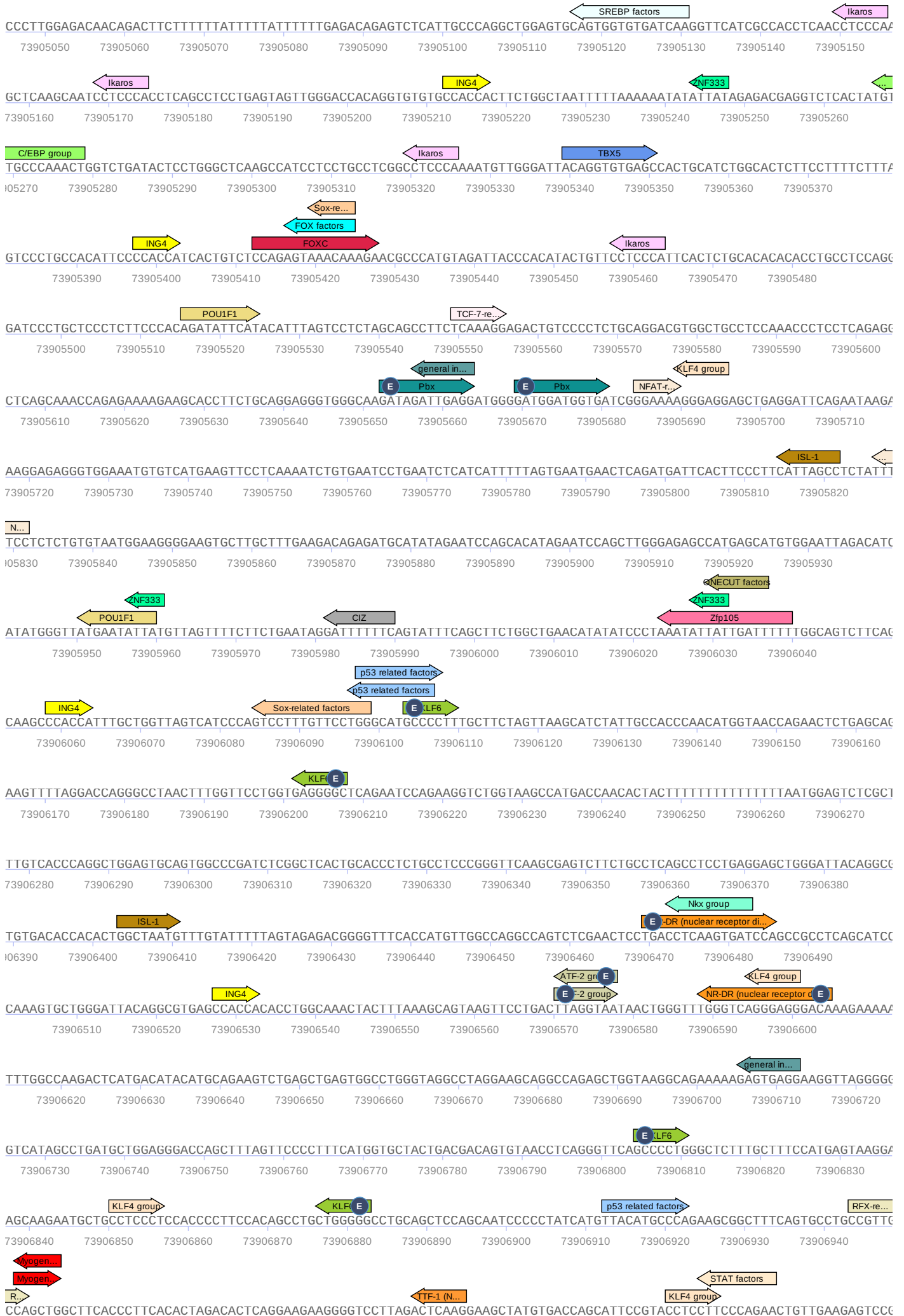
Term	Count	%	PValue	Fold Enrichm	FDR
extracellular exosome	155	71,76	4,60E-77	4,652226	6,13E-74
Signal	153	70,83	5,64E-58	3,504377	7,43E-55
Glycoprotein	158	73,15	1,99E-57	3,3079807	2,63E-54
