

## Group Meeting\_Fangjie\_5 (from 2017-02 )

### > Recent progress

#### Finished

- \_ Nucleosome CAP-SELEX (NCAP-SELEX) and dissociation assay (cyc5) with full-length TFs
- \_ HT-SELEX for FL TFs as a control
- \_ Comparison of TF affinity to motifs on free-DNA and on nucleosome
- \_ Nucleosome-linker SELEX with FL TFs for the linker length of 25bp

#### Ongoing

- \_ Nucleosome-linker SELEX for the linker length of 10 bp and 50 bp

#### Short-term plan

- \_ Nucleosome-SELEX with linker histone
- \_ SELEX for Pol I subunits
- \_ More zinc fingers for NCAP-SELEX

### > NCAP-SELEX related data

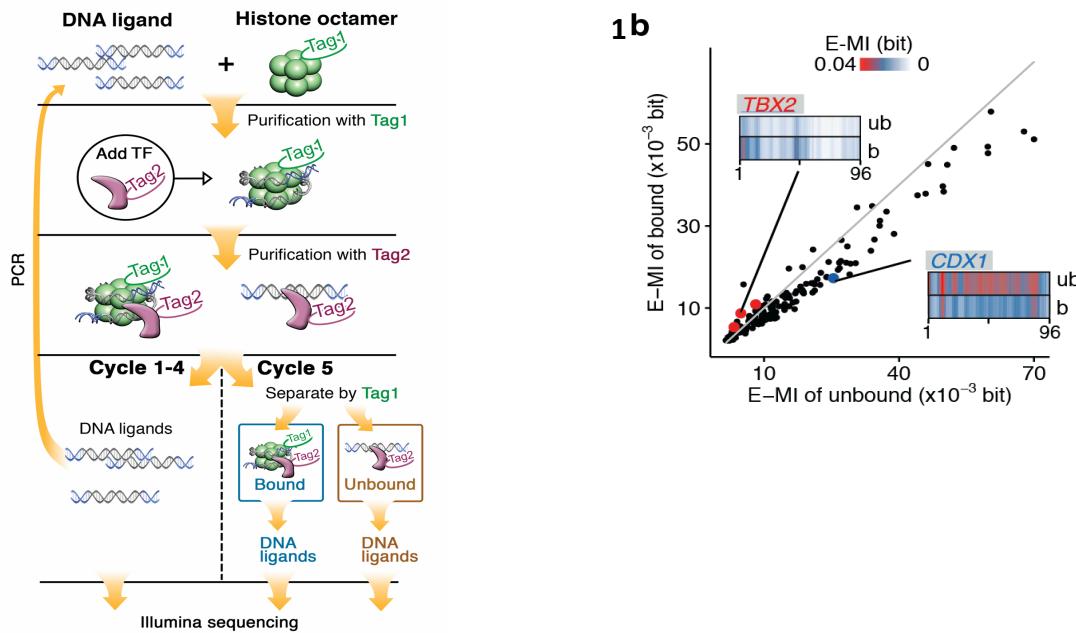
#### Comparison of TF's affinity to their motifs on free-DNA and on nucleosome

In the last cycle of NCAP-SELEX, after the incubation of nucleosome and TFs, the nucleosome-bound and unbound fractions were separated (Fig a). When we compare these two libraries (Fig b), most TFs have stronger signal on the unbound library. A few have stronger signal on the bound library.

If the dissociation of nucleosome is reaching an equilibrium in the experiment, the result would suggest that most TFs are having higher affinity to the free-DNA than to the nucleosomal DNA. However, the equilibrium assumption is less likely to be valid here, as the salt concentration (75 mM) won't allow a dissociated nucleosome to rebind. Thus the result tends to suggest that most TFs are catalyzing the nucleosome to dissociate if they find a good motif there.

