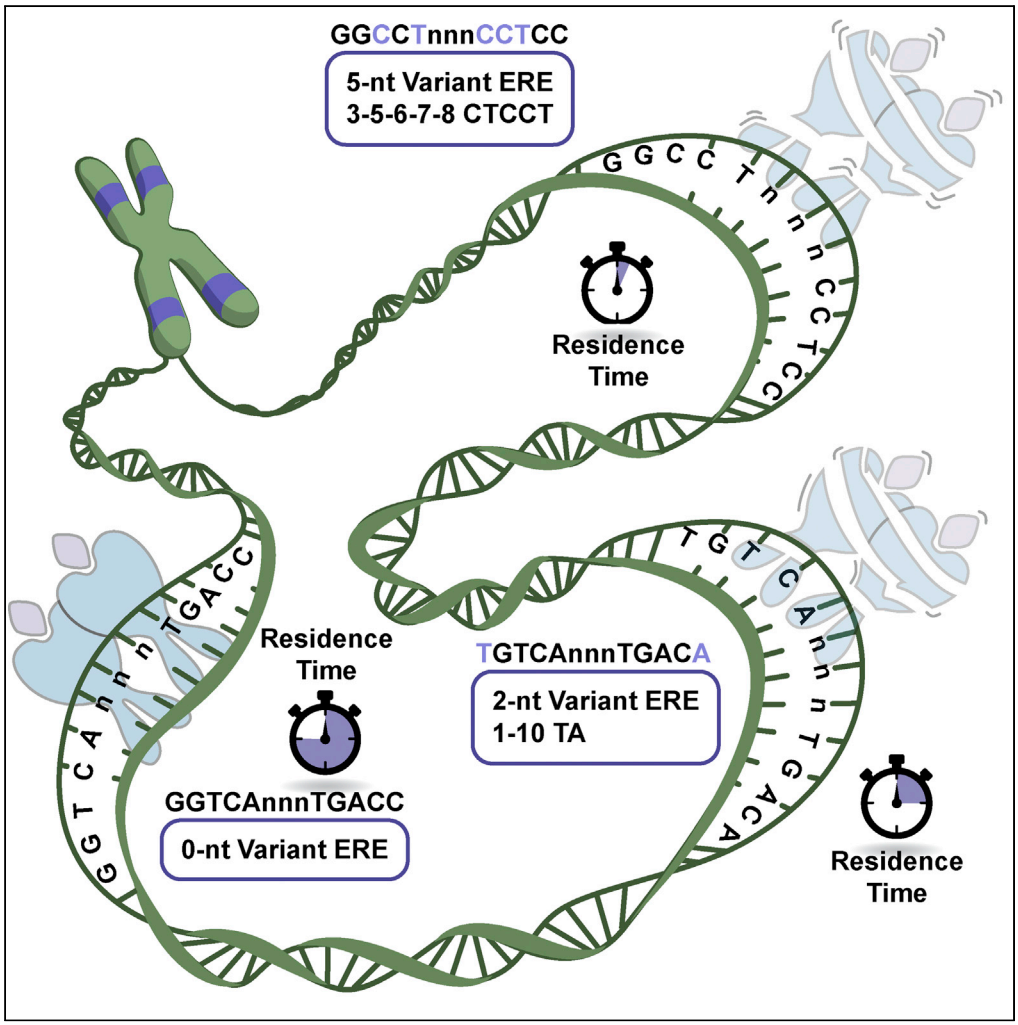
iScience

CellPress

Article

Decoding the Inversion Symmetry Underlying Transcription Factor DNA-Binding Specificity and Functionality in the Genome



Laurel A. Coons, Adam B. Burkholder, Sylvia C. Hewitt, Donald P. McDonnell, Kenneth S. Korach

korach@niehs.nih.gov

HIGHLIGHTS

Inversion symmetry (IS) is universal within the genome

Transcription factor binding in the genome follows IS

DNA elements where transcription factors bind are determined by internal IS

Functionality is determined by residence time (dictated by IS and DNA sequence constraints)

Coons et al., iScience *15*, 552–591
May 31, 2019
https://doi.org/10.1016/j.isci.2019.04.006

iScience

CellPress

Article

Decoding the Inversion Symmetry Underlying Transcription Factor DNA-Binding Specificity and Functionality in the Genome

Laurel A. Coons,[1,2] Adam B. Burkholder,[3] Sylvia C. Hewitt,[1] Donald P. McDonnell,[2] and Kenneth S. Korach[1,4,*]

SUMMARY

Understanding why a transcription factor (TF) binds to a specific DNA element in the genome and whether that binding event affects transcriptional output remains a great challenge. In this study, we demonstrate that TF binding in the genome follows inversion symmetry (IS). In addition, the specific DNA elements where TFs bind in the genome are determined by internal IS within the DNA element. These DNA-binding rules quantitatively define how TFs select the appropriate regulatory targets from a large number of similar DNA elements in the genome to elicit specific transcriptional and cellular responses. Importantly, we also demonstrate that these DNA-binding rules extend to DNA elements that do not support transcriptional activity. That is, the DNA-binding rules are obeyed, but the retention time of the TF at these non-functional DNA elements is not long enough to initiate and/or maintain transcription. We further demonstrate that IS is universal within the genome. Thus, IS is the DNA code that TFs use to interact with the genome and dictates (in conjunction with known DNA sequence constraints) which of those interactions are functionally active.

INTRODUCTION

The specificity of gene expression in all organisms depends on the ability of transcription factors (TFs) to interact with *cis*-regulatory DNA elements in the genome. Achievement of specific DNA element recognition is non-trivial as genomes harbor thousands of consensus and near-consensus binding sequences. Understanding (1) why a TF binds to a specific DNA element in the genome and (2) whether that binding event affects transcriptional output remains a great challenge.

In the past decade, there has been a dramatic increase in the use of genome-wide technologies, such as chromatin immunoprecipitation (ChIP) sequencing (ChIPSeq) and ChIP exonuclease (ChIPExo), to identify where a TF physically associates with the genome. These studies have revealed that TF binding is widespread with thousands to tens of thousands of chromatin-interacting events, also known as peaks, in the genome. Although these genomic technologies have provided large-scale snapshots (i.e., point-in-time images) of TF binding in the genome, a full understanding of the mechanistic and quantitative details of specific DNA element recognition by TFs in the context of gene regulation is lacking (Todeschini et al., 2014).

Steroid nuclear receptors (sNRs) are TFs that are activated by the binding of steroid hormones (i.e., ligand-activated TFs) (Tang et al., 2011). The sNR family includes two divisions: the estrogen receptors (ERs) and the ketosteroid receptors (KRs). The ERs are estrogen receptor α (ERα) and estrogen receptor β (ERβ). The KRs are the androgen receptor (AR), glucocorticoid receptor (GR), mineralocorticoid receptor (MR), and progesterone receptor (PR). The sNR family regulate gene expression by interacting with sequence-specific *cis*-regulatory DNA elements in the genome: the estrogen response element (ERE) used by the ERs and the hormone response element (HRE) used by the KRs (Cotnoir-White et al., 2011; Helsen et al., 2012). The evolutionary relationship among the sNRs has been deduced by the high conservation in their DNA-binding domains (DBDs), and not surprisingly they exhibit sequence conservation at the level of their cognate *cis*-regulatory DNA elements (Laudet, 1997).

In a previously published study, we demonstrated that not all TF DNA-binding events result in gene expression (Coons et al., 2017). Specifically, we showed that only a small fraction of all sNR chromatin-interacting events observed in ChIPSeq and ChIPExo experiments is associated with transcriptional output (Li et al., 2011; John et al., 2011; Vockley et al., 2016; Coons et al., 2017), and this functionality is restricted to

[1]Receptor Biology Section, Reproductive and Developmental Biology Laboratory, National Institute of Environmental Health Sciences/National Institutes of Health, 111 T.W. Alexander Dr., Research Triangle Park, NC 27709, USA

[2]Department of Pharmacology and Cancer Biology, Duke University School of Medicine, Durham, NC 27710, USA

[3]Integrative Bioinformatics, National Institute of Environmental Health Sciences/National Institutes of Health, Research Triangle Park, NC 27709, USA

[4]Lead Contact

*Correspondence: korach@niehs.nih.gov

https://doi.org/10.1016/j.isci.2019.04.006

DNA elements that vary from the consensus palindromic DNA element by one or two nucleotides (nts), named *nuclear receptor functional enhancers* (NRFEs) (Coons et al., 2017). Thus, DNA sequence constraints define which sNR chromatin-interacting events in the genome are functionally active. This raises the question as to the purpose or cause of the remaining non-functional (non-NRFE) chromatin-interacting events observed in ChIPSeq and ChIPExo experiments. Or, assuming a purpose, what information is contained in these structures.

In this study, using publicly available ChIPSeq and ChIPExo experiments, we decode the inversion (reverse-complement) symmetry underlying TF binding to DNA elements in the genome, and elucidate its role in distinguishing functional from non-functional chromatin-interacting events. That is, why TFs bind the specific DNA elements in the genome where they do and whether that binding event ultimately affects gene expression. Specifically, we demonstrate that TF binding in the genome follows inversion symmetry (i.e., the number of TF binding events at a particular DNA element in the genome is equivalent to the number of TF binding events at its reverse-complement DNA element in the genome). In addition, the specific DNA elements where TFs bind in the genome are determined by internal inversion symmetry within the DNA element (i.e., TF DNA-binding is determined by the position of variants within the DNA element and its reverse-complement position). The basis of these DNA-binding rules follows specific algebraic relationships. Thus, these DNA-binding rules quantitatively define how TFs select the appropriate regulatory targets from a large number of similar DNA elements in the genome to elicit specific transcriptional and cellular responses. Importantly, we also demonstrate that these DNA-binding rules extend (i.e., are applicable) to DNA elements that do not support transcriptional activity. That is, the DNA-binding rules are obeyed, but the retention time of the TF at these non-functional DNA elements is not long enough to initiate and/or maintain transcription. Thus, functionality is determined at the individual nucleotide level, and the residence time (or strength of binding) is dictated by the number and position of variants within the DNA element (i.e., inversion symmetry and DNA sequence constraints). We further demonstrate that the population of every DNA element in the single-stranded genome (i.e., 1.4 trillion DNA elements) is equivalent to the population of its reverse-complement DNA element in the single-stranded genome. Therefore, the inversion symmetry observed for TF binding in the genome represents an inherent inversion symmetry structure for all DNA elements in the genome. This property is maintained at the level of each individual chromosome. These findings suggest that the structural mechanisms (by which inversion symmetry ascribes TF DNA-binding and functionality) are universally applicable. Hence, analysis of TF binding in the genome has expanded our understanding as to why the genome is organized in an inversion symmetry structure (i.e., Chargaff's second parity rule). Inversion symmetry is the DNA code that TFs use to interact with the genome, and dictates (in conjunction with known DNA sequence constraints) which of those interactions are functionally active. In addition, we further demonstrate why the inversion symmetry that underlies TF DNA-binding specificity and functionality in the genome, observed in hundreds of ChIPSeq and ChIPExo experiments, has not been detected using current DNA motif identification algorithms.

RESULTS

Inversion Symmetry of sNR DNA-Binding at NRFEs in the Genome

Previously, we reported that NRFE EREs and HREs are composed of their consensus palindromic DNA element and those DNA elements that vary from their consensus palindromic DNA element by one or two nts (Coons et al., 2017). In the present study, we evaluated the DNA-binding preference of sNRs at each of the 30 DNA elements that make up the 1-nt variant group (i.e., DNA elements generated by varying the 0-nt variant consensus palindromic DNA element by one nt) in the genome. That is, the 3 alternative nucleotide possibilities (variants) in each of the 10 primary positions of the 13-nt consensus palindromic ERE (5′-GGTCAnnnTGACC-3′) and HRE (5′-GAACAnnnTGTTC-3′) DNA element.

First, the absolute number of times each 1-nt variant DNA element occurred in a ChIPSeq or ChIPExo experiment were counted. This analysis was completed by overlapping the location coordinates of every 1-nt variant DNA element in the genome and the location coordinates of the ChIPSeq or ChIPExo peaks in an experiment. The absolute number of times each 1-nt variant DNA element occurred in an experiment was then converted to a proportion or percent of the total number of 1-nt variant DNA elements contained within that ChIPSeq or ChIPExo experiment (i.e., the total number of 1-nt variant DNA elements contained within an experiment = 100%). Taking the average of ER DNA-binding at 1-nt variant ERE DNA elements in the genome from 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, reveals considerable variability in the ability of ER to bind the 30 DNA elements that make up the 1-nt

variant ERE group (Figure 1A). Of particular importance, this analysis demonstrates that ER DNA-binding at 1-nt variant ERE DNA elements in the genome follows inversion symmetry (Figure 1A). That is, the number of ER DNA-binding events at a particular 1-nt variant ERE DNA element in the genome is equivalent to the number of ER DNA-binding events at its reverse-complement DNA element in the genome. For example, the number of ER DNA-binding events at the 1T ERE DNA element (i.e., the variant thymidine [T] is in the first [1] position [5′-TGTCAnnnTGACC-3′]) is equivalent to the number of ER DNA-binding events at the 10A ERE DNA element (i.e., the variant adenine [A] is in the tenth [10] position [5′-GGTCAnnnTGACA-3′]) in the genome (Figure 1A). Accordingly, the number of ER DNA-binding events at DNA elements where a variant is in position 2 is equivalent to the number of ER DNA-binding events at DNA elements where a variant is in position 9 (Figure 1A). This equivalency between the number of ER DNA-binding events at a particular DNA element and its reverse-complement DNA element occurs for all five palindromic position pairs: 1-10, 2-9, 3-8, 4-7, 5-6 (Figure 1A). Furthermore, the number of ER DNA-binding events is greatest at 1-nt variant ERE DNA elements that (1) vary position 3 (or its reverse-complement position 8) to any nucleotide, and (2) vary position 1 (or its reverse-complement position 10) to either an A or a T (Figure 1A). The small standard deviation measured across these 157 ER experiments confirms that the inversion symmetry of ER DNA-binding at 1-nt variant ERE DNA elements in the genome is a universal property for all tissues, primary cell lines, and cancer cell lines tested to date (Figure 1A). Note: all detailed data and statistics associated with every figure are compiled in Data S1.

Similarly, taking the average of AR, GR, and PR DNA-binding at 1-nt variant HRE DNA elements in the genome from 193 KR experiments, representing a wide variety of mouse tissues and human cell lines, demonstrates that KR DNA-binding at 1-nt variant HRE DNA elements in the genome also follows inversion symmetry (Figure 1B). Furthermore, this analysis reveals a strikingly similar DNA-binding profile between the different KRs at 1-nt variant HRE DNA elements in the genome, suggesting a shared DNA-binding mechanism between all members of the KR family (Figure 1B). In contrast to ER DNA-binding at 1-nt variant ERE DNA elements, the number of KR DNA-binding events is greatest at 1-nt variant HRE DNA elements that (1) vary position 2 (or its reverse-complement position 9) to any nucleotide, (2) vary position 1 (or its reverse-complement position 10) to a 1A or 10T, and (3) vary position 5 (or its reverse-complement position 6) to a 5T or 6A (Figure 1B). Therefore, from an evolutionary perspective, the increased DNA-binding at 1-nt variant DNA elements in the genome changed from position 3 (or its reverse-complement position 8) in the ERE (by the ER) to position 2 (or its reverse-complement position 9) in the HRE (by the KRs), and the increased DNA-binding of ER when position 1 (or its reverse-complement position 10) is varied to a 1T or 10A in the ERE was transferred to position 5 (or its reverse-complement position 6) in the HRE by the KRs. The small standard deviation measured across these 193 KR experiments confirms that the inversion symmetry of KR DNA-binding at 1-nt variant HRE DNA elements in the genome is a universal property for all tissues, primary cell lines, and cancer cell lines tested to date (Figure 1B).

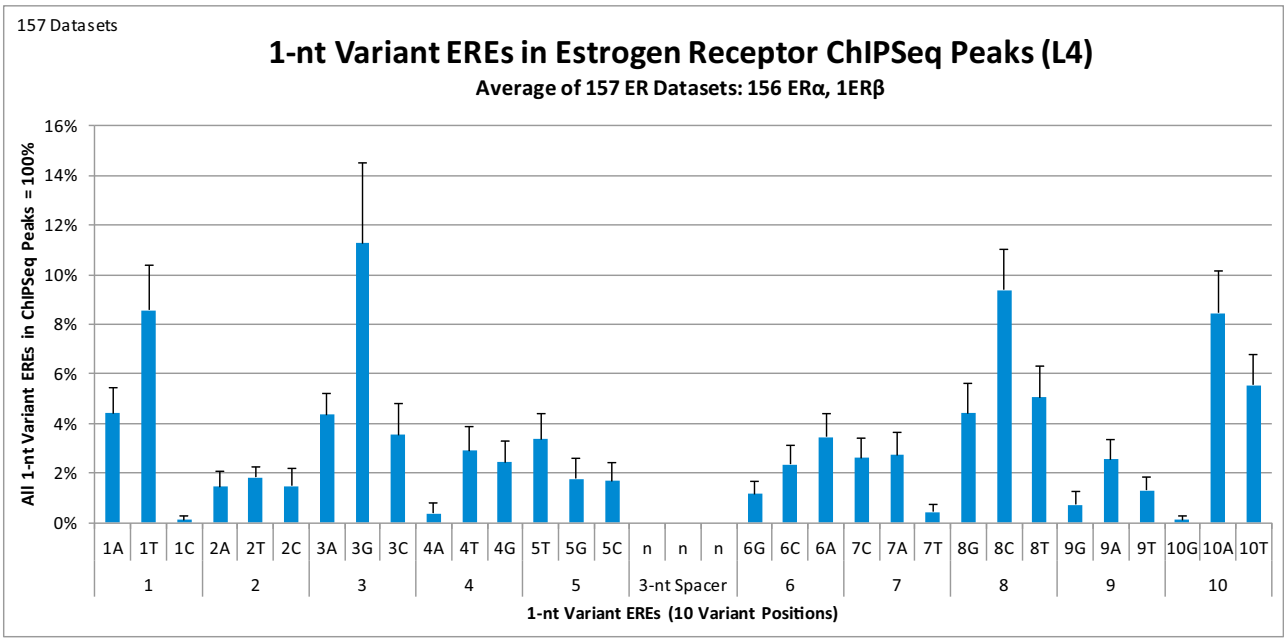
The inversion symmetry of sNR DNA-binding observed at 1-nt variant DNA elements in the genome (i.e., the number of sNR DNA-binding events at a particular 1-nt variant DNA element is equivalent to the number of sNR DNA-binding events at its reverse-complement DNA element) also occurs for sNR DNA- binding at 2-nt variant DNA elements in the genome (Figure S1). There are 405 DNA elements in the 2-nt variant group (i.e., DNA elements generated by varying the 0-nt variant consensus palindromic DNA element by two nts) (Figure S1). In addition, the increased DNA-binding of ER at DNA elements with variants in palindromic position pair 3-8 and 1-10 of the ERE, and increased DNA-binding of KR at DNA elements with variants in palindromic position pair 2-9 of the HRE also occurs for sNR DNA-binding at 2-nt variant ERE and HRE DNA elements in the genome (Figure S1). Thus, sNR DNA-binding at NRFEs in the genome follows inversion symmetry (i.e., the number of sNR DNA-binding events at a particular DNA element in the genome is equivalent to the number of sNR DNA-binding events at its reverse-complement DNA element in the genome). In addition, sNRs exhibit preferential DNA-binding at specific NRFE DNA elements in the genome; this preference is determined by internal inversion symmetry within the DNA element (i.e., sNR binding in the genome is determined by the position of variants within the DNA element and its reverse-complement position).

sNR DNA-Binding Is Most Predominant at NRFEs in the Genome

We previously demonstrated that NRFEs (composed of the 0-nt variant consensus palindromic DNA element, 1-nt variant DNA elements, and 2-nt variant DNA elements) constitute ~45% of the chromatin-interacting events in ER experiments and ~35% of the chromatin-interacting events in KR experiments

A

ERE = GGTCAnnnTGACC
1 2 3 4 5 6 7 8 9 10



B

HRE = GAACAnnnTGTTC
1 2 3 4 5 6 7 8 9 10

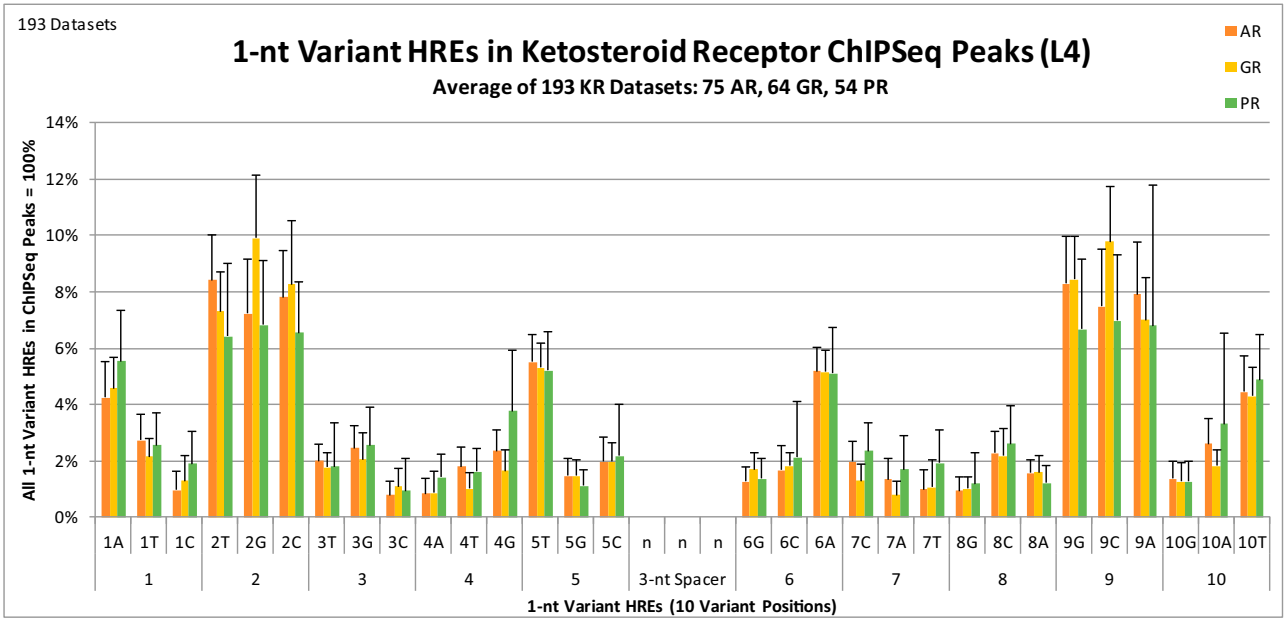


Figure 1. Inversion Symmetry of sNR DNA-Binding at 1-nt Variant DNA Elements in the Genome

(A) Average distribution of 1-nt variant ERE DNA elements in ER experiments (156 ERα and 1 ERβ) at the 30 1-nt variant ERE DNA elements in the genome (the total number of 1-nt variant ERE DNA elements contained within an experiment = 100%).

(B) Average distribution of 1-nt variant HRE DNA elements in KR experiments (75 AR, 64 GR, 54 PR) at the 30 1-nt variant HRE DNA elements in the genome (the total number of 1-nt variant HRE DNA elements contained within an experiment = 100%). These 30 1-nt variant EREs or HREs are defined by 10 variant positions (positions 1 through 5, followed by their reverse-complements).

(Coons et al., 2017). Thus, ~55% and ~65% of the chromatin-interacting events detected in ER and KR ChIPSeq and ChIPExo experiments are non-functional (i.e., non-NRFEs) (Coons et al., 2017). This raises the question as to the purpose or cause of this large amount of non-functional (i.e., non-NRFE) TF interaction with the genome. Or, assuming a purpose, what information is contained in these structures.

In the past, it has been shown that ERE and HRE "palindrome half-sites" (i.e., the left-hand or right-hand side of the symmetrical 0-nt variant consensus palindromic DNA element: GGTCA or TGACC in the ERE, GAACA or TGTTC in the HRE) are an enriched feature observed in sNR ChIPSeq and ChIPExo experiments, and that these sites can support hormone-dependent transcription (Fan et al., 1996; Norris et al., 1995). To further explore the interactions of sNRs with these DNA elements, we obtained the location coordinates of the 81,922 DNA elements that contain 0- to 5-nt variants relative to the 0-nt variant consensus palindromic DNA element in the mouse and human genome (i.e., a 5-nt variant = a half-site) (Table 1). Thus, we are expanding the concept of a "half-site" to include all 5-nt variant DNA elements, not just the "palindrome half-sites." The 0- to 5-nt variant DNA elements include the 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements, 405 2-nt variant DNA elements, 3,240 3-nt variant DNA elements, 17,010 4-nt variant DNA elements, and 61,236 5-nt variant DNA elements, for a total of 81,922 DNA elements (Table 1).

The analysis done in our previously published study counted the number of ChIPSeq or ChIPExo peaks that contained a 0- to 3-nt variant ERE or HRE DNA element, with peak assignment given to the 0-nt variant consensus palindromic DNA element or the DNA element with the least number of variants relative to the 0-nt variant consensus palindromic DNA element (i.e., defining each peak by a single DNA element with the total number of peaks in an experiment = 100%) (Coons et al., 2017). In this study, the absolute number of times a 0- to 5-nt variant DNA element occurred in a ChIPSeq or ChIPExo experiment were counted and used in a signal-to-noise ratio (S/N) analysis. Counting the absolute number of times every 0- to 5-nt variant DNA element occurred in an experiment eliminates any bias toward DNA elements with the least number of variants in individual peaks that contain multiple ERE or HRE DNA elements. That is, the occurrence of all 0- to 5-nt variant DNA elements was counted, because the preset hierarchal peak selection criteria based on the number of variants is eliminated.

The (S/N) is the absolute number of times a 0- to 5-nt variant DNA element occurs in an experiment (defined by the total number of peaks in the experiment and the peak length) compared with the random frequency expectation of that DNA element occurring in the genome (i.e., the probability that any 10-nt DNA element that has a maximum possibility of 4 nts in each position will occur in the genome is once every 1,048,576 nts (4^{10}) at random occurrence). (S/N) analysis demonstrates that the number of ER DNA-binding events (i.e., the ER DNA-binding signal) is greatest at the 0-nt variant consensus palindromic ERE DNA element, followed by DNA-binding at 1-nt variant ERE DNA elements, followed by DNA-binding at 2-nt variant ERE DNA elements in the genome (Figure S2). In contrast, there is little to no increase of ER DNA-binding at 3- to 5-nt variant ERE DNA elements above what would be expected with respect to random noise (i.e., random = 0 and 0 = log(1)) (Figure S2). This DNA-binding profile was observed in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines (Figure S2). This DNA-binding profile was also observed across multiple peak selection criteria (L4, L8, L10, L15, L20), where Lx represents an x-fold greater tag density at peaks than in the surrounding 10-kb region (i.e., performing a low-to-high stringency analysis of the ChIPSeq and ChIPExo data) (Figure S2). Likewise, this DNA-binding profile was observed in 194 KR experiments at 0- to 5-nt variant HRE DNA elements in the genome (Figure S3). This suggests that there is little to no increase in sNR DNA-binding at 3- to 5-nt variant ERE or HRE DNA elements in the genome (when analyzed by number of variants) above what would be expected to occur at random. Note: the relative (S/N) values of sNR DNA-binding signals are scale invariant (i.e., the relative ratios between the 0- to 5-nt variant groups are constant) (Figures S2 and S3).

Sequentially Track sNR DNA-Binding at 0- to 5-nt Variant DNA Elements in the Genome

To evaluate the DNA-binding of sNRs at every DNA element in the 0- to 5-nt variant groups, the absolute number of times each of the 81,922 0- to 5-nt variant DNA elements (Table 1) occurred in an experiment was counted. For display purposes, we have defined a 5-nt variant ERE and HRE DNA element by its five fixed positions, resulting in 252 half-site groups (i.e., categorizing the 61,236 5-nt variant DNA elements into 252 half-site groups, defined by the five positions that are fixed/not varied) (Table 1). Thus, the "palindrome

$$\mathbf{ERE = G\underline{GTC}AnnnT\underline{GAC}C}$$
1 2 3 4 5 6 7 8 9 10

$$\mathbf{HRE = G\underline{AAC}AnnnT\underline{GTT}C}$$
1 2 3 4 5 6 7 8 9 10

	DNA Element	Combinatorial Counts				DNA Element Frequency in Genome			
A	B	C	D	E	F	Mouse Genome (mm10)		Human Genome (hg19)	
k	ERE (*n* = 10)	Combinations $C = \binom{n}{k}$	4-nts $D = C \times 4^k$	Unique $E = C \times 3^k$	Total Unique $F = \Sigma E$	Unique	Total Unique	Unique	Total Unique
0	0-nt Variant ERE	1	1	1	1	2,367		2,194	
1	1-nt Variant ERE	10	40	30	31	71,428	73,795	60,313	62,507
2	2-nt Variant ERE	45	720	405	436	898,155	971,950	914,726	977,233
3	3-nt Variant ERE	120	7,680	3,240	3,676	6,750,607	7,722,557	7,516,184	8,493,417
4	4-nt Variant ERE	210	53,760	17,010	20,686	35,508,190	43,230,747	38,222,674	46,716,091
5	5-nt Variant ERE	252	258,048	61,236	81,922	134,732,965	177,963,712	141,280,093	187,996,184
A	B	C	D	E	F	Mouse Genome (mm10)		Human Genome (hg19)	
k	HRE (*n* = 10)	Combinations $C = \binom{n}{k}$	4-nts $D = C \times 4^k$	Unique $E = C \times 3^k$	Total Unique $F = \Sigma E$	Unique	Total Unique	Unique	Total Unique
0	0-nt Variant HRE	1	1	1	1	3,444		3,535	
1	1-nt Variant HRE	10	40	30	31	97,039	100,483	104,767	108,302
2	2-nt Variant HRE	45	720	405	436	1,337,516	1,437,999	1,339,543	1,447,845
3	3-nt Variant HRE	120	7,680	3,240	3,676	10,461,197	11,899,196	10,771,159	12,219,004
4	4-nt Variant HRE	210	53,760	17,010	20,686	49,391,434	61,290,630	54,136,564	66,355,568
5	5-nt Variant HRE	252	258,048	61,236	81,922	169,440,498	230,731,128	187,469,898	253,825,466
A	B	C	D	E	F				
k		Combinations $C = \binom{n}{k}$	4-nts $D = C \times 4^k$	Unique $E = C \times 3^k$	Total Unique $F = \Sigma E$				
0	0-nt Variant	10!/(0! x 10!)	1 x 1	1 x 1	1 + 0				
1	1-nt Variant	10!/(1! x 9!)	10 x 4	10 x 3	30 + 1				
2	2-nt Variant	10!/(2! x 8!)	45 x 16	45 x 9	405 + 31				
3	3-nt Variant	10!/(3! x 7!)	120 x 64	120 x 27	3,240 + 436				
4	4-nt Variant	10!/(4! x 6!)	210 x 256	210 x 81	17,010 + 3,676				
5	5-nt Variant	10!/(5! x 5!)	252 x 1,024	252 x 243	61,236 + 20,686				

Table 1. Number of 0- to 5-nt Variant 13-nt ERE and HRE DNA Elements in the Mouse and Human Genome

The number of 0- to 5-nt variants of the 13-nt ERE and HRE consensus palindromic DNA element on the positive or sense strand in the mouse (mm10) and human (hg19) genome. Here the 13-nt ERE is the estrogen response element (5′-GGTCAnnnTGACC-3′) and the 13-nt HRE is the hormone response element (5′-GAA CAnnnTGTTC-3′). The 0- to 5-nt variant 13-nt ERE and HRE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements (10 variant positions), 405 2-nt variant DNA elements (45 variant positions), 3,240 3-nt variant DNA elements (120 variant positions), 17,010 4-nt variant DNA elements (210 variant positions), and 61,236 5-nt variant DNA elements (252 variant positions), for a total of 81,922 DNA elements. The population count of each of the 81,922 0- to 5-nt variant 13-nt ERE and HRE DNA elements in the mouse (mm10) and human (hg19) genome can be found in Tables S55–S58.

half-sites'' make up two of these 252 half-site groups (i.e., positions 1 through 5 or positions 6 through 10 of the 0-nt variant consensus palindromic DNA element are fixed/not varied).

The 252 half-site groups are symmetrically split into a set of 126 groups and their 126 reverse-complements (Table 2). These 126 groups further split into three distinct subgroups (zero vacancies, one vacancy, two vacancies) depending on how many reverse-complement vacancies are in the DNA element (Table 2), that is, the information for which nucleotide occupies each of the 10 primary positions of the 0-nt variant consensus palindromic DNA element is missing/replaced by variants/vacant in the position and its reverse-complement position (i.e., its palindromic position pair). For example, the half-site group 1-2-3-4-5 (variants in positions 6-7-8-9-10) has zero vacancies because the information for which nucleotide occupies each position is present in either the position (positions 1 through 5) or its reverse-complement position (positions 6 through 10) (Table 2). In contrast, the half-site group 1-2-4-5-10 (variants in positions 3-6-7-8-9) has one vacancy because the information for which nucleotide occupies position 3 is missing/replaced by variants/vacant in the position (position 3) and its reverse-complement position (position 8) (Table 2). Of these 126 groups, 16 have zero vacancies, 80 have one vacancy, and 30 have two vacancies (Table 2). The reverse-complement vacancy position ID indicates which of the five palindromic position pairs (i.e., 1-10, 2-9, 3-8, 4-7, 5-6) is missing/replaced by variants/vacant (Table 2).

Furthermore, the number of reverse-complement vacancies equals the number of reverse-complement double occupants (zero double occupants, one double occupant, two double occupants) that occur in the DNA element (Table 2), that is, the information for which nucleotide occupies each of the 10 primary positions of the 0-nt variant consensus palindromic DNA element is occupied in the position and its reverse-complement position. For example, the half-site group 2-4-5-7-9 (variants in positions 1-3-6-8-10) has two reverse-complement vacancies because the information for which nucleotide occupies position 1 and position 3 is missing/replaced by variants/vacant in the position (position 1 and position 3) and its reverse-complement position (position 8 and position 10) (Table 2). This half-site group, 2-4-5-7-9, also has two reverse-complement double occupants because the information for which nucleotide occupies position 2 and position 4 is occupied in the position (position 2 and position 4) and its reverse-complement position (position 7 and position 9) (Table 2). The reverse-complement double occupant position ID indicates which of the five palindromic position pairs (i.e., 1-10, 2-9, 3-8, 4-7, 5-6) is doubly occupied (Table 2).