glycosylation site:N-linked (GlcNAc...)	149	68,98	1,30E-51	3,2687186	2,08E-48
extracellular space	96	44,44	1,76E-50	6,0130331	2,35E-47
Secreted	88	40,74	7,71E-34	4,2671002	1,02E-30

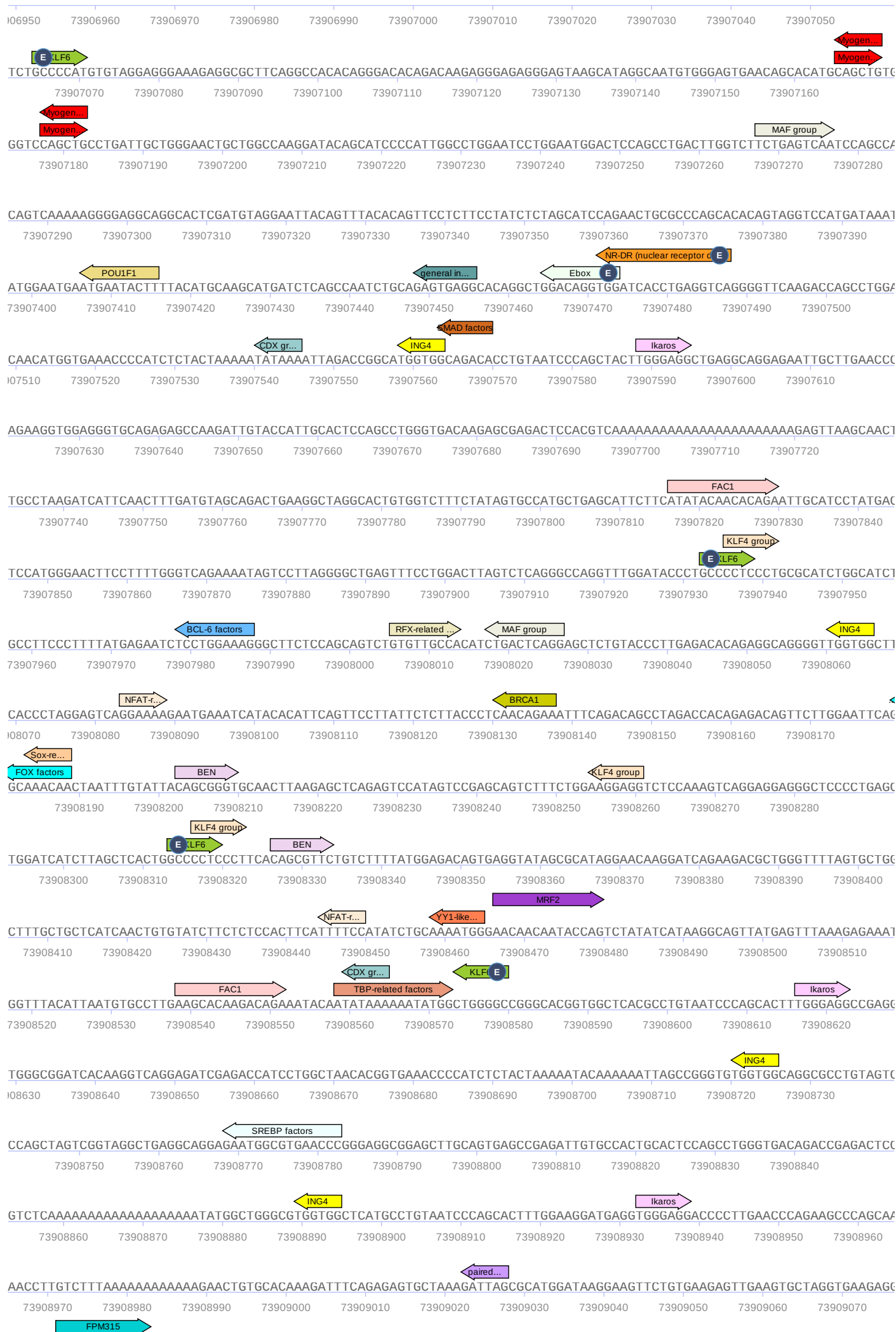
Supplementary Figure S3