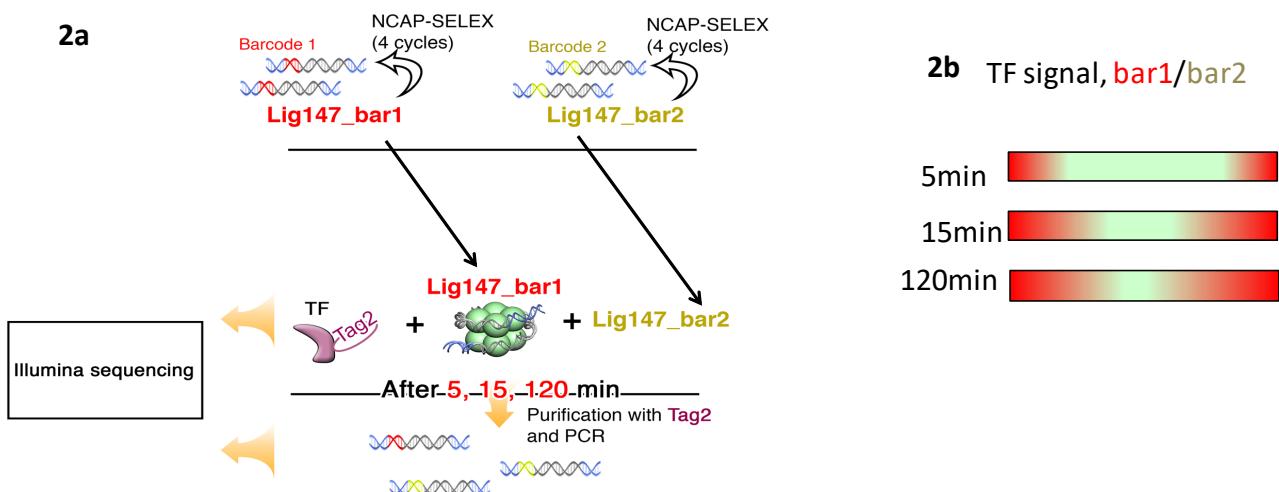
This is obvious as for the dissociation process of a nucleosome, the intermediate state would be the nucleosome with partially detached DNA. To stabilize such state, basically binding of any TF should work; TFs that prefer free DNA definitely works, but probably even a TF that prefers the bent nucleosomal DNA also works, because such TF has a larger probability to be found near a nucleosome, and it screens the charge of DNA there so that the partially detached DNA could be more stable.

Therefore, for the motifs enriched by NCAP-SELEX on lig147, we still lack enough information on whether individual TFs are

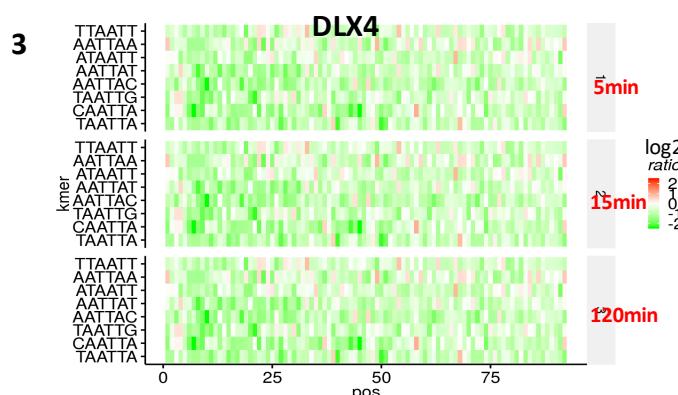


To address this issue, similar to the final round of methyl-SELEX, we designed the following experiment (Fig 2a) using FL TF. Fangjie Zhu

Pulling with tag2 at 3 consecutive time points was aiming to capture the different kinetics of TF binding to different positions of nucleosomal DNA (Fig. 2b, the expectation was that motifs at the center could have a slower kinetics for TF binding)



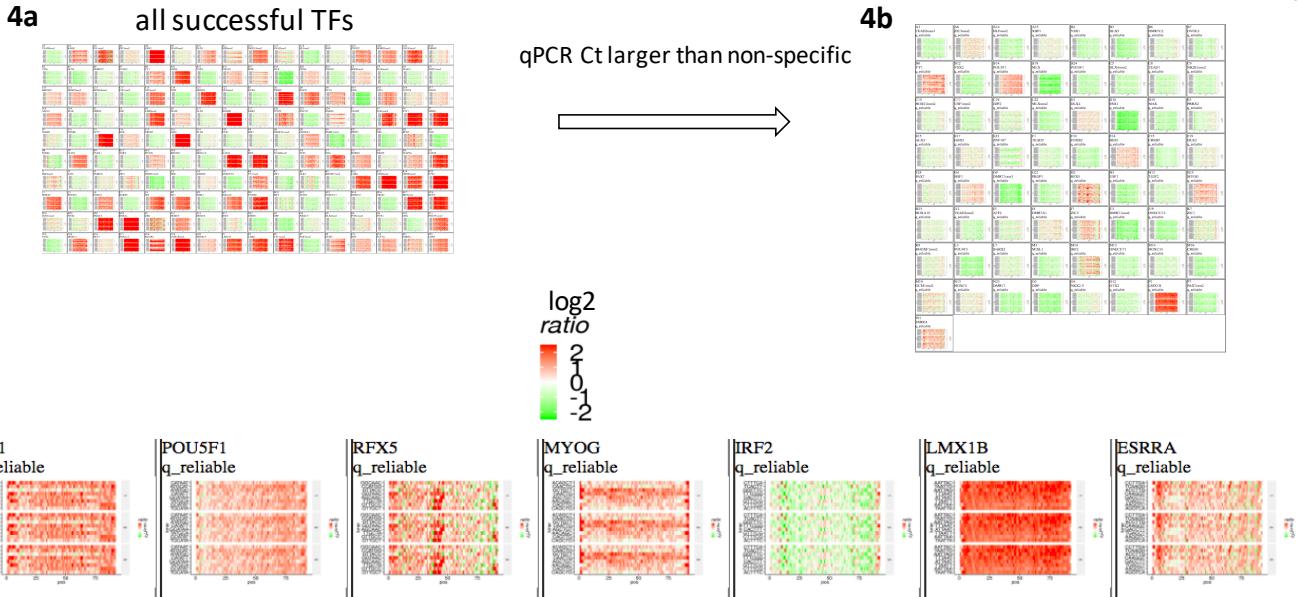
The mixture of bar1 and bar2 before selection ,and the 3 libraries after selection are both sequenced. For each TF, by comparing fold change of 6mers between its cyc 4 and cyc3 libraries, the top 8 most enriched 6mers were selected for mapping. Then, for each library, each of the 6mers were counted for both bar1 and bar2 with positional information, and the ratio of bar1\_count/bar2\_count were calculated also with positional information. Subsequently, the bar1\_count/bar2\_count ratios of "5min, 15min, 120min" libraries were normalized against the "before" library and log2 was taken. The result is visualized as follows for each TF (Fig3).



