Published in final edited form as:

Mol Genet Genomics. 2020 July ; 295(4): 825–835. doi:10.1007/s00438-020-01675-9.

Pioneer factors and their *in vitro* identification methods

Xinyang Yu¹, Michael J. Buck^{2,3}

¹Zhuhai Precision Medical Center, Zhuhai People's Hospital (Zhuhai Hospital Affiliated with Jinan University), Zhuhai, Guangdong 519000, P.R. China

²New York State Center of Excellence in Bioinformatics and Life Sciences and Department of Biochemistry, State University of New York at Buffalo, Buffalo, New York, 14203

³Department of Biomedical Informatics

Abstract

Pioneer transcription factors are a special group of transcription factors that can interact with nucleosomal DNA and initiate regulatory events. Their binding to regulatory regions is the first event in gene activation and can occur in silent or heterochromatin regions. Several research groups have worked to define pioneer factors and studied their binding characteristics using various techniques. In this review, we will describe the *in vitro* methods used to define and characterize pioneer factors with particular attention to differences in methodology and how these differences can affect the results.

Keywords

Pioneer factors; transcription factors; nucleosomes; binding assays

Eukaryotes such as humans are pluricellular organisms that are composed of multiple cell types, thus forming different tissues (Sapp, 2005). The human genome consists of more than 3 billion base pairs of DNAs packaged into 23 chromosomes, with approximately 30,000 protein-coding genes (Frazer, 2012; Pennisi, 2003). To establish competence for specific cell fates and guarantee the development of different cell lineages, cell lineage-specific spatial and temporal regulation of the expression of various genes is required (Pope & Medzhitov, 2018; Wittkopp, 2007). This cell type-specific regulation occurs largely at the transcriptional level, with transcription factor (TF) binding playing a dominant role. TFs regulate the transcription of genes via binding to specific regulatory DNA sequences, i.e., transcription factor binding sites (TFBS) (Spitz & Furlong, 2012). Typically, TFs recognize and bind to relatively short degenerative DNA sequences. Due to the degenerative nature of TFBS, they can appear millions of times in a genome. The majority of TFBS are never bound because most TFBS are inaccessible within chromatin.

Eukaryotic DNA is tightly packaged into a highly organized structure called chromatin. Chromatin is composed of nucleosomes that consist of nucleosome core particles of 147 bp

double-stranded DNA that wraps 1.67 turns around a histone core containing 2 units of H2A, H2B, H3 and H4 (Jiang & Pugh, 2009; Kornberg & Lorch, 1999). Nucleosomes are then connected by a short stretch of DNA called a linker and to form a ‘beads-on-a-string’ structure (Olins & Olins, 2003). Chromatin can be further compacted by H1, which helps to stabilize the higher-order chromatin fiber through binding to the linker DNA (Ronald Berezney; R. Berezney, Mortillaro, Ma, Wei, & Samarabandu, 1995). Importantly, all these histone proteins are positively charged, allowing strong interaction with the negatively charged phosphate backbone of DNA. As a result, the DNA is wrapped tightly around the histone proteins, rendering the nucleosomal DNA inaccessible to binding by most TFs (Luger, Dechassa, & Tremethick, 2012; Vermaak, Ahmad, & Henikoff, 2003).

Reading inaccessible regulatory regions and creating accessible DNA for gene regulation is a prerequisite for most developmental processes. Pioneer factors are a special class of TFs that can bind to their target sequences even buried inside chromatin. Pioneer factors have been shown to be essential regulatory proteins during organismal development, are often misregulated or mutated in cancer, and are the key TFs in generating induced pluripotent cells. In this review, we focus mainly on the *in vitro* analysis methods being applied to study TF binding capabilities and identify pioneer factors. We believe that improvement and progress on these *in vitro* studies will finally contribute to understanding more complicated gene regulation mechanisms *in vivo*.

Different pioneer factors display various nucleosome binding patterns

The concept of pioneer factor first arose in 2002, when researchers found that FOXA and GATA4 can bind to nucleosome arrays *in vitro*, while other TFs such as NF1, C/EBP, and GAL4-AH cannot (Cirillo et al., 2002). They proposed that these nucleosome-binding-capable TFs might represent a new functional class of regulatory proteins and named them ‘pioneer’ factors, indicating that their binding occurs prior to other events. To date, the well-characterized pioneer factors include FOXA proteins (Cirillo et al., 2002; Cirillo et al., 1998), Yamanaka factors such as SOX2, KLF4 and OCT4 (Soufi, Donahue, & Zaret, 2012), and some other TFs such as ASCL1, PAX7, PU.1, GATA 4 and P53 (Barozzi et al., 2014; Budry et al., 2012; Cirillo et al., 2002; Iwafuchi-Doi & Zaret, 2016; Laptenko, Beckerman, Freulich, & Prives, 2011; Lidor Nili et al., 2010; Soufi et al., 2015). Although these pioneer factors can overcome the barrier of nucleosomes and bind to nucleosomal DNA, they display various binding patterns.

The canonical pioneer factor FOXA (Cirillo et al., 2002; Cirillo et al., 1998) can bind the nucleosome dyad and displace the linker histone to maintain nucleosome accessibility (Iwafuchi-Doi et al., 2016). SOX2 and OCT4 can bind to sites in heterochromatin regions (Soufi et al., 2012), but heterochromatin does impede their binding to some extent (Onder et al., 2012). p53 prefers binding sites located near the nucleosome edge rather than the center (Laptenko et al., 2011; Yu & Buck, 2019), and it binds target sequences on chromatin with higher affinity than sites located in naked DNA (Espinosa & Emerson, 2001). Another noteworthy point is that even for canonical pioneer factors such as SOX2 and OCT4, they can bind to only approximately 4% of their total available recognition sizes in the whole genome (Soufi et al., 2012). Overall, as additional pioneer factors are characterized, it is

becoming clear that they are diverse in terms of nucleosome interaction and histone modification dependency.