The remaining 0- to 4-nt variant DNA elements (1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements, 405 2-nt variant DNA elements, 3,240 3-nt variant DNA elements, and 17,010 4-nt variant DNA elements, for a total of 20,686 DNA elements) can be categorized into these same 252 half-site groups (i.e., five positions are fixed, allowing for *up to* five positions to be varied) (Table 1). For example, the half-site group 1-2-3-4-5 contains all 1-nt variant DNA elements that vary positions 6 through 10 (Table 2). In the case of the ERE, this includes the following 1-nt variant ERE DNA elements: 6A, 6C, 6G, 7A, 7C, 7T, 8C, 8G, 8T, 9A, 9G, 9T, 10A, 10G, 10T (Tables S1 and S3). Thus, each of the 252 half-site groups contain 1 (0-nt variant consensus palindromic DNA element), 15 (1-nt variant DNA elements), 90 (2-nt variant DNA elements), 270 (3-nt variant DNA elements), 405 (4-nt variant DNA elements), and 243 (5-nt variant DNA elements), for a total of 1,024 0- to 5-nt variant DNA elements per half-site group (Tables S1–S4). This allows all 81,922 0-nt to 5-nt variant DNA elements to be categorized into the 252 half-site groups, thus providing the ability to sequentially track sNR DNA-binding at all 81,922 0-nt to 5-nt variant ERE or HRE DNA elements in the genome.

Inversion Symmetry of sNR DNA-Binding Continues through 5-nt Variant DNA Elements in the Genome

(S/N) analysis of sNR DNA-binding at 0- to 5-nt variant DNA elements in the genome (displayed by the 252 half-site groups) reveals a highly structured symmetrical DNA-binding profile that decays from 0- to 5-nt variants: ER at 0- to 5-nt variant ERE DNA elements (Figure 2) and KR at 0- to 5-nt variant HRE DNA elements (Figures 3, S4, and S5). The x-axis is ordered by the symmetrically split 126 groups (left-to-right: zero vacancies, one vacancy, two vacancies) followed by their 126 reverse-complements. The x-axis is labeled by the reverse-complement vacancy position ID (primary label) and reverse-complement double occupant position ID (secondary label) (Figures 2, 3, S4, and S5). The order of the reverse-complement vacancy position IDs for the ERE is 3-8 > 1-10 > 5-6 > 4-7 > 2-9 (Table S5), whereas that for the HRE is 2-9 > 5-6 > 1-10 >

	A	B	C	D	E
	Fixed Half-Site Positions	Reverse-Complement Vacancies & Double Occupants	Reverse-Complement Vacancy Position ID	Reverse-Complement Double Occupant Position ID	Positions That Can Be Varied
			Zero Vacancies		
1	1-2-3-4-5-----	0V			-----6-7-8-9-10
2	1-2--4-5---8--	0V			--3---6-7--9-10
3	-2-3-4-5-----10	0V			1-----6-7-8-9-
4	1-2-3-4--6----	0V			----5--7-8-9-10
5	1-2-3--5--7---	0V			---4--6--8-9-10
6	1--3-4-5----9-	0V			-2----6-7-8--10
7	-2--4-5---8--10	0V			1--3---6-7--9-
8	1-2--4--6--8--	0V			--3--5--7--9-10
9	1-2---5--7-8--	0V			--3-4--6---9-10
10	1---4-5---8-9-	0V			-2-3---6-7---10
11	-2-3-4--6----10	0V			1----5--7-8-9-
12	-2-3--5--7---10	0V			1---4--6--8-9-
13	--3-4-5----9-10	0V			1-2----6-7-8--
14	1-2-3---6-7---	0V			---4-5---8-9-10
15	1--3-4--6---9-	0V			-2---5--7-8--10
16	1--3--5--7--9-	0V			-2--4--6--8--10
			One Vacancy		
17	1-2--4-5-----10	1V	3	1	--3---6-7-8-9-
18	1-2--4--6----10	1V	3	1	--3--5--7-8-9-
19	1-2---5--7---10	1V	3	1	--3-4--6--8-9-
20	1---4-5----9-10	1V	3	1	-2-3---6-7-8--
21	1-2--4-5-6----	1V	3	5	--3----7-8-9-10
22	-2--4-5-6----10	1V	3	5	1--3----7-8-9-
23	1-2---5-6-7---	1V	3	5	--3-4----8-9-10
24	1---4-5-6---9-	1V	3	5	-2-3----7-8--10
25	1-2--4-5--7---	1V	3	4	--3---6--8-9-10
26	-2--4-5--7---10	1V	3	4	1--3---6--8-9-
27	1-2--4--6-7---	1V	3	4	--3--5---8-9-10
28	1---4-5--7--9-	1V	3	4	-2-3---6--8--10
29	1-2--4-5----9-	1V	3	2	--3---6-7-8--10
30	-2--4-5----9-10	1V	3	2	1--3---6-7-8--

Table 2. The 252 Half-Site Groups Symmetrically Split into 126 Groups and Their Reverse-Complements

(Continued on next page)

	A	B	C	D	E
	Fixed Half-Site Positions	Reverse-Complement Vacancies & Double Occupants	Reverse-Complement Vacancy Position ID	Reverse-Complement Double Occupant Position ID	Positions That Can Be Varied
31	1-2--4--6---9-	1V	3	2	--3--5--7-8--10
32	1-2---5--7--9-	1V	3	2	--3-4--6--8--10
33	-2-3-4-5---8--	1V	1	3	1-----6-7--9-10
34	-2-3-4--6--8--	1V	1	3	1----5--7--9-10
35	-2-3--5--7-8--	1V	1	3	1---4--6---9-10
36	--3-4-5---8-9-	1V	1	3	1-2----6-7---10
37	-2-3-4-5-6----	1V	1	5	1------7-8-9-10
38	-2--4-5-6--8--	1V	1	5	1--3----7--9-10
39	-2-3--5-6-7---	1V	1	5	1---4----8-9-10
40	--3-4-5-6---9-	1V	1	5	1-2-----7-8--10
41	-2-3-4-5--7---	1V	1	4	1-----6--8-9-10
42	-2--4-5--7-8--	1V	1	4	1--3---6---9-10
43	-2-3-4--6-7---	1V	1	4	1----5---8-9-10
44	--3-4-5--7--9-	1V	1	4	1-2----6--8--10
45	-2-3-4-5----9-	1V	1	2	1-----6-7-8--10
46	-2--4-5---8-9-	1V	1	2	1--3---6-7---10
47	-2-3-4--6---9-	1V	1	2	1----5--7-8--10
48	-2-3--5--7--9-	1V	1	2	1---4--6--8--10
49	1-2-3-4----8--	1V	5	3	----5-6-7--9-10
50	-2-3-4----8--10	1V	5	3	1----5-6-7--9-
51	1-2-3----7-8--	1V	5	3	---4-5-6---9-10
52	1--3-4----8-9-	1V	5	3	-2---5-6-7---10
53	1-2-3-4-----10	1V	5	1	----5-6-7-8-9-
54	1-2--4----8--10	1V	5	1	--3--5-6-7--9-
55	1-2-3----7---10	1V	5	1	---4-5-6--8-9-
56	1--3-4-----9-10	1V	5	1	-2---5-6-7-8--
57	1-2-3-4---7---	1V	5	4	----5-6--8-9-10
58	1-2--4---7-8--	1V	5	4	--3--5-6---9-10
59	-2-3-4---7---10	1V	5	4	1----5-6--8-9-
60	1--3-4---7--9-	1V	5	4	-2---5-6--8--10
61	1-2-3-4-----9-	1V	5	2	----5-6-7-8--10
62	1-2--4----8-9-	1V	5	2	--3--5-6-7---10
63	-2-3-4-----9-10	1V	5	2	1----5-6-7-8--

Table 2. *Continued*

(Continued on next page)

	A	B	C	D	E
	Fixed Half-Site Positions	Reverse-Complement Vacancies & Double Occupants	Reverse-Complement Vacancy Position ID	Reverse-Complement Double Occupant Position ID	Positions That Can Be Varied
64	1-2-3----7--9-	1V	5	2	---4-5-6--8--10
65	1-2-3--5---8--	1V	4	3	---4--6-7--9-10
66	-2-3--5---8--10	1V	4	3	1---4--6-7--9-
67	1-2-3---6--8--	1V	4	3	---4-5--7--9-10
68	1--3--5---8-9-	1V	4	3	-2--4--6-7---10
69	1-2-3--5-----10	1V	4	1	---4--6-7-8-9-
70	1-2---5---8--10	1V	4	1	--3-4--6-7--9-
71	1-2-3---6----10	1V	4	1	---4-5--7-8-9-
72	1--3--5----9-10	1V	4	1	-2--4--6-7-8--
73	1-2-3--5-6----	1V	4	5	---4---7-8-9-10
74	1-2---5-6--8--	1V	4	5	--3-4---7--9-10
75	-2-3--5-6----10	1V	4	5	1---4---7-8-9-
76	1--3--5-6---9-	1V	4	5	-2--4---7-8--10
77	1-2-3--5----9-	1V	4	2	---4--6-7-8--10
78	1-2---5---8-9-	1V	4	2	--3-4--6-7---10
79	-2-3--5----9-10	1V	4	2	1---4--6-7-8--
80	1-2-3---6---9-	1V	4	2	---4-5--7-8--10
81	1--3-4-5---8--	1V	2	3	-2----6-7--9-10
82	--3-4-5---8--10	1V	2	3	1-2----6-7--9-
83	1--3-4--6--8--	1V	2	3	-2---5--7--9-10
84	1--3--5--7-8--	1V	2	3	-2--4--6---9-10
85	1--3-4-5-----10	1V	2	1	-2----6-7-8-9-
86	1---4-5---8--10	1V	2	1	-2-3---6-7--9-
87	1--3-4--6----10	1V	2	1	-2---5--7-8-9-
88	1--3--5--7---10	1V	2	1	-2--4--6--8-9-
89	1--3-4-5-6----	1V	2	5	-2-----7-8-9-10
90	1---4-5-6--8--	1V	2	5	-2-3----7--9-10
91	--3-4-5-6----10	1V	2	5	1-2-----7-8-9-
92	1--3--5-6-7---	1V	2	5	-2--4----8-9-10
93	1--3-4-5--7---	1V	2	4	-2----6--8-9-10
94	1---4-5--7-8--	1V	2	4	-2-3---6---9-10
95	--3-4-5--7---10	1V	2	4	1-2----6--8-9-
96	1--3-4--6-7---	1V	2	4	-2---5---8-9-10

Table 2. *Continued*

(Continued on next page)

	A	B	C	D	E
	Fixed Half-Site Positions	Reverse-Complement Vacancies & Double Occupants	Reverse-Complement Vacancy Position ID	Reverse-Complement Double Occupant Position ID	Positions That Can Be Varied
			Two Vacancies		
97	-2--4-5--7--9-	2V	1–3	2–4	1--3---6--8--10
98	-2--4-5-6---9-	2V	1–3	2–5	1--3----7-8--10
99	-2--4-5-6-7---	2V	1–3	4–5	1--3-----8-9-10
100	1-2--4---7--9-	2V	3–5	2–4	--3--5-6--8--10
101	1-2--4-----9-10	2V	3–5	1–2	--3--5-6-7-8--
102	1-2--4---7---10	2V	3–5	1–4	--3--5-6--8-9-
103	1-2---5-6---9-	2V	3–4	2–5	--3-4---7-8--10
104	1-2---5----9-10	2V	3–4	1–2	--3-4--6-7-8--
105	1-2---5-6----10	2V	3–4	1–5	--3-4---7-8-9-
106	1---4-5-6-7---	2V	2–3	4–5	-2-3-----8-9-10
107	1---4-5--7---10	2V	2–3	1–4	-2-3---6--8-9-
108	1---4-5-6----10	2V	2–3	1–5	-2-3----7-8-9-
109	-2-3-4---7--9-	2V	1–5	2–4	1----5-6--8--10
110	-2-3-4----8-9-	2V	1–5	2–3	1----5-6-7---10
111	-2-3-4---7-8--	2V	1–5	3–4	1----5-6---9-10
112	-2-3--5-6---9-	2V	1–4	2–5	1---4---7-8--10
113	-2-3--5---8-9-	2V	1–4	2–3	1---4--6-7---10
114	-2-3--5-6--8--	2V	1–4	3–5	1---4---7--9-10
115	--3-4-5-6-7---	2V	1–2	4–5	1-2------8-9-10
116	--3-4-5--7-8--	2V	1–2	3–4	1-2----6---9-10
117	--3-4-5-6--8--	2V	1–2	3–5	1-2-----7--9-10
118	1-2-3------9-10	2V	4–5	1–2	---4-5-6-7-8--
119	1-2-3-----8-9-	2V	4–5	2–3	---4-5-6-7---10
120	1-2-3-----8--10	2V	4–5	1–3	---4-5-6-7--9-
121	1--3-4---7---10	2V	2–5	1–4	-2---5-6--8-9-
122	1--3-4---7-8--	2V	2–5	3–4	-2---5-6---9-10
123	1--3-4----8--10	2V	2–5	1–3	-2---5-6-7--9-
124	1--3--5-6----10	2V	2–4	1–5	-2--4---7-8-9-
125	1--3--5-6--8--	2V	2–4	3–5	-2--4---7--9-10
126	1--3--5---8—10	2V	2–4	1–3	-2--4--6-7--9-

Table 2. Continued

Column A: Categorizing the 61,236 5-nt variant DNA elements into 252 half-site groups, defined by the five positions that are fixed/not varied. The 252 half-site groups are symmetrically split into a set of 126 groups and their 126 reverse-complements. Column B: These 126 groups further split into three distinct subgroups (zero vacancies, one vacancy, two vacancies) depending on how many reverse-complement vacancies are in the DNA element (i.e., the information for which nucleotide occupies each of the 10 primary positions of the 0-nt variant consensus palindromic DNA element is missing/replaced by variants/vacant in the

position and its reverse-complement position [its palindromic position pair]). Of these 126 groups, 16 have zero vacancies, 80 have one vacancy, and 30 have two vacancies. Column C: The reverse-complement vacancy position ID indicates which of the five palindromic position pairs (i.e., 1-10, 2-9, 3-8, 4-7, 5-6) are missing/replaced by variants/vacant. Column B: The number of reverse-complement vacancies equals the number of reverse-complement double occupants in the DNA element (i.e., the information for which nucleotide occupies each of the 10 primary positions of the 0-nt variant consensus palindromic DNA element is occupied in the position and its reverse-complement position [its palindromic position pair]). Column D: The reverse-complement double occupant position ID indicates which of the five palindromic position pairs (i.e., 1-10, 2-9, 3-8, 4-7, 5-6) are doubly occupied. Column E: The positions that are replaced by variants or not fixed for 5-nt variant DNA elements, and may be replaced by variants for 0-nt to 4-nt variant DNA elements.

4-7 > 3-8 (Table S6). Note: this analysis includes upper and lower one-tailed Poisson significance thresholds at $p < 0.001$; thus the probability of the TF DNA-binding signal occurring outside these boundaries by chance is less than one in a thousand.

This highly structured symmetrical sNR DNA-binding profile (i.e., the left side [126 half-site groups] versus the right side [126 reverse-complements]) demonstrates that there is no sNR DNA-binding bias toward either the 5′ side (GGTCA or GAACA) or the 3′ side (TGACC or TGTTC) of the DNA element at any 1- to 5-nt variant ERE or HRE DNA elements in the genome (Figures 2, 3, S4, and S5). This symmetrical sNR DNA-binding profile also confirms that sNR binding in the genome follows inversion symmetry through 5-nt variant DNA elements (Figures 2, 3, S4, and S5). That is, the number of sNR DNA-binding events at a 1- to 5-nt variant DNA element in the genome is equivalent to the number of sNR DNA-binding events at its reverse-complement DNA element in the genome (Figures 2, 3, S4, and S5).

Furthermore, this highly structured symmetrical sNR DNA-binding profile demonstrates that the specific DNA elements where sNRs bind in the genome is determined by internal inversion symmetry within the DNA element, and this continues through 5-nt variant DNA elements (Figures 2, 3, S4, and S5). That is, sNRs exhibit preferential DNA-binding at particular DNA elements in the genome, and this preference is determined by internal inversion symmetry within the DNA element (i.e., sNR binding in the genome is determined by the position of variants within the DNA element and its reverse-complement position) (Figures 2, 3, S4, and S5). Thus, the specific 0- to 5-nt variant DNA elements that sNRs bind in the genome is determined by the position of variants within the DNA element and its reverse-complement position (Figures 2, 3, S4, and S5).

In addition, sNR DNA-binding at "palindrome half-sites" (half-site group #1 and half-side group #252 on the x-axis) are not particularly distinct compared with the other 250 half-site groups (Figures 2, 3, S4, and S5). Because sNR DNA-binding at 0- to 5-nt variant DNA elements in the genome is a progression that decays from 0- to 5-nt variant DNA elements, this confirms that sNR DNA-binding at 0- to 5-nt variant DNA elements in the genome requires all the information contained within the entire DNA element (i.e., on both sides of the 3-nt spacer) (Figures 2, 3, S4, and S5). Thus, the information is stored in the "complex" as a whole, not in simple physical subsets of the DNA element. This analysis therefore challenges previous suggestions that sNRs bind "palindrome half-sites" in the genome as individual monomers (Figures 2, 3, S4, and S5). The two endpoints (half-site group #1 and half-site group #252 on the x-axis) represent sNR DNA-binding at all DNA elements where variants are confined to the same side of the 3-nt spacer (i.e., the left-hand or right-hand side of the 0-nt variant consensus palindromic DNA element remains fixed) for all 1- to 5-nt variant DNA elements in the genome (Figures 2, 3, S4, and S5).

Moreover, this highly structured symmetrical sNR DNA-binding profile at 0- to 5-nt variant DNA elements in the genome was observed in 157 ER experiments and 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) (Tables S7 and S8). Of particular importance, there are minimal, if any, peaks in sNR ChIPSeq or ChIPExo experiments that do not contain a 0- to 5-nt variant DNA element, even at the least stringent (L4) peak selection criteria (Tables S7 and S8).

Quantify the Discrete States of sNR DNA-Binding at 1-nt Variant DNA Elements in the Genome

(S/N) analysis of sNR DNA-binding at 0- to 5-nt variant DNA elements in the genome (displayed by the 252 half-site groups) reveals a highly structured symmetrical DNA-binding profile associated with discrete (S/N) values that decay as the sNR DNA-binding signal decreases from 1- to 5-nt variant DNA elements (Figures 2, 3, S4, and S5). Specifically, (S/N) analysis of sNR DNA-binding at 1-nt variant DNA elements (displayed by the 252 half-site groups) produce a set of symmetrical discrete DNA-binding signals: five plateaus for ER DNA-binding

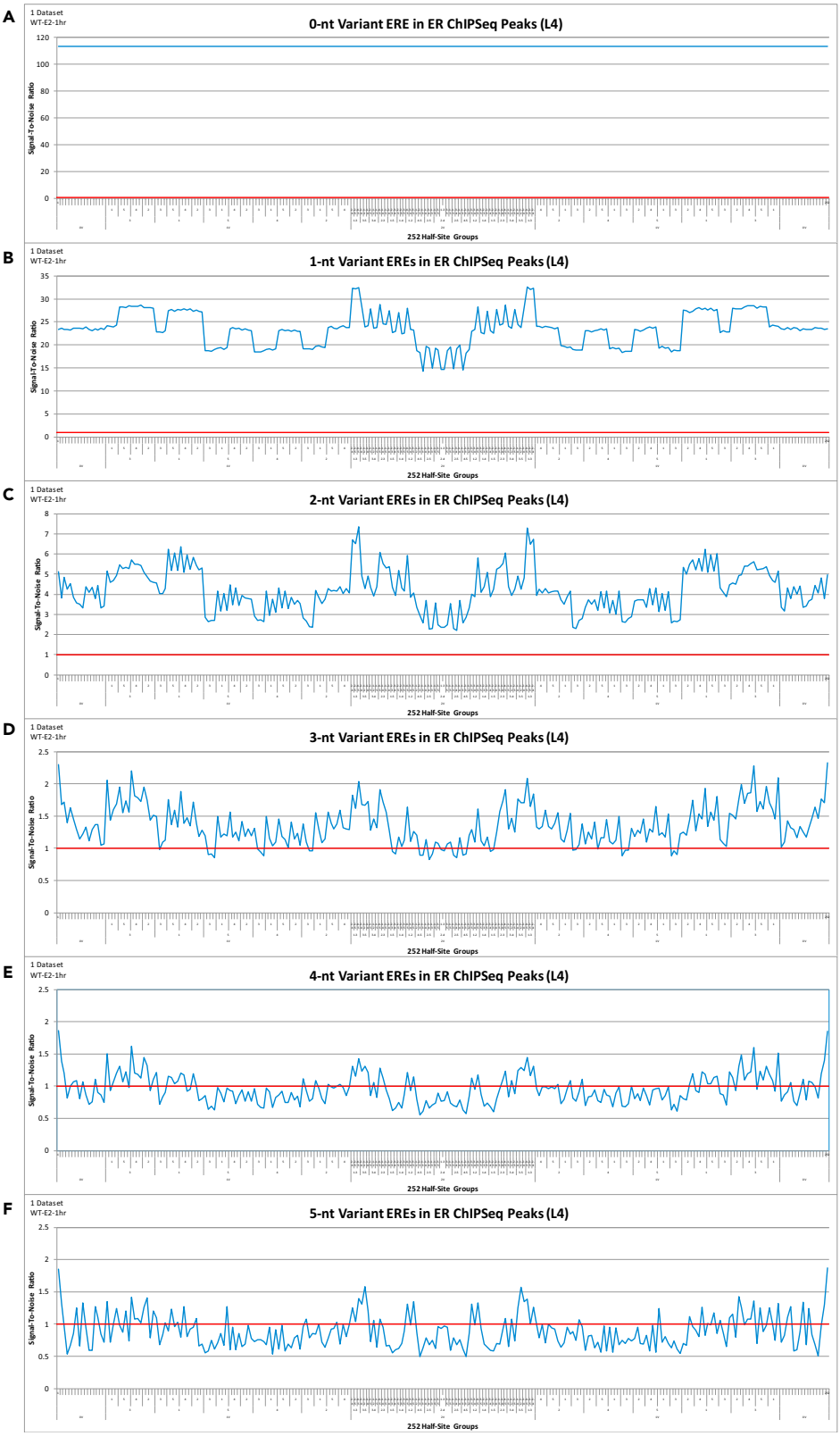


Figure 2. (S/N) Analysis of 0- to 5-nt Variant ERE DNA Elements in ER ChIPSeq Peaks

Figure 2. *Continued*

(A–F) (S/N) analysis of 0-nt to 5-nt variant ERE DNA elements (displayed by the 252 half-site groups) in ER (WT-E2-1hr) (76,163 peaks, 146-nt peak length) ChIPSeq peaks. Upper and lower one-tailed Poisson significance thresholds at $p < 0.001$ (A) 0.19, 2.07; (B) 0.77, 1.25; (C) 0.90, 1.10; (D) 0.94, 1.06; (E) 0.95, 1.05; and (F) 0.94, 1.06. X-axis order = reverse-complement vacancy position ID 3-8 > 1-10 > 5-6 > 4-7 > 2-9. See Table S5 for x-axis details. See Tables S64–S67 for step-by-step instructions of the data analysis from peak selection to (S/N) analysis of ER DNA-binding at 81,922 0-nt to 5-nt variant ERE DNA elements in the genome, displayed by the 252 half-site groups.

at 1-nt variant ERE DNA elements in the genome (Figure S6) and three plateaus for KR DNA-binding at 1-nt variant HRE DNA elements in the genome (Figure S7). These plateaus characterize the internal inversion symmetry contained within the DNA element, which determines the specific DNA elements that sNRs bind in the genome. These plateaus can be represented as (+2, +1, 0, −1, −2) for ER DNA-binding at 1-nt variant EREs (Figure S6) and (+1, 0, −1) for KR DNA-binding at 1-nt variant HREs in the genome (Figure S7).

ER DNA-binding at 1-nt variant ERE DNA elements in the genome produce five discrete DNA-binding signals across the 252 half-site groups: 6 groups reach the (+2) plateau, 60 groups reach the (+1) plateau, 120 groups reach the (0) plateau, 60 groups reach the (−1) plateau, and 6 groups reach the (−2) plateau (Figure S6). Thus, ER DNA-binding is greatest at the (+2) plateau, which are the six half-site groups that have two reverse-complement vacancies in palindromic position pair 3-8 and 1-10, and two reverse-complement double occupants in palindromic position pair (4-7 and 5-6), (2-9 and 5-6), or (2-9 and 4-7) (Figure S6). ER DNA-binding is least at the (−2) plateau, which are the six half-site groups that have two reverse-complement vacancies in palindromic position pair (4-7 and 5-6), (2-9 and 5-6), or (2-9 and 4-7) and two reverse-complement double occupants in palindromic position pair 3-8 and 1-10 (Figure S6).

In contrast, KR DNA-binding at 1-nt variant HRE DNA elements in the genome produce three discrete DNA-binding signals across the 252 half-site groups: 56 groups reach the (+1) plateau, 140 groups reach the (0) plateau, and 56 groups reach the (−1) plateau (Figure S7) Thus, KR DNA-binding is greatest at the (+1) plateau, which are the 56 half-site groups that have a reverse-complement vacancy in palindromic position pair 2-9 (Figure S7). KR DNA-binding is least at the (−1) plateau, which are the 56 half-site groups that have a reverse-complement double occupant in palindromic position pair 2-9 (Figure S7).

The symbolic representations of the discrete states (+2, +1, 0, −1, −2) for the ERE and (+1, 0, −1) for the HRE can be converted into experimental values (i.e., the (S/N) values) by taking the mean, standard deviation, and difference from the mean of each plateau (Table 3). That is, we replace the symbolic representations with actual (S/N) values determined from ChIPSeq and ChIPExo experiments (Table 3). The exact symmetry is demonstrated by calculating the difference from the mean of each plateau (Table 3). Thus, these calculations quantitatively demonstrate the sNR DNA-binding inversion symmetry of the discrete (S/N) spectrum (Table 3). The five discrete states for the ERE were observed in 157 ER experiments, and the three discrete states for the HRE were observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) (Tables S9 and S10).

The symbolic representations of the discrete states (+2, +1, 0, −1, −2) for the ERE and (+1, 0, −1) for the HRE can also be replaced with algebraic variables to formally define the methodology (Table 4; Figures S8 and S9). The 3-state HRE is generated by splitting the nucleotides into a (4,1,0,0,0) grouping (Table 4; Figure S8). That is, palindromic position pair 2-9 (symbolically represented by "A") versus palindromic position pairs 5-6, 1-10, 4-7, 3-8 (symbolically represented by "B") (Table 4; Figure S8). Similarly, the 5-state ERE is generated by splitting the nucleotides into a (3,2,0,0,0) grouping (Table 4; Figure S9). That is, palindromic position pair 3-8 and 1-10 (symbolically represented by "A") versus palindromic position pairs 5-6, 4-7, 2-9 (symbolically represented by "B") (Table 4; Figure S9). Therefore, if an A is replaced with a B, the impact is (–X) (i.e., the sNR DNA-binding affinity is decreased), whereas if a B is replaced with an A, the impact is exactly the opposite (+X) (i.e., the sNR DNA-binding affinity is increased) to the same magnitude (Table 4; Figures S8 and S9). Thus, this algebraic representation explains why sNR DNA-binding is quantitatively precise and also why there are three discrete DNA-binding states at 1-nt variant HRE DNA elements (by the KRs) and five discrete DNA-binding states at 1-nt variant ERE DNA elements (by the ERs) (Table 4; Figures S8 and S9). The algebraic complexities increase substantially beyond these (4,1,0,0,0) HRE and (3,2,0,0,0) ERE representations. For example, the (3,1,1,0,0) grouping (symbolically represented by "A, B, C") creates nine discrete states (Table 4, Figure S10). The (2,2,1,0,0) grouping (symbolically represented by "A, B, C") creates 11 discrete states, the (2,1,1,1,0) grouping (symbolically represented by "A, B, C,

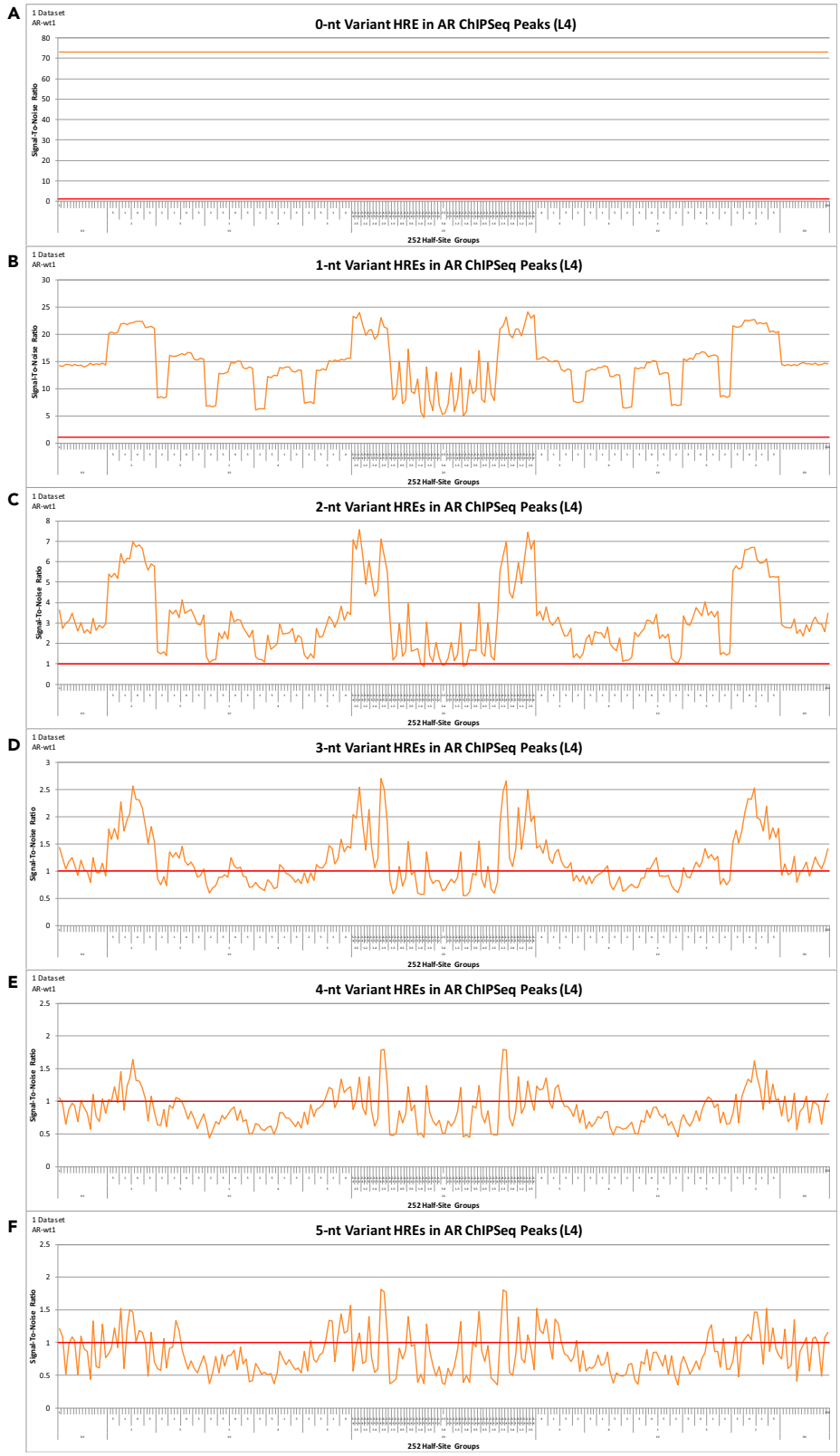


Figure 3. (S/N) Analysis of 0- to 5-nt Variant HRE DNA Elements in AR ChIPSeq Peaks

Figure 3. *Continued*

(A–F) (S/N) analysis of 0- to 5-nt variant HRE DNA elements (displayed by the 252 half-site groups) in AR (AR-wt1) (49,859 peaks, 136-nt peak length) ChIPSeq peaks. Upper and lower one-tailed Poisson significance thresholds at $p < 0.001$ (A) 0.00, 2.47; (B) 0.70, 1.33; (C) 0.87, 1.13; (D) 0.93, 1.08; (E) 0.94, 1.06; and (F) 0.92, 1.08. X-axis order = reverse-complement vacancy position ID 2-9 > 5-6 > 1-10 > 4-7 > 3-8. See Table S6 for x-axis details.

D") creates 25 discrete states, and the (1,1,1,1,1) grouping (symbolically represented by "A, B, C, D, E") creates 51 discrete states, which are all the remaining possibilities for a 5-nt DNA element in the genome (i.e., a 5-nt DNA element, followed by any arbitrary spacer, followed by its 5-nt reverse-complement DNA element) (Figure S11). Thus, we have provided a mathematical formula that represents the sNR DNA-binding profile observed in 157 ER experiments and 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) (Tables S9 and S10).