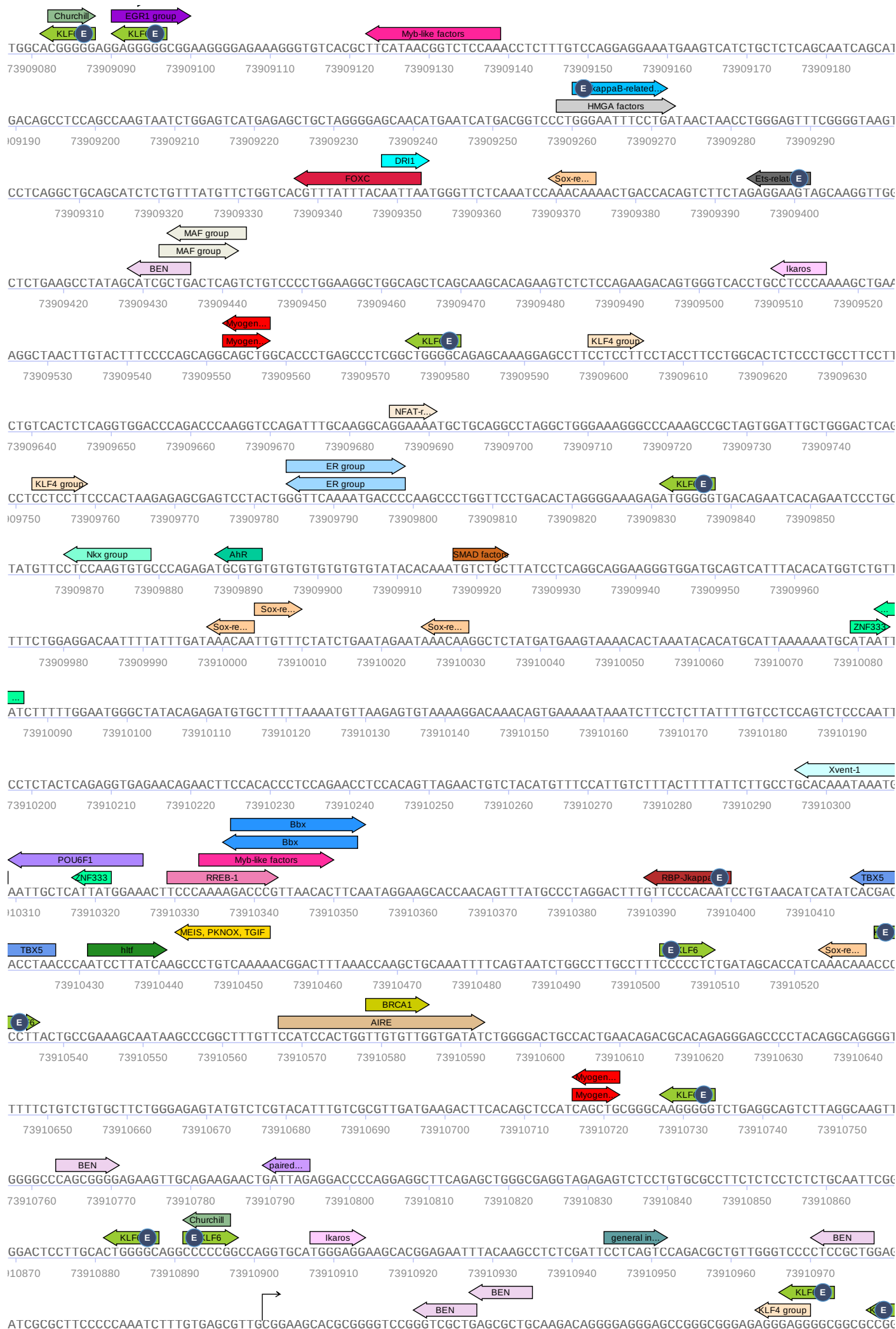
PLAU Promoter/Enhancer analysis

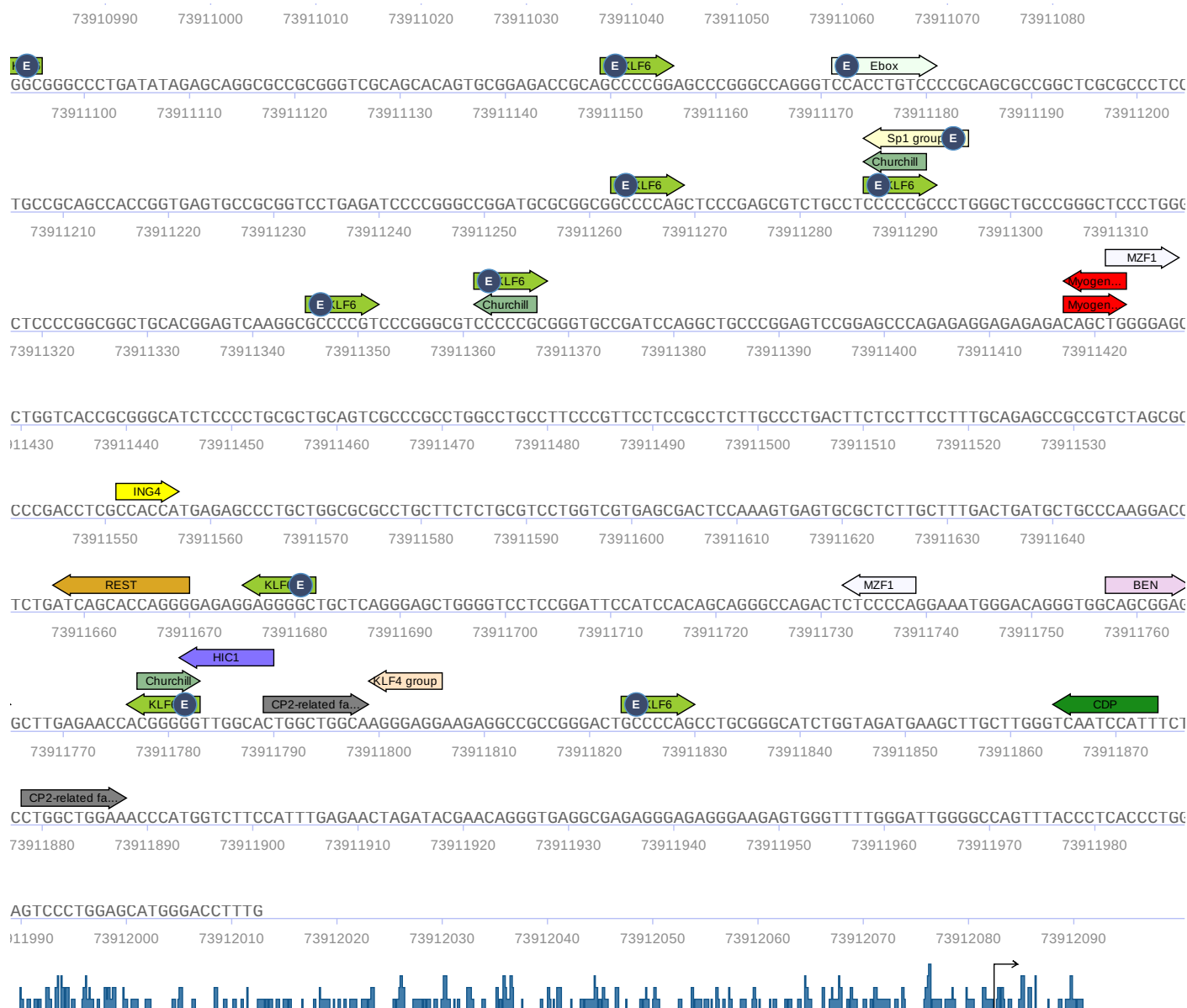






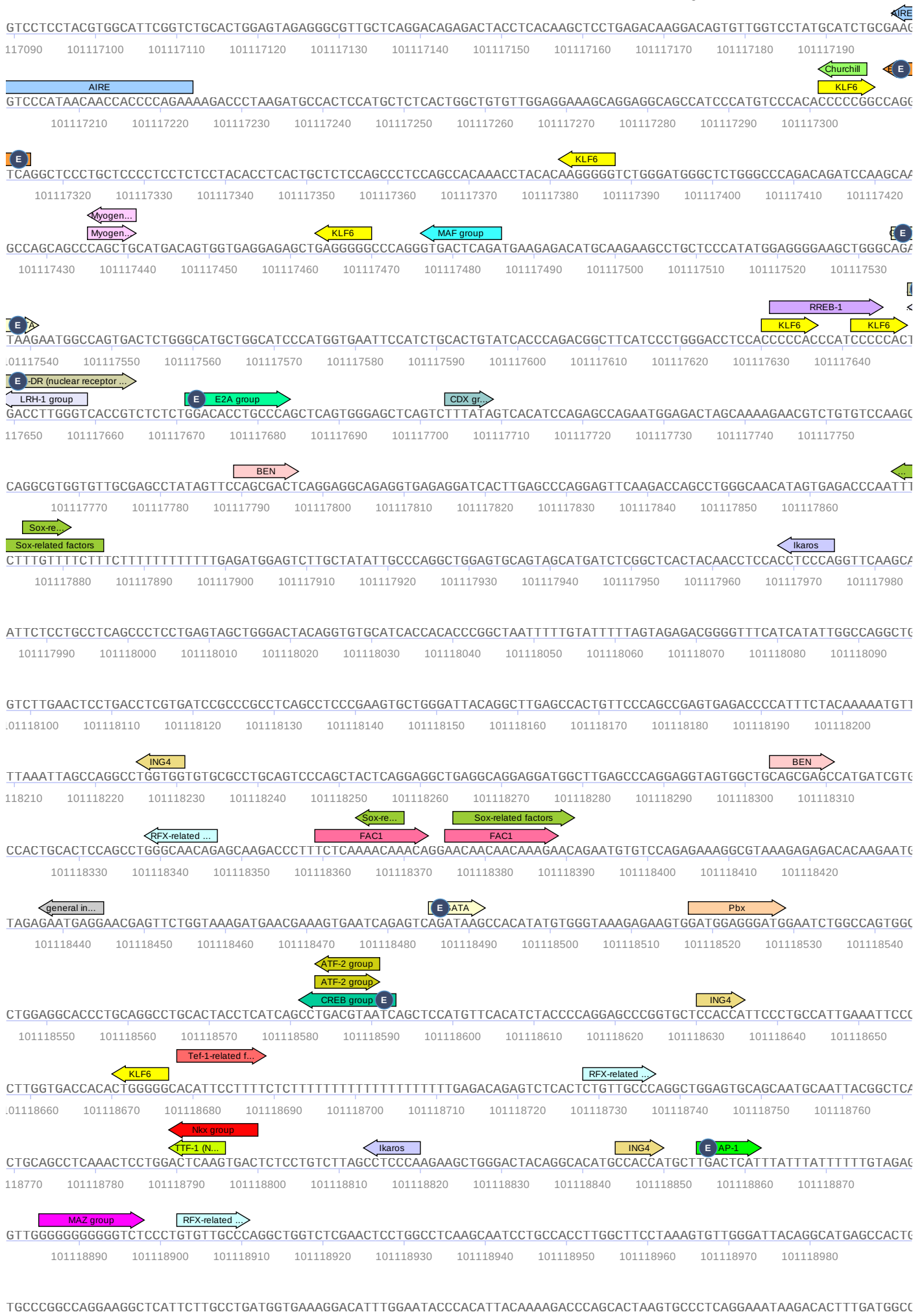


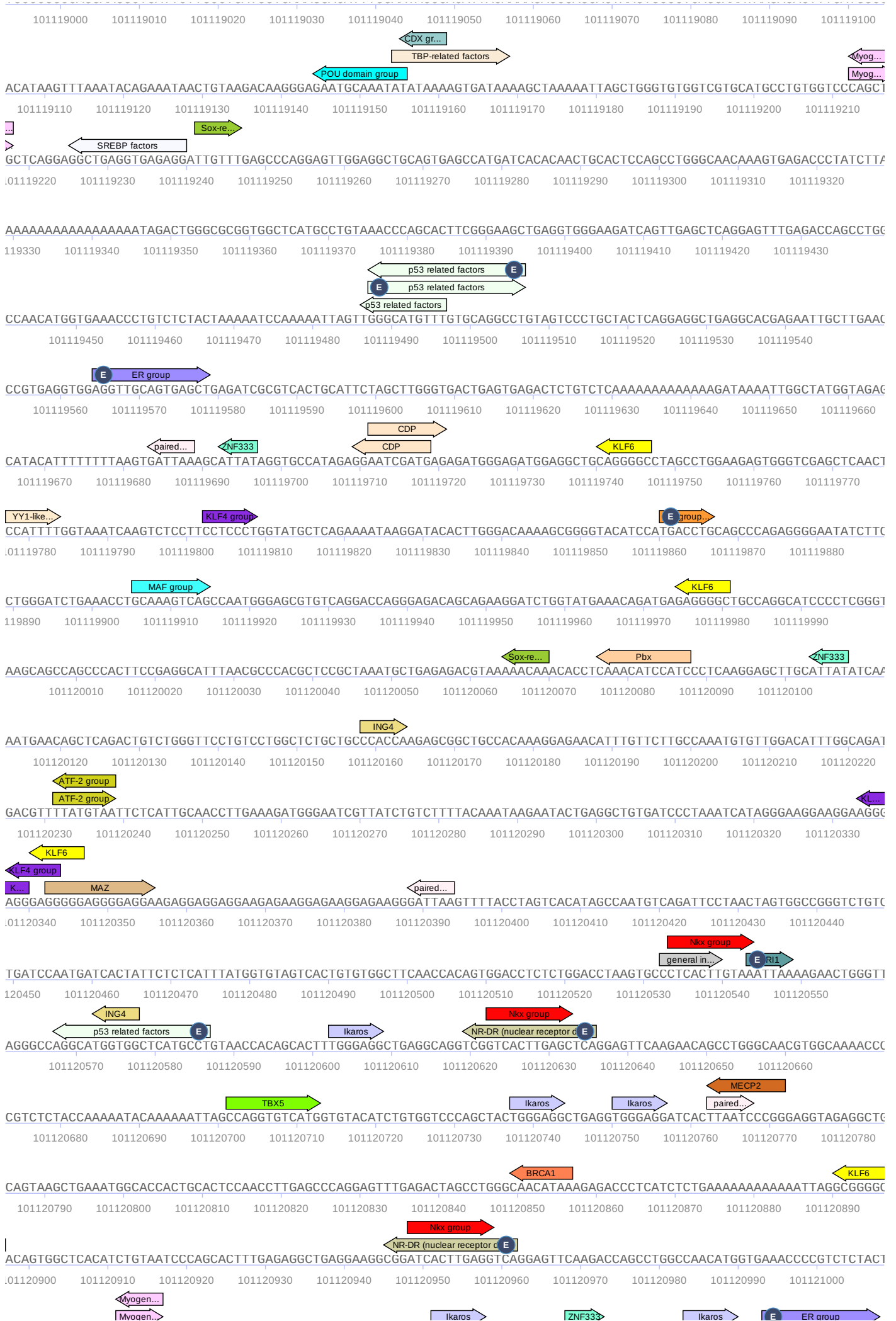