Recent studies have started to uncover inconsistencies in pioneer factor binding specificity to nucleosomes. Early investigations of FOXA1 discovered that FOXA protein can interact with target sequences located near the nucleosome dyad (Cirillo et al., 1998; McPherson, Shim, Friedman, & Zaret, 1993). In contrast, a recent approach utilizing a new high-throughput approach, NCAP-SELEX, found that FOXA protein cannot bind directly to the nucleosome dyad (Zhu et al., 2018). Moreover, the results indicated that nucleosomal DNA length will also affect TF binding to nucleosomes, which might be a caveat regarding *in vitro* nucleosome binding assays (Zhu et al., 2018). Studies on p53 protein also show divergence. An *in vitro* study indicated that the rotational positioning of the binding site affects p53 binding, and p53 can bind to target sequences located near the nucleosome dyad when located properly (Sahu et al., 2010). Meanwhile, our previous recent study using both *in vitro* and *in vivo* approaches demonstrated that p53 displays much higher binding affinity to nucleosome edges than the dyad (Yu & Buck, 2019). The difference between the findings of these studies is likely due to differences in key aspects of the experimental approach.

Here, we want to compare the methods that have been applied to test TF-nucleosome binding *in vitro* to determine whether any differences in the experimental design or conditions utilized might in turn affect the result. The techniques or materials applied in each experiment might differ from laboratory to laboratory, which might be why the results of pioneer factor-nucleosome *in vitro* binding assays vary among researchers. The important elements that should be taken into consideration include 1) the DNA template utilized as the nucleosome positioning sequence; 2) histone variants and modifications; 3) the composition of the binding buffer; 3) the final nucleosome concentration in the reaction system; and 4) the method applied to detect the binding complex. Below, we will discuss each of these elements in detail.

1. Nucleosomal DNA

The nucleosome will form on a piece of double-stranded DNA in a characteristic manner depending on the sequence (Kaplan et al., 2009; Yuan et al., 2005). Some specific sequences will consistently position the nucleosome in the same manner and thus have strong nucleosome positioning capability (Lowary & Widom, 1998). Other sequences of DNA can have varied nucleosome positions, causing the population of nucleosomes on that DNA to be located in different places. Currently, several *in vitro* nucleosome positioning sequences are in common use, from natural sequences such as 5S rDNA (Simpson & Stafford, 1983), MMTV LTR (Richard-Foy & Hager, 1987), alpha satellite sequences (Yang, Hansen, Oishi, Ryder, & Hamkalo, 1982) and sequences derived from *in vivo* regions that have been found to be occupied by nucleosomes (Darvekar, Johnsen, Eriksen, Johansen, & Sjøttem, 2012; Soufi et al., 2015) to artificial synthetic sequences such as Widom 601 DNA (Lowary & Widom, 1998), TG repeated units (Godde & Wolffe, 1996; Shrader & Crothers, 1989) and randomized DNA sequences (Zhu et al., 2018). Based on the goals of studying nucleosome behavior, different positioning sequences will be selected. Unlike other known nucleosome positioning templates that usually form labile nucleosomes that slide readily, Widom 601

DNA has the strongest nucleosome positioning ability and extremely high thermodynamic stability (as high as 55 degrees) (Senavirathne, Mahto, Hanne, O'Brian, & Fishel, 2017). Widom 601 DNA's structural stability means it is not an ideal template to study nucleosome dynamics, but it can generate a homogenous population of nucleosomes and thus serves as an ideal model to test how TFBS location on a nucleosome can affect TF binding. Different DNA sequences being applied to assemble nucleosomes *in vitro* are summarized in Table 1. Alternatively, the micrococcus nuclease (MNase) enzyme can be used to obtain naturally existing nucleosomes from cells (Darvekar et al., 2012).

2. Histone variants and modifications

The nucleosome template DNA is mixed with histone proteins at an optimized ratio, and nucleosomes are formed by salt gradient dialysis (Dyer et al., 2004; Lee & Narlikar, 2001), followed by sucrose gradient centrifuge or glycerol gradient centrifuge to pick purified nucleosomes from unbound DNA and histones (Lusser & Kadonaga, 2004; Noll & Noll, 1989; Schnitzler, 2001). In addition to commonly used histone proteins such as H2A, H2B, H3, and H4, other histone variants can also be applied to assemble nucleosomes *in vitro*. For example, H2A.Z and H3.3, which are variants that occupy approximately 10% of total H2A or H3 proteins *in vivo* (Pina & Suau, 1987), have been utilized to perform nucleosome reconstitution (Osakabe et al., 2018; Thakar et al., 2009). However, nucleosomes containing a combination of certain histone variants have been tested and found to be particularly unstable (Jin et al., 2009; Zlatanova & Thakar, 2008) and thus unsuitable for obtaining a homogeneous nucleosome population *in vitro*. For the purpose of examining histone posttranslational modifications, histone proteins with modified tails were also generated by applying chemically modified histone peptides (Chatterjee & Muir, 2010; Yun, Ruan, Huh, & Li, 2012).

3. Binding buffer

To mimic the biological environment where protein-DNA binding events occur *in vivo*, researchers utilized buffer systems such as Tris-HCl or HEPES to maintain a relatively stable physiological pH in the reaction mix (Durst & Staples, 1972; Vega & Bates, 1976). Salts are also included in the solution, but usually at low concentrations. As shown in Table 2, researchers usually included 50 to 100 mM sodium chloride or potassium chloride in the binding buffer to maintain sufficient ionic strength in solution to prevent nonspecific binding without breaking the specific interactions between proteins of interest and phosphate groups on the DNA backbone. In addition to monovalent cations, divalent salts were also added, such as Mg^{2+} and Zn^{2+} , which have been shown to stabilize the DNA duplex structure (Chiu & Dickerson, 2000; Nakano, Fujimoto, Hara, & Sugimoto, 1999; Owczarzy, Moreira, You, Behlke, & Walder, 2008; Tan & Chen, 2006). On the other hand, salts in an appropriate concentration range can also enhance protein solubility and prevent protein aggregation (Bondos & Bicknell, 2003). Reducing agents, namely, DTT or 2-mercaptoethanol, are also added to mitigate oxidation damage (Lebediker & Danieli, 2014; Wingfield, 2001). Moreover, additives such as BSA, glycerol, and nonionic detergents NP-40 can also be added to promote protein solubility and enhance protein-DNA stability (Lebediker & Danieli, 2014). Protease inhibitors such as aprotinin, leupeptin, PMSF, or EDTA can be

added when using cellular protein extracts. Some research groups also add poly(dI.dC), salmon sperm DNA or BSA to further reduce the nonspecific binding (Larouche, Bergeron, Leclerc, & Guerin, 1996). The table below summarizes the binding buffers utilized to test DNA-protein interactions (Table 2).