Internal Inversion Symmetry of sNR Binding in the Genome

(S/N) analysis of sNR DNA-binding at each of the 81,922 0- to 5-nt variant DNA elements in the genome (displayed by the 252 half-site groups) clearly demonstrates that sNR binding in the genome follows inversion symmetry (i.e., the number of sNR DNA-binding events at a particular DNA element in the genome is equivalent to the number of sNR DNA-binding events at its reverse-complement DNA element in the genome). The decay of the symmetrical sNR DNA-binding profile from 0- to 5-nt variant DNA elements in the genome confirms that the specific 0- to 5-nt variant DNA elements where sNRs bind in the genome is determined by internal inversion symmetry within the DNA element (i.e., sNR binding in the genome is determined by the position of variants within the DNA element and its reverse-complement position). Because it is difficult to translate the 252 half-site groups to specific DNA elements, the absolute counts of each of the 81,922 0- to 5-nt variant DNA elements in an experiment can also be displayed by the position of the variant within the DNA element, rather than by the 252 half-site groups. That is, the 0- to 5-nt variant DNA elements include the 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements (10 variant positions), 405 2-nt variant DNA elements (45 variant positions), 3,240 3-nt variant DNA elements (120 variant positions), 17,010 4-nt variant DNA elements (210 variant positions), and 61,236 5-nt variant DNA elements (252 variant positions), for a total of 81,922 DNA elements (Table 1).

Internal Inversion Symmetry of ER DNA-Binding at ERE DNA Elements in the Genome

As previously discussed, ER DNA-binding at 1-nt variant ERE DNA elements in the genome is increased at DNA elements with variants in palindromic position pair 3-8 or 1-10 (Figures 1 and S12). This ER DNA-binding profile at 1-nt variant ERE DNA elements in the genome was observed in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) (Table S11).

These 1-nt variant ERE DNA-binding rules (i.e., increased DNA-binding at DNA elements with variants in palindromic position pair 3-8 or 1-10) are also applicable to ER DNA-binding at 2-nt variant ERE DNA elements in the genome (Figure S13). Thus, ER DNA-binding at 2-nt variant ERE DNA elements is increased at DNA elements that have variants in positions 1, 3, 8, or 10 (in any combination with each other) in the genome (i.e., (3-8), (1-10), (1,3)-(8,10), (1,8)-(3,10)) (Figure S13). In addition, ER DNA-binding is also increased at DNA elements that have variants in palindromic position pair (2,3)-(8,9) (i.e., two of the four positions that distinguish the 0-nt variant consensus palindromic ERE DNA element from the 0-nt variant consensus palindromic HRE DNA element) (Figure S13). However, the increased DNA-binding of ER at these DNA elements is reduced (i.e., suppressed) at DNA elements where the variants "crossover" the 3-nt spacer in the genome (i.e., have variants on both sides of the 3-nt spacer) (i.e., (2,8)-(3,9)) (Figure S13). Furthermore, ER DNA-binding is increased at DNA elements with variants in palindromic position pair (3,5)-(6,8), as well as at DNA elements with variants in palindromic position pair (4,5)-(6,7) (Figure S13). Likewise, the increased DNA-binding of ER at these DNA elements is suppressed at DNA elements where the variants crossover the 3-nt spacer (i.e., (3,6)-(5,8), (4,6)-(5,7)) (Figure S13). Thus, the ER DNA-binding rules established at 1-nt variant ERE DNA elements in the genome (i.e., increased DNA-binding at DNA elements with variants in palindromic position pair 3-8 or 1-10) are also applicable to ER DNA-binding at 2-nt variant ERE DNA elements in the genome (i.e., increased DNA-binding at DNA elements with variants in positions 1, 3, 8, or 10 in any combination with each other). That is, the first set of DNA-binding rules (defined by ER DNA-binding at 1-nt variant ERE DNA elements in the genome) is based on which palindromic position pair the variant occupies (i.e., ER DNA-binding is increased at DNA elements with variants in palindromic

(A) Signal-To-Noise Ratio (S/N) Analysis of 1-nt Variant ERE DNA Elements in ER (WT-E2-1hr) ChIPSeq Peaks (Figure 2)					
5-State	Algebraic Variable	252 Half-Site Groups	Mean	Standard Deviation	Difference from Mean
+2	+2X	6	32.30	0.23	+8.83
+1	+X	60	27.89	0.42	+4.41
0	0	120	23.47	0.47	0.00
−1	-X	60	19.06	0.42	−4.41
−2	-2X	6	14.64	0.23	−8.83
(B) Signal-To-Noise Ratio (S/N) Analysis of 1-nt Variant HRE DNA Elements in AR (AR-wt1) ChIPSeq Peaks (Figure 3)					
3-State	Algebraic Variable	252 Half-Site Groups	Mean	Standard Deviation	Difference from Mean
+1	+X	56	21.51	1.16	+7.11
0	0	140	14.41	1.19	0.00
−1	-X	56	7.30	1.16	−7.11
(C) Signal-To-Noise Ratio (S/N) Analysis of 1-nt Variant HRE DNA Elements in GR (GR-WT-pred-6am-2) ChIPSeq Peaks (Figure S4)					
3-State	Algebraic Variable	252 Half-Site Groups	Mean	Standard Deviation	Difference from Mean
+1	+X	56	16.30	1.03	+5.00
0	0	140	11.29	1.06	0.00
−1	-X	56	6.29	1.03	−5.00
(D) Signal-To-Noise Ratio (S/N) Analysis of 1-nt Variant HRE DNA Elements in PR (Uterus-PGR-P4) ChIPSeq Peaks (Figure S5)					
3-State	Algebraic Variable	252 Half-Site Groups	Mean	Standard Deviation	Difference from Mean
+1	+X	56	15.38	0.92	+3.46
0	0	140	11.92	0.94	0.00
−1	-X	56	8.46	0.92	−3.46

Table 3. Quantify the Discrete States of sNR DNA-Binding at 1-nt Variant DNA Elements in the Genome

(S/N) analysis of sNR DNA-binding at 0-nt to 5-nt variant DNA elements in the genome (displayed by the 252 half-site groups) reveals a highly structured symmetrical DNA-binding profile associated with discrete (S/N) values that decay as the sNR DNA-binding signal decreases from 1-nt to 5-nt variant DNA elements (A) ER DNA-binding at 1-nt variant ERE DNA elements in the genome produce five (5) discrete DNA-binding signals across the 252 half-site groups: 6 groups reach the (+2) plateau, 60 groups reach the (+1) plateau, 120 groups reach the (0) plateau, 60 groups reach the (−1) plateau, 6 groups reach the (−2) plateau. (B-D) KR DNA-binding at 1-nt variant HRE DNA elements in the genome produce three (3) discrete DNA-binding signals across the 252 half-site groups: 56 groups reach the (+1) plateau, 140 groups reach the (0) plateau, 56 groups reach the (−1) plateau. The symbolic representations of the discrete states (+2, +1, 0, −1, −2) for the ERE and (+1, 0, −1) for the HRE can be converted into experimental values (i.e., the (S/N) values) by taking the mean, standard deviation, and difference from the mean of each plateau. The exact symmetry is demonstrated by calculating the difference from the mean of each plateau. Thus, these calculations quantitatively demonstrate the sNR DNA-binding inversion symmetry of the discrete (S/N) spectrum.

DNA Element	Discrete States	Algebraic Variables	Algebraic Equations
HRE	3-State (4,1,0,0,0)	(-X), 0, (+X)	A – B = X B – A = -X
ERE	5-State (3,2,0,0,0)	(-2X), (-X), 0, (+X), (+2X)	A – B = X B – A = -X AA – BB = 2X BB – AA = -2X AB – BB = A – B = X BB – AB = B – A = -X
	9-State (3,1,1,0,0)	(-X), (-Y), (-Z), -(Y + Z), 0, (+X), (+Y), (+Z), (Y + Z)	A – B = X B – C = Y A – C = Z AB – CC = Y + Z AC – BC = A – B = X BC – AC = B – A = -X AC – CC = A – C = Z CC – AC = C – A = -Z BC – CC = B – C = Y CC – BC = C – B = -Y CC – AB = -(Y + Z)

Table 4. Algebraic Equations of sNR DNA-Binding at 1-nt Variant DNA Elements in the Genome

The symbolic representations of the discrete states (+2, +1, 0, −1, −2) for the ERE and (+1, 0, −1) for the HRE can also be replaced with algebraic variables to formally define the methodology. The 3-state HRE is generated by splitting the nucleotides into a (4,1,0,0,0) grouping. That is, palindromic position pair 2-9 (symbolically represented by "A") versus palindromic position pairs 5-6, 1-10, 4-7, 3-8 (symbolically represented by "B"). The 5-state ERE is generated by splitting the nucleotides into a (3,2,0,0,0) grouping. That is, palindromic position pair 3-8 and 1-10 (symbolically represented by "A") versus palindromic position pairs 5-6, 4-7, 2-9 (symbolically represented by "B"). Therefore, if an A is replaced with a B, the impact is (–X) (i.e., the sNR DNA-binding affinity is decreased), whereas if a B is replaced with an A, the impact is exactly the opposite (+X) (i.e., the sNR DNA-binding affinity is increased) to the same magnitude. Thus, this algebraic representation explains why sNR DNA-binding is quantitatively precise and also why there are three discrete DNA-binding states at 1-nt variant HRE DNA elements (by the KRs) and five discrete DNA-binding states at 1-nt variant ERE DNA elements (by the ERs). The algebraic complexities increase substantially beyond these (4,1,0,0,0) HRE and (3,2,0,0,0) ERE representations. For example, (3,1,1,0,0) grouping (symbolically represented by "A, B, C") creates nine discrete states.

position pair 3-8 or 1-10). Furthermore, unlike 1-nt variant DNA elements, 2-nt variant DNA elements allow variants to crossover the 3-nt spacer, thus revealing a new DNA-binding pattern not applicable to 1-nt variant DNA elements. That is, the second set of DNA-binding rules (functionally subordinate to and independent of the DNA-binding rules defined by ER DNA-binding at 1-nt variant ERE DNA elements) defined by ER DNA-binding at 2-nt variant ERE DNA elements in the genome, is based on how the position of the variants relate to each other (i.e., ER DNA-binding is increased at DNA elements with variants that do not crossover the 3-nt spacer). To summarize, the second set of ER DNA-binding rules (functionally subordinate to and independent of the DNA-binding rules defined by ER DNA-binding at 1-nt variant DNA elements) defined by ER DNA-binding at 2-nt variant ERE DNA elements in the genome include increased ER DNA-binding at DNA elements with variants in palindromic position pair (2,3)-(8,9), (3,5)-(6,8), and (4,5)-(6,7); this increased DNA-binding of ER is suppressed at DNA elements where these variants crossover the 3-nt spacer (Figure S13). This ER DNA-binding profile at 2-nt variant ERE DNA elements in the genome was observed in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) (Table S12).

Analysis of ER DNA-binding at 3-nt variant ERE DNA elements in the genome demonstrates the same properties that were observed in ER DNA-binding at 1-nt variant ERE DNA elements and 2-nt variant ERE DNA elements in the genome (Figure S14). First, ER DNA-binding at 3-nt variant ERE DNA elements is increased at DNA elements with variants in palindromic position pair (1,3)-(8,10), and the third variant is on the same side of the 3-nt spacer (i.e., (1,3,5)-(6,8,10), (1,3,4)-(7,8,10), (1,2,3)-(8,9,10)) (Figure S14). Thus, confirming that the second set of DNA-binding rules, based on how the position of the variants relate to each other (i.e., ER DNA-binding is increased at DNA elements with variants that do not crossover the 3-nt spacer), is exhibited

in ER DNA-binding at 3-nt variant ERE DNA elements in the genome. Second, ER DNA-binding is increased at DNA elements with variants in palindromic position pair 3-8 and any other variant (i.e., (1,3,8)-(3,8,10), (3,5,8)-(3,6,8), (3,4,8)-(3,7,8), (2,3,8)-(3,8,9)) (Figure S14). Thus, confirming that the first set of DNA-binding rules, based on which palindromic position pair the variant occupies (i.e., ER DNA-binding is increased at DNA elements with variants in palindromic position pair 3-8 or 1-10), is exhibited in ER DNA-binding at 3-nt variant ERE DNA elements in the genome. However, this increased DNA-binding of ER was less apparent at DNA elements with variants in palindromic position pair 1-10 (i.e., (1,5,10)-(1,6,10), (1,3,10)-(1,8,10), (1,4,10)-(1,7,10), (1,2,10)-(1,9,10)) (Figure S14). Thus, ER prefers to bind 3-nt variant ERE DNA elements with variants in palindromic position pair (1,3)-(8,10) and the third variant being on the same side of the 3-nt spacer, rather than DNA elements with variants in palindromic position pair 1-10 (Figure S14). Furthermore, other than the increased binding of ER at DNA elements with variants in palindromic position pair 3-8, ER DNA-binding was suppressed if the third variant crossed over the 3-nt spacer for all remaining 1, 3, 8, or 10 variant combinations (i.e., (3,4,5)-(6,7,8), (2,3,5)-(6,8,9), (2,3,4)-(7,8,9), (1,4,5)-(6,7,10), (1,2,5)-(6,9,10), (1,2,4)-(7,9,10)) (Figure S14). This demonstrates that the ER DNA-binding rules established at 1-nt variant ERE DNA elements in the genome (i.e., increased DNA-binding at DNA elements with variants in palindromic position pair 3-8 and 1-10), and the distinct and functionally subordinate ER DNA-binding rules established at 2-nt variant ERE DNA elements in the genome (i.e., increased DNA-binding at DNA elements with variants on the same side of the 3-nt spacer) are co-existing and being maintained in ER DNA-binding at 3-nt variant ERE DNA elements in genome (Figure S14). This ER DNA-binding profile at 3-nt variant ERE DNA elements in the genome was observed in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) (Table S13).

These ER DNA-binding rules are maintained at 4-nt variant ERE DNA elements (Figure S15, Table S14) and 5-nt variant ERE DNE elements in the genome (Figure S16, Table S15). Thus, ER DNA-binding at 1- to 5-nt variant ERE DNA elements in the genome is determined by two distinct set of rules: (1) a first set established by ER DNA-binding at 1-nt variant ERE DNA elements, which favor binding at DNA elements with variants in palindromic position pair 3-8 and 1-10, and (2) a distinct and functionally subordinate second set, not applicable to 1- variant DNA elements, which favor binding at DNA elements with variants on the same side of the 3-nt spacer. These ER DNA-binding rules are applicable to ER DNA-binding from 1- to 5-nt variant ERE DNA elements in the genome (Figures 2 and S12–S16). Thus, the specific 1- to 5-nt variant ERE DNA elements where ER binds in the genome is determined by internal inversion symmetry within the DNA element (i.e., ER binding in the genome is determined by the position of variants within the DNA element and its reverse-complement position). Increased ER DNA-binding at specific 1- to 5-nt variant ERE DNA elements in the genome is determined by the reverse-complement vacancy in the palindromic position pair by the following hierarchy: 3-8 > 1-10 > 5-6 > 4-7 > 2-9 (Table S5).

Internal Inversion Symmetry of KR DNA-Binding at HRE DNA Elements in the Genome

As previously discussed, KR DNA-binding at 1-nt variant HRE DNA elements in the genome is increased at DNA elements with variants in palindromic position pair 2-9 (Figures 1B and S17). This KR DNA-binding profile at 1-nt variant HRE DNA elements in the genome was observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) (Table S16).

These 1-nt variant HRE DNA-binding rules (i.e., increased DNA-binding at DNA elements with variants in palindromic position pair 2-9) are also applicable to KR DNA-binding at 2-nt variant HRE DNA elements in the genome (Figure S18). Thus, KR DNA-binding at 2-nt variant HRE DNA elements is increased at DNA elements that have variants in palindromic position pair 2-9, followed by 2-5, 1-2, and 2-3 (and their reverse-complements 6-9, 9-10, 8-9) in the genome (Figure S18). However, the increased DNA-binding of KR at these DNA elements is reduced (i.e., suppressed) at DNA elements where the variants crossover the 3-nt spacer in the genome (i.e., 2-6, 1-9, 2-8, and their reverse-complements 5-9, 2-10, 3-9) (Figure S18). Similar to ER, KR DNA-binding is also increased at DNA elements with variants in palindromic position pair (4,5)-(6,7), potentially revealing a pre-evolved sNR DNA-binding mechanism (Figure S18). Thus, the KR DNA-binding rules established at 1-nt variant HRE DNA elements in the genome (i.e., increased DNA-binding at DNA elements with variants in position 2 or 9) are also applicable to KR DNA-binding at 2-nt variant HRE DNA elements in the genome (i.e., increased DNA-binding at DNA elements with variants in palindromic position pair 2-9). That is, the first set of DNA-binding rules (defined by KR DNA-binding at 1-nt

variant HRE DNA elements in the genome) is based on which palindromic position pair the variant occupies (i.e., KR DNA-binding is increased at DNA elements with variants in palindromic position pair 2-9). Furthermore, unlike 1-nt variant DNA elements, 2-nt variant DNA elements allow variants to crossover the 3-nt spacer, thus revealing a new DNA-binding pattern not applicable to 1-nt variant DNA elements. That is, the second set of DNA-binding rules (functionally subordinate to and independent of the DNA-binding rules defined by KR DNA-binding at 1-nt variant HRE DNA elements) defined by KR DNA-binding at 2-nt variant HRE DNA elements in the genome, is based on how the position of the variants relate to each other (i.e., KR DNA-binding is increased at DNA elements with variants that do not crossover the 3-nt spacer). To summarize, the second set of KR DNA-binding rules (functionally subordinate to and independent of the DNA-binding rules defined by KR DNA-binding at 1-nt variant HRE DNA elements) defined by KR DNA-binding at 2-nt variant HRE DNA elements in the genome include increased KR DNA-binding at DNA elements with variants in position 2 or 9 and the second variant is on the same side of the 3-nt spacer, and if variants are in palindromic position pair (4,5)-(6,7); this increased DNA-binding of KR is suppressed at DNA elements where these variants crossover the 3-nt spacer (Figure S18). This KR DNA-binding profile at 2-nt variant HRE DNA elements in the genome was observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) (Table S17).

Analysis of KR DNA-binding at 3-nt variant HRE DNA elements in the genome demonstrates the same properties that were observed in KR DNA-binding at 1-nt variant HRE DNA elements and 2-nt variant HRE DNA elements in the genome (Figure S19). First, KR DNA-binding at 3-nt variant HRE DNA elements is increased at DNA elements with variants in position 2 or 9, and the remaining two variants are on the same side of the 3-nt spacer (i.e., (1,2,5)-(6,9,10), (2,4,5)-(6,7,9), (2,3,5)-(6,8,9), (1,2,3)-(8,9,10)) (Figure S19). Thus, confirming that the second set of DNA-binding rules, based on how the position of the variants relate to each other (i.e., KR DNA-binding is increased at DNA elements with variants that do not crossover the 3-nt spacer), is exhibited in KR DNA-binding at 3-nt variant HRE DNA elements in the genome. Second, KR DNA-binding is increased at DNA elements with variants in palindromic position pair 2-9 with any other variant (i.e., (1,2,9)-(2,9,10), (2,5,9)-(2,6,9), (2,4,9)-(2,7,9), (2,3,9)-(2,8,9)) (Figure S19). Thus, confirming that the first set of DNA-binding rules, based on which palindromic position pair the variant occupies (i.e., KR DNA-binding is increased at DNA elements with variants in palindromic position pair 2-9), is exhibited in KR DNA-binding at 3-nt variant HRE DNA elements in the genome. This demonstrates that the KR DNA-binding rules established at 1-nt variant HRE DNA elements in the genome (i.e., increased DNA-binding at DNA elements with variants in palindromic position pair 2-9) and the distinct and functionally subordinate KR DNA-binding rules established at 2-nt variant HRE DNA elements in the genome (i.e., increased DNA-binding at DNA elements with variants on the same side of the 3-nt spacer) are co-existing and being maintained in KR DNA-binding at 3-nt variant HRE DNA elements in genome (Figure S19). This KR DNA-binding profile at 3-nt variant HRE DNA elements in the genome was observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) (Table S18).

These KR DNA-binding rules are maintained at 4-nt variant HRE DNA elements (Figure S20, Table S19) and 5-nt variant HRE DNA elements in the genome (Figure S21, Table S20). Thus, KR DNA-binding at 1- to 5-nt variant HRE elements in the genome is determined by two distinct set of rules: (1) a first set established by KR DNA-binding at 1-nt variant HRE DNA elements, which favor binding at DNA elements with variants in palindromic position pair 2-9 and (2) a distinct and functionally subordinate second set, not applicable to 1-nt variant DNA elements, which favor binding at DNA elements with variants on the same side of the 3-nt spacer. These KR DNA-binding rules are applicable to KR DNA-binding from 1-nt to 5-nt variant HRE DNA elements in the genome and for all members of the KR family (Figures 3, S4, S5, and S17–S21). Thus, the specific 1- to 5-nt variant HRE DNA elements where KR binds in the genome is determined by internal inversion symmetry within the DNA element (i.e., KR binding in the genome is determined by the position of variants within the DNA element and its reverse-complement position). Increased KR DNA-binding at specific 1- to 5-nt variant HRE DNA elements in the genome is determined by the reverse-complement vacancy in the palindromic position pair by the following hierarchy: 2-9 > 5-6 > 1-10 > 4-7 > 3-8 (Table S6).

Distinguishing NRFE (Functional) from Non-NRFE (Non-Functional) DNA Elements in the Genome

In a previously published study, we determined that only a small fraction of all sNR chromatin-interacting events observed in ChIPSeq and ChIPExo experiments is associated with transcriptional activity, and this

functionality is restricted to DNA elements that vary from the consensus palindromic DNA element by one or two nts (i.e., NRFEs) (Coons et al., 2017). This finding was observed in a wide variety of models including DBD mutant models, cancerous tumors that were sensitive to hormone deprivation versus tumors that had acquired resistance, constitutively active models, SUMOylation mutant models, etc. (Coons et al., 2017). This functionality was further confirmed via hormone-mediated recruitment of a variety of genomic features that have been correlated with enhancer activation, including enhancer RNA (eRNA) transcription, RNAPII occupancy, and coregulator and TF recruitment (Coons et al., 2017). This study included analysis of over 1300 experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) (Coons et al., 2017).

One of the models used in our previous study to dissect the relative biological roles of *cis*-regulatory DBD-dependent transcriptional regulation was GR-Dim, a GR DBD mutant mouse model (A458T) (Coons et al., 2017). Tissues derived from GR-Dim are refractory to endogenous glucocorticoids (i.e., no hormone-mediated transcriptional response, neither induction nor repression) (Kleiman et al., 2012; Frijters et al., 2010; Lim et al., 2015; Berg, 1998; Schiller et al., 2014). Using GR-Dim as a specific example from the numerous models analyzed in the previous study, analysis of GR-Dim DNA-binding at 0- to 5-nt variant HRE DNA elements in the genome demonstrates a dramatic reduction in DNA-binding at 0- to 2-nt variant HRE DNA elements in the genome compared to wtGR, consistent with the conclusions observed in our previous study using a different analysis methodology (Coons et al., 2017) (Figures 4 and S22–S24). Other than a slight decrease in DNA-binding at a few specific 3-nt variant HRE DNA elements, the DNA-binding profiles were identical between GR-Dim and wtGR at 3- to 5-nt variant HRE DNA elements in the genome, including at "palindrome half-sites" (i.e., half-site group #1 and half-site group #252 on the x-axis) (Figures 4 and S22–S24). This confirms that GR DNA-binding at "palindrome half-sites" is not functionally active (Figures 4 and S22–S24). The two endpoints (half-site group #1 and half-side group #252 on the x-axis) represent GR DNA-binding at all DNA elements where variants are confined to the same side of the 3-nt spacer (i.e., the left-hand or right-hand side of the 0-nt variant consensus palindromic DNA element remains fixed) for all 1- to 5-nt variant DNA elements in the genome (Figures 4 and S22–S24). Furthermore, although GR-Dim has a significantly reduced DNA-binding signal at NRFE HRE DNA elements, the DNA-binding profiles of GR-Dim and wtGR are highly structured and converge onto the same ((S/N) value = 1) noise profile (Figures 4 and S22–S24). This would not occur if this DNA-binding was random (i.e., two random signals do not map on top of each other), and thus these identical DNA-binding profiles demonstrate inherent structure in DNA-binding through 5-nt variant DNA elements in the genome (Figures 4 and S22–S24). Note that the slight decrease in DNA-binding at a few specific 3-nt variant HRE DNA elements by GR-Dim is made apparent by the DNA-binding enhancement of a particular experiment at 0- to 2-nt variant HRE DNA elements as the relative (S/N) values of sNR DNA-binding signals are scale invariant (i.e., the relative ratios between the 0- to 5-nt variant groups are constant) (Figures S2 and S3).

Of particular importance, the DNA-binding curve of GR-Dim has a nearly identical structure as wtGR at 0- to 5-nt variant HRE DNA elements in the genome, with significantly reduced amplitude compared with wtGR at 0- to 2-nt variant HRE DNA elements in the genome (Figures 4 and S22–S24). This demonstrates that the effect of the DBD mutation is a loss in GR residence time at the DNA element, rather than the inability to follow the same DNA-binding rules as wtGR (Figures 4 and S22–S24). A loss in residence time of GR-Dim on DNA compared with wtGR has also been identified via reflected light-sheet microscopy (Gebhardt et al., 2013). When we display this same data by the position of the variants, rather than by the 252 half-site groups, the impact of the DBD mutation is most pronounced at DNA elements where wtGR DNA-binding is most increased (Figures 5 and S25). For example, the DNA-binding signal of GR-Dim at 1-nt variant HRE DNA elements is most reduced at DNA elements that vary position 2 (or its reverse-complement position 9), position 5 (or its reverse-complement position 6), and position 1 (or its reverse-complement position 10), the DNA elements where wtGR DNA-binding is most increased (Figures 5 and S25). Likewise, GR-Dim's decreased DNA-binding signal at 2-nt variant HRE DNA elements in the genome occurs specifically at DNA elements where wtGR DNA-binding is most increased (i.e., at DNA elements with variants in palindromic position pair 2-9 followed by 2-5, 1-2 and 2-3 [and their reverse-complements 6-9, 9-10, 8-9]) (Figures 5 and S25). Similarly, GR-Dim's slight decrease in DNA-binding at 3-nt variant HRE DNA elements occurs specifically at DNA elements where wtGR DNA-binding is most increased: at DNA elements with variants in palindromic position pair 2-9 and positions 5, 1, or 3 (i.e., (2,5,9)-(2,6,9), (1,2,9)-(2,9,10), (2,3,9)-(2,8,9)) (Figures 5 and S25). Thus, these are the DNA elements that define functionally active binding sites in chromatin, consistent with the conclusions observed in our previous study using a different analysis

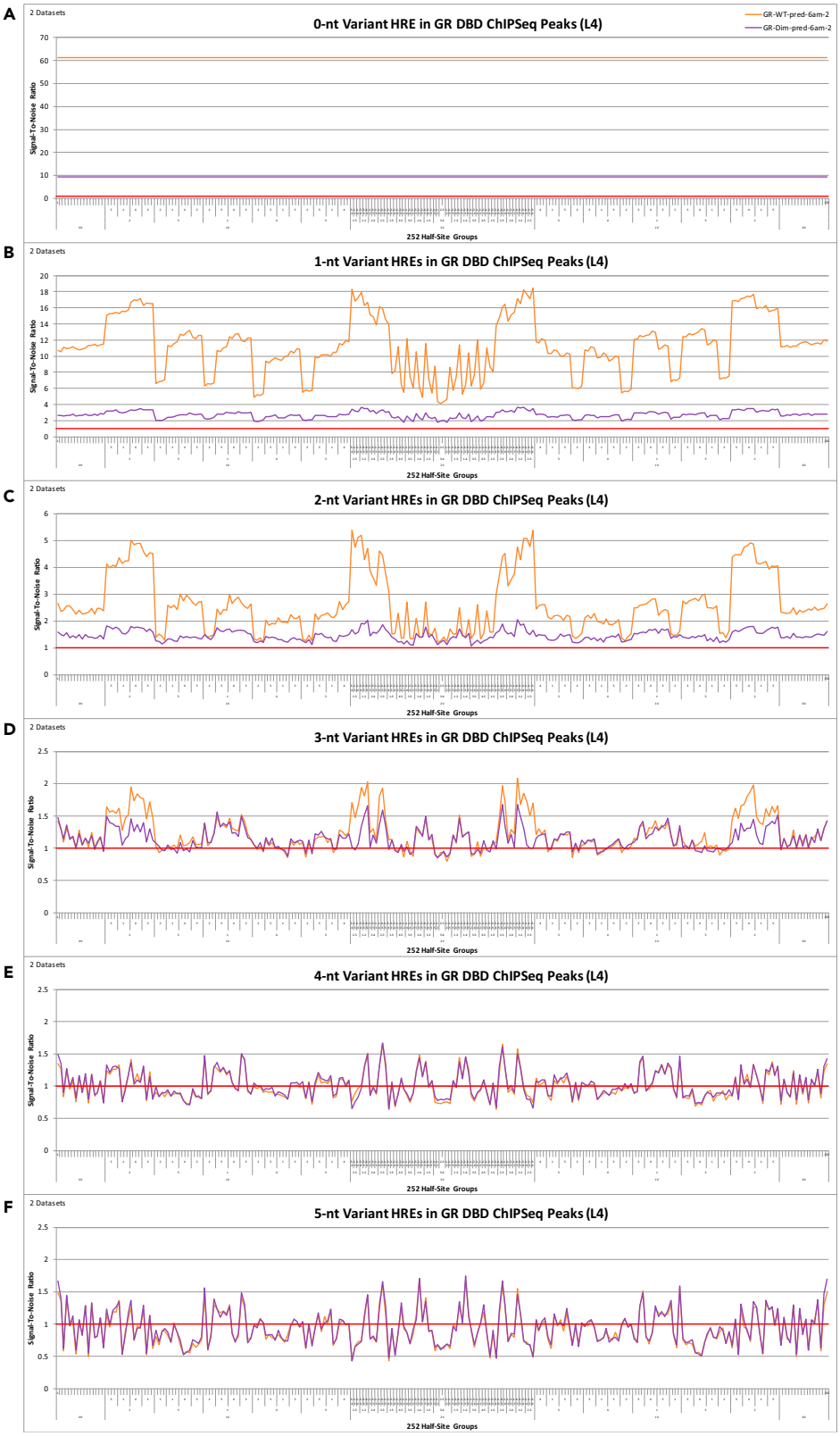


Figure 4. (S/N) Analysis of 0- to 5-nt Variant HRE DNA Elements in GR and GR-Dim ChIPSeq Peaks

Figure 4. *Continued*

(A–F) (S/N) analysis of 0- to 5-nt variant HRE DNA elements (displayed by the 252 half-site groups) in GR (GR-WT-pred-6am-2) (23,742 peaks, 142-nt peak length) ChIPSeq peaks and GR-Dim (GR-Dim-pred-6am-2) (34,966 peaks, 148-nt peak length) ChIPSeq peaks. Upper and lower one-tailed Poisson significance thresholds at $p < 0.001$: (A) 0.00, 3.11 [GR-WT] 0.00, 2.63 [GR-Dim]; (B) 0.58, 1.47 [GR-WT] 0.66, 1.38 [GR-Dim]; (C) 0.82, 1.19 [GR-WT] 0.86, 1.15 [GR-Dim]; (D) 0.90, 1.11 [GR-WT] 0.92, 1.09 [GR-Dim]; (E) 0.92, 1.09 [GR-WT] 0.93, 1.07 [GR-Dim]; and (F) 0.89, 1.11 [GR-WT] 0.91, 1.09 [GR-Dim]. X-axis order = reverse-complement vacancy position ID 2-9 > 5-6 > 1-10 > 4-7 > 3-8. See Table S6 for x-axis details.

methodology (Coons et al., 2017). Collectively, this suggests that the retention time of sNR DNA-binding at 3- to 5-nt variant DNA elements (non-NRFEs), observed in both the wtGR and GR-Dim, is not long enough to initiate and/or maintain transcription, and hence, is non-functional.

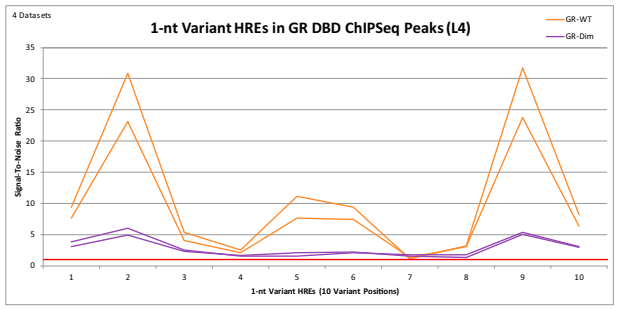
To summarize, the DNA-binding of sNRs in the genome is determined by (1) a set of rules defined by sNR DNA-binding at 1-nt variant DNA elements based on which palindromic position pair the variant occupies (i.e., ER DNA-binding is increased at DNA elements with variants in palindromic position pair 3-8 or 1-10 of the ERE, whereas KR DNA-binding is increased at DNA elements with variants in palindromic position pair 2-9 of the HRE) and (2) a second distinct set of rules (functionally subordinate to and independent of the DNA-binding rules defined by sNR DNA-binding at 1-nt variant DNA elements) defined by sNR DNA-binding at 2-nt variant DNA elements based on how the position of the variants relate to each other (i.e., sNR DNA-binding is increased at DNA elements with variants that do not crossover the 3-nt spacer). These DNA-binding rules, dictated by inversion symmetry, continue through 5-nt variant DNA elements in the genome, even though the sNR DNA-binding signal decays as the number of variants increases from 0- to 5-nt variant DNA elements. Furthermore, the sNR DBD mutant mouse model, GR-Dim, confirms that although DNA-binding at NRFEs (i.e., the 0-nt variant consensus palindromic DNA element, 1-nt variant DNA elements, and 2-nt variant DNA elements) are the only functional DNA-binding sites, DNA-binding at non-NRFEs are subject to the same DNA-binding rules, dictated by inversion symmetry, governing NRFE DNA-binding sites in the genome. Thus, the sNR DNA-binding rules extend (i.e., are applicable) to DNA elements that do not support transcriptional activity. That is, the DNA-binding rules are obeyed, but the retention time of the sNR at these non-functional DNA elements is not long enough to initiate and/or maintain transcription. Therefore, functionality is determined at the individual nucleotide level, and the residence time (or strength of binding) is dictated by the number and position of variants within the DNA element (i.e., inversion symmetry and DNA sequence constraints).