In this analysis, the high non-specific adsorption of nucleosome to His beads again became a strong disturbing factor. If looking at all the successful TFs (Fig. 4a), it seems to suggest that after 4 cycles of NCAP-SELEX, the sites enriched for some TFs indeed has higher affinity to the TF when it is wrapped onto the nucleosome. However, after we filter for the wells where the Ct in qPCR is larger than the non-specific adsorption level of nucleosome, as in Fig 4b, then the result suggest that even for the sites enriched by NCAP-SELEX, they still have higher affinity to TFs when they are not wrapped onto a nucleosome.

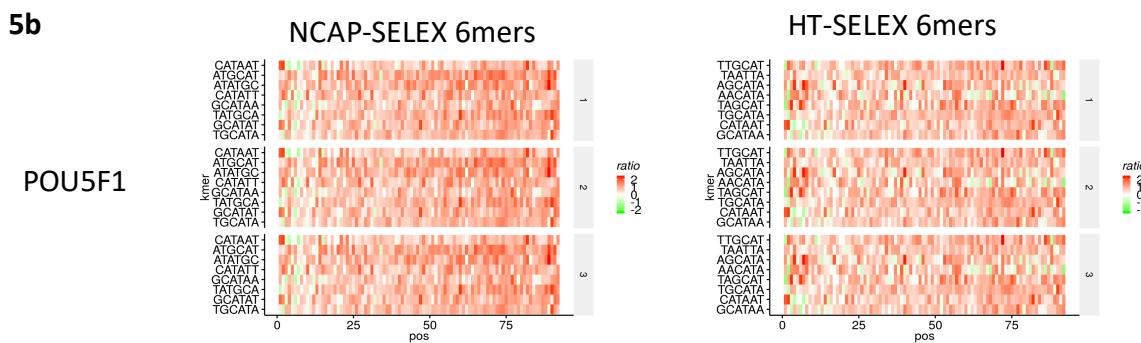
The red wells in Fig 4b usually has a qPCR Ct value closer to the cutoff ns-adsorption Ct, so still hard to tell if they really prefer nucleosomal DNA sites better than the same sites on free DNA. Some red wells are indicated in Fig 4c. LMX1B is a really distinct hit and maybe worth further validation. Also POU5F1 is there, it would be nice if it is true.

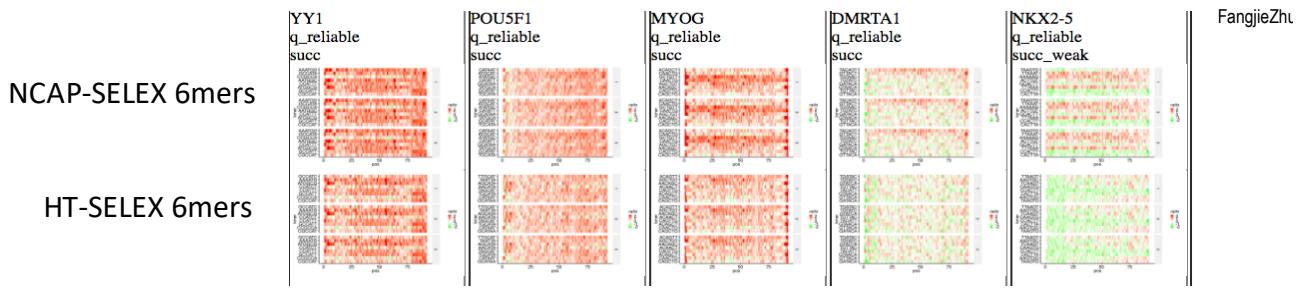
In Fig 4b we also see that some TFs affinity ratio also show positional dependence. For example YY1, IRF2 and RFX5. Interestingly , the binding of RFX5 towards the dyad region of nucleosomal DNA is more favored than binding to the free DNA, which corresponds to its positional preference of binding. Likely at the dyad there RFX5 can form additional contacts with the nucleosome to stabilize the binding.