4. Nucleosome concentration

The nucleosome concentration in the binding assay is another important element that affects binding and nucleosome stability. Previously, many studies utilized radiolabeled probes to detect supershifted TF-nucleosome complexes. Radioactive techniques have high sensitivity, allowing small quantities of nucleosomes to be used in the assays. For example, Angelov's group applied as little as 0.5 nM nucleosomes to bind NF- κ B (Angelov et al., 2004), and another group used nucleosomes at the picomolar level to interact with GR proteins (Li & Wrangé, 1993, 1995). However, nucleosomes tend to partially disassemble at subnanomolar concentrations, and subnucleosomal structures such as hexasomes or tetrasomes appear (Y. Chen et al., 2017; Lai & Pugh, 2017; Rhee, Bataille, Zhang, & Pugh, 2014). These subtypes can also be obtained by adjusting the final salt concentration (Arimura, Tachiwana, Oda, Sato, & Kurumizaka, 2012; Krajewski, Li, & Dou, 2018).

A few studies have demonstrated the occurrence of subnucleosomal structures at low nucleosome concentrations (Claudet, Angelov, Bouvet, Dimitrov, & Bednar, 2005; Kelbauskas, Sun, Woodbury, & Lohr, 2008). The researchers observed the binding of NF- κ B to its recognition site located near the nucleosome dyad at a nucleosome concentration of only ~7.5 nM, but they failed to detect binding at 40 nM nucleosome (Lone et al., 2013). They obtained partially disassembled subnucleosomal particles by diluting nucleosomes to approximately 10 nM. Under these conditions, H2A/H2B dimers were eliminated, and the researchers believed that this substructure decreased the hindrance to NF- κ B access, thus allowing interaction to occur. Taking all these results into consideration suggests 10 nM as a threshold concentration for *in vitro* reconstituted nucleosome core particles to remain intact and relatively stable (Claudet et al., 2005; Kelbauskas et al., 2008; Lone et al., 2013).

5. Quantifying TF-nucleosome binding

The electrophoretic mobility shift assay (EMSA), a traditional technique that observes and quantifies protein-DNA interactions, can also be applied to detect TF-nucleosome binding *in vitro* (Dey et al., 2012). It is based on the principle that protein-DNA complexes are larger and thus run more slowly than unbound free DNA, allowing easy separation into at least two different-sized bands on the gel. Researchers have been using either agarose gel or nondenaturing polyacrylamide gel to visualize the TF-nucleosome complexes. Polyacrylamide gel is a better band-separating system than agarose gel because it has greater resolving power, presents sharper bands, and generates very pure recovered DNA (Guilliatt, 2002; Tietz, 1998). On the other hand, the nucleosome itself is already a DNA-protein complex, and when TFs bind, even more complicated protein-DNA-protein supercomplexes will be generated. The "cage effect" of the polyacrylamide gel contributes to stabilizing this large complex, thus efficiently preventing dissociation as the supercomplex moves through the gel (Nelson, Hendy, Reid, & Cavanagh, 2002; Vossen & Fried, 1997). This technique has

been further enhanced to determine the affinity of different proteins to their corresponding target by applying radiolabeled or biotin-labeled probes. By measuring the signal emitted from each band, which corresponds to the concentration of probes being shifted away, researchers can quantify the templates bound by the tested protein, calculate the K_D values, and finally deduce the binding affinity.

One major limitation of EMSA is that it can test the TF binding to only one nucleosome type at a time. To test thousands of different nucleosomal DNA templates, researchers must perform thousands of binding assays and run thousands of gels, which is time consuming and not feasible. Recently, a few laboratories have used high-throughput sequencing to investigate TF binding to large nucleosome libraries simultaneously (Yu & Buck, 2019; Zhu et al., 2018). Our study tested p53 protein binding to multiple nucleosome types simultaneously through just one binding reaction and then utilized next generation sequencing (NGS) to obtain the binding frequency of each nucleosome type (Yu & Buck, 2019). Taipale's group filtered randomized DNA sequences and retained the ones that were favorable for nucleosome formation and then tested the binding of 220 TFs to those random nucleosomes individually and performed computational analysis to discover motifs on those nucleosomal DNAs to define where the TFBS were located on each nucleosome (Zhu et al., 2018).

Future questions for studying pioneer factors

The overall goal of investigating pioneer TF-nucleosome binding *in vitro* is to address the following question: what are the rules that dictate pioneer factor binding? In general, several elements could affect pioneer factor binding to the nucleosome, including the following: 1) TFBS positioning. Every single nucleotide has a specific location within a nucleosome known as its translational position (Simpson, 1991), and the rotational position describes the orientation of each nucleotide relative to the histone core (Lowary & Widom, 1997; Simpson, 1991). 2) Histone modification on the histone core. Certain modifications that occur on histone tails, such as histone acetylation, can loosen histone-DNA contacts, destabilize nucleosomes, and finally facilitate TF binding (Fenley, Anandakrishnan, Kidane, & Onufriev, 2018; Henikoff, 2008). 3) Presence of cofactors or chromatin remodelers. The interactions between TFs and cofactors are important for transcription regulation (Goi, Little, & Xie, 2013; Siggers, Duyzend, Reddy, Khan, & Bulyk, 2011), while chromatin remodelers regulate chromatin accessibility and affect nucleosome positioning (Langst & Manelyte, 2015; Tyagi, Imam, Verma, & Patel, 2016).