Furthermore, the fact that ChIPSeq and ChIPExo analysis of sNR binding in the genome follows inversion symmetry at 0- to 5-nt variant DNA elements for six different sNRs, in hundreds of experiments, and at the least stringent peak selection criteria (L4) demonstrates a previously underappreciated level of accuracy of these technologies. Much of the criticism of these technologies relates to the practice of comparing individual peaks between two experiments (which are highly variable from sample to sample) and the use of position weight matrices (PWMs) for DNA motif identification (which does not detect this level of detail in the sequence analysis) (Coons et al., 2017).

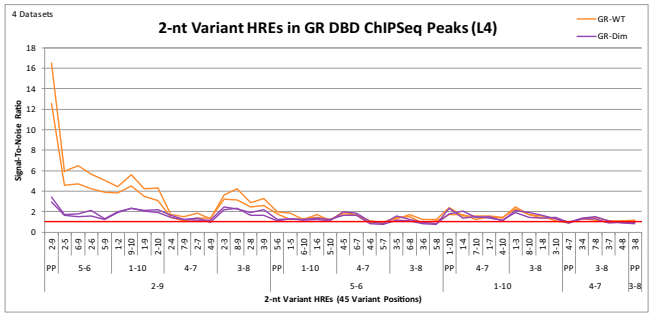
Inversion Symmetry Detection Methodology

Heuristic-based search engines and expert systems are limited to algorithms that are empirically derived from the currently available information database. For example, scoring DNA sequences against PWMs is a widely adopted method to identify TF *cis*-regulatory DNA elements in ChIPSeq or ChIPExo experiments. PWMs are used to scan a DNA sequence for the presence of DNA sequences that are significantly more similar to the PWMs than to the background (Aerts, 2012). DNA motif identification using PWMs has suggested that ER interacts with many "ERE-like" motifs, including PPAR, TR, and VDR motifs in the genome (Hewitt et al., 2012). These observations have led many to conclude that ER can indiscriminately bind to the sequence-specific *cis*-regulatory DNA element of other nuclear receptors, thus promoting the misconception that there is considerable promiscuity in DNA-binding between members of the NR superfamily. Furthermore, different *cis*-regulatory DNA elements have been delineated for the different members of the KR family, a conclusion not observed in our analysis of KR DNA-binding at 81,922 0- to 5-nt variant HRE DNA elements in the genome from 194 KR experiments (representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20)) (Geserick et al., 2005). To assess the resolution capabilities of PWM-based DNA motif identification, we grouped the ChIPSeq or ChIPExo peaks from each experiment by the number of variants each ERE or HRE DNA element within the peak contained (i.e., peaks that contain a 0-nt variant consensus palindromic DNA element, 1-nt variant DNA

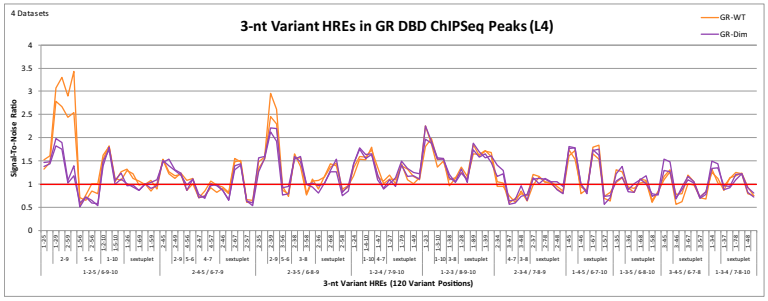
A



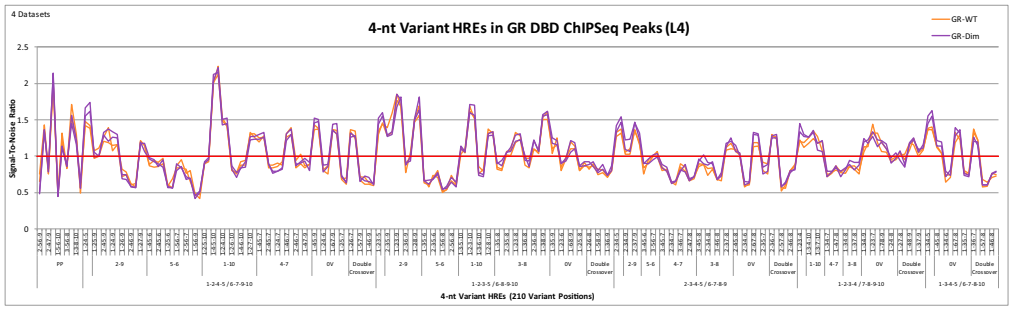
B



C



D



E

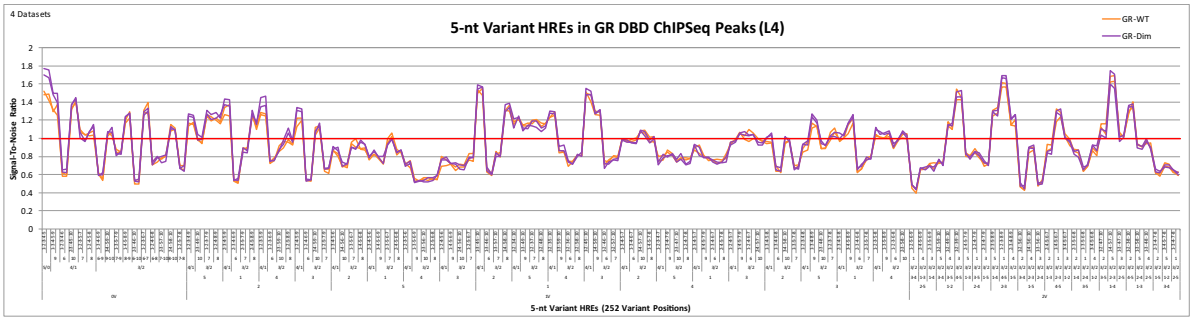


Figure 5. (S/N) Analysis of 1- to 5-nt Variant HRE DNA Elements in GR and GR-Dim ChIPSeq Peaks

Figure 5. *Continued*

(A–E) (S/N) analysis of 1- to 5-nt variant HRE DNA elements (displayed by variant position) in GR (GR-WT-pred-6am-1) (34,758 peaks, 151-nt peak length) (GR-WT-pred-6am-2) (23,742 peaks, 142-nt peak length) ChIP-Seq peaks and GR-Dim (GR-Dim-pred-6am-1) (22,130 peaks, 163-nt peak length) (GR-Dim-pred-6am-2) (34,966 peaks, 148-nt peak length) ChIPSeq peaks. Upper and lower one-tailed Poisson significance thresholds at $p < 0.001$: (A) 0.00, 2.60; 0.00, 3.11 [GR-WT] 0.00, 2.91; 0.00, 2.63 [GR-Dim]; (B) 0.58, 1.49; 0.48, 1.62 [GR-WT] 0.48, 1.61; 0.56, 1.49 [GR-Dim]; (C) 0.75, 1.27; 0.69, 1.35 [GR-WT] 0.70, 1.34; 0.74, 1.28 [GR-Dim]; (D) 0.85, 1.16; 0.81, 1.20 [GR-WT] 0.82, 1.19; 0.85, 1.16 [GR-Dim]; (E) 0.91, 1.09; 0.89, 1.11 [GR-WT] 0.89, 1.11; 0.91, 1.09 [GR-Dim].

element, 2-nt variant DNA element, 3-nt variant DNA element, 4-nt variant DNA element, or 5-nt variant DNA element) and analyzed these grouped peaks using PWMs.

Of 157 ER experiments, on average, PWMs identified the ERE motif (VAGGTCACNSTGACC) in 99%–99% (L4-L20) of the peaks that contain a 0-nt variant consensus palindromic ERE DNA element, 95%–97% (L4-L20) of the peaks that contain a 1-nt variant ERE DNA element, 50%–58% (L4-L20) of the peaks that contain a 2-nt variant ERE DNA element, 26%–38% (L4-L20) of the peaks that contain a 3-nt variant ERE DNA element, 27%–41% (L4-L20) of the peaks that contain a 4-nt variant ERE DNA element, and 26%–41% (L4-L20) of the peaks that contain a 5-nt variant ERE DNA element (Figure S26). Because peaks can contain multiple ERE DNA elements, we then assigned the peaks to the ERE DNA element with the least number of variants relative to the 0-nt consensus palindromic ERE DNA element, thus defining each ChIPSeq or ChIPExo peak by a single ERE DNA element (i.e., unique). This secondary analysis allowed us to determine whether the ERE motif identified (i.e., selected by PWMs in ~33% of the peaks that contain a 3- to 5-nt variant ERE DNA element) was due to a 0- to 2-nt variant ERE DNA element also occurring in those peaks. Now, PWMs identified the ERE motif in 99%–99% (L4-L20) of the peaks that contain a 0-nt variant consensus palindromic ERE DNA element, 95%–97% (L4-L20) of the peaks that contain a 1-nt variant ERE DNA element, 46%–52% (L4-L20) of the peaks that contain a 2-nt variant ERE DNA element, 3%–4% (L4-L20) of the peaks that contain a 3-nt variant ERE DNA element, 1%–1% (L4-L20) of the peaks that contain a 4-nt variant ERE DNA element, and 1%–1% (L4-L20) of the peaks that contain a 5-nt variant ERE DNA element (Figure S26). Thus, PWMs identified the ERE motif in the majority of peaks that contain a 0- or a 1-nt variant ERE DNA element, about half of the peaks that contain a 2-nt variant ERE DNA element, and did not identify the ERE motif in peaks that contain a 3-nt, 4-nt, or 5-nt variant ERE DNA element (Figure S26). Therefore, PWM analyses would lead one to conclude that the majority of ER DNA-binding events in the genome are being driven by mechanisms other than the ERE sequence, because peaks that contain a 0-nt variant consensus palindromic ERE DNA element or 1-nt variant ERE DNA element constitute, on average, 16%–27% (L4-L20) of all ER DNA-binding events in an experiment, a proportion observed in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) (Coons et al., 2017).

Alternatively, of 194 KR experiments, there was a wide distribution in the number of peaks identified by PWMs as containing the ARE motif (RGRACASNSTGTYCYB), GRE motif (NRGVACABNVTGTYCY), GRE 2 motif (VAGRACAKWCTGTYC), PRE motif (VAGRACAKNCTGTBC), or ARE palindrome half-site motif (CCAGGAACAG) (Table 5, Figures S27–S32). For example, of the peaks that contain a 1-nt variant HRE DNA element, on average, 64%–72% (L4-L20) of the peaks were identified by PWMs as containing the GRE motif, whereas 99%–99% (L4-L20) of those exact same peaks were identified by PWMs as containing the PRE motif (Table 5, Figure S27–S32). Or, of the peaks that contain a 2-nt variant HRE DNA element, 31%–39% (L4-L20) of the peaks were identified by PWMs as containing the GRE motif, whereas 86%–90% (L4-L20) of those exact same peaks were identified by PWMs as containing the PRE motif (Table 5, Figure S27–S32). Thus, it is apparent why the use of PWMs for DNA motif identification has led investigators to delineate different *cis*-regulatory DNA elements for the different KRs. However, analysis of KR DNA-binding at 81,922 0- to 5-nt variant HRE DNA elements in the genome quantitatively demonstrates that all members of the KR family follow the same DNA-binding rules and thus bind the same DNA elements (Tables S8 and S16–S20). These DNA-binding rules at 0- to 5-nt variant HRE DNA elements in the genome was observed in 194 KR ChIPSeq experiments, representing a wide variety of mouse tissues and human cell lines and across multiple peak selection criteria (L4-L20) (Tables S8 and S16–S20).

It is important to recognize the caveats associated with current PWM-based DNA motif identification algorithms, and careful consideration(s) should be given with regard to any conclusions drawn from their output. PWMs are adequate preliminary analysis tools to use when you have a TF with an unknown *cis*-regulatory DNA element, but once this information is identified it is not a useful tool for discriminating between similar DNA elements. Our analysis of the resolution capabilities of PWM-based DNA motif identification explains why the majority of TF DNA-binding events in the genome have previously "appeared"

The Number of DNA Motifs Detected by Position Weight Matrices					
All Peaks That Contain 0- to 5-nt Variant HRE DNA Elements					
	DNA Motifs Detected by Position Weight Matrices				
A	B	C	D	E	F
DNA Element	ARE	GRE	GRE 2	PRE	ARE Palindrome Half-Site
	RGRACASNSTGTYCYB	NRGVACABNVTGTYCY	VAGRACAKWCTGTYC	VAGRACAKNCTGTBC	CCAGGAACAG
0-nt Variant HRE	96%–96%	96%–97%	99%–100%	99%–100%	88%–89%
1-nt Variant HRE	76%–81%	64%–72%	84%–88%	99%–99%	77%–80%
2-nt Variant HRE	41%–50%	31%–39%	43%–53%	86%–90%	69%–73%
3-nt Variant HRE	19%–29%	14%–24%	19%–31%	59%–69%	60%–64%
4-nt Variant HRE	18%–29%	14%–24%	19%–31%	50%–64%	57%–63%
5-nt Variant HRE	18%–30%	14%–24%	19%–31%	50%–63%	56%–63%
Unique Peaks That Contain 0- to 5-nt Variant HRE DNA Elements					
	DNA Motifs Detected by Position Weight Matrices				
A	B	C	D	E	F
DNA Element	ARE	GRE	GRE 2	PRE	ARE Palindrome Half-Site
	RGRACASNSTGTYCYB	NRGVACABNVTGTYCY	VAGRACAKWCTGTYC	VAGRACAKNCTGTBC	CCAGGAACAG
0-nt Variant HRE	96%–96%	96%–97%	99%–100%	99%–100%	88%–89%
1-nt Variant HRE	76%–81%	64%–71%	84%–88%	99%–99%	77%–80%
2-nt Variant HRE	39%–47%	29%–37%	42%–51%	85%–89%	68%–72%
3-nt Variant HRE	9%–14%	6%–10%	8%–13%	49%–57%	56%–59%
4-nt Variant HRE	1%–2%	1%–2%	1%–2%	19%–25%	46%–50%
5-nt Variant HRE	1%–2%	1%–1%	1%–1%	6%–8%	33%–39%

Table 5. Inversion Symmetry Detection Methodology

Scoring DNA sequences against position weight matrices (PWMs) is a widely adopted method to identify TF *cis*-regulatory DNA elements in ChIPSeq or ChIPExo experiments. To assess the resolution capabilities of PWM-based DNA motif identification, we grouped the peaks from each KR experiment (194 experiments) by the number of variants each HRE DNA element within the peak contained (i.e., peaks that contain a 0-nt variant consensus palindromic DNA element, 1-nt variant DNA element, 2-nt variant DNA element, 3-nt variant DNA element, 4-nt variant DNA element, or 5-nt variant DNA element) and analyzed these grouped peaks using PWMs for detection of the (Column B) ARE motif (RGRACASNSTGTYCYB), (Column C) GRE motif (NRGVACABNVTGTYCY), (Column D) GRE 2 motif (VAGRACAKWCTGTYC), (Column E) PRE motif (VAGRACAKNCTGTBC), and (Column F) ARE palindrome half-site motif (CCAGGAACAG). The data are displayed as the number of peaks identified by PWMs as containing each DNA motif, out of the total number of peaks that contain a 0- to 5-nt variant HRE DNA element. For example, of the peaks that contain a 1-nt variant HRE DNA element, on average, 64%–72% (L4-L20) of the peaks were identified by PWMs as containing the GRE motif, whereas 99%–99% (L4-L20) of those exact same peaks were identified as containing the PRE motif. Because peaks can contain multiple HRE DNA elements, we then assigned the peaks to the HRE DNA element with the least number of variants relative to the 0-nt consensus palindromic HRE DNA element, thus defining each ChIPSeq or ChIPExo peak by a single HRE DNA element (i.e., unique). This secondary analysis allowed us to determine whether the DNA motif identified (i.e., selected by PWMs in peaks that contain a 3- to 5-nt variant HRE DNA element) was due to a 0- to 2-nt variant HRE DNA element also occurring in those peaks.

to be driven by mechanisms other than the DNA sequence: the methods being used for DNA motif identification are limited in their detection capabilities. Inversion symmetry needs to be a priority in the design of future DNA motif identification search engines.

Inversion Symmetry of sNR DNA-Binding at 15-nt DNA Elements in the Genome

Although the core 13-nt ERE and HRE DNA elements for sNRs have been well defined, we are also interested in the contribution of the two nucleotides surrounding the 13-nt DNA elements in the genome,

that is, the impact of these two additional nts on sNR binding in the genome (i.e., expanding the 13-nt ERE or HRE DNA element with 10 primary positions to a 15-nt ERE or HRE DNA element with 12 primary positions). Because the analysis of sNR DNA-binding at 13-nt ERE and HRE DNA elements was done by overlapping the location coordinates of each 0- to 5-nt variant ERE or HRE DNA element in the genome and the location coordinates of the ChIPSeq or ChIPExo peaks from each experiment, the counts of all sNR DNA-binding events at 15-nt ERE or HRE DNA elements in the genome are included in the 13-nt ERE and HRE DNA element analysis (i.e., every 13-nt ERE and HRE DNA element in the genome is part of a 15-nt ERE and HRE DNA element). For example, the 0-nt variant consensus palindromic 13-nt ERE DNA element (5′-GGTCAnnnTGACC-3′) occurred 1,202 times in a particular ER experiment (WT-E2-1hr.L4) (Figure S2). Each of those (1,202) 0-nt variant consensus palindromic 13-nt ERE DNA elements is part of a 15-nt ERE DNA element. Thus, analysis of ER DNA-binding at all possible 15-nt ERE DNA elements splits those (1,202) 0-nt variant consensus palindromic 13-nt ERE DNA elements into 16 categories (i.e., 4 nt possibilities in position 1 and 4 nt possibilities in position 12): 109 (1T), 155 (1G), 19 (1C), 13 (12G), 151 (12C), 110 (12A), 11 (1–12 TG), 67 (1–12 TC), 17 (1–12 GG), 305 (1–12 AT), 29 (1–12 TA), 107 (1–12 GC), 5 (1–12 CG), 14 (1–12 CC), 88 (1–12 GA), 2 (1–12 CA), totaling 1,202 (Figure S33). Thus, assessing the role of the flanking nucleotides in a 15-nt ERE or HRE DNA element on sNR binding in the genome requires segregating the counts from the 13-nt ERE and HRE DNA element analysis and displaying them separately.

First, we analyzed the DNA-binding preference of sNRs at the 16 possible 15-nt ERE or HRE DNA elements in the genome (i.e., 4 nt possibilities in position 1 and 4 nt possibilities in position 12) (Figure S33). sNR binding in the genome was most predominant when an adenine (A) was in position 1 and a thymine (T) was in position 12 of the 15-nt ERE DNA element (5′-AGGTCAnnnTGACCT-3′) for ERs and of the 15-nt HRE DNA element (5′-AGAACAnnnTGTTCT-3′) for KRs (Figure S33). This was observed in 154 ER experiments and 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) (Figure S33). Thus, these 15-nt ERE and HRE DNA elements (i.e., adenine [A] in position 1 and thymine [T] in position 12) represent the 0-nt variant consensus palindromic 15-nt ERE and HRE DNA elements for ERs and KRs (Figure S33).

Next, we repeated the same sNR DNA-binding analysis (as was done for sNR DNA-binding at 0- to 5-nt variant 13-nt ERE and HRE DNA elements in the genome) for the newly identified 0-nt variant consensus palindromic 15-nt ERE and HRE DNA elements and extending through 6-nt variant DNA elements (i.e., a 6-nt variant = a half-site for a 15-nt DNA element with 12 primary positions) in the genome (Figure S34). Analysis of sNR DNA-binding at 13-nt ERE and HRE DNA elements included the 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements (10 variant positions), 405 2-nt variant DNA elements (45 variant positions), 3,240 3-nt variant DNA elements (120 variant positions), 17,010 4-nt variant DNA elements (210 variant positions), and 61,236 5-nt variant DNA elements (252 variant positions), for a total of 81,922 DNA elements (Table 1; Figure S34). Analysis of sNR DNA-binding at 15-nt ERE and HRE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 36 1-nt variant DNA elements (12 variant positions), 594 2-nt variant DNA elements (66 variant positions), 5,940 3-nt variant DNA elements (220 variant positions), 40,095 4-nt variant DNA elements (495 variant positions),192,456 5-nt variant DNA elements (792 variant positions), and 673,596 6-nt variant DNA elements (924 variant positions), for a total of 912,718 DNA elements (Figure S34). All data analyses completed for sNR DNA-binding at 0- to 5-nt variant 13-nt ERE and HRE DNA elements in the genome were repeated for sNR DNA-binding at 0- to 6-nt variant 15-nt ERE and HRE DNA elements in the genome (Figures S35–S61; Tables S21–S40). Note: a comparative figure summary between the 13-nt ERE and HRE DNA element analysis and the 15-nt ERE and HRE DNA element analysis can be found in the Supplemental Information.

As demonstrated for the 13-nt NRFE ERE and HRE DNA elements, sNR DNA-binding at 15-nt NRFE ERE and HRE DNA elements in the genome follows inversion symmetry (i.e., the number of sNR DNA-binding events at a particular DNA element in the genome is equivalent to the number of sNR DNA-binding events at its reverse-complement DNA element in the genome) (Figures S35 and S36). In addition, sNRs exhibit preferential DNA-binding at certain NRFE ERE and HRE DNA elements, this preference is determined by internal inversion symmetry within the DNA element (i.e., sNR binding in the genome is determined by the position of variants within the DNA element and its reverse-complement position) (Figures S35 and S36). Also, as with the 13-nt NRFE ERE and HRE DNA elements, sNR DNA-binding is most predominant at 15-nt NRFE ERE and HRE DNA elements versus at non-NRFE DNA elements in the genome (Figures S37 and S38).

Next, to sequentially track sNR DNA-binding at 0-nt to 6-nt variant 15-nt ERE and HRE DNA elements in the genome, we define a 6-nt variant DNA element by its six fixed positions, resulting in 924 half-site groups (i.e., categorizing the 673,596 6-nt variant ERE and HRE DNA elements into 924 half-site groups, defined by the six positions that are fixed/not varied) (Figure S34). The 924 half-site groups are symmetrically split into a set of 452 groups and their 452 reverse-complements, plus 20 groups that are "innate palindromes" (i.e., they are their own reverse-complement) (Figure S39). These 452 groups further split into four distinct subgroups (zero vacancies, one vacancy, two vacancies), plus the 20 innate palindromes that have three vacancies (Figure S39). The remaining 0-nt to 5-nt variant 15-nt DNA elements (1 0-nt variant consensus palindromic DNA element, 36 1-nt variant DNA elements, 594 2-nt variant DNA elements, 5,940 3-nt variant DNA elements, 40,095 4-nt variant DNA elements, and 192,456 5-nt variant DNA elements, for a total of 239,122 DNA elements) can be categorized into these same 924 half-site groups (i.e., six positions are fixed, allowing for *up to* six positions to be varied) (Figure S39). Each of the 924 half-site groups contain 1 (0-nt variant consensus palindromic DNA element), 18 (1-nt variant DNA elements), 135 (2-nt variant DNA elements), 540 (3-nt variant DNA elements), 1,215 (4-nt variant DNA elements), 1,458 (5-nt variant DNA elements), and 729 (6-nt variant DNA elements), for a total of 4,096 0- to 6-nt variant DNA elements per half-site group (Tables S21–S24). This allows all 912,718 0- to 6-nt variant DNA elements to be categorized into the 924 half-site groups, thus providing the ability to sequentially track sNR DNA-binding at 0- to 6-nt variant 15-nt ERE or HRE DNA elements in the genome. The x-axis is ordered by the symmetrically split 452 groups (left-to-right: zero vacancies, one vacancy, two vacancies) followed by their 452 reverse-complements, with the 20 half-site groups that are innate palindromes in the middle (Figures S40–S43; Tables S25 and S26). The x-axis is labeled by the reverse-complement vacancy position ID (primary label) and reverse-complement double occupant position ID (secondary label) (Figures S40–S43; Tables S25 and S26). The order of the reverse-complement vacancy position IDs for the ERE is 4-9 > 1-12 > 2-11 > 6-7 > 5-8 > 3-10 (Table S25), whereas that for the HRE is 3-10 > 1-12 > 6-7 > 2-11 > 5-8 > 4-9 (Table S26).

As expected, (S/N) analysis of sNR DNA-binding at 0- to 6-nt variant 15-nt ERE or HRE DNA elements in the genome (displayed by the 924 half-site groups) reveals a highly structured symmetrical DNA-binding profile that decays from 0- to 6-nt variant DNA elements (Figures S40–S43). This highly structured symmetrical sNR DNA-binding profile demonstrates (1) there is no sNR DNA-binding bias toward either the 5′ side (AGGTCA or AGAACA) or the 3′ side (TGACCT or TGTTCT) of the DNA element at any 1- to 6-nt variant 15-nt ERE or HRE DNA elements in the genome, (2) sNR binding in the genome follows inversion symmetry (i.e., the number of sNR DNA-binding events at a 1- to 6-nt variant DNA element in the genome is equivalent to the number of sNR DNA-binding events at its reverse-complement DNA element in the genome), (3) the specific DNA elements where sNRs bind in the genome is determined by internal inversion symmetry within the DNA element and this continues through 6-nt variant DNA elements (i.e., sNR binding in the genome is determined by the position of variants within the DNA element and its reverse-complement position), (4) sNR DNA-binding at 0- to 6-nt variant DNA elements in the genome requires all the information contained within the entire DNA element (i.e., on both sides of the 3-nt spacer), challenging previous suggestions that sNRs bind "palindrome half-sites" in the genome as individual monomers (Figures S40–S43). Moreover, this highly structured symmetrical sNR DNA-binding profile at 0- to 6-nt variant DNA elements in the genome was observed in 157 ER experiments and 194 KR experiments, representing a wide variety of mouse tissues and human cell lines and across multiple peak selection criteria (L4-L20) (Tables S27 and S28).

Because it is difficult to translate the 924 half-site groups to specific DNA elements, the absolute counts of each of the 912,718 0- to 6-nt variant ERE and HRE DNA element in an experiment can also be displayed by the position of the variant within the DNA element, rather than by the 924 half-site groups (Figures S44–S55). Analysis of sNR DNA-binding at 0- to 6-nt variant 15-nt ERE and HRE DNA elements in the genome shows a predominance of DNA-binding at DNA elements that have variants in these flanking nucleotides (positions 1 and 12) (Figures S44–S55). Thus, having fixed nucleotides in position 1 or position 12 is the least important of the 12 primary positions in a 15-nt ERE or HRE DNA element (Figures S44–S55). This sNR DNA-binding profile was observed in 157 ER experiments and 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) (Tables S29–S40).

As with the 13-nt DNA element analysis, GR-Dim has a significantly reduced DNA-binding signal at 15-nt NRFE HRE DNA elements, whereas the DNA-binding profiles of GR-Dim and wtGR are highly structured and converge onto the same ((S/N) value = 1) noise profile (demonstrating inherent structure in DNA-binding through 6-nt variant 15-nt HRE DNA elements in the genome) (Figures S56–S61). Likewise, the nearly

identical structure of the DNA-binding curve of GR-Dim compared with wtGR at 0- to 6-nt variant HRE DNA elements in the genome (albeit with a significantly reduced amplitude at NRFE HRE DNA elements) demonstrates that the effect of the DBD mutation is a loss in residence time at the DNA element, rather than the inability to follow the same DNA-binding rules as wtGR (Figures S56–S61). Thus, the sNR DNA-binding rules, dictated by inversion symmetry, extend (i.e., are applicable) to DNA elements that do not support transcriptional activity. That is, the DNA-binding rules are obeyed, but the retention time of the sNR at these non-functional DNA elements is not long enough to initiate and/or maintain transcription. Therefore, functionality is determined at the individual nucleotide level, and the residence time (or strength of binding) is dictated by the number and position of variants within the DNA element (i.e., inversion symmetry and DNA sequence constraints).

Of importance, comparison of the sNR DNA-binding profile at the 81,922 0- to 5-nt variant 13-nt ERE or HRE DNA elements in the genome versus at the 912,718 0-nt to 6-nt variant 15-nt ERE or HRE DNA elements in the genome is almost equivalent, confirming that the DNA-binding rules established by the 13-nt DNA element analysis also apply to the 15-nt DNA element analysis (Figures S62–S71). That is, the information content in positions 1 and 12 of the 15-nt ERE and HRE DNA elements is less important than in the other 10 primary positions of the 15-nt ERE and HRE DNA elements, and thus the analysis completed for sNR DNA-binding at 13-nt ERE and HRE DNA elements was adequate to describe the DNA-binding mechanisms underlying sNR binding in the genome (Figures S62–S71). At best, fixing the flanking nucleotides of the 15-nt HRE or ERE DNA elements (i.e., adenine [A] in position 1 and thymine [T] in position 12) causes a few very subtle effects on the relative amplitudes of sNR DNA-binding at certain DNA elements (with variants in specific palindromic position pairs) in the genome (Figures S72–S77; Tables S41–S46).

To summarize, sNR binding in the genome is determined by (1) a set of rules defined by sNR DNA-binding at 1-nt variant ERE and HRE DNA elements based on which palindromic position pair the variant occupies (i.e., ER DNA-binding is increased at DNA elements with variants in palindromic position pair 3-8 or 1-10 [4-9 or 2-11 for the 15-nt DNA elements] of the ERE, whereas KR DNA-binding is increased at DNA elements with variants in palindromic position pair 2-9 [3-10 for the 15-nt DNA elements] of the HRE), and (2) a second distinct set of rules (functionally subordinate to and independent of the DNA-binding rules defined by sNR DNA-binding at 1-nt variant DNA elements) defined by sNR DNA-binding at 2-nt variant DNA elements based on how the position of the variants relate to each other (i.e., sNR DNA-binding is increased at DNA elements with variants that do not crossover the 3-nt spacer). These DNA-binding rules, dictated by inversion symmetry, continue through 5-nt variant DNA elements (6-nt variant DNA elements for the 15-nt DNA elements) in the genome, even though the sNR DNA-binding signal decays as the number the variants increases from 0- to 5-nt variant DNA elements (0- to 6-nt variant DNA elements for the 15-nt DNA elements). These sNR DNA-binding rules were observed in 157 ER experiments and 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20). Of particular importance, there are numerous time course trials in these 351 experiments, confirming that these DNA-binding rules determine sNR binding in the genome for both initial sNR DNA-binding events (i.e., early time points) and subsequent sNR DNA-binding events (i.e., late time points) (Figures S78 and S79). The time course trials further demonstrate that hormone-mediated sNR DNA-binding at NRFE sites in the genome is retained the longest over time, in agreement with previous studies and hormone-mediated eRNA transcription (Figures S78 and S79) (Coons et al., 2017; Coons, 2017).

To further explain the importance of the progression and decay of sNR DNA-binding from 0- to 5-nt variant DNA elements (0-nt to 6-nt variant DNA elements for the 15-nt DNA elements) in the genome, we first looked at (S/N) analysis of sNR DNA-binding at the 0-nt variant consensus palindromic DNA element. (S/N) analysis of sNR DNA-binding at the 0-nt variant consensus palindromic DNA element results in the largest DNA-binding signal, but because this is a single DNA element there is no information to be extracted because all the states are equal. This is the state of equilibrium or maximal entropy. Because this is the state with the largest DNA-binding signal, the DNA-binding signal gives us a relative measure of the integrity of the data with respect to random noise. When a variant is added to the 0-nt variant consensus palindromic DNA element, this introduces an additional variable into the possible DNA element representations. Thus, this introduction of additional variables breaks the symmetry of the equilibrium state, thereby reducing the entropy of the system and the DNA-binding signal as the symmetry of the system is sequentially degraded. The reduction of entropy with each broken symmetry and the introduction of

additional variables to the system is interpreted as an equivalent gain in information. Thus, the primary information generating process is the introduction of additional variants in certain palindromic position pairs of the DNA element (i.e., the first set of DNA-binding rules), with the relative positions of these variants playing a secondary, but not insignificant role (i.e., the second set of DNA-binding rules). For example, (S/N) analysis of sNR DNA-binding at 1-nt variant DNA elements in the genome introduces a variable (i.e., the position of the varied nucleotide) that is strictly associated with the position of the variant and its reverse-complement position (i.e., the palindromic position pairs). (S/N) analysis of sNR DNA-binding at 2-nt variant DNA elements in the genome introduces another variable (i.e., the relative position of the two variants) that is associated with how the position of the variants relate to each other. Thus, 2-nt variant DNA elements introduce another variable and additional restrictions (i.e., subrules) that generate an additional set of substructures, which are qualifiers to the superstructure rules defined by sNR DNA-binding at the 0-nt variant consensus palindromic DNA element and at the 1-nt variant DNA elements in the genome. That is, the primary DNA-binding rules are still fully intact and operational at each subsequent level (through 5-/6-nt variant DNA elements); the introduction of another variant at each level introduces additional variables. Thus, every subsequent level provides the opportunity to deduce more subrules that are subordinate to the superstructure rules (defined by sNR DNA-binding at the 0-nt variant consensus palindromic DNA element and at the 1-nt variant DNA elements in the genome). That is, the primary sNR DNA-binding rules, dictated by inversion symmetry, continue through 5-nt variant DNA elements (6-nt variant DNA elements for the 15-nt DNA elements) in the genome. Additional DNA-binding rules, generated by the introduction of additional variables, are subordinate to the primary sNR DNA-binding rules.

Inversion Symmetry of the Single-Stranded Genome

The observation that sNR binding in the genome follows inversion symmetry was unexpected, prompting us to consider why this was the case. Therefore, we evaluated the number of times every DNA element (1- to 20-nt) occurs in the single-stranded mouse and human genome (Tables 6, S47, and S48). Thus, this includes all DNA elements from the four 1-nt DNA elements to the 1 trillion (1,099,511,627,776) 20-nt DNA elements (Tables 6, S47, and S48). Comparing the population count of every DNA element to the population count of its reverse-complement DNA element revealed a near-perfect inversion symmetry (i.e., 0.999 correlation coefficient) for all 1- to 20-nt DNA elements in the single-stranded mouse and human genome (Tables 6, S47, and S48). This result is Chargaff's second parity rule. This inversion symmetry structure is also maintained at the level of each individual chromosome, except for chromosome M (i.e., the mitochondrial DNA) (Tables S49 and S50).