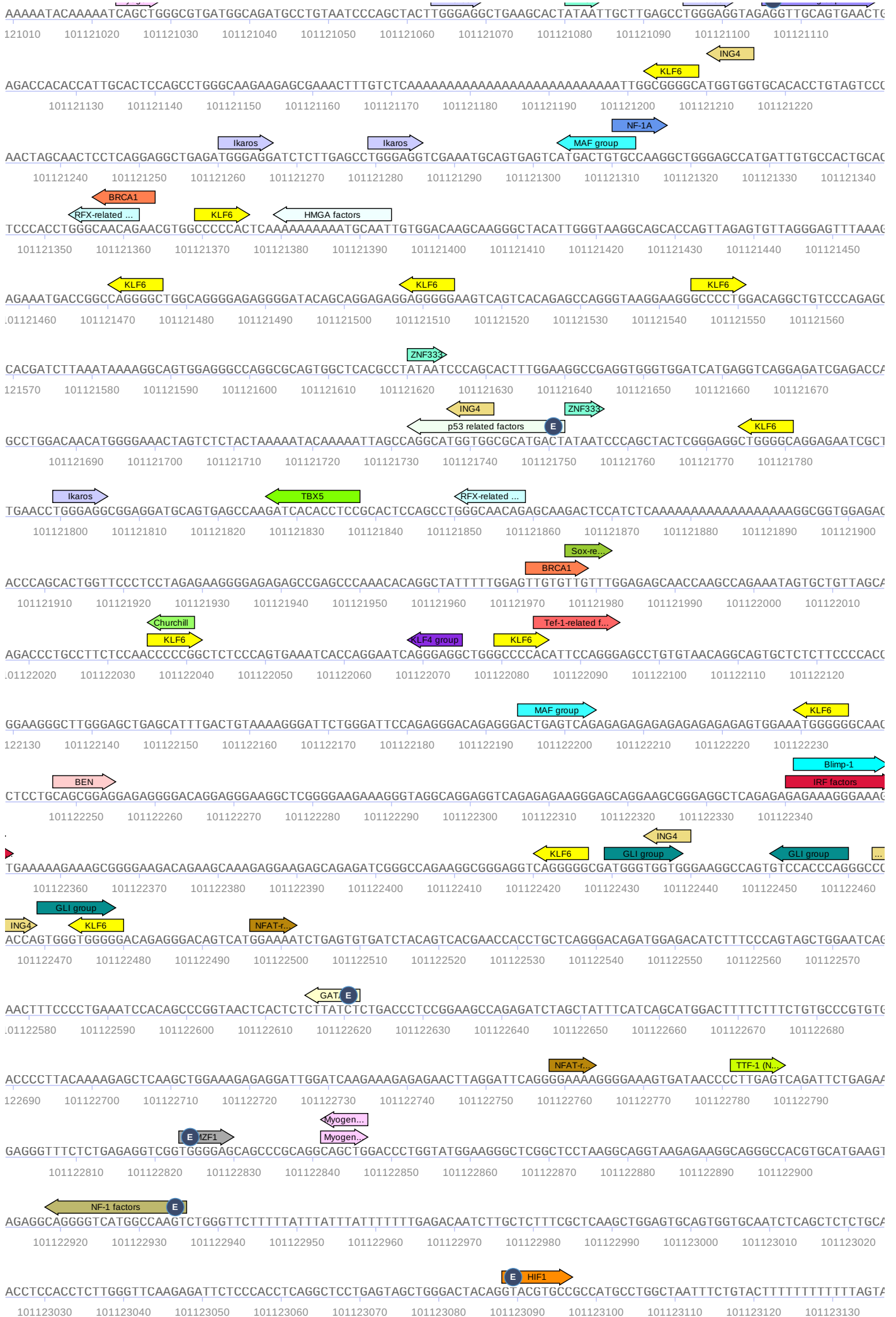


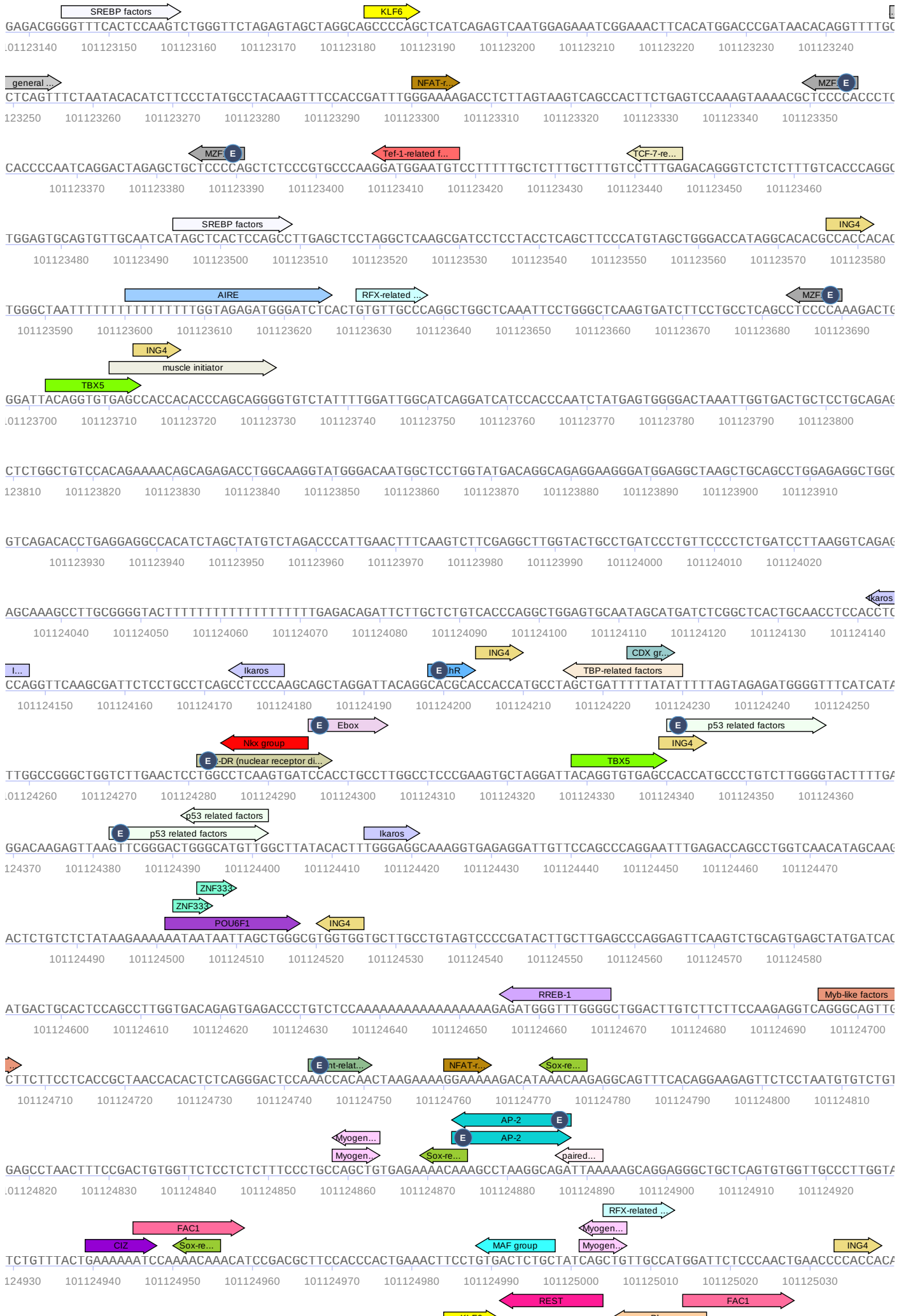
Supplementary Figure S4

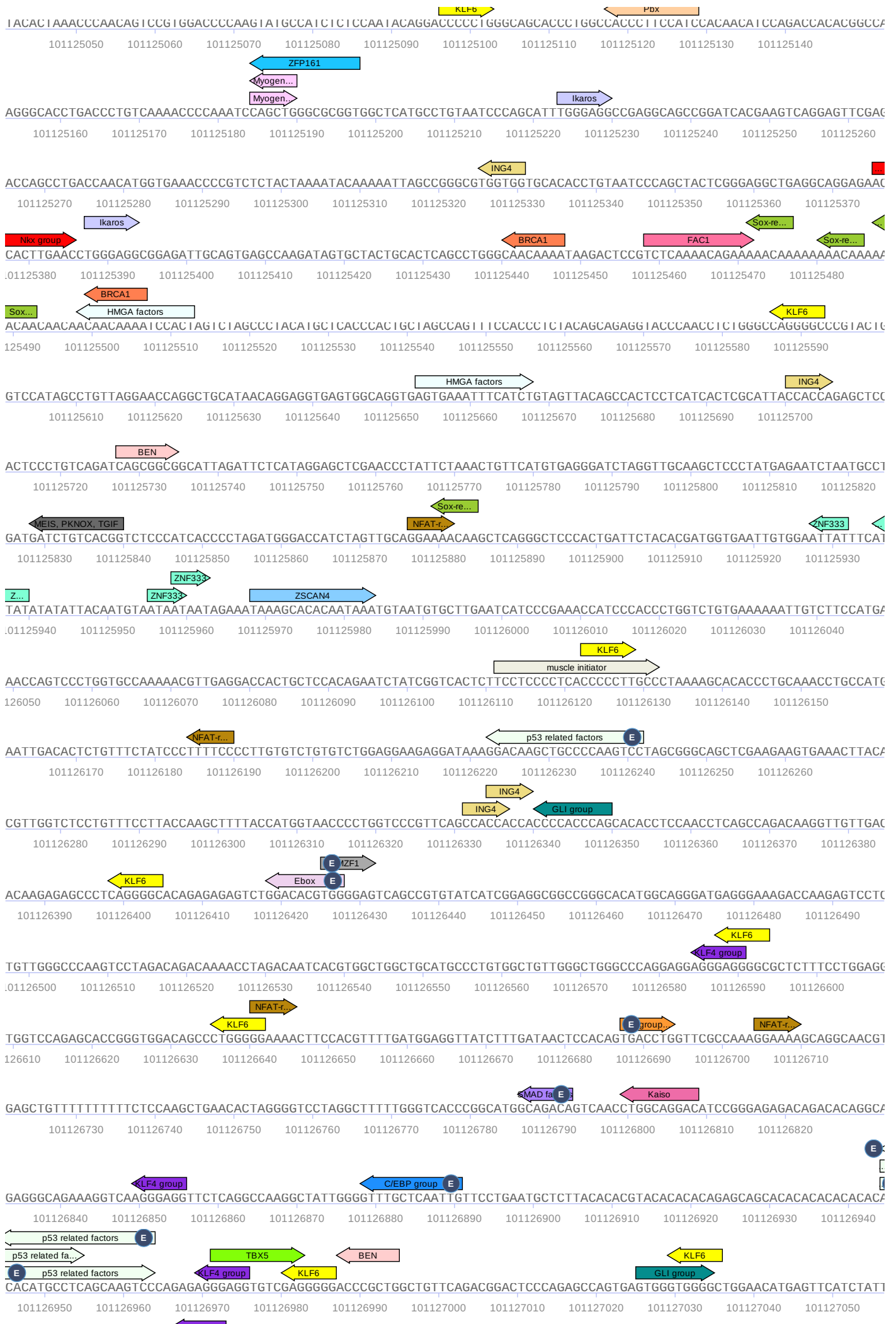
SERPINE1 Promoter/Enhancer analysis

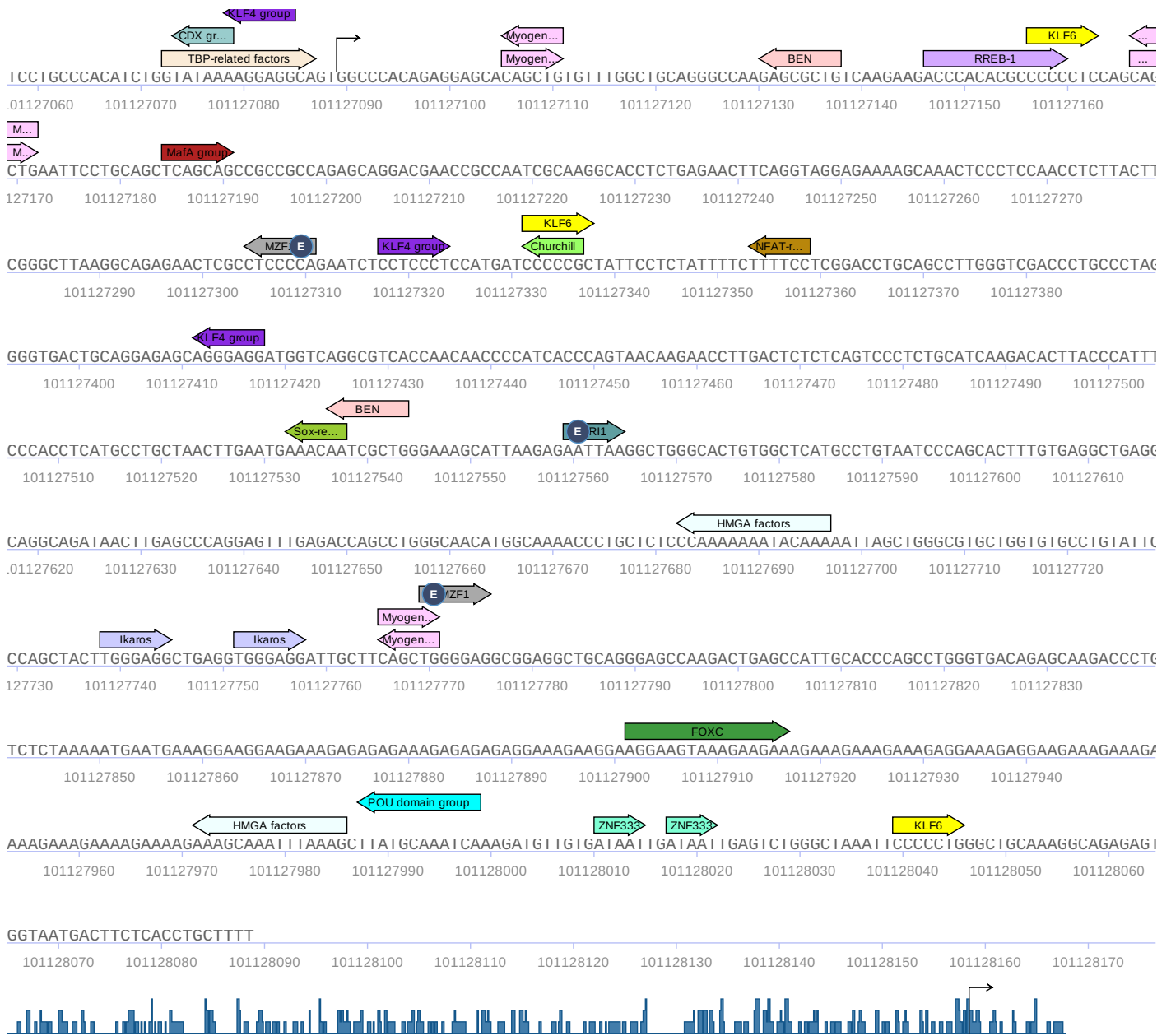