Instead of using top 6mers enriched in NCAP-SELEX, top 6mers enriched in HT-SELEX also gives a similar result (Fig 5a), suggesting TFs do not like sites on nucleosomal DNA. Just for some TFs, their affinity to HT-SELEX sites on nucleosome is weaker than their affinity to NCAP-SELEX sites on nucleosome (Fig 5b), e.g., POU5F1. This suggests their specificity change when binding the nucleosome.

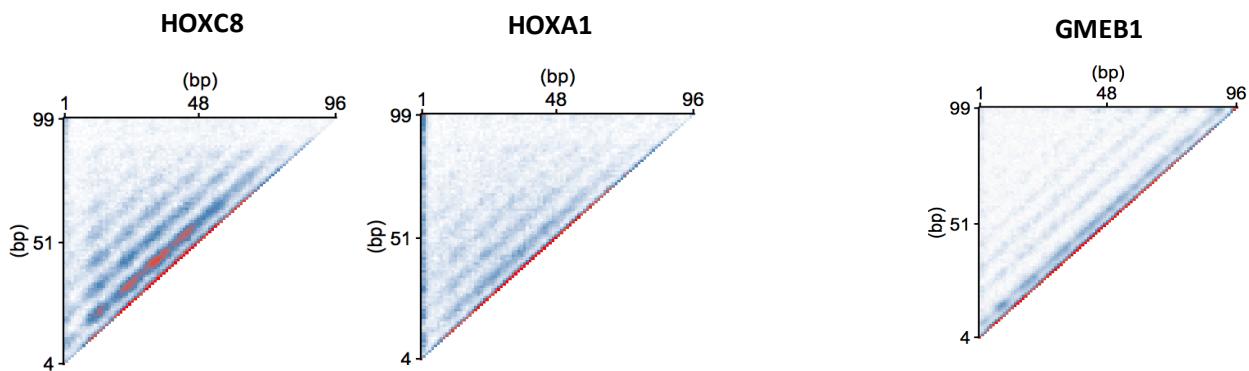
Finally, a bit unexpected but no kinetic effect has been observed





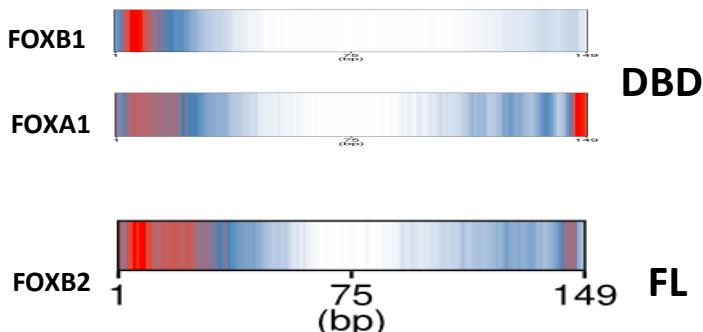
## &gt; preferred spacing of TF FL without nucleosome

in the control HT-SELEX using FL TF proteins, some TFs, like HOXC8, HOXA1, and GMEB1 showed stripes in there E-MI pattern, indicating their preferred spacing of dimer binding which is likely due to the reported allosteric effect mediated by DNA.



## &gt; fox bind to nucleosome better with full length?

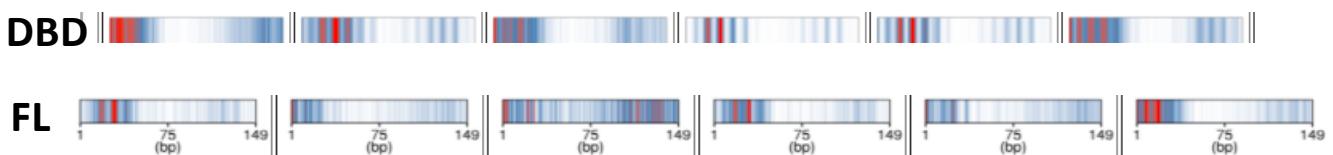
the result with lig200 seems to suggest FL FOX is binding better the nucleosomal DNA than DBD, but sample size is too small and needs further validation



## &gt; some of the signals not validated

- \_ 40bp and 80bp spaced dimer bindings of T\_box and Ets
- \_ binding towards the center by SOX

all FL SOXs failed in lig147 nucleosome SELEX, by most of them are successful for TFctrl and lig200. With lig200 the signal seems to correspond to DBD signal (Fig), but the penetration is a bit shallower which suggests that DBD binds better the nucleosomal DNA.



### > long distance MI of nucleosome comes from periodic repeats

While TFs usually raise MI in neighbouring base pairs, nucleosome is able to induce MI across long distances, and in a 10 bp periodicity (**Fig A**). Because MI is a very general measure and may include all kinds of bias, it is worth to further look into the bias actually involved.