In addition to demonstrating the inherent inversion symmetry structure for all DNA elements in the single-stranded genome, we further demonstrate the absence of analog symmetries between reverse and complement pairs of DNA elements in the single-stranded genome (Tables S47–S50). However, the inversion symmetry structure should be viewed as a parity conservation law. The reverse pairs and the complement pairs are both binary or parity operations. That is, the reverse-of-the-reverse and the complement-of-the-complement leave the DNA element unchanged. Thus, an odd number of operations change the parity, and an even number of operations conserve the parity. For every DNA element, four sets of DNA elements can be generated (i.e., the original DNA element, the reverse-complement DNA element, the reverse DNA element, the complement DNA element), and there are six pairwise comparisons that can be assessed. In both cases where two parity operations are applied (i.e., the original DNA element versus the reverse-complement DNA element and the reverse DNA element versus the complement DNA element) and parity is conserved, the population counts have a correlation coefficient of unity (1) (Tables S47–S50), whereas when the comparisons differ by only a single (1) parity operation (i.e., the original DNA element versus the reverse DNA element, the original DNA element versus the complement DNA element, the reverse DNA element versus the reverse-complement DNA element, and the complement DNA element versus the reverse-complement DNA element) the results are significantly uncorrelated (Tables S47–S50). Yet, all four comparisons are identical numerically (Tables S47–S50). This suggests that the genome is governed by a very general parity conservation law.

This inversion symmetry structure for all DNA elements in the single-stranded genome holds for many species, both eukaryotes and prokaryotes (Shporer et al., 2016; Albrecht-Buehler, 2006). These findings suggest that the structural mechanisms (by which inversion symmetry ascribes TF DNA-binding and functionality) are universally applicable. That is, this analysis accounts for every possible *cis*-regulatory DNA element (for any TF) in the genome with 2- to 40-nt primary positions (i.e., a 20-nt element, followed by any arbitrary

Mouse Genome (mm10)								
DNA Element	Total DNA Elements in Genome ($B = 4^A$)					Population Comparison		
A	B	C	D	E	F	G	H	I
Length (nt)	Expected	Actual	Missing	Actual	Missing	Correlation Coefficient	Slope	Y-Intercept
1	4	4	–	100%	0%	0.999997	0.999985	9,944.94
2	16	16	–	100%	0%	0.999991	0.999991	1,482.34
3	64	64	–	100%	0%	0.999992	0.999992	327.33
4	256	256	–	100%	0%	0.999992	0.999992	85.53
5	1,024	1,024	–	100%	0%	0.999990	0.999990	24.84
6	4,096	4,096	–	100%	0%	0.999986	0.999986	8.80
7	16,384	16,384	–	100%	0%	0.999977	0.999977	3.67
8	65,536	65,536	–	100%	0%	0.999958	0.999958	1.70
9	262,144	262,144	–	100%	0%	0.999927	0.999927	0.74
10	1,048,576	1,048,575	1	100%	0%	0.999891	0.999891	0.28
11	4,194,304	4,192,822	1,482	100%	0%	0.999858	0.999859	0.09
12	16,777,216	16,583,844	193,372	99%	1%	0.999831	0.999832	0.01
13	67,108,864	61,933,424	5,175,440	92%	8%	0.999805	0.999814	(0.07)
14	268,435,456	200,322,339	68,113,117	75%	25%	0.999782	0.999806	(0.20)
15	1,073,741,824	537,352,842	536,388,982	50%	50%	0.999767	0.999803	(0.34)
16	4,294,967,296	1,093,135,359	3,201,831,937	25%	75%	0.999776	0.999824	(0.59)
17	17,179,869,184	1,574,548,194	15,605,320,990	9%	91%	0.999800	0.999848	(0.81)
18	68,719,476,736	1,837,040,301	66,882,436,435	3%	97%	0.999810	0.999856	(0.92)
19	274,877,906,944	1,959,582,013	272,918,324,931	1%	99%	0.999808	0.999853	(0.96)
20	1,099,511,627,776	2,021,349,628	1,097,490,278,148	0%	100%	0.999799	0.999845	(0.97)
Human Genome (hg19)								
DNA Element	Total DNA Elements in Genome ($B = 4^A$)					Population Comparison		
A	B	C	D	E	F	G	H	I
Length (nt)	Expected	Actual	Missing	Actual	Missing	Correlation Coefficient	Slope	Y-Intercept
1	4	4	–	100%	0%	0.999994	0.999976	17,191.60
2	16	16	–	100%	0%	0.999984	0.999984	2,908.31
3	64	64	–	100%	0%	0.999986	0.999986	643.10
4	256	256	–	100%	0%	0.999986	0.999986	161.07
5	1,024	1,024	–	100%	0%	0.999985	0.999985	42.31
6	4,096	4,096	–	100%	0%	0.999983	0.999983	12.00
7	16,384	16,384	–	100%	0%	0.999978	0.999978	3.88

Table 6. Inversion Symmetry of the Single-Stranded Mouse and Human Genome

(Continued on next page)

Human Genome (hg19)								
DNA Element	Total DNA Elements in Genome ($B = 4^A$)					Population Comparison		
A	B	C	D	E	F	G	H	I
Length (nt)	Expected	Actual	Missing	Actual	Missing	Correlation Coefficient	Slope	Y-Intercept
8	65,536	65,536	–	100%	0%	0.999967	0.999967	1.46
9	262,144	262,144	–	100%	0%	0.999947	0.999947	0.58
10	1,048,576	1,048,576	–	100%	0%	0.999919	0.999919	0.22
11	4,194,304	4,193,313	991	100%	0%	0.999887	0.999887	0.08
12	16,777,216	16,609,017	168,199	99%	1%	0.999851	0.999852	0.01
13	67,108,864	62,296,994	4,811,870	93%	7%	0.999809	0.999820	(0.07)
14	268,435,456	202,655,673	65,779,783	75%	25%	0.999761	0.999797	(0.19)
15	1,073,741,824	546,364,018	527,377,806	51%	49%	0.999715	0.999780	(0.33)
16	4,294,967,296	1,130,819,799	3,164,147,497	26%	74%	0.999708	0.999807	(0.56)
17	17,179,869,184	1,682,092,708	15,497,776,476	10%	90%	0.999739	0.999854	(0.77)
18	68,719,476,736	2,015,311,971	66,704,164,765	3%	97%	0.999765	0.999885	(0.88)
19	274,877,906,944	2,181,567,011	272,696,339,933	1%	99%	0.999768	0.999898	(0.93)
20	1,099,511,627,776	2,265,704,990	1,097,245,922,786	0%	100%	0.999752	0.999899	(0.95)

Table 6. Continued

Evaluation of the number of times every DNA element (1- to 20-nt) occurs in the single-stranded mouse and human genome. Column B: This includes all DNA elements from the four 1-nt DNA elements to the 1 trillion (1,099,511,627,776) 20-nt DNA elements. Column C and D: All possible 1- to 9-nt DNA elements exist in the mouse and human genome, and only one 10-nt DNA element is missing from the mouse genome. Column E: The absence of multiple DNA elements in the mouse and human genome occurs at 11-nt DNA elements and continues through 20-nt DNA elements, where less than 0.2% of all possible 20-nt DNA elements occur in the genome. Columns G–I: The equivalence of the population count of every DNA element to the population count of its reverse-complement DNA element in the single-stranded mouse and human genome was determined using the correlation coefficient, slope and y-intercept. See Tables S47–S50 for the population count of every 1- to 20-nt DNA elements in the single-stranded mouse and human genome, and at the level of each individual chromosome.

spacer, followed by its 20-nt reverse-complement element). For example, the 13-nt ERE and HRE DNA elements are a 5-nt DNA element (5-nt element, 3-nt spacer, 5-nt reverse-complement element), whereas the 15-nt ERE and HRE DNA elements are a 6-nt DNA element (6-nt element, 3-nt spacer, 6-nt reverse-complement element) (Tables S47–S50). Thus, the analysis of TF binding in the genome has expanded our understanding as to why the genome is organized in an inversion symmetry structure (i.e., Chargaff's second parity rule). Inversion symmetry is the DNA code that TFs use to interact with the genome, and dictates (in conjunction with known DNA sequence constraints) which of those interactions are functionally active.

Of particular importance is the observation that all possible 1- to 9-nt DNA elements exist in the mouse and human genome, and only one 10-nt DNA element is missing from the mouse genome (Tables 6, S47, and S48). The absence of multiple DNA elements in the mouse and human genome occurs at 11-nt DNA elements and continues through 20-nt DNA elements (less than 0.2% of all possible 20-nt DNA elements occur in the genome) (Tables 6, S47, and S48). That is, the number of DNA elements that occur in the genome, out of the total number of possible DNA elements, decreases from ~100% to ~0% from 11- to 20-nt DNA elements (Tables 6, S47, and S48). This is due to the fact that the sample size is limited to 3 billion nts, the size of the single-stranded genome. Hence, the inversion symmetry would be expected to be universal (beyond 20-nt DNA elements) if the DNA sample size were significantly larger than 3 billion nts.

To further illustrate that the structural mechanisms (by which inversion symmetry ascribes TF DNA-binding and functionality) are universally applicable, we repeated the same DNA-binding analysis as was done with the six sNRs (at ERE and HRE DNA elements) on the unrelated TF p53 (at p53RE DNA elements) in the

genome. The tumor suppressor p53 is a TF that controls cellular stress responses (Tonelli et al., 2017). Once activated, p53 binds DNA and regulates gene expression programs that contribute to apoptosis, senescence, or cell-cycle arrest, preventing the dissemination of damaged cells (Tonelli et al., 2017). These processes are involved in tumor suppression, setting the selective pressure for p53 inactivation in tumors (Tonelli et al., 2017). We evaluated the DNA-binding of p53 at 0- to 5-nt variant DNA elements of its 10-nt consensus palindromic DNA element (p53RE) (5′-TGCCCGGGCA-3′) in the genome (Figure S80). Analysis of the 0- to 5-nt p53RE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements (10 variant positions), 405 2-nt variant DNA elements (45 variant positions), 3,240 3-nt variant DNA elements (120 variant positions), 17,010 4-nt variant DNA elements (210 variant positions), and 61,236 5-nt variant DNA elements (252 variant positions), for a total of 81,922 DNA elements (Figure S80). All 81,922 0- to 5-nt variant p53RE DNA elements can be categorized into the 252 half-site groups, providing the ability to sequentially track p53 DNA-binding at 0- to 5-nt variant p53RE DNA elements in the genome (Tables S51–S52). As expected, (S/N) analysis of p53 DNA-binding at 0- to 5-nt variant p53RE DNA elements in the genome reveals a highly structured symmetrical DNA-binding profile that decays from 0- to 5-nt variant p53RE DNA elements (Figure S81). This highly structured symmetrical p53 DNA-binding profile (i.e., the left side [126 half-site groups] versus the right side [126 reverse-complements]) demonstrates (1) that p53 binding in the genome follows inversion symmetry (i.e., the number of p53 DNA-binding events at a particular 1- to 5-nt variant DNA element in the genome is equivalent to the number of p53 DNA-binding events at its reverse-complement DNA element in the genome), and (2) the specific DNA elements where p53 binds in the genome is determined by internal inversion symmetry within the DNA element (i.e., p53 binding in the genome is determined by the position of variants within the DNA element and its reverse-complement position) (Figure S81). These p53 DNA-binding rules, dictated by inversion symmetry, continue through 5-nt variant p53RE DNA elements in the genome, even though the p53 DNA-binding signal decays as the number of variants increase from 0- to 5-nt variant p53RE DNA elements (Figure S81). Increased p53 DNA-binding at specific 1- to 5-nt variant p53RE DNA elements in the genome is determined by the reverse-complement vacancy in the palindromic position pair by the following hierarchy: 5-6 > 1-10 > 3-8 > 4-7 > 2-9 (Table S53). Moreover, this highly structured symmetrical p53 DNA-binding profile at 0- to 5-nt variant p53RE DNA elements in the genome was observed in 22 p53 experiments and across multiple peak selection criteria (L4-L20) (Table S54).

DISCUSSION

The information about when and where a gene is going to be expressed resides in DNA elements called *cis*-regulatory elements (Yáñez-Cuna et al., 2013). TFs direct gene expression by binding to *cis*-regulatory elements in the genome. Knowing where TFs bind in the genome and if that binding event will affect transcription is key to understanding gene regulation and biological responses. Despite there being a great deal of interest in understanding how TFs control gene expression, we still do not know how TFs select the appropriate regulatory targets from a large number of near-consensus DNA elements in the genome to elicit specific transcriptional and cellular responses. Knowing precisely how the binding of TFs to the genome turns genes on and off is vital to understand how different cell types function in health and disease.

It is commonly thought that variations in *cis*-regulatory DNA elements are responsible for changes in gene expression and phenotype. A recent study explored whether single nucleotide variants (SNV) in ER DNA-binding sites contribute to ER action through changes in the ER cistrome, thereby affecting disease progression (Bahreini et al., 2016). In this study, analysis of ER ChIPSeq data from MCF7 cells identified an intronic SNV in the IGF1R gene (rs62022087) that increased ER DNA-binding to this DNA element (compared with the wt-allele) (Bahreini et al., 2016). The wt-allele is a 2-nt variant ERE DNA element with variants in positions 7 and 8 (Bahreini et al., 2016). The SNV (rs62022087) is a 1-nt variant ERE DNA element, with a variant in position 8 (the nucleotide in position 7 has been mutated to the original nucleotide in position 7 of the 0-nt variant consensus palindromic ERE DNA element) (Bahreini et al., 2016). The conclusions in our study explain why the identified SNV results in increased ER DNA-binding at that element. The 2-nt variant ERE (wt-allele) that has variants in positions 7 and 8 exhibits a weak ER DNA-binding signal in 157 ER experiments (i.e., an average (S/N) value less than 4) (Figure S13; Table S12). This is because sNR DNA-binding is reduced when variants are in palindromic position pair 4-7 of the ERE or HRE DNA elements; its variation results in a deleterious effect (Tables S7 and S8). Thus, the reduced ER DNA-binding at the wt-allele (as shown in their study) is due to a variant occurring in position 7 of the ERE (Bahreini et al., 2016). In contrast, ER DNA-binding is increased at DNA elements that have a variant in palindromic position pair 3-8 in 157 ER experiments (i.e., an average (S/N) value above 50) (Figure S12; Table S11). Thus, the increased ER DNA-binding at the intronic

SNV (rs62022087) (as shown in their study) is due to the 1-nt variant in position 8 of the ERE (Bahreini et al., 2016). In another study, ER^{enh588} has been shown to interact with the promoter region of the CyclinD1 (CCND1) gene, and regulates CCND1 expression by ER binding at a 1-nt variant ERE DNA element (a variant in position 1) in enhancer 588 (Korkmaz et al., 2016). According to the study, mutating this 1-nt variant ERE DNA element in enhancer 588 to a 2-nt variant ERE (inserting a variant in position 10) still allowed for estradiol-mediated activity of CCND1. However, this activity was diminished when the 1-nt variant ERE DNA element in enhancer 588 was mutated to a 5-nt variant ERE DNA element (Korkmaz et al., 2016). These studies demonstrate why decoding the inversion symmetry underlying TF DNA-binding specificity and functionality in the genome is key to understanding how TFs select the appropriate regulatory targets from a large number of similar DNA elements in the genome to elicit specific transcriptional and cellular responses.

In our study, we demonstrate that TF binding in the genome follows inversion symmetry (i.e., the number of TF DNA-binding events at a particular DNA element in the genome is equivalent to the number of TF DNA-binding events at its reverse-complement DNA element in the genome). In addition, the specific DNA elements where TFs bind in the genome is determined by internal inversion symmetry within the DNA element (i.e., TF binding in the genome is determined by the position of variants within the DNA element and its reverse-complement position), hence requiring all the information contained within the entire DNA element (i.e., on both sides of the 3-nt spacer). This confirms that the majority of TF DNA-binding events that contain "nonspecific" (low-affinity) binding sites are in fact DNA sequence dependent and account for the large excess of non-NRFE (non-functional) binding sites observed in ChIPSeq and ChIPExo experiments (~55% and ~65%) (Coons et al., 2017). Thus, TF DNA-binding at non-NRFEs reflects more than just "opportunistic" DNA-binding (i.e., interacting in areas of spontaneous accessible chromatin) or artifacts of the ChIPSeq or ChIPExo technology. TF DNA-binding at non-NRFEs is a direct extension of TF DNA-binding at NRFEs, but occurs in a domain that has insufficient binding strength to initiate and/or maintain transcription.

TF binding in the genome is determined by (1) a set of rules defined by TF DNA-binding at 1-nt variant DNA elements based on which palindromic position pair the variant occupies (e.g., ER DNA-binding is increased at DNA elements with variants in palindromic position pair 3-8 or 1-10 of the ERE, whereas KR DNA-binding is increased at DNA elements with variants in palindromic position pair 2-9 of the HRE), and (2) a second distinct set of rules (functionally subordinate to and independent of the DNA-binding rules defined by TF DNA-binding at 1-nt variant DNA elements) defined by TF DNA-binding at 2-nt variant DNA elements based on how the position of the variants relate to each other (i.e., TF DNA-binding is increased at DNA elements with variants that do not crossover the 3-nt spacer). These DNA-binding rules, dictated by inversion symmetry, continue through 5- and 6-nt variant DNA elements in the genome, even though the TF DNA-binding signal decays as the number the variants increases from 0- to 5- and 6-nt variant DNA elements. The basis of these DNA-binding rules follows specific algebraic relationships. Thus, these DNA-binding rules quantitatively define how TFs select the appropriate regulatory targets from a large number of similar DNA elements in the genome to elicit specific transcriptional and cellular responses. Furthermore, the variables that define these algebraic relationships assume specific numerical values for each different or distinct TF, and these algebraic variables can be quantified or measured in laboratory experiments. Thus, this study goes beyond a theoretical analysis and defines a measurable experimental scientific methodology. That is, we have defined a mathematical formula that represents the outcome observed in hundreds of ChIPSeq and ChIPExo experiments.

Historically, the "half-site" has been defined as the two physical nucleotide sequences on the left-hand or right-hand side of the symmetrical 0-nt variant consensus palindromic DNA element (e.g., GGTCA or TGACC in the ERE, GAACA or TGTTC in the HRE), which we call "palindrome half-sites," that is, a linear string of DNA nucleotides defining the physical entity. Our analysis shows that "half-sites" do not need to be confined to a contiguous set of nucleotides, because replacing any reverse-complement nucleotide from either the left-hand or right-hand half-site results in comparable TF DNA-binding to the "palindrome half-sites." This generalizes the definition of a "half-site" because we retain the DNA-binding characteristics of the "palindrome half-sites" without the constraint of nucleotide proximity. Thus, specific focus on TF DNA-binding at "palindrome half-sites" is an elementary example of a more complex TF DNA-binding structure, and continued fixation on these anomalous data points limits ones understanding of the processes and properties of TF binding in the genome.

It is perplexing that the inversion symmetry that underlies TF DNA-binding specificity and functionality in the genome, observed in hundreds of ChIPSeq and ChIPExo experiments, has not been detected using

current DNA motif identification algorithms. Our analysis reveals that the resolution capability of PWM-based DNA motif identification is limited. An overreliance on these PWM-based algorithms has promoted the misconception that the majority of TF DNA-binding events in the genome are driven by mechanisms other than the DNA sequence. This overreliance has also led to the misconception that sNRs can indiscriminately bind to the *cis*-regulatory DNA element of other nuclear receptors. Furthermore, different *cis*-regulatory DNA elements have been delineated for the different members of the KR family. Analysis of KR DNA-binding at 81,922 0- to 5-nt variant 13-nt HRE DNA elements and at 912,718 0- to 6-nt variant 15-nt HRE DNA elements in the genome (in 194 KR ChIPSeq and ChIPExo experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria [L4-L20]), quantitatively demonstrates that all members of the KR family follow the same DNA-binding rules and thus bind the same DNA elements. This is further illustrated by the equivalent structure of the DNA-binding curves observed in AR and an AR DBD mouse model (SPARKI), where the second zinc finger in the AR DBD is replaced by that of the GR, at 81,922 0- to 5-nt variant 13-nt HRE DNA elements and at 912,718 0- to 6-nt variant 15-nt HRE DNA elements in the genome (Figures S82–S91). It is important to recognize the caveats associated with current PWM-based DNA motif identification algorithms, and careful considerations should be given with regard to any conclusions drawn from their output. Furthermore, inversion symmetry needs to be a priority in the design of future DNA motif identification search engines.

The concept of sNRs binding to specific so-called "negative response elements" in the genome as a mode of transcriptional repression has been previously described: nHRE is TYACNnnnTGAYC and nERE is TAACAnnnCGTCC (Beato, 1989) (Chen et al., 2001). Both of these negative response elements represent 5-nt variant 13-nt DNA elements: the nHRE has variants in positions 1-2-5-8-9 (positions 3-4-6-7-10 are fixed) and the nERE has variants in positions 1-2-3-6-8 (positions 4-5-7-9-10 are fixed). Thus, there are a total of 243 "negative response element" sequences for the nHRE (Tables S2 and S4) and 243 "negative response element" sequences for the nERE (Tables S1 and S3), that is, the three alternative nucleotide possibilities (variants) at each of those five variant positions. (S/N) analysis of KR DNA-binding in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) results in an average KR DNA-binding signal of 0.7 (standard deviation = 0.2) at these 243 nHRE DNA elements in the genome (Table S20). Likewise, (S/N) analysis of ER DNA-binding in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) results in an average ER DNA-binding signal of 1.2 (standard deviation = 0.4) at these 243 nERE DNA elements in the genome (Table S15). This confirms that sNR DNA-binding does not occur at "negative response elements" in the genome above (in any statistically significant measure) that which would be expected to occur at random. Furthermore, GR-Dim, a GR DBD mutant mouse model that exhibits neither hormone-mediated gene induction nor repression, has a DNA-binding signal (S/N) of 1.1 (standard deviation = 0.1) at these 243 nHRE DNA elements, whereas wtGR has a DNA-binding signal (S/N) of 0.9 (standard deviation = 0.1) at these 243 nHRE DNA elements (Figure 5). This confirms that sNRs do not bind "negative response elements" in the genome above what would be expected to occur at random, and even if one were to look at the "negative response elements" bound by sNRs that occur at random in the genome, GR-Dim is bound there as well. Thus, these DNA elements are unlikely to be responsible for hormone-mediated gene repression. In addition, in a previous study we showed that all hormone-mediated gene expression, both gene induction and repression, requires sNR DNA-binding to NRFEs (Coons et al., 2017). We further demonstrated that hormone-mediated NRFE binding by sNRs in the genome results directly in gene induction; the coregulators that are recruited to those NRFE sites are being redistributed from sites that are undergoing active transcription (Coons et al., 2017). Thus, this redistribution process consequently results in an effective repression of genes driven by those sites (Coons et al., 2017). This suggests that hormone-mediated sNR DNA-binding at NRFE DNA elements in the genome is a 2-fold mechanistic process. First, it directly causes gene induction, and second, indirectly, it causes gene repression.

Genome-wide ChIP studies provide a snapshot of TF occupancy in the genome. In a previous study we determined that only a small fraction of all TF chromatin-interacting events observed in ChIPSeq and ChIPExo experiments is associated with transcriptional activity, and this functionality is restricted to DNA elements that vary from the consensus palindromic DNA element by one or two nts (i.e., NRFEs) (Coons et al., 2017). sNR DBD mutant mouse models confirm that whereas DNA-binding at NRFE DNA elements (i.e., the 0-nt variant consensus palindromic DNA element, 1-nt variant DNA elements, and 2-nt variant DNA elements) are the only functional DNA-binding sites, DNA-binding at non-NRFE DNA elements in the genome are subject to the

same DNA-binding rules (dictated by inversion symmetry) that govern NRFE DNA elements. This DNA-binding structure, observed at non-NRFE DNA elements, is independent of peaks that contain NRFE DNA elements (Tables S67 and S71). Thus, these DNA-binding rules extend (i.e., are applicable) to DNA elements that do not support transcriptional activity. Furthermore, because the DNA-binding curve of sNR DBD mutants maintain the same structure compared with wild-type (albeit with a significantly reduced amplitude at NRFEs), this demonstrates that the effect of the DBD mutation is a loss in residence time on the DNA. That is, the DNA-binding rules are obeyed, but the retention time of the TF at non-NRFE DNA elements is not long enough to initiate and/or maintain transcription. Thus, functional (NRFE) versus non-functional (non-NRFE) DNA-binding is determined at the individual nucleotide level, and the residence time (or strength of binding) is dictated by the number and position of variants within the DNA element (i.e., inversion symmetry and DNA sequence constraints). Collectively, this suggests that the retention time of sNR DNA-binding at non-NRFE DNA elements in the genome is not long enough to initiate and/or maintain transcription, and hence, is non-functional. Thus, although the "non-specific" (low-affinity) non-NRFE DNA elements may not be capable of retaining the TF long enough to ensure activation, they may instead be used to maintain a high concentration of the TF in the vicinity of the functional binding sites (Boeva, 2016). The recent emergence of the CRISPR/Cas9 and CRISPRi genome editing technology to assess the effects of mutating many enhancers simultaneously, while maintaining the native genomic context of the binding site, provides a mechanism to explore this new dimension of TF action.

In addition to decoding the inversion symmetry underlying TF DNA-binding specificity and functionality in the genome, we also demonstrate that the population of every DNA element (1- to 20-nt) in the single-stranded genome is equivalent to the population of its reverse-complement DNA element in the single-stranded genome. Thus, the inversion symmetry observed for TF binding in the genome represents an inherent inversion symmetry structure for all DNA elements in the genome. This property is maintained at the level of each individual chromosome. These findings suggest that the structural mechanisms (by which inversion symmetry ascribes TF DNA-binding and functionality) are universally applicable. To further illustrate the generality of the mechanism, we demonstrate that DNA-binding of the unrelated TF p53 in the genome also follows inversion symmetry (i.e., the number of TF DNA-binding events at a particular DNA element in the genome is equivalent to the number of TF DNA-binding events at its reverse-complement DNA element in the genome), and the specific DNA elements where p53 binds in the genome is determined by internal inversion symmetry within the DNA element (i.e., TF binding in the genome is determined by the position of variants within the DNA element and its reverse-complement position). Thus, analysis of TF binding in the genome has expanded our understanding as to why the genome is organized in an inversion symmetry structure (i.e., Chargaff's second parity rule). Inversion symmetry is the DNA code that TFs use to interact with the genome and dictates (in conjunction with known DNA sequence constraints) which of those interactions are functionally active.

Conclusion

Gene regulatory programs are encoded in the sequence of the DNA. As in any other language, decoding the instructions between DNA and gene expression is the key for understanding transcriptional regulation. Without understanding the grammar of transcriptional regulation, one cannot tell which DNA sequence changes affect gene expression and how. The generalization of the second Chargaff rule states that the counts of any string of nucleotides on a single chromosomal strand equals the counts of its reverse-complement string of nucleotides on that same single chromosomal strand. This inversion symmetry holds for many species, both eukaryotes and prokaryotes. Our study quantitatively defines how TFs use inversion symmetry to select the appropriate regulatory targets from a large number of similar DNA elements in the genome to elicit specific transcriptional and cellular responses, that is, why a TF binds to the specific DNA elements in the genome where it does and whether that binding event ultimately affects transcriptional output. These data demonstrate that functional versus non-functional DNA-binding is dictated at the individual nucleotide level, and the residence time (or strength of binding) is determined by the number and position of variants within the DNA element (i.e., inversion symmetry and DNA sequence constraints). Furthermore, the inversion symmetry observed for TF binding in the genome represents an inherent inversion symmetry structure for all DNA elements in the genome. These findings suggest that the structural mechanisms (by which inversion symmetry ascribes TF DNA-binding and functionality) are universally applicable. Hence, analysis of TF binding in the genome has expanded our understanding as to why the genome is organized in an inversion symmetry structure (i.e., Chargaff's

second parity rule). Inversion symmetry is the DNA code that TFs use to interact with the genome and dictates (in conjunction with known DNA sequence constraints) which of those interactions are functionally active. The use of information theory and techniques in genomic analysis and modeling is therefore vital in future genomic studies.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

DATA AND SOFTWARE AVAILABILITY

Data available from the following links https://data.mendeley.com/datasets/8ks7pm2jw6/1 and https://archive.org/details/Laurel_Coons_Research_Data_Set_2019_05.

ACKNOWLEDGMENT OF DATA RESOURCES

157 ER experiments: (Hewitt et al., 2012) (Gertz et al., 2013) (Gordon et al., 2014) (Miranda et al., 2013) (Joseph et al., 2010) (Kong et al., 2011) (Guertin et al., 2014) (Welboren et al., 2009) (Liu et al., 2014) (Mohammed et al., 2015) (Franco et al., 2015) (Nagarajan et al., 2014) (Jansen et al., 2013) (Le et al., 2013) (Swinstead et al., 2016) (Palierne et al., 2016) (Korkmaz et al., 2016) (Xue et al., 2016) (Singhal et al., 2016) (Vareslija et al., 2016) (Li et al., 2013).

194 KR experiments: (Rubel et al., 2012) (Stavreva et al., 2015) (Lain et al., 2013) (Sahu et al., 2014) (Lim et al., 2015) (Miranda et al., 2013) (Grontved et al., 2013) (Langlais et al., 2012) (Siersbaek et al., 2011) (Steger et al., 2010) (Mohammed et al., 2015) (Vasquez et al., 2015) (Kaya et al., 2015) (Yin et al., 2012) (Clarke and Graham, 2012) (Jing et al., 2015) (Paakinaho et al., 2014) (Sahu et al., 2011) (Chan et al., 2015) (Puc et al., 2015) (Toropainen et al., 2015) (Chen et al., 2015a) (Sutinen et al., 2014) (Jin et al., 2014) (Le Billan et al., 2015) (Urbanucci et al., 2012) (Chng et al., 2012) (Sahu et al., 2013) (Wu et al., 2014) (Zhu et al., 2012) (Ni et al., 2013) (Gertz et al., 2013) (Lim et al., 2015) (Schiller et al., 2014) (Chen et al., 2015b) (Swinstead et al., 2016) (Thomas et al., 2015) (Chodankar et al., 2014) (Stelloo et al., 2015) (Nevedomskaya et al., 2016) (Castro et al., 2016) (Wang et al., 2016) (Handle et al., 2016)

8 GR and GR-Dim experiments: (Lim et al., 2015)

7 ER (time-course) experiments: (Guertin et al., 2014)

6 AR and AR-SPARKI experiments: (Sahu et al., 2014)

22 p53 experiments: (Tonelli et al., 2017).

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2019.04.006.

ACKNOWLEDGMENTS

The authors thank Drs. Frank Day and David Fargo at NIEHS Integrative Bioinformatics and NIEHS computational biology facility for their help with writing script codes and server utilization. We also thank Drs. Franco Demayo, Leping Li, and Ching-yi Chang for critical reading of our study. Studies in this article were supported by the Division of Intramural Research of the National Institute of Environmental Health Sciences, National Institutes of Health [Project 1ZIAES070065 (to K.S.K.)]. D.P.M is funded by DK48807 (NIH). The funders had no role in study design.

AUTHOR CONTRIBUTIONS

L.A.C, A.B.B, and S.C.H designed and performed the analyses. L.A.C. wrote the manuscript. D.P.M and K.S.K supervised and funded the research. All authors reviewed the manuscript before submission.

DECLARATION OF INTERESTS

The authors have no competing financial interests with this work.

Received: June 4, 2018
Revised: November 26, 2018
Accepted: December 4, 2018
Published: May 31, 2019

REFERENCES

Aerts, S. (2012). Chapter five - computational strategies for the genome-wide identification of cis-regulatory elements and transcriptional targets. In Current Topics in Developmental Biology, S. Plaza and F. Payre, eds. (Academic Press), pp. 121–145.

Albrecht-Buehler, G. (2006). Asymptotically increasing compliance of genomes with Chargaff's second parity rules through inversions and inverted transpositions. Proc. Natl. Acad. Sci. U S A *103*, 17828–17833.

Bahreini, A., Levine, K., Santana-Santos, L., Benos, P.V., Wang, P., Andersen, C., Oesterreich, S., and Lee, A.V. (2016). Non-coding single nucleotide variants affecting estrogen receptor binding and activity. Genome Med. *8*, 128.

Beato, M. (1989). Gene regulation by steroid hormones. Cell *56*, 335–344.

Berg, J.P. (1998). Glucocorticoid-stimulated gene expression knocked out by knock-in mutation. Eur. J. Endocrinol. *139*, 479–480.

Boeva, V. (2016). Analysis of genomic sequence motifs for deciphering transcription factor binding and transcriptional regulation in eukaryotic cells. Front. Genet. *7*, 24.

Castro, M.A., De Santiago, I., Campbell, T.M., Vaughn, C., Hickey, T.E., Ross, E., Tilley, W.D., Markowetz, F., Ponder, B.A., and Meyer, K.B. (2016). Regulators of genetic risk of breast cancer identified by integrative network analysis. Nat. Genet. *48*, 12–21.

Chan, S.C., Selth, L.A., Li, Y., Nyquist, M.D., Miao, L., Bradner, J.E., Raj, G.V., Tilley, W.D., and Dehm, S.M. (2015). Targeting chromatin binding regulation of constitutively active AR variants to overcome prostate cancer resistance to endocrine-based therapies. Nucleic Acids Res. *43*, 5880–5897.

Chen, Z., Lan, X., Thomas-Ahner, J.M., Wu, D., Liu, X., Ye, Z., Wang, L., Sunkel, B., Grenade, C., Chen, J., et al. (2015a). Agonist and antagonist switch DNA motifs recognized by human androgen receptor in prostate cancer. EMBO J. *34*, 502–516.

Chen, Z., Lan, X., Wu, D., Sunkel, B., Ye, Z., Huang, J., Liu, Z., Clinton, S.K., Jin, V.X., and Wang, Q. (2015b). Ligand-dependent genomic function of glucocorticoid receptor in triple-negative breast cancer. Nat. Commun. *6*, 8323.

Chen, Z., Zheng, H., and Dong, K.W. (2001). Identification of negative and positive estrogen response elements in human GnRH upstream promoter in the placental JEG-3 cells. Mol. Cell Endocrinol. *184*, 125–134.