During the past decades, the protein-DNA binding assay has been applied by different labs to test proteins of interest. This traditional technique has been further revised to test pioneer factor-nucleosome binding. However, the experimental conditions used in different labs are slightly different. As noted above, nucleosomes should be maintained above 10 nM in solution to prevent partial dissociation. It is also recommended to use a strong nucleosome positioning sequence to obtain a homogenous nucleosome population when studying the effect of TFBS location on the nucleosome. Taking all these points into consideration, a standard criterion for *in vitro* TF-nucleosome binding assays might be required in the future.

Overall, we believe that these *in vitro* studies will enhance the understanding of pioneer factors and pave the road for future investigation.

References

- Angelov D, Charra M, Seve M, Cote J, Khochbin S, & Dimitrov S (2000). Differential remodeling of the HIV-1 nucleosome upon transcription activators and SWI/SNF complex binding. *J Mol Biol*, 302(2), 315–326. doi:10.1006/jmbi.2000.4069 [PubMed: 10970736]
- Angelov D, Lenouvel F, Hans F, Muller CW, Bouvet P, Bednar J, ... Dimitrov S (2004). The histone octamer is invisible when NF-kappaB binds to the nucleosome. *J Biol Chem*, 279(41), 42374–42382. doi:10.1074/jbc.M407235200 [PubMed: 15269206]
- Arimura Y, Tachiwana H, Oda T, Sato M, & Kurumizaka H (2012). Structural analysis of the hexasome, lacking one histone H2A/H2B dimer from the conventional nucleosome. *Biochemistry*, 51(15), 3302–3309. doi:10.1021/bi300129b [PubMed: 22448809]
- Barozzi I, Simonatto M, Bonifacio S, Yang L, Rohs R, Ghisletti S, & Natoli G (2014). Coregulation of transcription factor binding and nucleosome occupancy through DNA features of mammalian enhancers. *Mol Cell*, 54(5), 844–857. doi:10.1016/j.molcel.2014.04.006 [PubMed: 24813947]
- Berezney R Structural and functional organization of the nuclear matrix [A] (1995) [A] (1995). San Diego [u.a.: Acad. Pr.
- Berezney R, Mortillaro MJ, Ma H, Wei X, & Samarabandu J (1995). The nuclear matrix: a structural milieu for genomic function. *Int Rev Cytol*, 162A, 1–65. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8575878> [PubMed: 8575878]
- Bondos SE, & Bicknell A (2003). Detection and prevention of protein aggregation before, during, and after purification. *Anal Biochem*, 316(2), 223–231. doi:10.1016/s0003-2697(03)00059-9 [PubMed: 12711344]
- Bowers SR, Calero-Nieto FJ, Valeaux S, Fernandez-Fuentes N, & Cockerill PN (2010). Runx1 binds as a dimeric complex to overlapping Runx1 sites within a palindromic element in the human GM-CSF enhancer. *Nucleic Acids Res*, 38(18), 6124–6134. doi:10.1093/nar/gkq356 [PubMed: 20483917]
- Budry L, Balsalobre A, Gauthier Y, Khetchoumian K, Honore A, Vallette S, ... Drouin J (2012). The selector gene Pax7 dictates alternate pituitary cell fates through its pioneer action on chromatin remodeling. *Genes Dev*, 26(20), 2299–2310. doi:10.1101/gad.200436.112 [PubMed: 23070814]
- Chatterjee C, & Muir TW (2010). Chemical approaches for studying histone modifications. *J Biol Chem*, 285(15), 11045–11050. doi:10.1074/jbc.R109.080291 [PubMed: 20147749]
- Chen H, Li B, & Workman JL (1994). A histone-binding protein, nucleoplasmin, stimulates transcription factor binding to nucleosomes and factor-induced nucleosome disassembly. *EMBO J*, 13(2), 380–390. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/8313883> [PubMed: 8313883]
- Chen Y, Tokuda JM, Topping T, Meisburger SP, Pablit SA, Gloss LM, & Pollack L (2017). Asymmetric unwrapping of nucleosomal DNA propagates asymmetric opening and dissociation of the histone core. *Proc Natl Acad Sci U S A*, 114(2), 334–339. doi:10.1073/pnas.1611118114 [PubMed: 28028239]
- Chiu TK, & Dickerson RE (2000). 1 A crystal structures of B-DNA reveal sequence-specific binding and groove-specific bending of DNA by magnesium and calcium. *J Mol Biol*, 301(4), 915–945. doi:10.1006/jmbi.2000.4012 [PubMed: 10966796]
- Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, & Zaret KS (2002). Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol Cell*, 9(2), 279–289. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/11864602> [PubMed: 11864602]
- Cirillo LA, McPherson CE, Bossard P, Stevens K, Cherian S, Shim EY, ... Zaret KS. (1998). Binding of the winged-helix transcription factor HNF3 to a linker histone site on the nucleosome. *EMBO J*, 17(1), 244–254. doi:10.1093/emboj/17.1.244 [PubMed: 9427758]