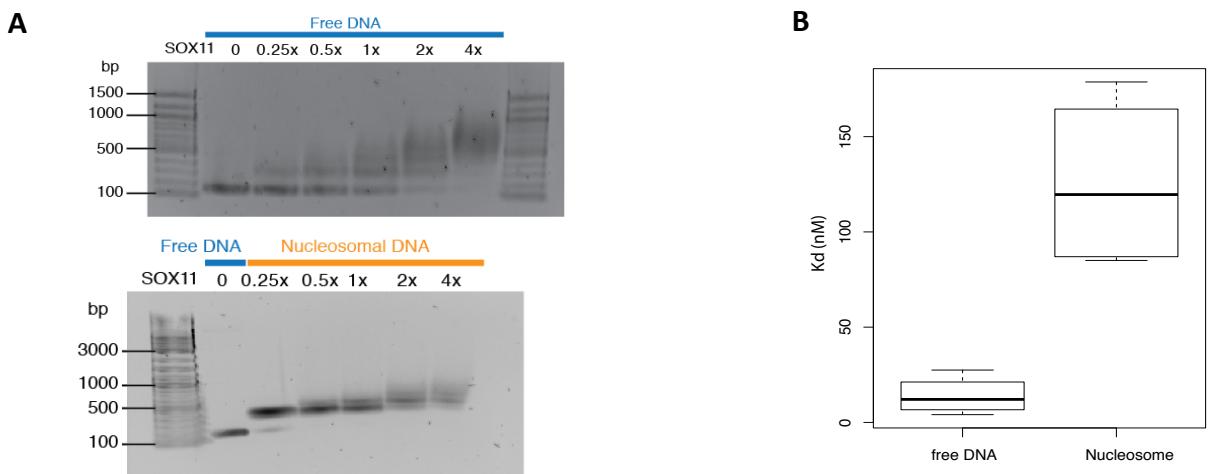
As shown in **Figure B**, when we focus on the bias from pairs of same 3-mers, the 10-bp stripes are even more distinct than observed for the total MI (**Fig A**), therefore much of the MI contrast may come from the oligonucleotide repetition in 10 bp intervals. It is understandable as the positions on nucleosomal DNA are almost identical after a translation in multiples of 10 bp.



### > Kd from EMSA

Svetlana from Patrick lab offered EMSA pictures of SOX11-DNA and SOX11-nucleosome. From the EMSA gel (**Fig A**), and with the provided DNA conc (1 $\mu$ g in 40  $\mu$ L) and SOX/DNA ratio, the Kd value is estimated by the intensities of the shifted and the original bands. A rough estimation (using individual lanes rather than curve fitting of all lanes) indicates the Kd of SOX11 to free DNA and to Nucleosome differs by 1 order (**Fig B**).

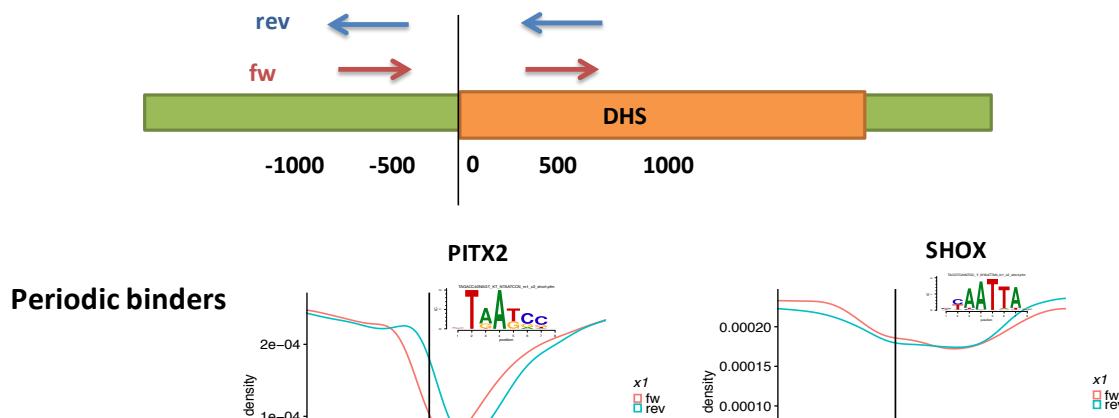
It could be that the estimation is not accurate, or also could be that the nucleosome generally inhibits the binding of SOX11 DBD, and the dyad is only relatively more accessible due to that there is only 1 DNA helix there.