Chng, K.R., Chang, C.W., Tan, S.K., Yang, C., Hong, S.Z., Sng, N.Y., and Cheung, E. (2012). A transcriptional repressor co-regulatory network governing androgen response in prostate cancers. EMBO J. *31*, 2810–2823.

Chodankar, R., Wu, D.-Y., Schiller, B.J., Yamamoto, K.R., and Stallcup, M.R. (2014). Hic-5 is a transcription coregulator that acts before and/or after glucocorticoid receptor genome occupancy in a gene-selective manner. Proc. Natl. Acad. Sci. U S A *111*, 4007–4012.

Clarke, C.L., and Graham, J.D. (2012). Non-overlapping progesterone receptor cistromes contribute to cell-specific transcriptional outcomes. PLoS One *7*, e35859.

Coons, L.A. (2017). Elucidation of the Molecular Mechanisms Underlying Estrogen-Mediated Estrogen Receptor Activation, Dissertation (Duke University), https://dukespace.lib.duke.edu/dspace/handle/10161/16283.

Coons, L.A., Hewitt, S.C., Burkholder, A.B., McDonnell, D.P., and Korach, K.S. (2017). DNA sequence constraints define functionally active steroid nuclear receptor binding sites in chromatin. Endocrinology *158*, 3212–3234.

Cotnoir-White, D., Laperriere, D., and Mader, S. (2011). Evolution of the repertoire of nuclear receptor binding sites in genomes. Mol. Cell. Endocrinol. *334*, 76–82.

Fan, J.D., Wagner, B.L., and Mcdonnell, D.P. (1996). Identification of the sequences within the human complement 3 promoter required for estrogen responsiveness provides insight into the mechanism of tamoxifen mixed agonist activity. Mol. Endocrinol. *10*, 1605–1616.

Franco, H.L., Nagari, A., and Kraus, W.L. (2015). TNFalpha signaling exposes latent estrogen receptor binding sites to alter the breast cancer cell transcriptome. Mol. Cell *58*, 21–34.

Frijters, R., Fleuren, W., Toonen, E.J., Tuckermann, J.P., Reichardt, H.M., Maaden, H., Elsas, A., Lierop, M.-J., Dokter, W., Vlieg, J., and Alkema, W. (2010). Prednisolone-induced differential gene expression in mouse liver carrying wild type or a dimerization-defective glucocorticoid receptor. BMC Genomics *11*, 1–14.

Gebhardt, J.C.M., Suter, D.M., Roy, R., Zhao, Z.W., Chapman, A.R., Basu, S., Maniatis, T., and Xie, X.S. (2013). Single molecule imaging of transcription factor binding to DNA in live mammalian cells. Nat. Methods *10*, 421–426.

Gertz, J., Savic, D., Varley, K.E., Partridge, E.C., Safi, A., Jain, P., Cooper, G.M., Reddy, T.E., Crawford, G.E., and Myers, R.M. (2013). Distinct properties of cell-type-specific and shared transcription factor binding sites. Mol. Cell *52*, 25–36.

Geserick, C., Meyer, H.A., and Haendler, B. (2005). The role of DNA response elements as allosteric modulators of steroid receptor function. Mol. Cell. Endocrinol. *236*, 1–7.

Gordon, F.K., Vallaster, C.S., Westerling, T., Iyer, L.K., Brown, M., and Schnitzler, G.R. (2014). Research resource: aorta- and liver-specific erα-binding patterns and gene regulation by estrogen. Mol. Endocrinol. *28*, 1337–1351.

Grontved, L., John, S., Baek, S., Liu, Y., Buckley, J.R., Vinson, C., Aguilera, G., and Hager, G.L. (2013). C/EBP maintains chromatin accessibility in liver and facilitates glucocorticoid receptor recruitment to steroid response elements. EMBO J. *32*, 1568–1583.

Guertin, M.J., Zhang, X., Coonrod, S.A., and Hager, G.L. (2014). Transient estrogen receptor binding and p300 redistribution support a squelching mechanism for estradiol-repressed genes. Mol. Endocrinol. *28*, 1522–1533.

Handle, F., Erb, H.H., Luef, B., Hoefer, J., Dietrich, D., Parson, W., Kristiansen, G., Santer, F.R., and Culig, Z. (2016). SOCS3 modulates the response to enzalutamide and is regulated by androgen receptor signaling and CpG methylation in prostate cancer cells. Mol. Cancer Res. *14*, 574–585.

Helsen, C., Kerkhofs, S., Clinckemalie, L., Spans, L., Laurent, M., Boonen, S., Vanderschueren, D., and Claessens, F. (2012). Structural basis for nuclear hormone receptor DNA binding. Mol. Cell. Endocrinol. *348*, 411–417.

Hewitt, S.C., Li, L., Grimm, S.A., Chen, Y., Liu, L., Li, Y., Bushel, P.R., Fargo, D., and Korach, K.S. (2012). Research resource: whole-genome estrogen receptor alpha binding in mouse uterine tissue revealed by ChIP-Seq. Mol. Endocrinol. *26*, 887–898.

Jansen, M.P., Knijnenburg, T., Reijm, E.A., Simon, I., Kerkhoven, R., Droog, M., Velds, A., Van Laere, S., Dirix, L., Alexi, X., et al. (2013). Hallmarks of aromatase inhibitor drug resistance revealed by epigenetic profiling in breast cancer. Cancer Res. *73*, 6632–6641.

Jin, H.-J., Zhao, J.C., Wu, L., Kim, J., and Yu, J. (2014). Cooperativity and equilibrium with FOXA1 define the androgen receptor transcriptional program. Nat. Commun. *5*, 3972.

Jing, D., Bhadri, V.A., Beck, D., Thoms, J.A., Yakob, N.A., Wong, J.W., Knezevic, K., Pimanda, J.E., and Lock, R.B. (2015). Opposing regulation of BIM and BCL2 controls glucocorticoid-induced apoptosis of pediatric acute lymphoblastic leukemia cells. Blood *125*, 273–283.

John, S., Sabo, P.J., Thurman, R.E., Sung, M.H., Biddie, S.C., Johnson, T.A., Hager, G.L., and Stamatoyannopoulos, J.A. (2011). Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. Nat. Genet. *43*, 264–268.

Joseph, R., Orlov, Y.L., Huss, M., Sun, W., Kong, S.L., Ukil, L., Pan, Y.F., Li, G., Lim, M., Thomsen, J.S., et al. (2010). Integrative model of genomic factors for determining binding site selection by estrogen receptor-alpha. Mol. Syst. Biol. *6*, 456.

Kaya, H.S., Hantak, A.M., Stubbs, L.J., Taylor, R.N., Bagchi, I.C., and Bagchi, M.K. (2015). Roles of progesterone receptor A and B isoforms during human endometrial decidualization. Mol. Endocrinol. *29*, 882–895.

Kleiman, A., Hubner, S., Rodriguez Parkitna, J.M., Neumann, A., Hofer, S., Weigand, M.A., Bauer, M., Schmid, W., Schutz, G., Libert, C., et al. (2012). Glucocorticoid receptor dimerization is required for survival in septic shock via suppression of interleukin-1 in macrophages. FASEB J. *26*, 722–729.

Kong, S.L., Li, G., Loh, S.L., Sung, W.-K., and Liu, E.T. (2011). Cellular reprogramming by the conjoint action of ERα, FOXA1, and GATA3 to a ligand-inducible growth state. Mol. Syst. Biol. *7*, 526.

Korkmaz, G., Lopes, R., Ugalde, A.P., Nevedomskaya, E., Han, R., Myacheva, K., Zwart, W., Elkon, R., and Agami, R. (2016). Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9. Nat. Biotechnol. *34*, 192–198.

Lain, A.R., Creighton, C.J., and Conneely, O.M. (2013). Research resource: progesterone receptor targetome underlying mammary gland branching morphogenesis. Mol. Endocrinol. *27*, 1743–1761.

Langlais, D., Couture, C., Balsalobre, A., and Drouin, J. (2012). The Stat3/GR interaction code: predictive value of direct/indirect DNA recruitment for transcription outcome. Mol. Cell *47*, 38–49.

Laudet, V. (1997). Evolution of the nuclear receptor superfamily: early diversification from an ancestral orphan receptor. J. Mol. Endocrinol. *19*, 207–226.

Le Billan, F., Khan, J.A., Lamribet, K., Viengchareun, S., Bouligand, J., Fagart, J., and Lombes, M. (2015). Cistrome of the aldosterone-activated mineralocorticoid receptor in human renal cells. FASEB J. *29*, 3977–3989.

Le, T.P., Sun, M., Luo, X., Kraus, W.L., and Greene, G.L. (2013). Mapping ERbeta genomic binding sites reveals unique genomic features and identifies EBF1 as an ERbeta interactor. PLoS One *8*, e71355.

Li, W., Notani, D., Ma, Q., Tanasa, B., Nunez, E., Chen, A.Y., Merkurjev, D., Zhang, J., Ohgi, K., Song, X., et al. (2013). Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature *498*, 516–520.

Li, X.Y., Thomas, S., Sabo, P.J., Eisen, M.B., Stamatoyannopoulos, J.A., and Biggin, M.D. (2011). The role of chromatin accessibility in directing the widespread, overlapping patterns of Drosophila transcription factor binding. Genome Biol. *12*, R34.

Lim, H.W., Uhlenhaut, N.H., Rauch, A., Weiner, J., Hubner, S., Hubner, N., Won, K.J., Lazar, M.A., Tuckermann, J., and Steger, D.J. (2015). Genomic redistribution of GR monomers and dimers mediates transcriptional response to exogenous glucocorticoid in vivo. Genome Res. *25*, 836–844.

Liu, Z., Merkurjev, D., Yang, F., Li, W., Oh, S., Friedman, M.J., Song, X., Zhang, F., Ma, Q., Ohgi, K.A., et al. (2014). Enhancer activation requires trans-recruitment of a mega transcription factor complex. Cell *159*, 358–373.

Miranda, T.B., Voss, T.C., Sung, M.H., Baek, S., John, S., Hawkins, M., Grontved, L., Schiltz, R.L., and Hager, G.L. (2013). Reprogramming the chromatin landscape: interplay of the estrogen and glucocorticoid receptors at the genomic level. Cancer Res. *73*, 5130–5139.

Mohammed, H., Russell, I.A., Stark, R., Rueda, O.M., Hickey, T.E., Tarulli, G.A., Serandour, A.A., Birrell, S.N., Bruna, A., Saadi, A., et al. (2015). Progesterone receptor modulates ERalpha action in breast cancer. Nature *523*, 313–317.

Nagarajan, S., Hossan, T., Alawi, M., Najafova, Z., Indenbirken, D., Bedi, U., Taipaleenmaki, H., Ben-Batalla, I., Scheller, M., Loges, S., et al. (2014). Bromodomain protein BRD4 is required for estrogen receptor-dependent enhancer activation and gene transcription. Cell Rep. *8*, 460–469.

Nevedomskaya, E., Stelloo, S., van der Poel, H.G., de Jong, J., Wessels, L.F.A., Bergman, A.M., and Zwart, W. (2016). Androgen receptor DNA binding and chromatin accessibility profiling in prostate cancer. Genom. Data *7*, 124–126.

Ni, M., Chen, Y., Fei, T., Li, D., Lim, E., Liu, X.S., and Brown, M. (2013). Amplitude modulation of androgen signaling by c-MYC. Genes Dev. *27*, 734–748.

Norris, J., Fan, D., Aleman, C., Marks, J.R., Futreal, P.A., Wiseman, R.W., Iglehart, J.D., Deininger, P.L., and Mcdonnell, D.P. (1995). Identification of a new subclass of Alu DNA repeats which can function as estrogen receptor-dependent transcriptional enhancers. J. Biol. Chem. *270*, 22777–22782.

Paakinaho, V., Kaikkonen, S., Makkonen, H., Benes, V., and Palvimo, J.J. (2014). SUMOylation regulates the chromatin occupancy and anti-proliferative gene programs of glucocorticoid receptor. Nucleic Acids Res. *42*, 1575–1592.

Palierne, G., Fabre, A., Solinhac, R., Le Peron, C., Avner, S., Lenfant, F., Fontaine, C., Salbert, G., Flouriot, G., Arnal, J.F., and Metivier, R. (2016). Changes in gene expression and estrogen receptor cistrome in mouse liver upon acute E2 treatment. Mol. Endocrinol. *30*, 709–732.

Puc, J., Kozbial, P., Li, W., Tan, Y., Liu, Z., Suter, T., Ohgi, K.A., Zhang, J., Aggarwal, A.K., and Rosenfeld, M.G. (2015). Ligand-dependent enhancer activation regulated by topoisomerase-I activity. Cell *160*, 367–380.

Rubel, C.A., Lanz, R.B., Kommagani, R., Franco, H.L., Lydon, J.P., and Demayo, F.J. (2012). Research resource: genome-wide profiling of progesterone receptor binding in the mouse uterus. Mol. Endocrinol. *26*, 1428–1442.

Sahu, B., Laakso, M., Ovaska, K., Mirtti, T., Lundin, J., Rannikko, A., Sankila, A., Turunen, J.P., Lundin, M., Konsti, J., et al. (2011). Dual role of FoxA1 in androgen receptor binding to chromatin, androgen signalling and prostate cancer. EMBO J. *30*, 3962–3976.

Sahu, B., Laakso, M., Pihlajamaa, P., Ovaska, K., Sinielnikov, I., Hautaniemi, S., and Janne, O.A. (2013). FoxA1 specifies unique androgen and glucocorticoid receptor binding events in prostate cancer cells. Cancer Res. *73*, 1570–1580.

Sahu, B., Pihlajamaa, P., Dubois, V., Kerkhofs, S., Claessens, F., and Janne, O.A. (2014). Androgen receptor uses relaxed response element stringency for selective chromatin binding and transcriptional regulation in vivo. Nucleic Acids Res. *42*, 4230–4240.

Schiller, B.J., Chodankar, R., Watson, L.C., Stallcup, M.R., and Yamamoto, K.R. (2014). Glucocorticoid receptor binds half sites as a monomer and regulates specific target genes. Genome Biol. *15*, 418.

Shporer, S., Chor, B., Rosset, S., and Horn, D. (2016). Inversion symmetry of DNA k-mer counts: validity and deviations. BMC Genomics *17*, 696.

Siersbaek, R., Nielsen, R., John, S., Sung, M.H., Baek, S., Loft, A., Hager, G.L., and Mandrup, S. (2011). Extensive chromatin remodelling and establishment of transcription factor 'hotspots' during early adipogenesis. EMBO J. *30*, 1459–1472.

Singhal, H., Greene, M.E., Tarulli, G., Zarnke, A.L., Bourgo, R.J., Laine, M., Chang, Y.F., Ma, S., Dembo, A.G., Raj, G.V., et al. (2016). Genomic agonism and phenotypic antagonism between estrogen and progesterone receptors in breast cancer. Sci. Adv. *2*, e1501924.

Stavreva, D.A., Coulon, A., Baek, S., Sung, M.H., John, S., Stixova, L., Tesikova, M., Hakim, O., Miranda, T., Hawkins, M., et al. (2015). Dynamics of chromatin accessibility and long-range interactions in response to glucocorticoid pulsing. Genome Res. *25*, 845–857.

Steger, D.J., Grant, G.R., Schupp, M., Tomaru, T., Lefterova, M.I., Schug, J., Manduchi, E., Stoeckert, C.J., Jr., and Lazar, M.A. (2010). Propagation of adipogenic signals through an epigenomic transition state. Genes Dev. *24*, 1035–1044.

Stelloo, S., Nevedomskaya, E., Van Der Poel, H.G., De Jong, J., Van Leenders, G.J., Jenster, G., Wessels, L.F., Bergman, A.M., and Zwart, W. (2015). Androgen receptor profiling predicts prostate cancer outcome. EMBO Mol. Med. *7*, 1450–1464.

Sutinen, P., Malinen, M., Heikkinen, S., and Palvimo, J.J. (2014). SUMOylation modulates the transcriptional activity of androgen receptor in a target gene and pathway selective manner. Nucleic Acids Res. *42*, 8310–8319.

Swinstead, E.E., Miranda, T.B., Paakinaho, V., Baek, S., Goldstein, I., Hawkins, M., Karpova, T.S., Ball, D., Mazza, D., Lavis, L.D., et al. (2016). Steroid receptors reprogram FoxA1 occupancy through dynamic chromatin transitions. Cell *165*, 593–605.

Tang, Q., Chen, Y., Meyer, C., Geistlinger, T., Lupien, M., Wang, Q., Liu, T., Zhang, Y., Brown, M., and Liu, X.S. (2011). A comprehensive view of nuclear receptor cancer cistromes. Cancer Res. *71*, 6940–6947.

Thomas, A.L., Coarfa, C., Qian, J., Wilkerson, J.J., Rajapakshe, K., Krett, N.L., Gunaratne, P.H., and Rosen, S.T. (2015). Identification of potential glucocorticoid receptor therapeutic targets in multiple myeloma. Nucl. Recept Signal. *13*, e006.

Todeschini, A.-L., Georges, A., and Veitia, R.A. (2014). Transcription factors: specific DNA binding and specific gene regulation. Trends Genet. *30*, 211–219.

Tonelli, C., Morelli, M.J., Sabo, A., Verrecchia, A., Rotta, L., Capra, T., Bianchi, S., Campaner, S., and Amati, B. (2017). Genome-wide analysis of p53-regulated transcription in Myc-driven lymphomas. Oncogene *36*, 2921–2929.

Toropainen, S., Malinen, M., Kaikkonen, S., Rytinki, M., Jaaskelainen, T., Sahu, B., Janne, O.A., and Palvimo, J.J. (2015). SUMO ligase PIAS1 functions as a target gene selective androgen receptor coregulator on prostate cancer cell chromatin. Nucleic Acids Res. *43*, 848–861.

Urbanucci, A., Sahu, B., Seppala, J., Larjo, A., Latonen, L.M., Waltering, K.K., Tammela, T.L., Vessella, R.L., Lahdesmaki, H., Janne, O.A., and Visakorpi, T. (2012). Overexpression of androgen receptor enhances the binding of the receptor to the chromatin in prostate cancer. Oncogene *31*, 2153–2163.

Vareslija, D., Mcbryan, J., Fagan, A., Redmond, A.M., Hao, Y., Sims, A.H., Turnbull, A., Dixon, J.M., Ó Gaora, P., Hudson, L., et al. (2016). Adaptation to AI Therapy in breast cancer can induce dynamic alterations in er activity resulting in estrogen-independent metastatic tumors. Clin. Cancer Res. *22*, 2765–2777.

Vasquez, Y.M., Mazur, E.C., Li, X., Kommagani, R., Jiang, L., Chen, R., Lanz, R.B., Kovanci, E., Gibbons, W.E., and Demayo, F.J. (2015). FOXO1 is required for binding of PR on IRF4, novel transcriptional regulator of endometrial stromal decidualization. Mol. Endocrinol. *29*, 421–433.

Vockley, C.M., D'Ippolito, A.M., Mcdowell, I.C., Majoros, W.H., Safi, A., Song, L., Crawford, G.E., and Reddy, T.E. (2016). Direct GR binding sites potentiate clusters of TF binding across the human genome. Cell *166*, 1269–1281.e19.

Wang, J., Zou, J.X., Xue, X., Cai, D., Zhang, Y., Duan, Z., Xiang, Q., Yang, J.C., Louie, M.C., Borowsky, A.D., et al. (2016). ROR-gamma drives androgen receptor expression and represents a therapeutic target in castration-resistant prostate cancer. Nat. Med. *22*, 488–496.

Welboren, W.J., Van Driel, M.A., Janssen-Megens, E.M., Van Heeringen, S.J., Sweep, F.C., Span, P.N., and Stunnenberg, H.G. (2009). ChIP-Seq of ERalpha and RNA polymerase II defines genes differentially responding to ligands. EMBO J. *28*, 1418–1428.

Wu, D., Sunkel, B., Chen, Z., Liu, X., Ye, Z., Li, Q., Grenade, C., Ke, J., Zhang, C., Chen, H., et al. (2014). Three-tiered role of the pioneer factor GATA2 in promoting androgen-dependent gene expression in prostate cancer. Nucleic Acids Res. *42*, 3607–3622.

Xue, X., Yang, Y.A., Zhang, A., Fong, K.W., Kim, J., Song, B., Li, S., Zhao, J.C., and Yu, J. (2016). LncRNA HOTAIR enhances ER signaling and confers tamoxifen resistance in breast cancer. Oncogene *35*, 2746–2755.

Yáñez-Cuna, J.O., Kvon, E.Z., and Stark, A. (2013). Deciphering the transcriptional cis-regulatory code. Trends Genet. *29*, 11–22.

Yin, P., Roqueiro, D., Huang, L., Owen, J.K., Xie, A., Navarro, A., Monsivais, D., Coon, J.S.T., Kim, J.J., Dai, Y., and Bulun, S.E. (2012). Genome-wide progesterone receptor binding: cell type-specific and shared mechanisms in T47D breast cancer cells and primary leiomyoma cells. PLoS One *7*, e29021.

Zhu, Z., Shi, M., Hu, W., Estrella, H., Engebretsen, J., Nichols, T., Briere, D., Hosea, N., Los, G., Rejto, P.A., and Fanjul, A. (2012). Dose-dependent effects of small-molecule antagonists on the genomic landscape of androgen receptor binding. BMC Genomics *13*, 355.

ISCI, Volume *15*

Supplemental Information

Decoding the Inversion Symmetry Underlying Transcription Factor DNA-Binding Specificity and Functionality in the Genome

Laurel A. Coons, Adam B. Burkholder, Sylvia C. Hewitt, Donald P. McDonnell, and Kenneth S. Korach

Supplemental Information

Decoding the Inversion Symmetry Underlying Transcription Factor DNA-Binding Specificity and Functionality in the Genome

This document contains the following supplemental information:

Contents

Figure S1-S25 Descriptions: 13-nt ERE and HRE DNA Element Analysis

****All detailed data and statistics associated with every figure are compiled in Data S1****

The analyses in this section were completed by overlapping the location coordinates of each 0-nt to 5-nt variant 13-nt ERE or HRE DNA element in the genome and the location coordinates of the ChIPSeq or ChIPExo peaks in an experiment. The 0-nt variant consensus palindromic 13-nt ERE DNA element (5′-GGTCAnnnTGACC-3′) and 0-nt variant consensus palindromic 13-nt HRE DNA element (5′-GAACAnnnTGTTC-3′) have ten (10) primary positions. The 0-nt to 5-nt variant ERE or HRE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements (10 variant positions), 405 2-nt variant DNA elements (45 variant positions), 3,240 3-nt variant DNA elements (120 variant positions), 17,010 4-nt variant DNA elements (210 variant positions), and 61,236 5-nt variant DNA elements (252 variant positions), for a total of 81,922 DNA elements.

Figure S1. Inversion Symmetry of sNR DNA-Binding at 2-nt Variant EREs and HREs in the Genome (%)
(A) Average distribution of 2-nt variant EREs in ER experiments (156 ERα and 1 ERβ) at the four hundred and five (405) 2-nt variant EREs in the genome (the total number of 2-nt variant EREs contained within a ChIPSeq experiment =100%) **(B)** Average distribution of 2-nt variant HREs in 193 KR experiments (75 AR, 64 GR, 54 PR) at the four hundred and five (405) 2-nt variant HREs in the genome (the total number of 2-nt variant HREs contained within a ChIPSeq experiment =100%). These four hundred and five (405) 2-nt variant EREs or HREs are defined by forty-five (45) variant positions. Five (5) of those are when the variants are in the palindromic position pairs (PP): 1-10, 2-9, 3-8, 4-7, 5-6. The remaining forty (40) are symmetrically split into 20 groups and their reverse-complements.

*****

Figure S2. (S/N) analysis of 0-nt to 5-nt Variant EREs in ER ChIPSeq Peaks (displayed by number of variants)
(S/N) analysis of ER DNA-binding at 0-nt to 5-nt variant EREs in 157 ER experiments (156 ERα and 1 ERβ). Experiments are ordered by the experiment with highest (S/N) value at the 0-nt variant consensus palindromic ERE DNA element in the genome. The signal-to-noise ratio (S/N) is the absolute number of times a 0-nt to 5-nt variant DNA element occurs within an experiment [defined by the total number of peaks in the experiment and the peak length] compared to the random frequency expectation of that DNA element occurring in the genome (i.e., the probability that any 10-nt DNA element that has a maximum possibility of 4 nucleotides at each of the 10 primary positions will occur is once every 1,048,576 nucleotides (4^{10}) at random occurrence). Thus, the expected noise is calculated by multiplying the total number of peaks in the experiment by the length of the peaks over 4^{10}. For (S/N) analysis of 0-nt to 5-nt variant DNA elements (displayed by the number of variants in the DNA element), the expected noise is multiplied by the number of DNA elements in each group: 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements, 405 2-nt variant DNA elements, 3,240 3-nt variant DNA elements, 17,010 4-nt variant DNA elements, and 61,236 5-nt variant DNA elements, for a total of 81,922 0-nt to 5-nt variant DNA elements. **Note**: the relative enhancement values of the sNR DNA-binding signals are scale invariant (i.e., the relative ratios between the 0-nt to 5-nt variant groups are constant). Y-axis =the natural log of the (S/N) value (i.e., random =0 and 0 =log(1)).

Figure S3. (S/N) analysis of 0-nt to 5-nt Variant HREs in KR ChIPSeq Peaks (displayed by number of variants)
(S/N) analysis of KR DNA-binding at 0-nt to 5-nt variant HREs in 194 KR experiments (75 AR, 64 GR, 1 MR, 54 PR). Experiments are grouped by KR, and then ordered by the experiment with highest (S/N) value at the 0-nt variant consensus palindromic HRE DNA element in the genome. The signal-to-noise ratio (S/N) is the absolute number of times a 0-nt to 5-nt variant DNA element occurs within an experiment [defined by the total number of peaks in the experiment and the peak length] compared to the random frequency expectation of that DNA element occurring in the genome (i.e., the probability that any 10-nt DNA element that has a maximum possibility of 4 nucleotides at each of the 10 primary positions will occur is once every 1,048,576 nucleotides (4^{10}) at random occurrence). Thus, the expected noise is calculated by multiplying the total number of peaks in the experiment by the length of the peaks over 4^{10}. For (S/N) analysis of 0-nt to 5-nt variant DNA elements (displayed by the number of variants in the DNA element), the expected noise is multiplied by the number of DNA elements in each group: 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements, 405 2-nt variant DNA elements, 3,240 3-nt variant DNA elements, 17,010 4-nt variant DNA elements, and 61,236 5-nt variant DNA elements, for a total of 81,922 0-nt to 5-nt variant DNA elements. **Note**: the relative enhancement values of the sNR DNA-binding signals are scale invariant (i.e., the relative ratios between the 0-nt to 5-nt variant groups are constant). Y-axis =the natural log of the (S/N) value (i.e., random =0 and 0 =log(1)).

*****

Figure S4. (S/N) analysis of 0-nt to 5-nt Variant HREs in GR ChIPSeq Peaks (displayed by 252 half-site groups)
(S/N) analysis of 0-nt to 5-nt variant HREs (displayed by the 252 half-site groups) in GR (Liver-GR-WT-pred-6am-2) [23,742 peaks, 142-nt peak length] ChIPSeq peaks. X-axis order =Reverse-Complement Vacancy Position ID 2-9 > 5-6 > 1-10 > 4-7 > 3-8. See **Table S6** for x-axis details. This DNA-binding profile was observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S8)**.

Figure S5. (S/N) analysis of 0-nt to 5-nt Variant HREs in PR ChIPSeq Peaks (displayed by 252 half-site groups)

(S/N) analysis of 0-nt to 5-nt variant HREs (displayed by the 252 half-site groups) in PR (Uterus-PGR-P4) [66,545 peaks, 154-nt peak length] ChIPSeq peaks. X-axis order =Reverse-Complement Vacancy Position ID 2-9 > 5-6 > 1-10 > 4-7 > 3-8. See **Table S6** for x-axis details. This DNA-binding profile was observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S8)**.

*****

Figure S6. Quantify the 5 Discrete States of ER DNA-Binding at 1-nt Variant EREs in the Genome (3,2,0,0,0)
(S/N) analysis of 1-nt variant EREs (displayed by the 252 half-site groups) in ER (WT-E2-1hr) [76,163 peaks, 146-nt peak length] ChIPSeq peaks produce five (5) discrete binding signals: 6 groups reach the (+2) plateau, 60 groups reach the (+1) plateau, 120 groups reach the (0) plateau, 60 groups reach the (-1) plateau, 6 groups reach the (-2) plateau. Thus, ER DNA-binding is greatest at the (+2) plateau, which are the 6 half-site groups that have two reverse-complement vacancies in palindromic position pair 3-8 and 1-10 and two reverse-complement double occupants in palindromic position pair (4-7 and 5-6), (2-9 and 5-6), or (2-9 and 4-7). ER DNA-binding is least at the (-2) plateau, which are the 6 half-site groups that have two reverse-complement vacancies in palindromic position pair (4-7 and 5-6), (2-9 and 5-6), or (2-9 and 4-7), and two reverse-complement double occupants in palindromic position pair 3-8 and 1-10. See **Table S5** for x-axis details. This DNA-binding profile was observed in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S9)**.

Figure S7. Quantify the 3 Discrete States of KR DNA-Binding at 1-nt Variant HREs in the Genome (4,1,0,0,0)
(S/N) analysis of 1-nt variant HREs (displayed by the 252 half-site groups) in AR (AR-wt1) [49,859 peaks, 136-nt peak length] ChIPSeq peaks produce three (3) discrete binding signals: 56 groups reach the (+1) plateau, 140 groups reach the (0) plateau, 56 groups reach the (-1) plateau. Thus, KR DNA-binding is greatest at the (+1) plateau, which are the 56 half-site groups that have a reverse-complement vacancy in palindromic position pair 2-9. KR DNA-binding is least at the (-1) plateau, which are the 56 half-site groups that have a reverse-complement double occupant in palindromic position pair 2-9. See **Table S6** for x-axis details. This DNA-binding profile was observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S10)**.

Figure S8. Algebraic Equations of the 3-State DNA Element (4,1,0,0,0)
The symbolic representations of the discrete states (+1, 0, -1) for the HRE can also be replaced with algebraic variables to formally define the methodology. The 3-state HRE is generated by splitting the nucleotides into a (4,1,0,0,0) grouping. That is, palindromic position pair 2-9 (symbolically represented by "A") versus palindromic position pairs 5-6, 1-10, 4-7, 3-8 (symbolically represented by "B"). Therefore, if an A is replaced with a B, the impact is (–X) (i.e., the sNR DNA-binding affinity is decreased); while if a B is replaced with an A, the impact is the exactly the opposite (+X) (i.e., the sNR DNA-binding affinity is increased) to the same magnitude. Thus, this algebraic representation explains why sNR DNA-binding is quantitatively precise, and also why there are three (3) discrete DNA-binding states at 1-nt variant HRE DNA elements (by the KRs).

Figure S9. Algebraic Equations of the 5-State DNA Element (3,2,0,0,0)
The symbolic representations of the discrete states (+2, +1, 0, -1, -2) for the ERE can also be replaced with algebraic variables to formally define the methodology. The 5-state ERE is generated by splitting the nucleotides into a (3,2,0,0,0) grouping. That is, palindromic position pair 3-8 and 1-10 (symbolically represented by "A") versus palindromic position pairs 5-6, 4-7, 2-9 (symbolically represented by "B"). Therefore, if an A is replaced with a B, the impact is (–X) (i.e., the sNR DNA-binding affinity is decreased); while if a B is replaced with an A, the impact is exactly the opposite (+X) (i.e., the sNR DNA-binding affinity is increased) to the same magnitude. Thus, this algebraic representation explains why sNR DNA-binding is quantitatively precise, and also why there are five (5) discrete DNA-binding states at 1-nt variant ERE DNA elements (by the ERs).

Figure S10. Algebraic Equations of the 9-State DNA Element (3,1,1,0,0)
The algebraic complexities increase substantially beyond the (4,1,0,0,0) HRE and (3,2,0,0,0) ERE representations. The (3,1,1,0,0) grouping (symbolically represented by "A, B, C") creates 9 discrete states.

Figure S11. Algebraic Equations of the 11-State, 25-State, 51-State DNA Elements (2,2,1,0,0) (2,1,1,1,0) (1,1,1,1,1)
The algebraic complexities increase substantially beyond the (4,1,0,0,0) HRE and (3,2,0,0,0) ERE representations. The (2,2,1,0,0) grouping (symbolically represented by "A, B, C") creates 11 discrete states, the (2,1,1,1,1) grouping (symbolically represented by "A, B, C, D") creates 25 discrete states, and the (1,1,1,1,1) grouping (symbolically represented by "A, B, C, D, E") creates 51 discrete states; which are all of the remaining possibilities for a 5-nt DNA element in the genome (i.e., a 5-nt element, followed by any arbitrary spacer, followed by its 5-nt reverse-complement element).

*****

Figure S12. (S/N) analysis of 1-nt Variant EREs in ER ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 1-nt variant EREs in ER (WT-E2-1hr) ChIPSeq peaks [76,163 peaks, 146-nt peak length].
Displayed by the 10 variant positions of the 30 1-nt variant EREs. This DNA-binding profile was observed in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S11)**.

Figure S13. (S/N) analysis of 2-nt Variant EREs in ER ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 2-nt variant EREs in ER (WT-E2-1hr) ChIPSeq peaks [76,163 peaks, 146-nt peak length].
Displayed by the 45 variant positions of the 405 2-nt variant EREs. 5 of the 45 variant positions are the five (5) palindromic position pairs: 1-10, 2-9, 3-8, 4-7, 5-6 (PP, dark red solid bars). 20 of the 45 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (blue solid bars). 20 of the 45 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (blue diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome). This DNA-binding profile was observed in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S12)**.

Figure S14. (S/N) analysis of 3-nt Variant EREs in ER ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 3-nt variant EREs in ER (WT-E2-1hr) ChIPSeq peaks [76,163 peaks, 146-nt peak length].
Displayed by the 120 variant positions of the 3,240 3-nt variant EREs. 20 of the 120 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (blue solid bars). 100 of the 120 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (blue diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome). This DNA-binding profile was observed in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S13)**.