- Claudet C, Angelov D, Bouvet P, Dimitrov S, & Bednar J (2005). Histone octamer instability under single molecule experiment conditions. *J Biol Chem*, 280(20), 19958–19965. doi:10.1074/jbc.M500121200 [PubMed: 15772075]
- Darvekar S, Johnsen SS, Eriksen AB, Johansen T, & Sjøttem E (2012). Identification of two independent nucleosome-binding domains in the transcriptional co-activator SPBP. *Biochem J*, 442(1), 65–75. doi:10.1042/BJ20111230 [PubMed: 22081970]
- Dey B, Thukral S, Krishnan S, Chakrobarty M, Gupta S, Manghani C, & Rani V (2012). DNA-protein interactions: methods for detection and analysis. *Mol Cell Biochem*, 365(1-2), 279–299. doi:10.1007/s11010-012-1269-z [PubMed: 22399265]
- Durst RA, & Staples BR (1972). Tris-tris-HCl: a standard buffer for use in the physiologic pH range. *Clin Chem*, 18(3), 206–208. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/5020814> [PubMed: 5020814]
- Dyer PN, Edayathumangalam RS, White CL, Bao Y, Chakravarthy S, Muthurajan UM, & Luger K (2004). Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol*, 375, 23–44. doi:10.1016/s0076-6879(03)75002-2 [PubMed: 14870657]
- Espinosa JM, & Emerson BM (2001). Transcriptional regulation by p53 through intrinsic DNA/chromatin binding and site-directed cofactor recruitment. *Mol Cell*, 8(1), 57–69. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/11511360> [PubMed: 11511360]
- Fenley AT, Anandakrishnan R, Kidane YH, & Onufriev AV (2018). Modulation of nucleosomal DNA accessibility via charge-altering post-translational modifications in histone core. *Epigenetics Chromatin*, 11(1), 11. doi:10.1186/s13072-018-0181-5 [PubMed: 29548294]
- Frazer KA (2012). Decoding the human genome. *Genome Res*, 22(9), 1599–1601. doi:10.1101/gr.146175.112 [PubMed: 22955971]
- Galati A, Rossetti L, Pisano S, Chapman L, Rhodes D, Savino M, & Cacchione S (2006). The human telomeric protein TRF1 specifically recognizes nucleosomal binding sites and alters nucleosome structure. *J Mol Biol*, 360(2), 377–385. doi:10.1016/j.jmb.2006.04.071 [PubMed: 16756990]
- Godde JS, & Wolffe AP (1996). Nucleosome assembly on CTG triplet repeats. *J Biol Chem*, 271(25), 15222–15229. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8663027> [PubMed: 8663027]
- Goi C, Little P, & Xie C (2013). Cell-type and transcription factor specific enrichment of transcriptional cofactor motifs in ENCODE ChIP-seq data. *BMC Genomics*, 14 Suppl 5, S2. doi:10.1186/1471-2164-14-S5-S2
- Guillatt AM (2002). Agarose and polyacrylamide gel electrophoresis. *Methods Mol Biol*, 187, 1–11. doi:10.1385/1-59259-273-2:001 [PubMed: 12013737]
- Henikoff S (2008). Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat Rev Genet*, 9(1), 15–26. doi:10.1038/nrg2206 [PubMed: 18059368]
- Howe L, & Ausio J (1998). Nucleosome translational position, not histone acetylation, determines TFIIIA binding to nucleosomal *Xenopus laevis* 5S rRNA genes. *Mol Cell Biol*, 18(3), 1156–1162. doi:10.1128/mcb.18.3.1156 [PubMed: 9488430]
- Iwafuchi-Doi M, Donahue G, Kakumanu A, Watts JA, Mahony S, Pugh BF, ... Zaret KS. (2016). The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation. *Mol Cell*, 62(1), 79–91. doi:10.1016/j.molcel.2016.03.001 [PubMed: 27058788]
- Iwafuchi-Doi M, & Zaret KS (2016). Cell fate control by pioneer transcription factors. *Development*, 143(11), 1833–1837. doi:10.1242/dev.133900 [PubMed: 27246709]
- Jiang C, & Pugh BF (2009). Nucleosome positioning and gene regulation: advances through genomics. *Nat Rev Genet*, 10(3), 161–172. doi:10.1038/nrg2522 [PubMed: 19204718]
- Jin C, Zang C, Wei G, Cui K, Peng W, Zhao K, & Felsenfeld G (2009). H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat Genet*, 41(8), 941–945. doi:10.1038/ng.409 [PubMed: 19633671]
- Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y, ... Segal E (2009). The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature*, 458(7236), 362–366. doi:10.1038/nature07667 [PubMed: 19092803]