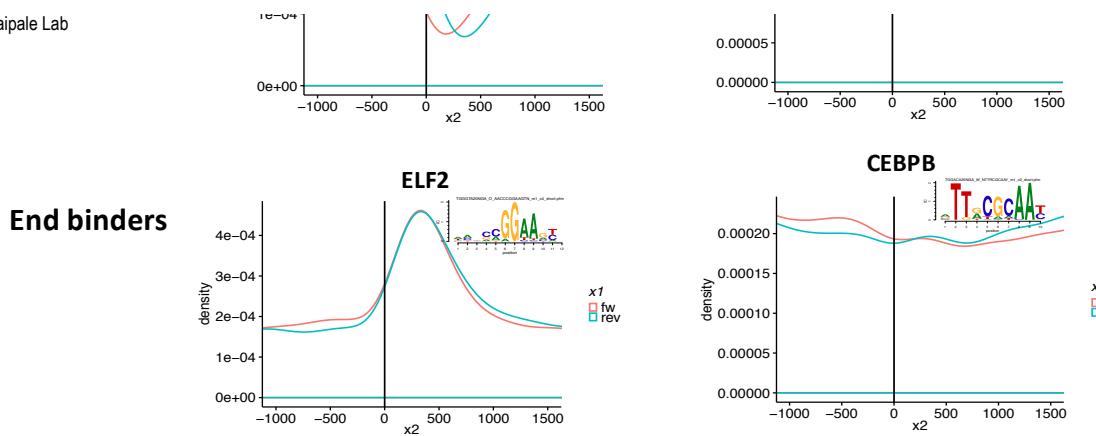


### > relative positioning of TF and DHS

As the relative positioning of TF and nucleosome may correlate with the positioning of TF binding sites (TFBS) in the DHS, e.g., the TFs that are less compatible to bind nucleosomal DNA may define the boundary of DHS, and prevent nucleosome from diffusing into the DHS. Thus the relative positioning of TFBS to the DHSs in LoVo cells were checked.

Might not be meaningful, I first checked the distribution of MOODS hits of each TF's motif against the DHSs. The following figures show a few examples. Some of the TFs show patterns at the DHS borders (ELF2 and PITX2), but this is more likely due to the oligonucleotide bias across the DHS borders, rather than the actual TF positioning.