Figure S15. (S/N) analysis of 4-nt Variant EREs in ER ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 4-nt variant EREs in ER (WT-E2-1hr) ChIPSeq peaks [76,163 peaks, 146-nt peak length].
Displayed by the 210 variant positions of the 17,010 4-nt variant EREs. 10 of the 210 variant positions are the five (5) palindromic position pairs: 1-3-8-10, 3-5-6-8, 3-4-7-8, 2-3-8-9, 1-5-6-10, 1-4-7-10, 1-2-9-10, 4-5-6-7, 2-5-6-9, 2-4-7-9 (PP, dark red solid bars). 10 of the 210 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (blue solid bars). 190 of the 45 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (blue diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome). This DNA-binding profile was observed in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S14)**.

Figure S16. (S/N) analysis of 5-nt Variant EREs in ER ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 5-nt variant EREs in ER (WT-E2-1hr) ChIPSeq peaks [76,163 peaks, 146-nt peak length].
Displayed by the 252 variant positions of the 61,236 5-nt variant EREs. 2 of the 252 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (blue solid bars). 250 of the 252 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (blue diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome). This DNA-binding profile was observed in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S15)**.

*****

Figure S17. (S/N) analysis of 1-nt Variant HREs in KR ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 1-nt variant HREs in AR (AR-wt1) ChIPSeq peaks [49,859 peaks, 136-nt peak length].
Displayed by the 10 variant positions of the 30 1-nt variant HREs. This DNA-binding profile was observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S16)**.

Figure S18. (S/N) analysis of 2-nt Variant HREs in KR ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 2-nt variant HREs in AR (AR-wt1) ChIPSeq peaks [49,859 peaks, 136-nt peak length].
Displayed by the 45 variant positions of the 405 2-nt variant HREs. 5 of the 45 variant positions are the five (5) palindromic position pairs: 1-10, 2-9, 3-8, 4-7, 5-6 (PP, dark red solid bars). 20 of the 45 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (orange solid bars). 20 of the 45 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (orange diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome). This DNA-binding profile was observed in 194 KR

experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S17)**.

Figure S19. (S/N) analysis of 3-nt Variant HREs in KR ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 3-nt variant HREs in AR (AR-wt1) ChIPSeq peaks [49,859 peaks, 136-nt peak length].
Displayed by the 120 variant positions of the 3,240 3-nt variant HREs. 20 of the 120 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (orange solid bars). 100 of the 120 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (orange diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome). This DNA-binding profile was observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S18)**.

Figure S20. (S/N) analysis of 4-nt Variant HREs in KR ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 4-nt variant HREs in AR (AR-wt1) ChIPSeq peaks [49,859 peaks, 136-nt peak length].
Displayed by the 210 variant positions of the 17,010 4-nt variant HREs. 10 of the 210 variant positions are the five (5) palindromic position pairs: 2-5-6-9, 1-2-9-10, 2-4-7-9, 2-3-8-9, 1-5-6-10, 4-5-6-7, 3-5-6-8, 1-4-7-10, 1-3-8-10, 3-4-7-8 (PP, dark red solid bars). 10 of the 210 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (orange solid bars). 190 of the 45 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (orange diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome). This DNA-binding profile was observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S19)**.

Figure S21. (S/N) analysis of 5-nt Variant HREs in KR ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 5-nt variant HREs in AR (AR-wt1) ChIPSeq peaks [49,859 peaks, 136-nt peak length].
Displayed by the 252 variant positions of the 61,236 5-nt variant HREs. 2 of the 252 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (orange solid bars). 250 of the 252 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (orange diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome). This DNA-binding profile was observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S20)**.

*****

Here we show replicate experiments for GR and GR-Dim to illustrate that the slight decrease in DNA-binding at a few specific 3-nt variant HRE DNA elements by GR-Dim is made apparent by the DNA-binding enhancement of a particular experiment at 0-nt to 2-nt variant HRE DNA elements as the relative (S/N) values of sNR DNA-binding signals are scale invariant (i.e., the relative ratios between the 0-nt to 5-nt variant groups are constant) **(Figure S2-S3)**. Thus, the reduced amplitude in DNA-binding at a few specific 3-nt variant HRE DNA elements only occurs in experiments with the highest (S/N) values at 0-nt to 2-nt variant HRE DNA elements.

Figure S22. (S/N) analysis of 0-nt to 5-nt Variant HREs in GR and GR-Dim ChIPSeq Peaks #2 (displayed by 252 half-site groups)
(S/N) analysis of 0-nt to 5-nt variant HREs (displayed by the 252 half-site groups) in GR (Liver-GR-WT-pred-6am-1) [34,758 peaks, 151-nt peak length] ChIPSeq peaks and GR-Dim (Liver-GR-Dim-pred-6am-1) [22,130 peaks, 163-nt peak length] ChIPSeq peaks. X-axis order =Reverse-Complement Vacancy Position ID 2-9 > 5-6 > 1-10 > 4-7 > 3-8. See **Table S6** for x-axis details.

Figure S23. (S/N) analysis of 0-nt to 5-nt Variant HREs in GR and GR-Dim ChIPExo Peaks #1 (displayed by 252 half-site groups)
(S/N) analysis of 0-nt to 5-nt variant HREs (displayed by the 252 half-site groups) in GR (Liver-GR-WT-pred-6am-1) [20,966 peaks, 445-nt peak length] ChIPExo peaks and GR-Dim (Liver-GR-Dim-pred-6am-1) [19,480 peaks, 391-nt peak length] ChIPExo peaks. X-axis order =Reverse-Complement Vacancy Position ID 2-9 > 5-6 > 1-10 > 4-7 > 3-8. See **Table S6** for x-axis details.

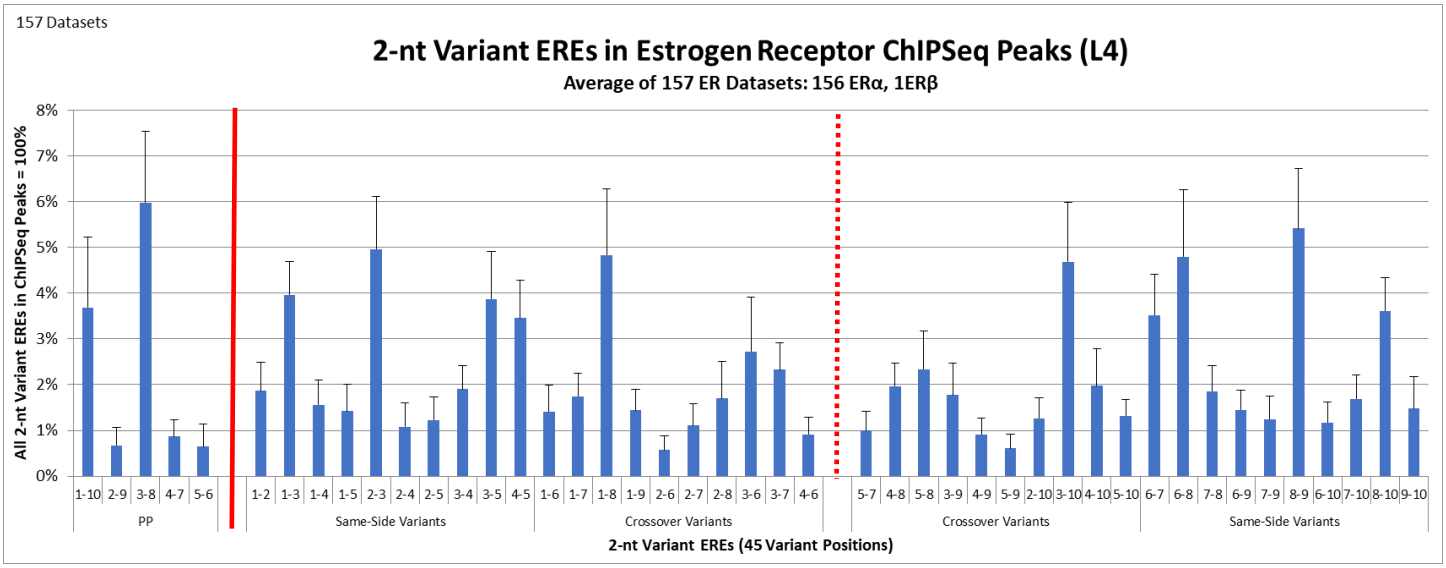
Figure S24. (S/N) analysis of 0-nt to 5-nt Variant HREs in GR and GR-Dim ChIPExo Peaks #2 (displayed by 252 half-site groups)
(S/N) analysis of 0-nt to 5-nt variant HREs (displayed by the 252 half-site groups) in GR (Liver-GR-WT-pred-6am-2) [22,129 peaks, 338-nt peak length] ChIPExo peaks and GR-Dim (Liver-GR-Dim-pred-6am-2) [25,670 peaks, 408-nt peak length] ChIPExo peaks. X-axis order =Reverse-Complement Vacancy Position ID 2-9 > 5-6 > 1-10 > 4-7 > 3-8. See **Table S6** for x-axis details.

Figure S25. (S/N) analysis of 1-nt to 5-nt Variant HREs in GR and GR-Dim ChIPExo Peaks (displayed by variant position) (S/N) analysis of 1-nt to 5-nt variant HREs (displayed by variant position) in GR (Liver-GR-WT-pred-6am-1) [20,966 peaks, 445-nt peak length] (Liver-GR-WT-pred-6am-2) [22,129 peaks, 338-nt peak length] ChIPExo peaks and GR-Dim (Liver-GR-Dim-pred-6am-1) [19,480 peaks, 391-nt peak length] (Liver-GR-Dim-pred-6am-2) [25,670 peaks, 408-nt peak length] ChIPExo peaks.

Figure S1. Inversion Symmetry of sNR DNA-Binding at 2-nt Variant EREs and HREs in the Genome (%)

(A) 2-nt Variant EREs in ER ChIPSeq Peaks (157 Experiments)

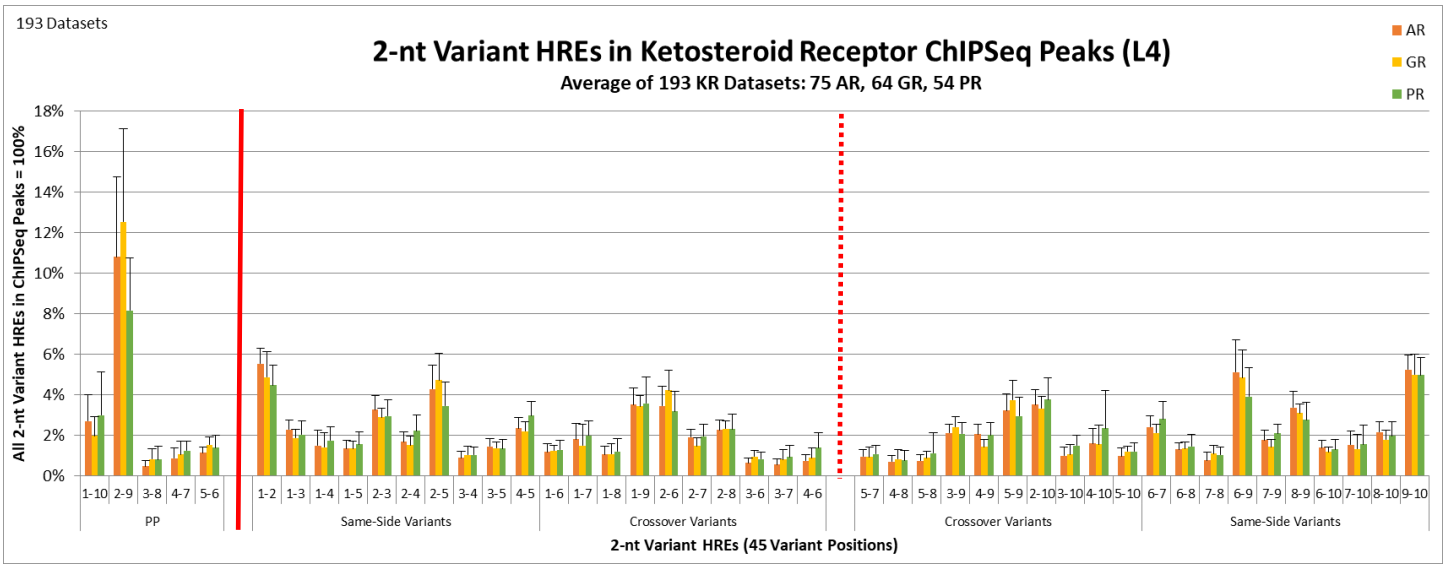
ERE = GGTCAnnnTGACC
1 2 3 4 5 6 7 8 9 10



---- **Symmetry Divider**

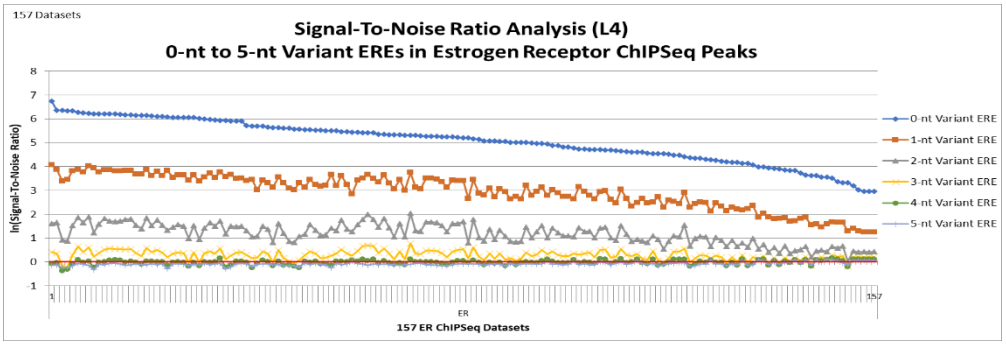
(B) 2-nt Variant HREs in KR ChIPSeq Peaks (193 Experiments)

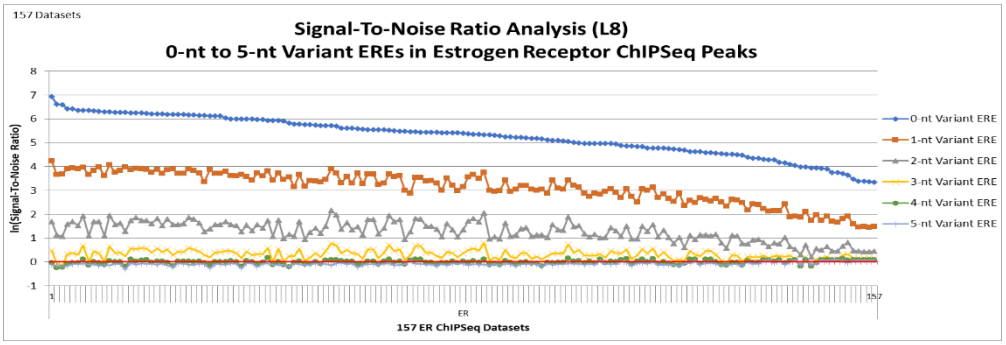
HRE = GAACAnnnTGTTC
1 2 3 4 5 6 7 8 9 10

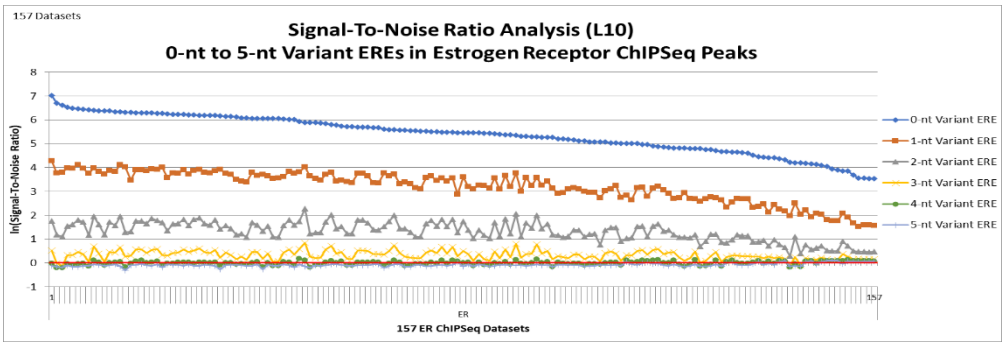


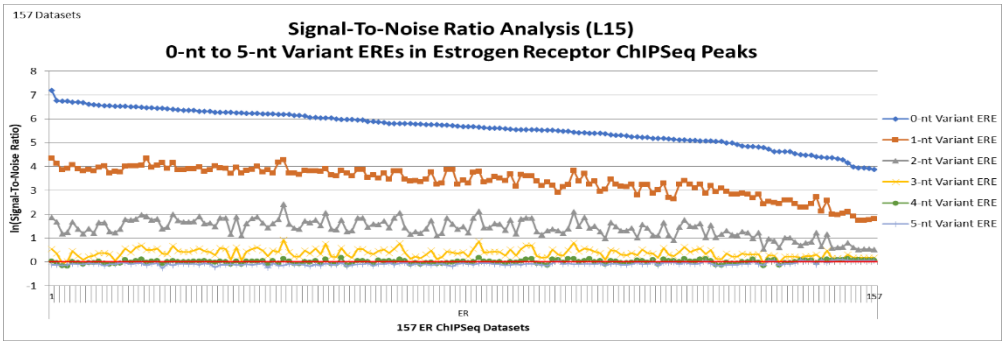
---- **Symmetry Divider**

Figure S2. (S/N) analysis of 0-nt to 5-nt Variant EREs in ER ChIPSeq Peaks (# of Variants)









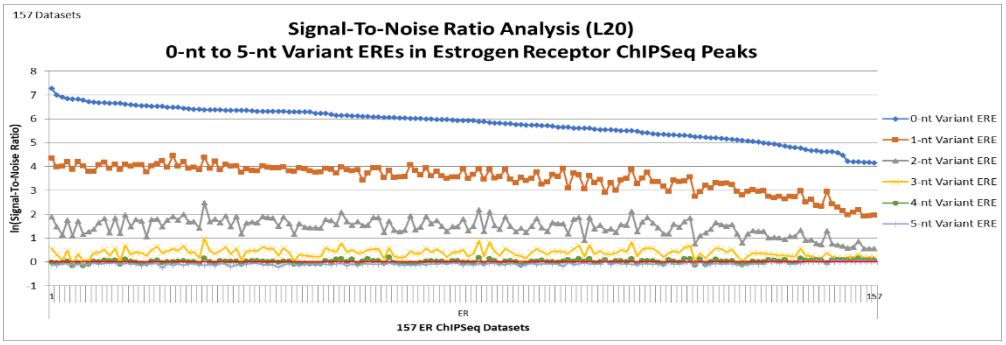
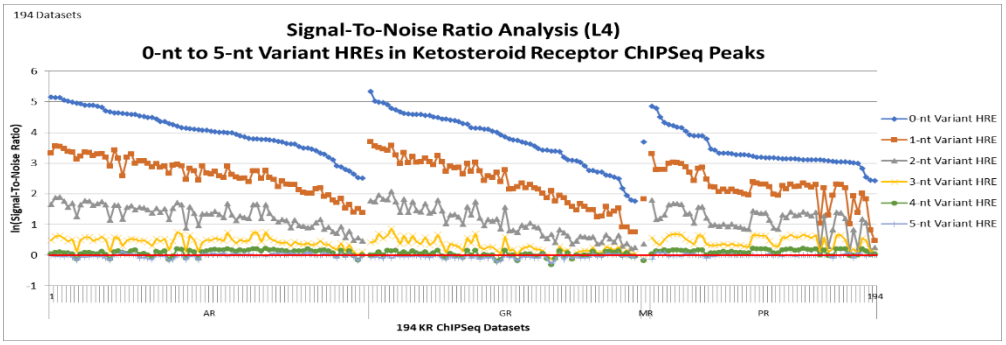
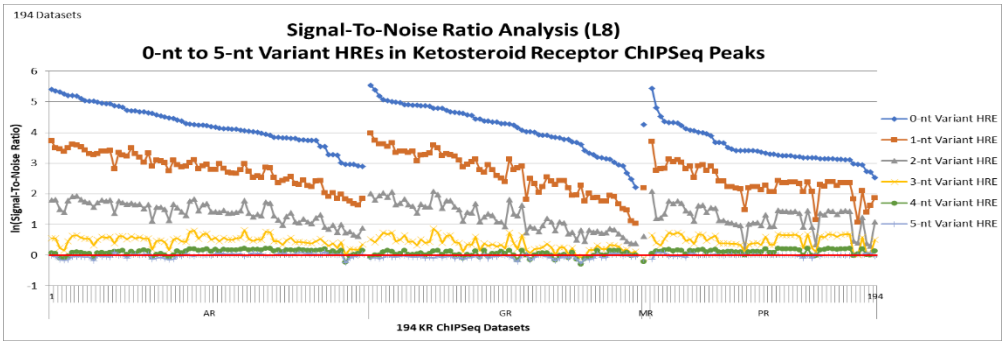
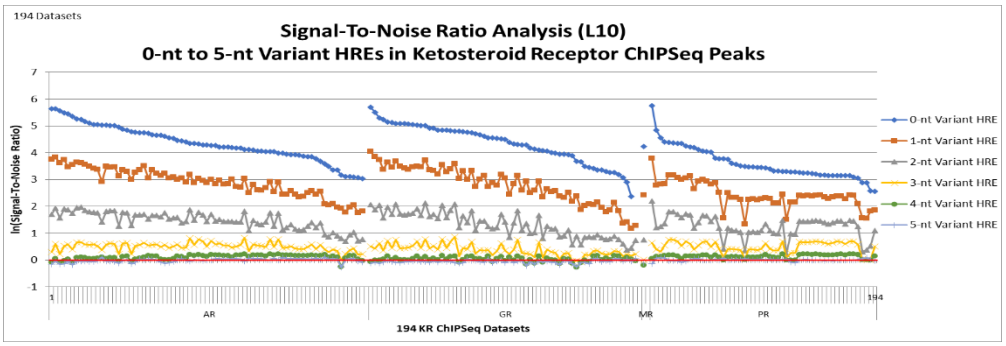
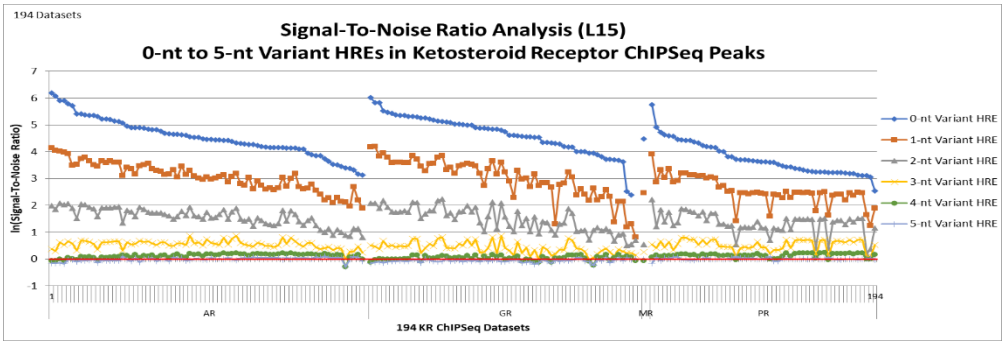


Figure S3. (S/N) analysis of 0-nt to 5-nt Variant HREs in KR ChIPSeq Peaks (# of Variants)









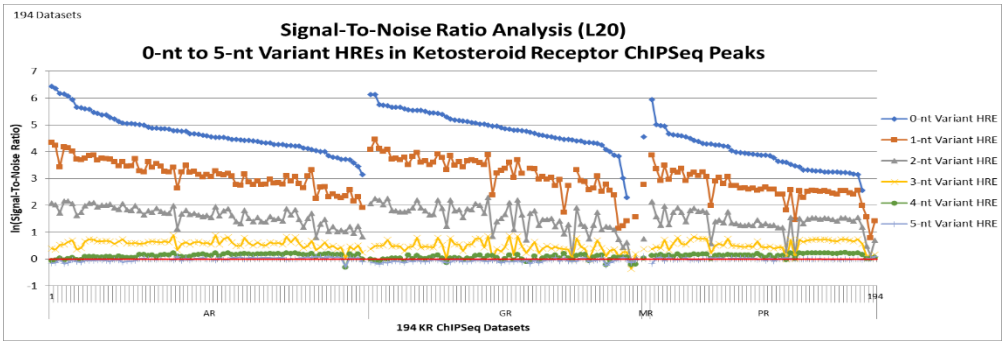


Figure S4. (S/N) analysis of 0-nt to 5-nt Variant HREs in GR ChIPSeq Peaks (252 Half-Site Groups)

Figure S5. (S/N) analysis of 0-nt to 5-nt Variant HREs in PR ChIPSeq Peaks (252 Half-Site Groups)

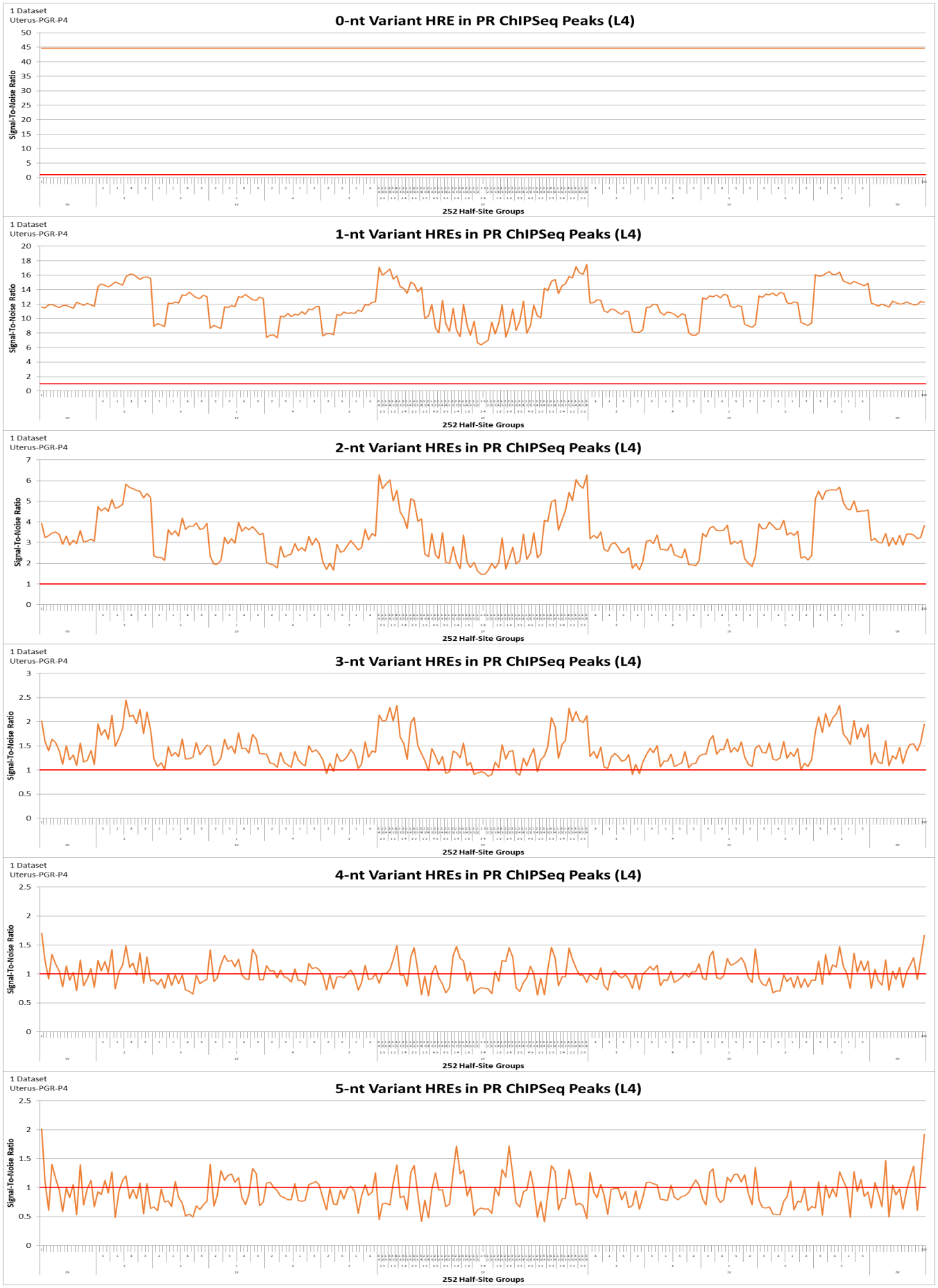
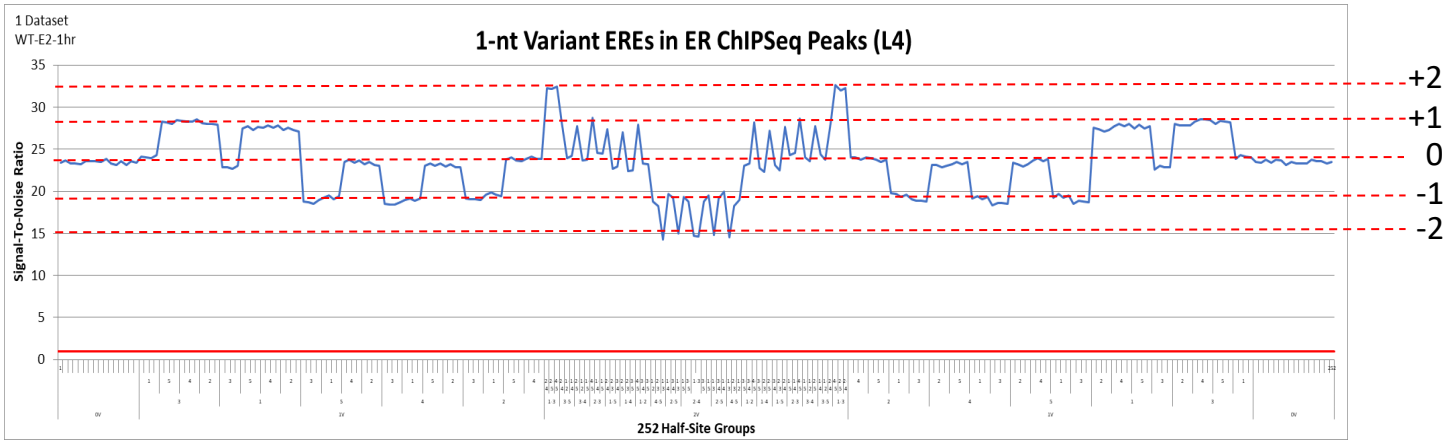


Figure S6. Quantify the 5 Discrete States of ER DNA-Binding at 1-nt Variant EREs in the Genome (3,2,0,0,0)

ERE = GGTCAnnnTGACC

1 2 3 4 5 6 7 8 9 10



Five (5) Discrete DNA-Binding Plateaus

One Vacancy

RC Double Occupant Position ID \ ERE	3-8	1-10	5-6	4-7	2-9
3-8		0	-1	-1	-1
1-10	0		-1	-1	-1
5-6	+1	+1		0	0
4-7	+1	+1	0		0
2-9	+1	+1	0	0	

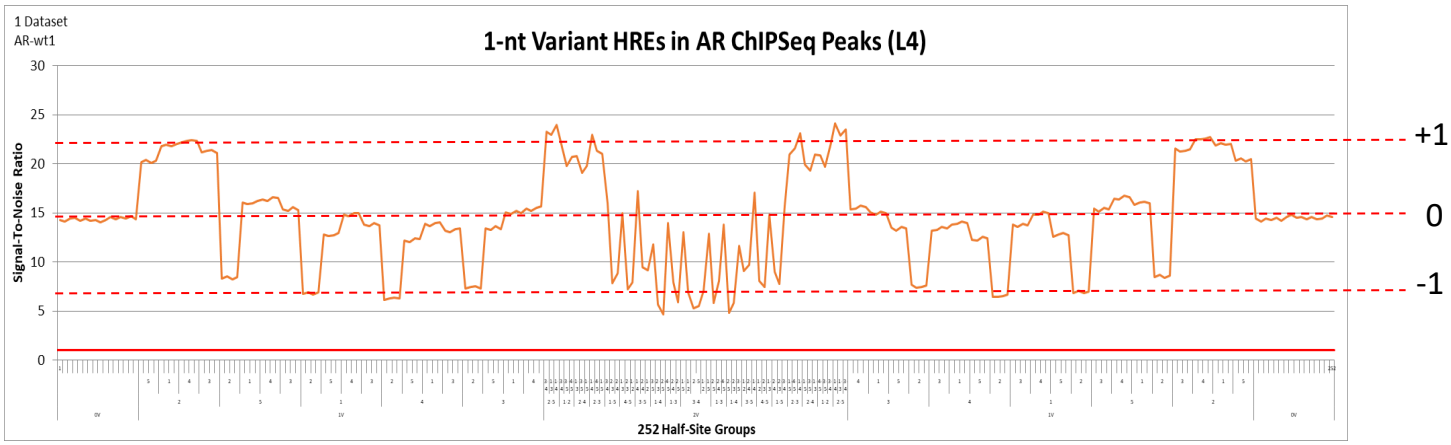
(S/N) Value	# of 252 Half-Site Groups
+2	6
+1	60
0	120
-1	60
-2	6

Two Vacancies

RC Double Occupant Position ID \ ERE	(1,3)-(8,10)	(3,5)-(6,8)	(3,4)-(7,8)	(2,3)-(8,9)	(1,5)-(6,10)	(1,4),(7,10)	(1,2)-(9,10)	(4,5)-(6,7)	(2,5)-(6,9)	(2,4)-(7,9)
(1,3)-(8,10)								-2	-2	-2
(3,5)-(6,8)						0	0			-1
(3,4)-(7,8)					0		0		-1	
(2,3)-(8,9)					0	0		-1		
(1,5)-(6,10)			0	0						-1
(1,4),(7,10)		0		0					-1	
(1,2)-(9,10)		0	0					-1		
(4,5)-(6,7)	+2			+1			+1			
(2,5)-(6,9)	+2		+1			+1				
(2,4)-(7,9)	+2	+1			+1					

Figure S7. Quantify the 3 Discrete States of KR DNA-Binding at 1-nt Variant HREs in the Genome (4,1,0,0,0)