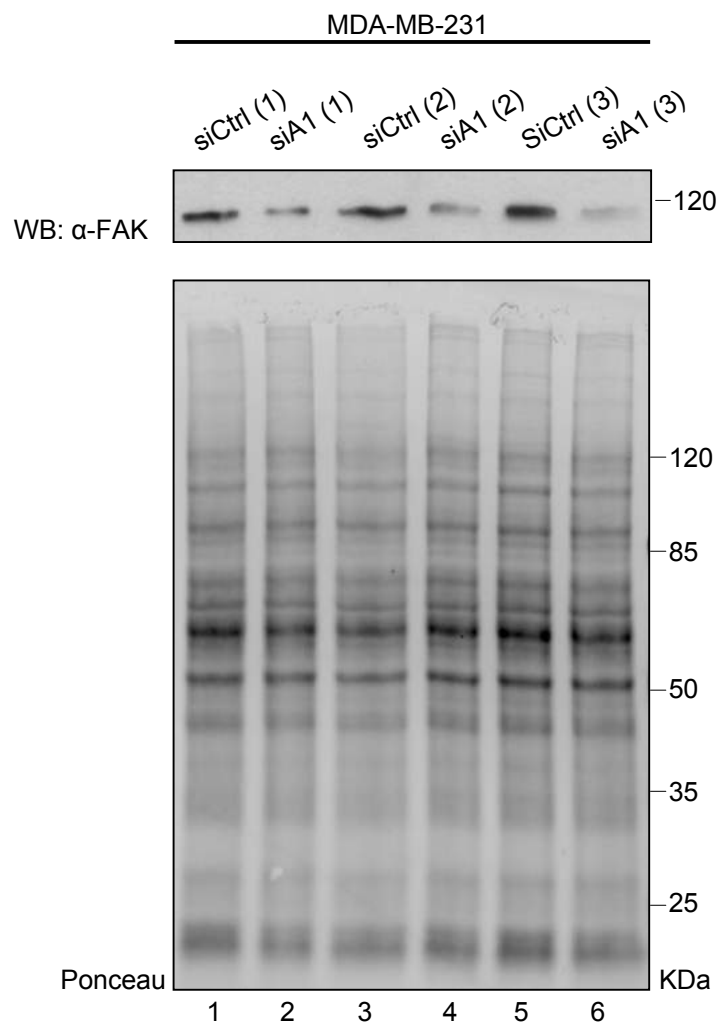








Supplementary figure 5



Supplementary Data – Methods

qRT–PCR primers	
Primer Name	Oligo sequence (5' → 3')
NRP2 rev	CGATGTTCCCACAGTGT TTG
NRP2 for	CGAACCCAACCAGAAGATTG
CTSC rev	CCAGAATTGCCAAGGTCATC