- Kelbauskas L, Sun J, Woodbury N, & Lohr D (2008). Nucleosomal stability and dynamics vary significantly when viewed by internal versus terminal labels. *Biochemistry*, 47(36), 9627–9635. doi:10.1021/bi8000775 [PubMed: 18702521]
- Kornberg RD, & Lorch Y (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell*, 98(3), 285–294. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/10458604> [PubMed: 10458604]
- Krajewski WA, Li J, & Dou Y (2018). Effects of histone H2B ubiquitylation on the nucleosome structure and dynamics. *Nucleic Acids Res*, 46(15), 7631–7642. doi:10.1093/nar/gky526 [PubMed: 29931239]
- Lai WKM, & Pugh BF (2017). Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nat Rev Mol Cell Biol*, 18(9), 548–562. doi:10.1038/nrm.2017.47 [PubMed: 28537572]
- Langst G, & Manelyte L (2015). Chromatin Remodelers: From Function to Dysfunction. *Genes* (Basel), 6(2), 299–324. doi:10.3390/genes6020299 [PubMed: 26075616]
- Laptenko O, Beckerman R, Freulich E, & Prives C (2011). p53 binding to nucleosomes within the p21 promoter in vivo leads to nucleosome loss and transcriptional activation. *Proc Natl Acad Sci U S A*, 108(26), 10385–10390. doi:10.1073/pnas.1105680108 [PubMed: 21606339]
- Larouche K, Bergeron MJ, Leclerc S, & Guerin SL (1996). Optimization of competitor poly(dI-dC).poly(dI-dC) levels is advised in DNA-protein interaction studies involving enriched nuclear proteins. *Biotechniques*, 20(3), 439–444. doi:10.2144/19962003439 [PubMed: 8679204]
- Lebendiker M, & Danieli T (2014). Production of prone-to-aggregate proteins. *FEBS Lett*, 588(2), 236–246. doi:10.1016/j.febslet.2013.10.044 [PubMed: 24211444]
- Lee KM, & Narlikar G (2001). Assembly of nucleosomal templates by salt dialysis. *Curr Protoc Mol Biol*, Chapter 21, Unit 21 26. doi:10.1002/0471142727.mb2106s54
- Li Q, & Wrangé O (1993). Translational positioning of a nucleosomal glucocorticoid response element modulates glucocorticoid receptor affinity. *Genes Dev*, 7(12A), 2471–2482. doi:10.1101/gad.7.12a.2471 [PubMed: 8253391]
- Li Q, & Wrangé O (1995). Accessibility of a glucocorticoid response element in a nucleosome depends on its rotational positioning. *Mol Cell Biol*, 15(8), 4375–4384. doi:10.1128/mcb.15.8.4375 [PubMed: 7623832]
- Lidor Nili E, Field Y, Lubling Y, Widom J, Oren M, & Segal E (2010). p53 binds preferentially to genomic regions with high DNA-encoded nucleosome occupancy. *Genome Res*, 20(10), 1361–1368. doi:10.1101/gr.103945.109 [PubMed: 20716666]
- Lone IN, Shukla MS, Charles Richard JL, Peshev ZY, Dimitrov S, & Angelov D (2013). Binding of NF-kappaB to nucleosomes: effect of translational positioning, nucleosome remodeling and linker histone H1. *PLoS Genet*, 9(9), e1003830. doi:10.1371/journal.pgen.1003830 [PubMed: 24086160]
- Lowary PT, & Widom J (1997). Nucleosome packaging and nucleosome positioning of genomic DNA. *Proc Natl Acad Sci U S A*, 94(4), 1183–1188. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/9037027> [PubMed: 9037027]
- Lowary PT, & Widom J (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J Mol Biol*, 276(1), 19–42. doi:10.1006/jmbi.1997.1494 [PubMed: 9514715]
- Luger K, Dechassa ML, & Tremethick DJ (2012). New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nat Rev Mol Cell Biol*, 13(7), 436–447. doi:10.1038/nrm3382 [PubMed: 22722606]
- Lusser A, & Kadonaga JT (2004). Strategies for the reconstitution of chromatin. *Nat Methods*, 1(1), 19–26. doi:10.1038/nmeth709 [PubMed: 15789029]
- McPherson CE, Shim EY, Friedman DS, & Zaret KS (1993). An active tissue-specific enhancer and bound transcription factors existing in a precisely positioned nucleosomal array. *Cell*, 75(2), 387–398. doi:10.1016/0092-8674(93)80079-t [PubMed: 8402920]
- Nakano S, Fujimoto M, Hara H, & Sugimoto N (1999). Nucleic acid duplex stability: influence of base composition on cation effects. *Nucleic Acids Res*, 27(14), 2957–2965. doi:10.1093/nar/27.14.2957 [PubMed: 10390539]

- Nelson C, Hendy S, Reid K, & Cavanagh J (2002). Acrylamide capture of DNA-bound complexes: electrophoretic purification of transcription factors. *Biotechniques*, 32(4), 808–810, 812, 814–805. doi:10.2144/02324st07 [PubMed: 11962603]
- Noll H, & Noll M (1989). Sucrose gradient techniques and applications to nucleosome structure. *Methods Enzymol*, 170, 55–116. doi:10.1016/0076-6879(89)70043-4 [PubMed: 2770551]
- Olins DE, & Olins AL (2003). Chromatin history: our view from the bridge. *Nat Rev Mol Cell Biol*, 4(10), 809–814. doi:10.1038/nrm1225 [PubMed: 14570061]
- Onder TT, Kara N, Cherry A, Sinha AU, Zhu N, Bernt KM, ... Daley GQ (2012). Chromatin-modifying enzymes as modulators of reprogramming. *Nature*, 483(7391), 598–602. doi:10.1038/nature10953 [PubMed: 22388813]
- Osakabe A, Lorkovic ZJ, Kobayashi W, Tachiwana H, Yelagandula R, Kurumizaka H, & Berger F (2018). Histone H2A variants confer specific properties to nucleosomes and impact on chromatin accessibility. *Nucleic Acids Res*, 46(15), 7675–7685. doi:10.1093/nar/gky540 [PubMed: 29945241]
- Owczarzy R, Moreira BG, You Y, Behlke MA, & Walder JA (2008). Predicting stability of DNA duplexes in solutions containing magnesium and monovalent cations. *Biochemistry*, 47(19), 5336–5353. doi:10.1021/bi702363u [PubMed: 18422348]
- Pennisi E (2003). Human genome. A low number wins the GeneSweep Pool. *Science*, 300(5625), 1484. doi:10.1126/science.300.5625.1484b
- Perlmann T, & Wrange O (1988). Specific glucocorticoid receptor binding to DNA reconstituted in a nucleosome. *EMBO J*, 7(10), 3073–3079. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/2846275> [PubMed: 2846275]
- Pina B, & Suau P (1987). Changes in histones H2A and H3 variant composition in differentiating and mature rat brain cortical neurons. *Dev Biol*, 123(1), 51–58. doi:10.1016/0012-1606(87)90426-x [PubMed: 3622934]
- Pope SD, & Medzhitov R (2018). Emerging Principles of Gene Expression Programs and Their Regulation. *Mol Cell*, 71(3), 389–397. doi:10.1016/j.molcel.2018.07.017 [PubMed: 30075140]
- Rhee HS, Bataille AR, Zhang L, & Pugh BF (2014). Subnucleosomal structures and nucleosome asymmetry across a genome. *Cell*, 159(6), 1377–1388. doi:10.1016/j.cell.2014.10.054 [PubMed: 25480300]
- Richard-Foy H, & Hager GL (1987). Sequence-specific positioning of nucleosomes over the steroid-inducible MMTV promoter. *EMBO J*, 6(8), 2321–2328. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2822386> [PubMed: 2822386]
- Sahu G, Wang D, Chen CB, Zhurkin VB, Harrington RE, Appella E, ... Nagaich AK. (2010). p53 binding to nucleosomal DNA depends on the rotational positioning of DNA response element. *J Biol Chem*, 285(2), 1321–1332. doi:10.1074/jbc.M109.081182 [PubMed: 19887449]
- Sapp J (2005). The prokaryote-eukaryote dichotomy: meanings and mythology. *Microbiol Mol Biol Rev*, 69(2), 292–305. doi:10.1128/MMBR.69.2.292-305.2005 [PubMed: 15944457]
- Schnitzler GR (2001). Isolation of histones and nucleosome cores from mammalian cells. *Curr Protoc Mol Biol*, Chapter 21, Unit 21 25. doi:10.1002/0471142727.mb2105s50
- Senavirathne G, Mahto SK, Hanne J, O'Brian D, & Fishel R (2017). Dynamic unwrapping of nucleosomes by HsRAD51 that includes sliding and rotational motion of histone octamers. *Nucleic Acids Res*, 45(2), 685–698. doi:10.1093/nar/gkw920 [PubMed: 27738136]
- Shrader TE, & Crothers DM (1989). Artificial nucleosome positioning sequences. *Proc Natl Acad Sci U S A*, 86(19), 7418–7422. doi:10.1073/pnas.86.19.7418 [PubMed: 2798415]
- Siggers T, Duyzend MH, Reddy J, Khan S, & Bulky ML (2011). Non-DNA-binding cofactors enhance DNA-binding specificity of a transcriptional regulatory complex. *Mol Syst Biol*, 7, 555. doi:10.1038/msb.2011.89 [PubMed: 22146299]
- Simpson RT (1991). Nucleosome positioning: occurrence, mechanisms, and functional consequences. *Prog Nucleic Acid Res Mol Biol*, 40, 143–184. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/2031082> [PubMed: 2031082]
- Simpson RT, & Stafford DW (1983). Structural features of a phased nucleosome core particle. *Proc Natl Acad Sci U S A*, 80(1), 51–55. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/6572008> [PubMed: 6572008]