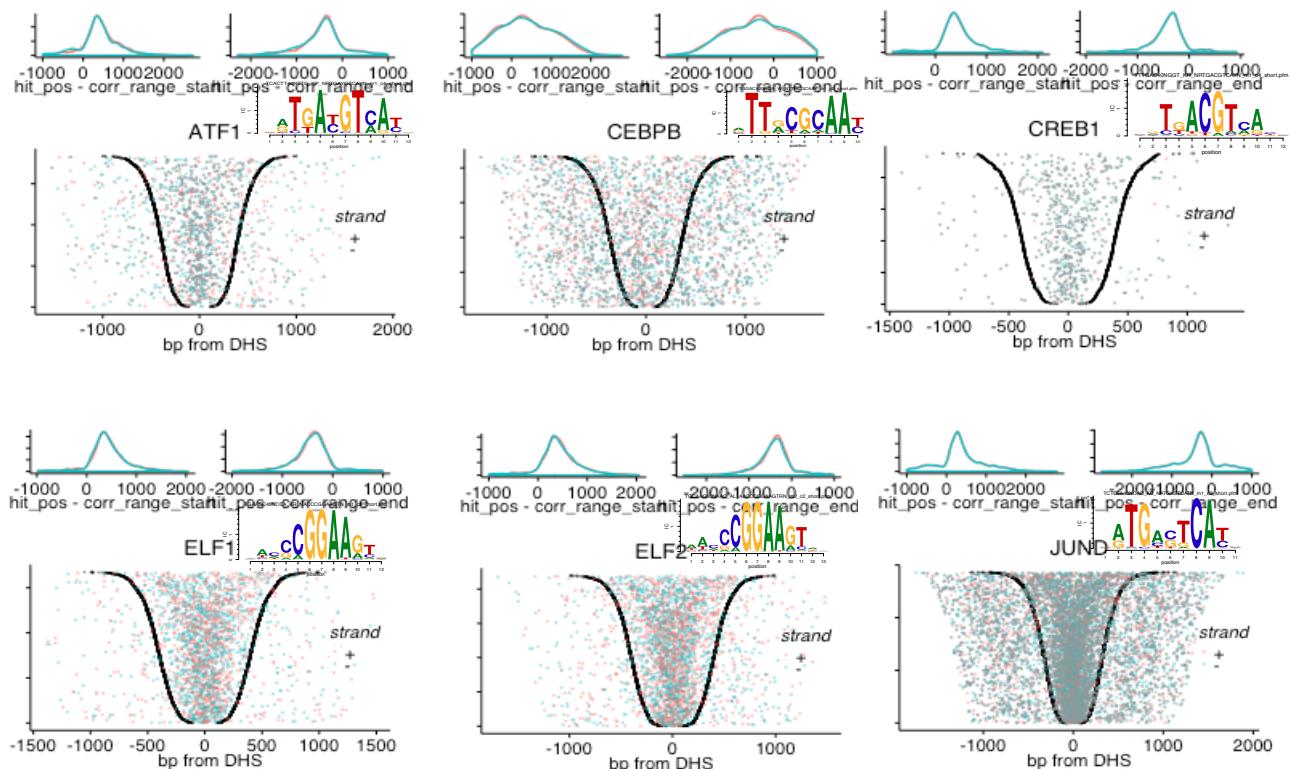


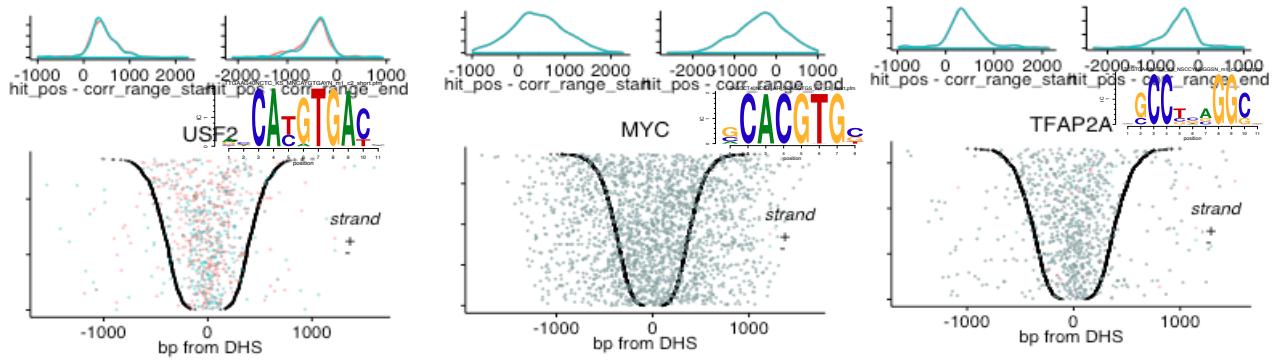
A better strategy is to use ChIP-seq data in combination, that is, to consider only the motif-hits in ChIP-seq peaks, and examine their relative positioning towards the border of DHSs.

With our previous ChIP-seq data, I picked the good-quality ChIPs for TFs which are also successful in NCAP-SELEX, and checked the relative positioning of the DHSs and the motif hits inside ChIP peaks. As shown in the following fig., the black lines indicate the border of DHSs, and the red/blue dots represent TF motifs in ChIP peaks, many TFs just have more hits inside the DHSs, rather than defining its border. Thus the border is less likely maintained by these TFs.

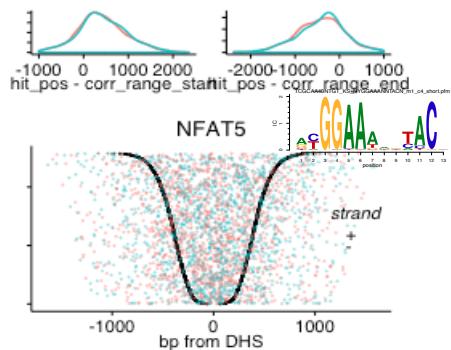
We hoped to observe a difference between the end binder TFs and the periodic binder TFs, however, as listed in the table, we only have ChIP data for 1 plausible periodic binder (NFAT5, which is not successful in SELEX but all other NFAT5 bind periodically). There seems to be more motif hits close to the border of DHS, but the sample size is too small and hard to conclude. We can try to also look at ChIP & DHS data for other cell lines to test more periodic binders. Plotting the accumulated marginal distribution of hit positions against both boundary of DHS revealed a little asymmetric effect for ELF1, hit very weak

### End binders





### Periodic binders



batch	TF	end_or_start	family
LoVo12	ATF1	end	1.1.7 CREB-related
LoVo12	CEBPB	end	1.1.8 C/EBP-related
LoVo12	ELF1	end	3.5.2 Ets-related
LoVo12	ELF2	end	3.5.2 Ets-related
LoVo12	JUND	end	1.1.1 Jun-related
LoVo12	MYC	end	1.2.6 bHLH-ZIP
LoVo12	NFAT5	periodic	6.1.3 NFAT-related
LoVo12	TFAP2A	end	1.3.1 AP-2
LoVo12	USF2	end	1.2.6 bHLH-ZIP
LoVo3	FOXA2	end	3.3.1 Forkhead box (FOX)
LoVo3	MAFG	end	1.1.3 Maf-related
LoVo4	CREB1	end	1.1.7 CREB-related

ChIP-seq data usually contains limited number of peaks, which may lead to noisy result. Bussemaker's paper {Kribelbauer, 2017 #1790} suggested a method to pick all motif hits with considerable ChIP-seq signal even not in a peak for analysis. Such method gives robust statistics and may worth a try.