CTSC for	TGTCAACTGCTCGGTTATGG
SLC1A5 rev	AAACCCACATCCTCCATCTC
SLC1A5 for	TCCGCTTCTTCAACTCCTTC
LGMN rev	GGAGTGGGATTGTCTTCAGAGT
LGMN for	GCAGGTTCAAATGGCTGGTAT
PLAUR rev	TGATGAGCCACAGGAAATGC
PLAUR for	TGGGTTAGACTTGTGCAACC
PLAU rev	AGCAAGGCAATGTCGTTGTG
PLAU for	GCTCAAGGCTTAACTCCAACAC
SERPINE1 rev	AAGGGGCAGCAATGAACATG
SERPINE1 for	TGGCTCAGACCAACAAGTTC
STC1 rev	TCCTTTGGAAAGTGGAGCAC
STC1 for	CTCAGGGAAAAGCATTTCGTC
ADAM9 rev	AACATCTGGCTGACAGAACTGA
ADAM9 for	CTTGCTGCGAAGGAAGTACC
HMGA1 rev	AGCCCCTCTTCCCCACAAAGAGT
HMGA1 for	ACCAGCGCCAAATGTTTCATCCTCA

siRNA sequences	
siRNA name	siRNA sequence (5' → 3')
siNRP2	GCUUCUCUGCGCGUUACUATT
siLGMN	GCAUUGGUGCCGUUCCUAUTT
siA1_3	ACUGGAGAAGGAGGAAGAGTT
siCTRL	ACAGUCGCGUUUGCGACUGTT