- Soufi A, Donahue G, & Zaret KS (2012). Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell*, 151(5), 994–1004. doi:10.1016/j.cell.2012.09.045 [PubMed: 23159369]
- Soufi A, Garcia MF, Jaroszewicz A, Osman N, Pellegrini M, & Zaret KS (2015). Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell*, 161(3), 555–568. doi:10.1016/j.cell.2015.03.017 [PubMed: 25892221]
- Spitz F, & Furlong EE (2012). Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet*, 13(9), 613–626. doi:10.1038/nrg3207 [PubMed: 22868264]
- Steger DJ, & Workman JL (1997). Stable co-occupancy of transcription factors and histones at the HIV-1 enhancer. *EMBO J*, 16(9), 2463–2472. doi:10.1093/emboj/16.9.2463 [PubMed: 9171359]
- Tan ZJ, & Chen SJ (2006). Nucleic acid helix stability: effects of salt concentration, cation valence and size, and chain length. *Biophys J*, 90(4), 1175–1190. doi:10.1529/biophysj.105.070904 [PubMed: 16299077]
- Thakar A, Gupta P, Ishibashi T, Finn R, Silva-Moreno B, Uchiyama S, ... Zlatanova J (2009). H2A.Z and H3.3 histone variants affect nucleosome structure: biochemical and biophysical studies. *Biochemistry*, 48(46), 10852–10857. doi:10.1021/bi901129e [PubMed: 19856965]
- Tietz D (1998). *Nucleic acid electrophoresis*. Heidelberg: Springer.
- Tyagi M, Imam N, Verma K, & Patel AK (2016). Chromatin remodelers: We are the drivers!! *Nucleus*, 7(4), 388–404. doi:10.1080/19491034.2016.1211217 [PubMed: 27429206]
- Vega CA, & Bates RG (1976). Buffers for the physiological pH range: thermodynamic constants of four substituted aminoethanesulfonic acids from 5 to 50 degrees C. *Anal Chem*, 48(9), 1293–1296. doi:10.1021/ac50003a010 [PubMed: 182042]
- Vermaak D, Ahmad K, & Henikoff S (2003). Maintenance of chromatin states: an open-and-shut case. *Curr Opin Cell Biol*, 15(3), 266–274. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/12787767> [PubMed: 12787767]
- Vettese-Dadey M, Walter P, Chen H, Juan LJ, & Workman JL (1994). Role of the histone amino termini in facilitated binding of a transcription factor, GAL4-AH, to nucleosome cores. *Mol Cell Biol*, 14(2), 970–981. doi:10.1128/mcb.14.2.970 [PubMed: 8289837]
- Vossen KM, & Fried MG (1997). Sequestration stabilizes lac repressor-DNA complexes during gel electrophoresis. *Anal Biochem*, 245(1), 85–92. doi:10.1006/abio.1996.9944 [PubMed: 9025972]
- Wingfield PT (2001). Use of protein folding reagents. *Curr Protoc Protein Sci*, Appendix 3, Appendix 3A. doi:10.1002/0471140864.psa03as00
- Wittkopp PJ (2007). Variable gene expression in eukaryotes: a network perspective. *J Exp Biol*, 210(Pt 9), 1567–1575. doi:10.1242/jeb.002592 [PubMed: 17449821]
- Yang TP, Hansen SK, Oishi KK, Ryder OA, & Hamkalo BA (1982). Characterization of a cloned repetitive DNA sequence concentrated on the human X chromosome. *Proc Natl Acad Sci U S A*, 79(21), 6593–6597. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/6959140> [PubMed: 6959140]
- Yu X, & Buck MJ (2019). Defining TP53 pioneering capabilities with competitive nucleosome binding assays. *Genome Res*, 29(1), 107–115. doi:10.1101/gr.234104.117 [PubMed: 30409772]
- Yuan GC, Liu YJ, Dion MF, Slack MD, Wu LF, Altschuler SJ, & Rando OJ (2005). Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science*, 309(5734), 626–630. doi:10.1126/science.1112178 [PubMed: 15961632]
- Yun M, Ruan C, Huh JW, & Li B (2012). Reconstitution of modified chromatin templates for in vitro functional assays. *Methods Mol Biol*, 833, 237–253. doi:10.1007/978-1-61779-477-3_15 [PubMed: 22183598]
- Zhu F, Farnung L, Kaasinen E, Sahu B, Yin Y, Wei B, ... Taipale J (2018). The interaction landscape between transcription factors and the nucleosome. *Nature*, 562(7725), 76–81. doi:10.1038/s41586-018-0549-5 [PubMed: 30250250]
- Zlatanova J, & Thakar A (2008). H2A.Z: view from the top. *Structure*, 16(2), 166–179. doi:10.1016/j.str.2007.12.008 [PubMed: 18275809]