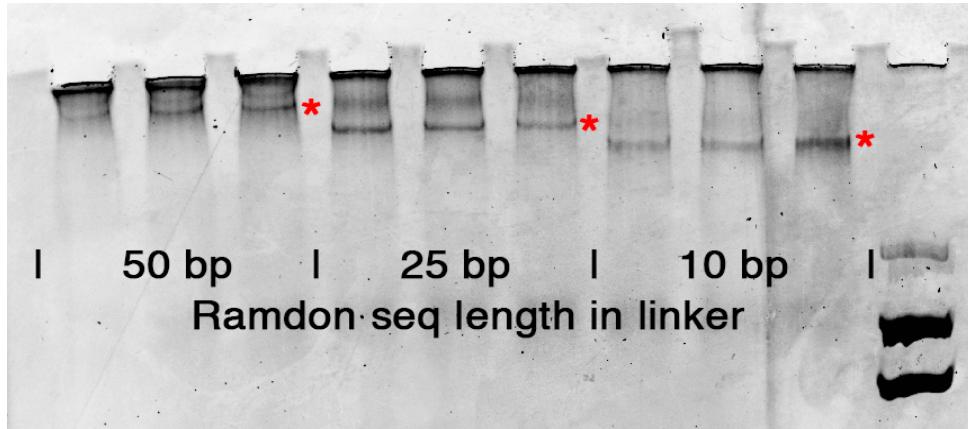
### > Nucleosome Linker-DNA SELEX

The nucleosome *in vivo* is locked by the linker histones. The ability of TFs to compete with the linker histones correlates with their ability to open the chromatin. Such mechanism was proposed especially for pioneer factors for example FOXA1, which has a structure similar to the linker histones. And validated both *in vitro* and *in vivo* {Iwafuchi-Doi, 2016, Mol Cell} {Cirillo, 1998, EMBO J}. Therefore, we hoped to systematically study the ability of TFs to bind to the linker-DNA region of a nucleosome.

We designed SELEX ligands with 10, 25, 50 bp in random at the linker region when reconstituted into a nucleosome. The center region of these ligands consists of the Widom601 sequence which serves to position the nucleosome.

The synthesis of dsDNA from ss oligo for these ligands took a bit effort but not too much. The reconstitution of nucleosome onto these ligands were done following the previous protocol. The result of reconstitution is quite promising, as the nucleosome is very well positioned on these ligands by Widom601. As can be seen on the following gel, the nucleosome band (\*) is very sharp although the ligands length is well above 147bp (213/243/293 bp, respectively).

Now SELEX is done for the ligand with 25bp linker and the remaining is ongoing



### > Future plan

> **For NCAP-SELEX:** Perform ChIP-seq using antibodies both with MNase digestion and without MNase digestion. For structural evidence, maybe we can also consider using NMR. The methyl-TROSY NMR allows huge molecules to generate signals only in a few points as reporters to avoid confusing and signal broadening. We can have a few reporters on TF and histones/DNA, then we should be able to tell which position of the nucleosomal DNA is mostly interacted with TF, thus verifying the NCAP-SELEX result.

> **Linker-DNA SELEX for mono, di, tetra-nucleosome**

> **SELEX of Linker-histones specificity.** The crystal structure suggested a lot of DNA contacts of linker histone with the entrance/exit/dyad of the nucleosome, might be interesting to look into this and compare between different var of linker histones.

> **SELEX of TFs in the presence of linker nucleosome**

> **SELEX of nucleosome remodeler.** In the nucleosome meeting, many research were about the chromatin remodelers, both structural and functional. What we can do is to address the sequence preference of the remodeler themselves, and which TF promotes the remodeling event. There is a nice way to separate the remodeled ligand from unchanged by their different EMSA shift (this is a standard in vitro function assay for the remodelers). Also LNA-based approach might also be promising for HT. LNA (locked nucleic acid) has been used to purify DNA with specific sequence by invading into dsDNA, probably for the nucleosome-wrapped ligand such invasion can be inhibited to some extent. If we load nucleosome initially on the fixed part of a SELEX ligand and then if it is remodeled away, we should be able to purify only the ligands with shifted nucleosome in HT.

Similar remodeling assay probably can also be carried out using Bei's nuclear extract. Less related but probably also interesting is, with nuclear extract, to test if any TF can load nucleosome onto free-DNA SELEX ligand. Technically, after incubating free-DNA ligand with nuclear extract, we can pull down using histone antibody, and then amplify with the adapter seq on ligand.

> **experiment/data analysis of dissociation data, and equilibrium binding assay**

> **minor-groove-binders' specificity in mammalian cells and in insect cells**

> **address pioneer factor more clearly**

– Test if DNA curvature affects TF binding. I discussed with people in Bejorn's lab, they have available protocol to make sharply looped ds-DNA (70-80 bp circle), and they think the bending will be relatively homogeneously distributed across the ligand due to the persistent length of DNA.

– X-ray/light scattering of TF/nucleosome mixture to monitor the whole process of pioneer factor action.

– Play with polyamine to design artificial pioneer TFs. We can change the sequence-specificities, length, numbers of binding sites, spacing between binding sites etc. of the artificial polyamine, and see which one best bind/dissociate a nucleosome.

> **Deep learning model for nucleosome signal**

– the plan is to first try to discriminate the bound and unbound library of nucleosome.

> **RNA Pol selectivity, and the in-vitro version of starr-seq**