Table. 1.

DNA templates applied to reconstitute nucleosomes *in vitro*

DNA template	Sequence information	Sequence composition	Publication
In vivo fragments	Derived from genomic sequences	LIN28B	1. (Soufi et al., 2015)
		5S rRNA	2. (Angelov et al., 2004)
			3. (Howe & Ausio, 1998)
			4. (Steger & Workman, 1997)
	Derived from virus	Albumin enhancer	5. (Cirillo et al., 1998)
		Chromatins from HeLa cells	6. (Darvekar et al., 2012)
		HIV-1 5' LTR	7. (Angelov et al., 2000)
			4. (Steger & Workman, 1997)
Synthetic fragments	Containing repeated or phased units	MTV LTR	8. (Perlmann & Wrangle, 1988)
		TG motif	9. (H. Chen, Li, & Workman, 1994)
			10. (Li & Wrangle, 1993)
			11. (Li & Wrangle, 1995)
		TTAGGG repeat	12. (Galati et al., 2006)
		A-tracts	13. (Sahu et al., 2010)
	Artificial nucleosome positioning sequence	pBend vector	14. (Vettese-Dadey, Walter, Chen, Juan, & Workman, 1994)
		Widom 601 DNA	15. (Yu & Buck, 2019)
		Randomized DNA sequences	16. (Zhu et al., 2018)

Table. 2.

Experimental conditions of various TF-nucleosome binding assays (Part I)

Reaction	Buffer (pH 7.0-8.0)		Monovalent salt (mM)				Divalent salt		Stabilization	
	Tris-HCl	HEPES	NaCl	KCl	Na ₂ EDTA	K ₂ HPO ₄	MgCl ₂ or MgSO ₄	ZnCl ₂ or ZnSO ₄	Glycerol	NP-40
1. (Soufi et al., 2015)	10 mM			10			1 mM	10 μ M	5%	
2. (H. Chen et al., 1994)	10 mM	10 mM	100	100				10 μ M		0.10%
3. (Bowers, Calero-Nieto, Valeaux, Fernandez-Fuentes, & Cockerill, 2010)		20 mM	50	80			3 mM		10%	
4. (Darvekar et al., 2012)		10 mM	100						5%	0.030%
5. (Angelov et al., 2000)		10 mM		50					5%	
6. (Angelov et al., 2004)		20 mM	50						5%	
7. (Cirillo et al., 1998)	10 mM			40			1 mM		0.50%	
8. (Li & Wrangé, 1993)	20 mM		50		1				10%	
9. (Vettese-Dadey et al., 1994)	10 mM		100						20%	0.10%
10. (Perlmann & Wrangé, 1988)	20 mM		50		1				10%	
11. (Li & Wrangé, 1995)	20 mM		50		1				10%	
12. (Howe & Ausio, 1998)	20 mM		70					10 μ M	6%	0.07%
13. (Sahu et al., 2010)	10 mM		50				2 mM		5%	0.025%
14. (Steger & Workman, 1997)		10 mM		50					5%	
15. (Galati et al., 2006)		20 mM	100	50			1 mM		5%	0.10%
16. (Yu & Buck, 2019)	10 mM		50				2 mM		5%	0.0250%
17. (Zhu et al., 2018)		20 mM	5	50-140		1	2 mM	3 μ M		

Table. 2.

Experimental conditions of various TF-nucleosome binding assays (Part II)

Reaction	Reagents to reduce nonspecific binding						Protease inhibitors				Reducing agents	
	Ficoll	BSA (mg/ml)	Poly(dIdC)	Pork insulin	Spermidine	EDTA or EGTA (pH 8.0)	PMSF (mM)	Aprotinin	Leupeptin	2-mercaptoethanol	DTT	
1. (Soufi et al., 2015)		0.5									1 mM	
2. (H. Chen et al., 1994)		1				0.4 mM	0.1			10 mM		
3. (Bowers et al., 2010)		0.1	6.25 ng/μl				0.1	5 μg/ml	5 μg/ml		1%	
4. (Darvekar et al., 2012)		0.15									1 mM	
5. (Angelov et al., 2000)		0.2					0.5				5 mM	
6. (Angelov et al., 2004)		100										
7. (Cirillo et al., 1998)	1%	3									5 mM	
8. (Li & Wrangle, 1993)				0.1 mg/ml						1 mM	5 mM	
9. (Vettese-Dadey et al., 1994)		0.1				1 mM						
10. (Perlmann & Wrangle, 1988)				0.1 mg/ml							5 mM	
11. (Li & Wrangle, 1995)				0.1 mg/ml							5 mM	
12. (Howe & Ausio, 1998)		0.1	40 ng/μl								2.5 mM	
13. (Sahu et al., 2010)					2 mM	1 mM					1 mM	
14. (Steger & Workman, 1997)		0.25					0.5				5 mM	
15. (Galati et al., 2006)		0.5				0.1 mM					1 mM	
16. (Yu & Buck, 2019)		0.25									1 mM	
17. (Zhu et al., 2018)						100 μM						