HRE = GAACAnnnTGTTC
1 2 3 4 5 6 7 8 9 10



Three (3) Discrete DNA-Binding Plateaus:

One Vacancy

RC Double Occupant Position ID \ RC Vacancy Position ID	HRE	2-9	5-6	1-10	4-7	3-8
	2-9		-1	-1	-1	-1
	5-6	+1		0	0	0
	1-10	+1	0		0	0
	4-7	+1	0	0		0
	3-8	+1	0	0	0	

(S/N) Value	# of 252 Half-Site Groups
+1	56
0	140
-1	56

Two Vacancies

RC Double Occupant Position ID \ RC Vacancy Position ID	HRE	(2,5)-(6,9)	(1,2)-(9,10)	(2,4)-(7,9)	(2,3)-(8,9)	(1,5)-(6,10)	(4,5)-(6,7)	(3,5)-(6,8)	(1,4),(7,10)	(1,3)-(8,10)	(3,4)-(7,8)
	(2,5)-(6,9)								-1	-1	-1
	(1,2)-(9,10)						-1	-1			-1
	(2,4)-(7,9)					-1		-1		-1	
	(2,3)-(8,9)					-1	-1		-1		
	(1,5)-(6,10)			+1	+1						0
	(4,5)-(6,7)		+1		+1					0	
	(3,5)-(6,8)		+1	+1					0		
	(1,4),(7,10)	+1			+1			0			
	(1,3)-(8,10)	+1		+1			0				
	(3,4)-(7,8)	+1	+1			0					

Figure S8. Algebraic Equations of the 3-State DNA Element (4,1,0,0,0)

Hormone Response Element (HRE) = 3 Discrete States (4,1,0,0,0)

One Vacancy

			RC Vacancy Position ID				
			N_1	N_2	N_3	N_4	N_5
		Rel Charge	A	B	B	B	B
RC Double Occupant Position ID	N_1	A		-X	-X	-X	-X
	N_2	B	+X		0	0	0
	N_3	B	+X	0		0	0
	N_4	B	+X	0	0		0
	N_5	B	+X	0	0	0	

Two Vacancies

			RC Vacancy Position ID									
			N_1-N_2	N_1-N_3	N_1-N_4	N_1-N_5	N_2-N_3	N_2-N_4	N_2-N_5	N_3-N_4	N_3-N_5	N_4-N_5
		Rel Charge	AB	AB	AB	AB	BB	BB	BB	BB	BB	BB
RC Double Occupant Position ID	N_1-N_2	AB								-X	-X	-X
	N_1-N_3	AB						-X	-X			-X
	N_1-N_4	AB					-X		-X		-X	
	N_1-N_5	AB					-X	-X		-X		
	N_2-N_3	BB			+X	+X						0
	N_2-N_4	BB		+X		+X					0	
	N_2-N_5	BB		+X	+X					0		
	N_3-N_4	BB	+X			+X			0			
	N_3-N_5	BB	+X		+X			0				
	N_4-N_5	BB	+X	+X			0					

Algebraic Equations:

A – B = X

B – A = -X

Figure S9. Algebraic Equations of the 5-State DNA Element (3,2,0,0,0)

Estrogen Response Element (ERE) = 5 Discrete States (3,2,0,0,0)

One Vacancy

			RC Vacancy Position ID				
			N_1	N_2	N_3	N_4	N_5
		Rel Charge	A	A	B	B	B
RC Double Occupant Position ID	N_1	A		0	-X	-X	-X
	N_2	A	0		-X	-X	-X
	N_3	B	+X	+X		0	0
	N_4	B	+X	+X	0		0
	N_5	B	+X	+X	0	0	

Algebraic Equations:

A – B = X

B – A = -X

Two Vacancies

			RC Vacancy Position ID									
			N_1-N_2	N_1-N_3	N_1-N_4	N_1-N_5	N_2-N_3	N_2-N_4	N_2-N_5	N_3-N_4	N_3-N_5	N_4-N_5
		Rel Charge	AA	AB	AB	AB	AB	AB	AB	BB	BB	BB
RC Double Occupant Position ID	N_1-N_2	AA								-2X	-2X	-2X
	N_1-N_3	AB						0	0			-X
	N_1-N_4	AB					0		0		-X	
	N_1-N_5	AB					0	0		-X		
	N_2-N_3	AB			0	0						-X
	N_2-N_4	AB		0		0					-X	
	N_2-N_5	AB		0	0					-X		
	N_3-N_4	BB	+2X			+X			+X			
	N_3-N_5	BB	+2X		+X			+X				
	N_4-N_5	BB	+2X	+X			+X					

Algebraic Equations:

AA – BB = 2X

BB – AA = -2X

AB – BB = A – B = X

BB – AB = B – A = -X

Figure S10. Algebraic Equations of the 9-State DNA Element (3,1,1,0,0)

9 Discrete States (3,1,1,0,0)

One Vacancy

			RC Vacancy Position ID				
			N_1	N_2	N_3	N_4	N_5
		Rel Charge	A	B	C	C	C
RC Double Occupant Position ID	N_1	A		-X	-Z	-Z	-Z
	N_2	B	+X		-Y	-Y	-Y
	N_3	C	+Z	+Y		0	0
	N_4	C	+Z	+Y	0		0
	N_5	C	+Z	+Y	0	0	

Algebraic Equations:

A – B = X

B – C = Y

A – C = Z

Two Vacancies

			RC Vacancy Position ID									
			N_1-N_2	N_1-N_3	N_1-N_4	N_1-N_5	N_2-N_3	N_2-N_4	N_2-N_5	N_3-N_4	N_3-N_5	N_4-N_5
		Rel Charge	AB	AC	AC	AC	BC	BC	BC	CC	CC	CC
RC Double Occupant Position ID	N_1-N_2	AB								-(Y+Z)	-(Y+Z)	-(Y+Z)
	N_1-N_3	AC						-X	-X			-Z
	N_1-N_4	AC					-X		-X		-Z	
	N_1-N_5	AC					-X	-X		-Z		
	N_2-N_3	BC			+X	+X						-Y
	N_2-N_4	BC		+X		+X					-Y	
	N_2-N_5	BC		+X	+X					-Y		
	N_3-N_4	CC	Y+Z			+Z			+Y			
	N_3-N_5	CC	Y+Z		+Z			+Y				
	N_4-N_5	CC	Y+Z	+Z			+Y					

Algebraic Equations:

AB – CC = Y + Z

AC – BC = A – B = X

BC – AC = B – A = -X

AC – CC = A – C = Z

CC – AC = C –A = -Z

BC – CC = B – C = Y

CC – BC = C-B = -Y

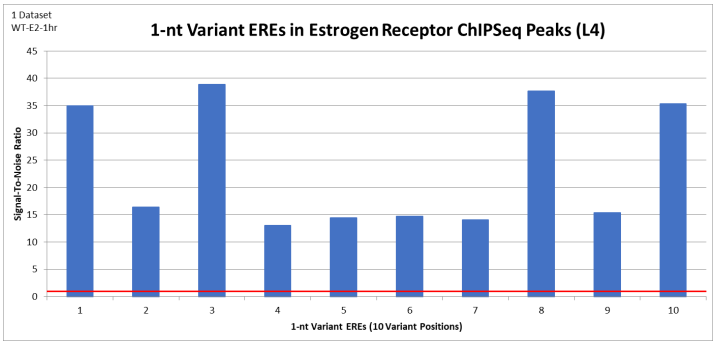
CC – AB = -(Y+Z)

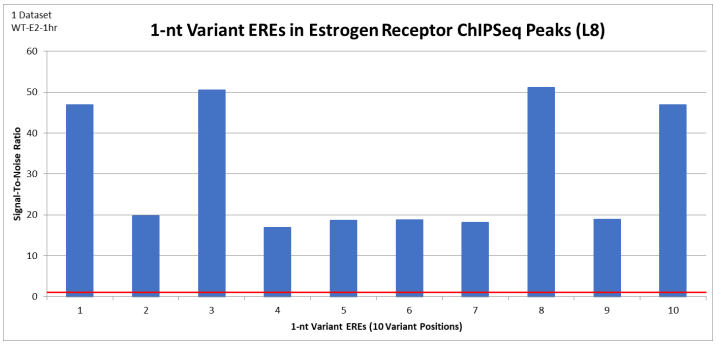
Figure S11. Algebraic Equations of the 11-State, 25-State, 51-State DNA Elements (2,2,1,0,0) (2,1,1,1,0) (1,1,1,1,1)

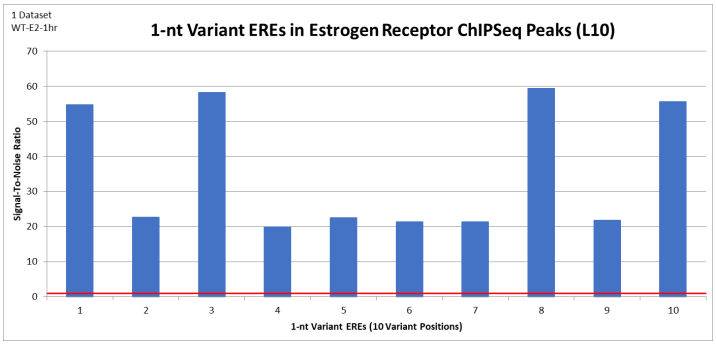
All Possible Discrete States For a 5-nt DNA Element in the Genome
(a 5-nt DNA element, followed by any arbitrary spacer, followed by its 5-nt reverse-complement DNA element)

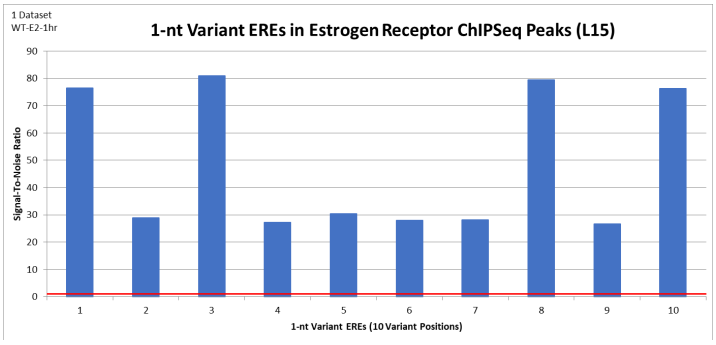
Independent Variables	DNA Element	Position N_1	Position N_2	Position N_3	Position N_4	Position N_5	Total	Plateaus
1		5	0	0	0	0	5	1
2	HRE	4	1				5	3
	ERE	3	2				5	5
3		3	1	1			5	9
		2	2	1			5	11
4		2	1	1	1		5	25
5		1	1	1	1	1	5	51

Figure S12. (S/N) analysis of 1-nt Variant EREs in ER ChIPSeq Peaks (Variant Position)









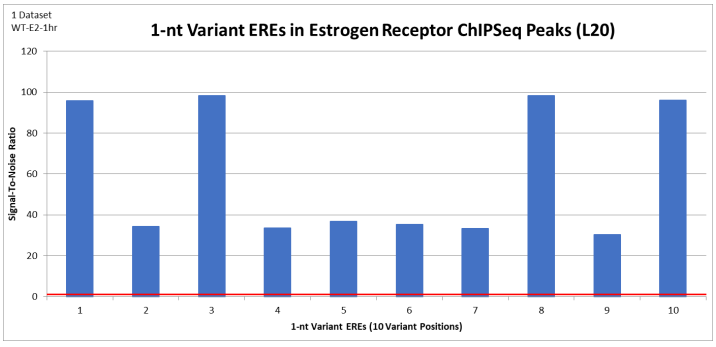
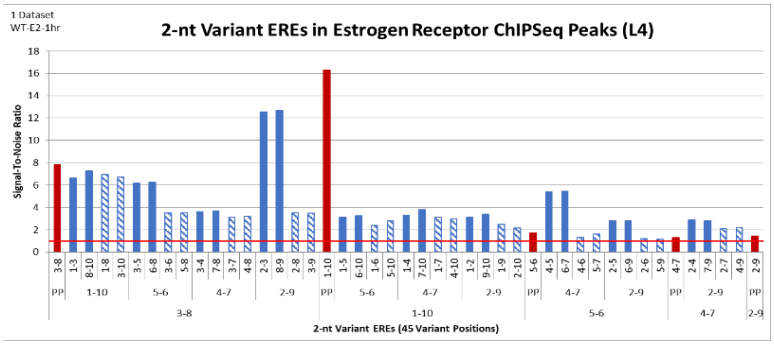
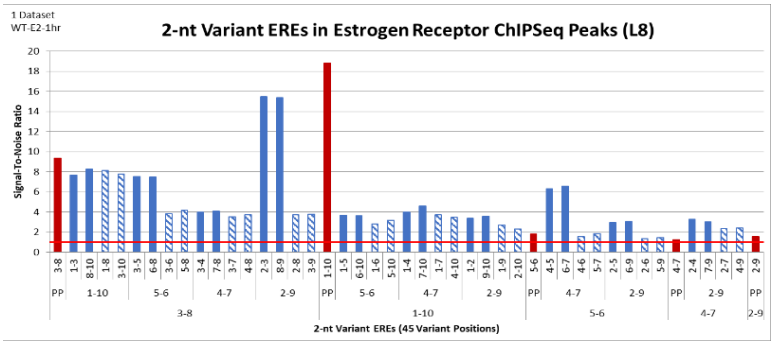
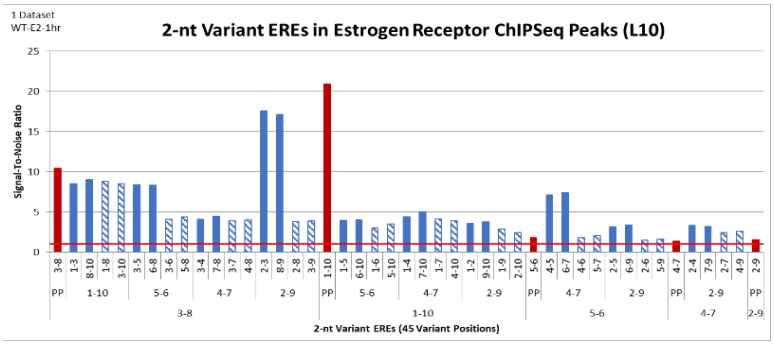
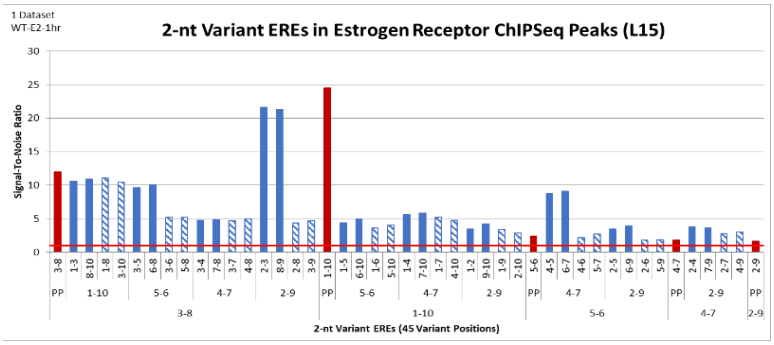


Figure S13. (S/N) analysis of 2-nt Variant EREs in ER ChIPSeq Peaks (Variant Position)









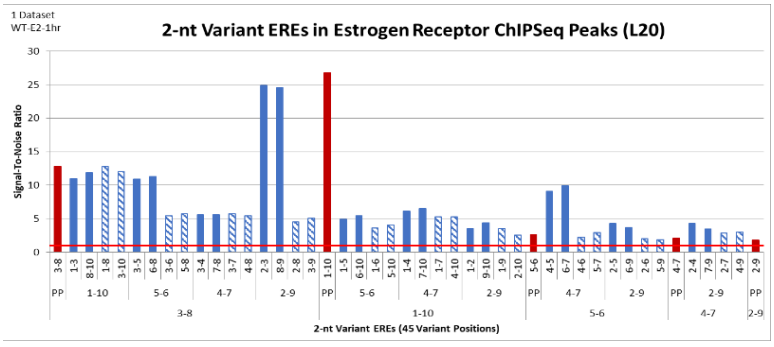
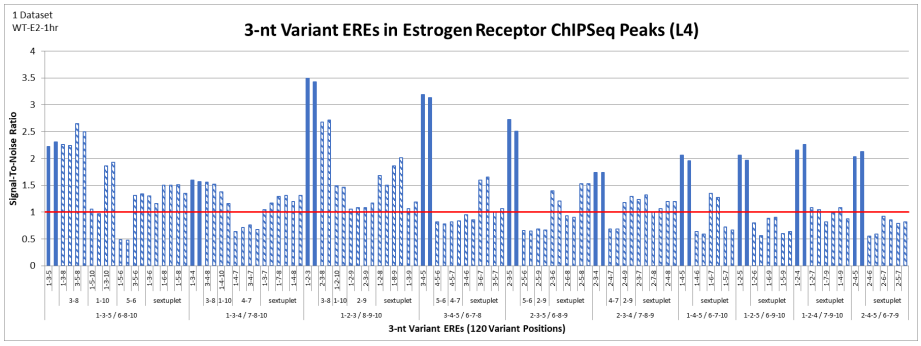
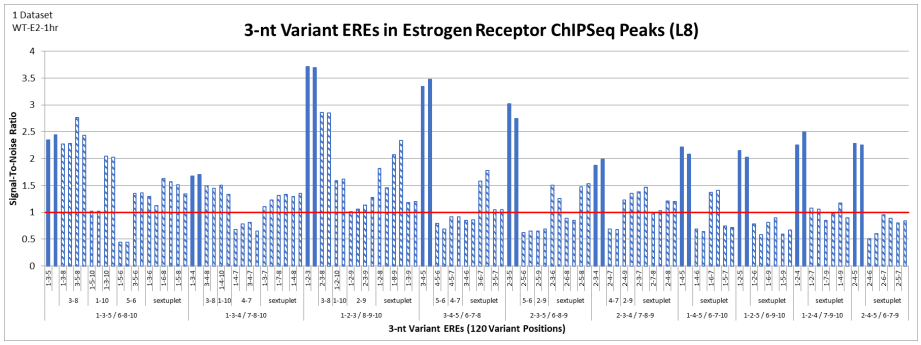
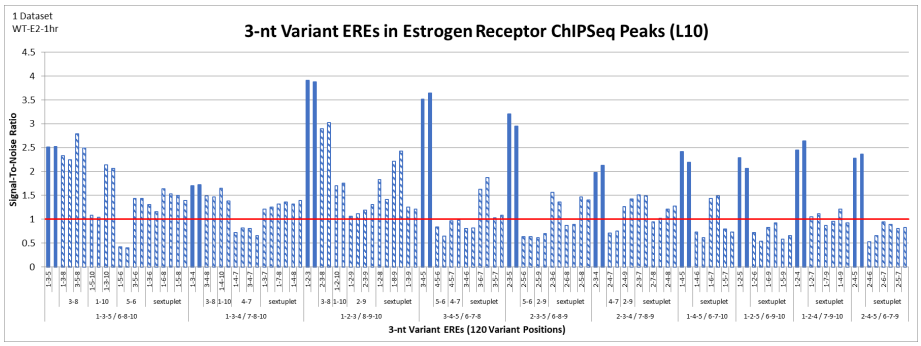
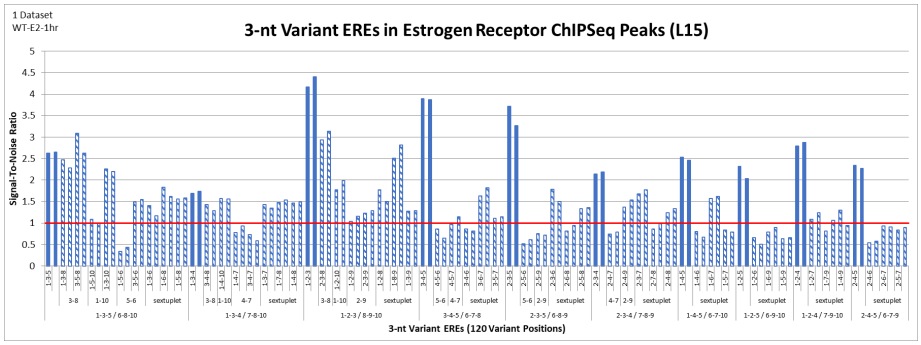


Figure S14. (S/N) analysis of 3-nt Variant EREs in ER ChIPSeq Peaks (Variant Position)









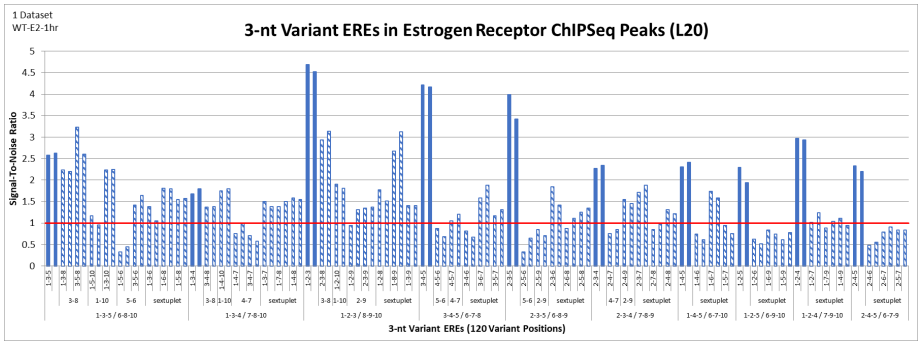
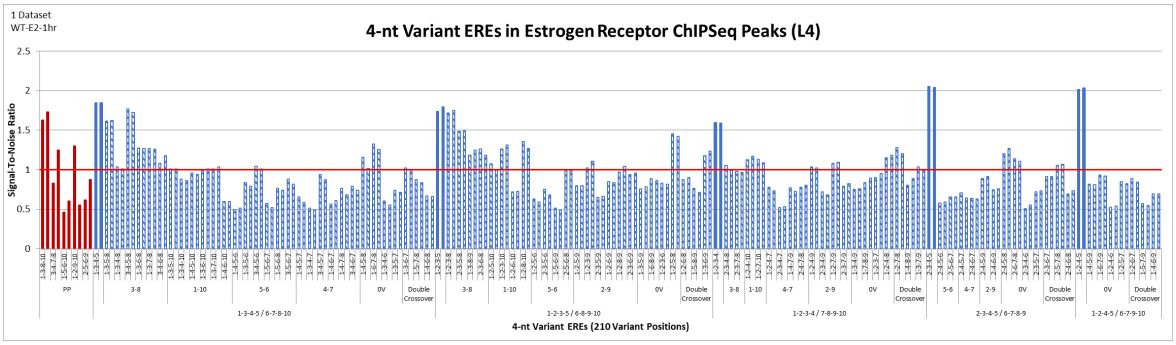
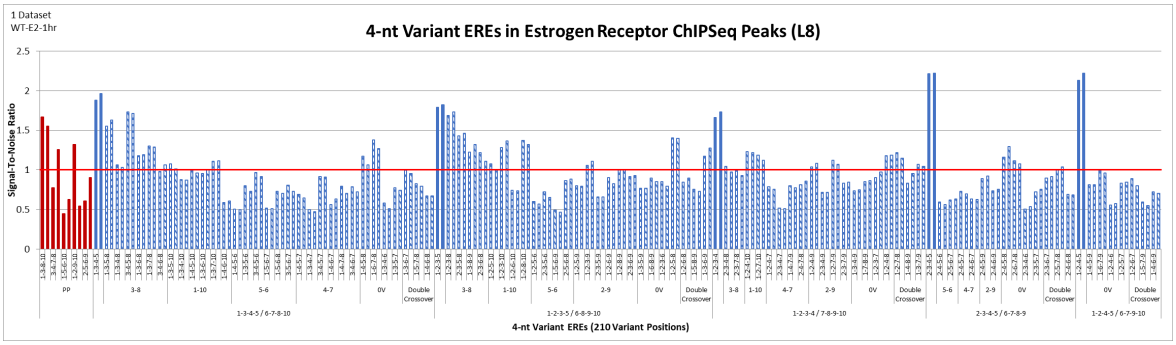
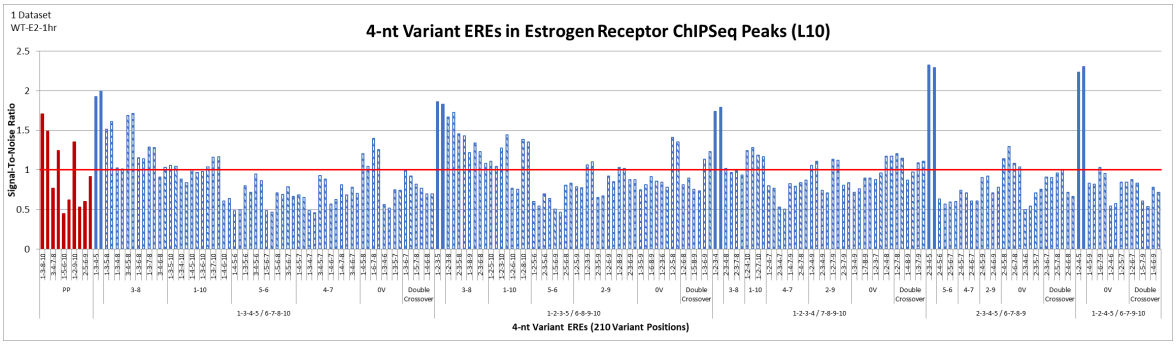
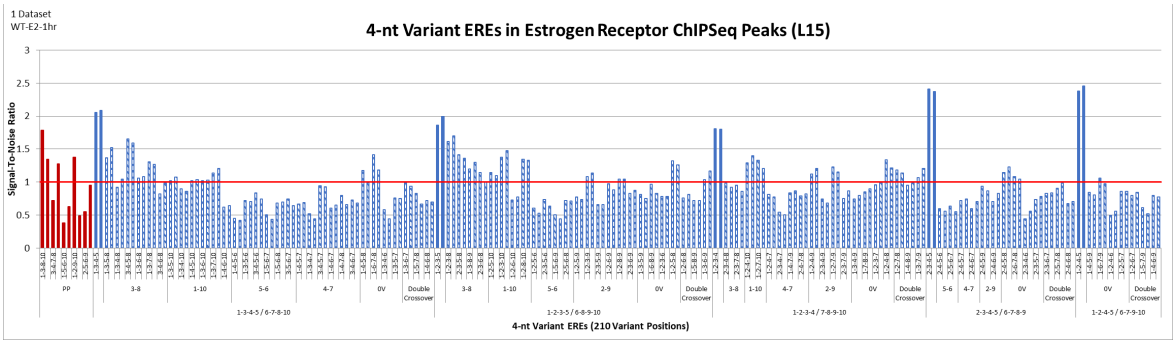


Figure S15. (S/N) analysis of 4-nt Variant EREs in ER ChIPSeq Peaks (Variant Position)









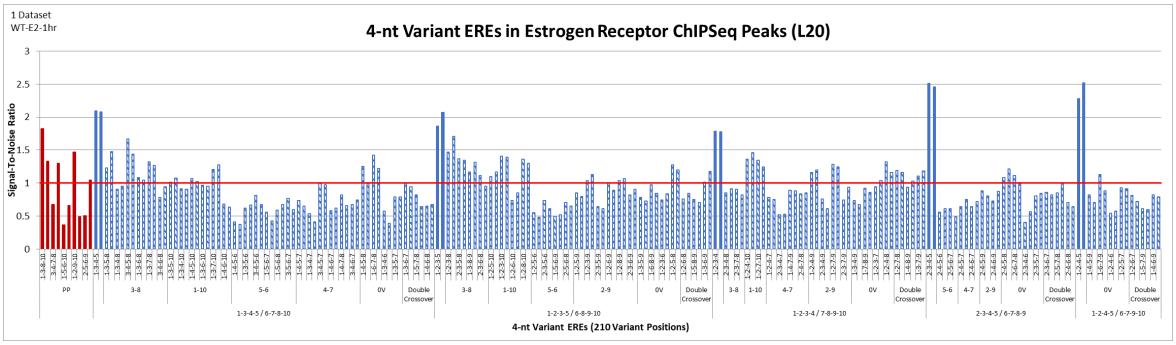
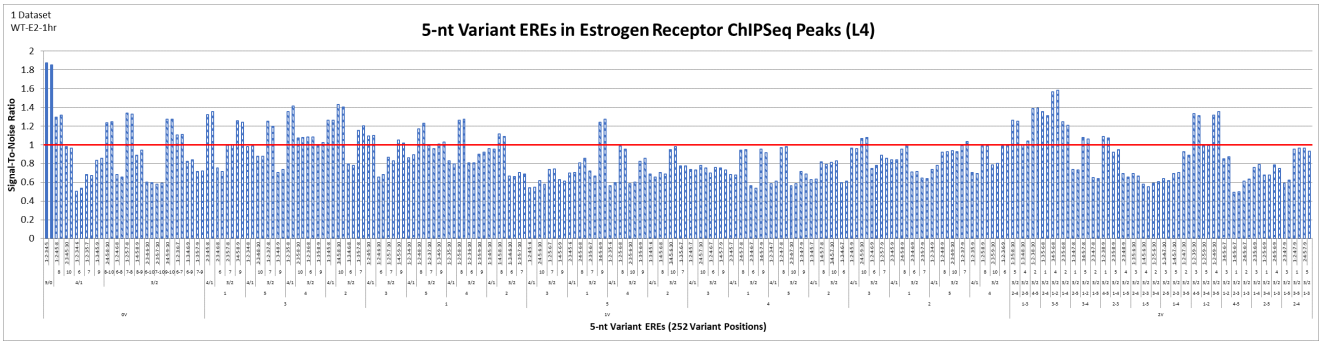
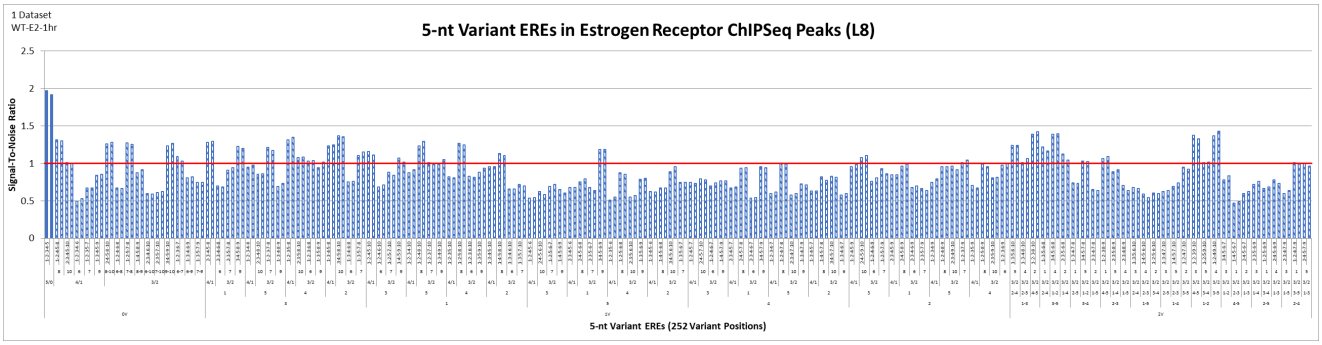
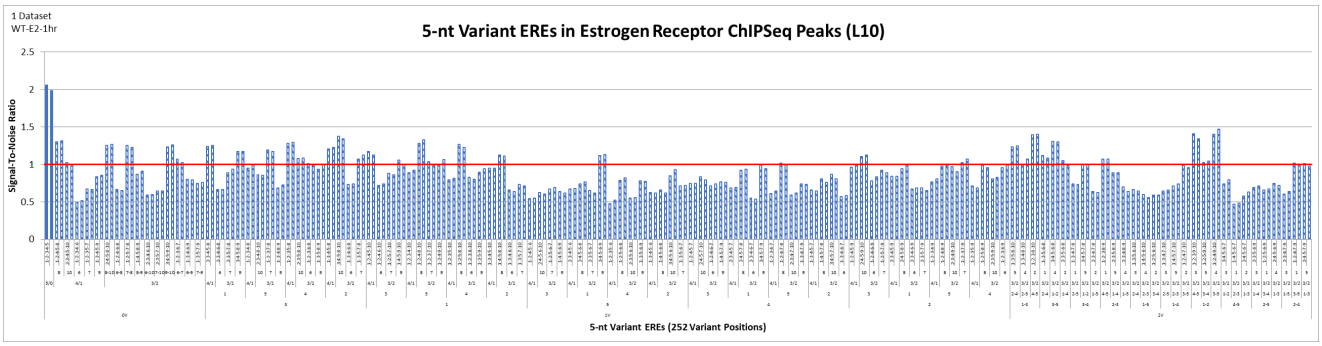
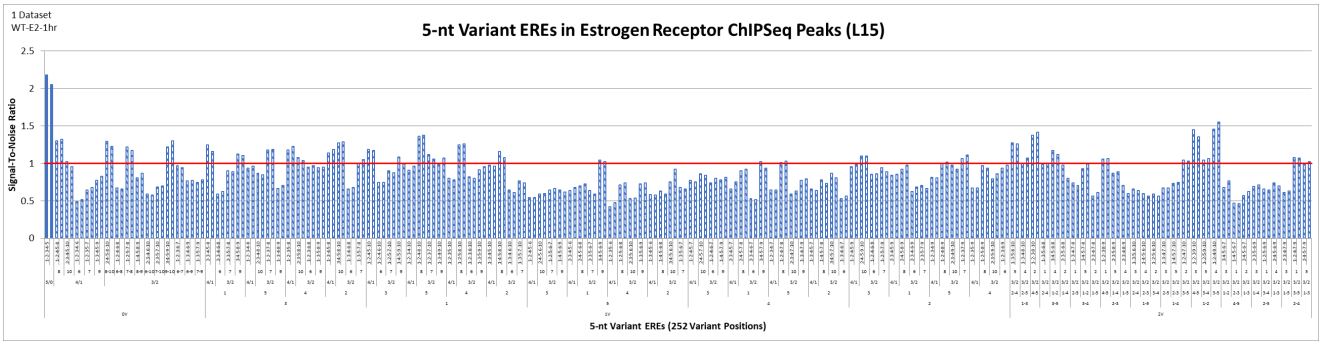


Figure S16. (S/N) analysis of 5-nt Variant EREs in ER ChIPSeq Peaks (Variant Position)









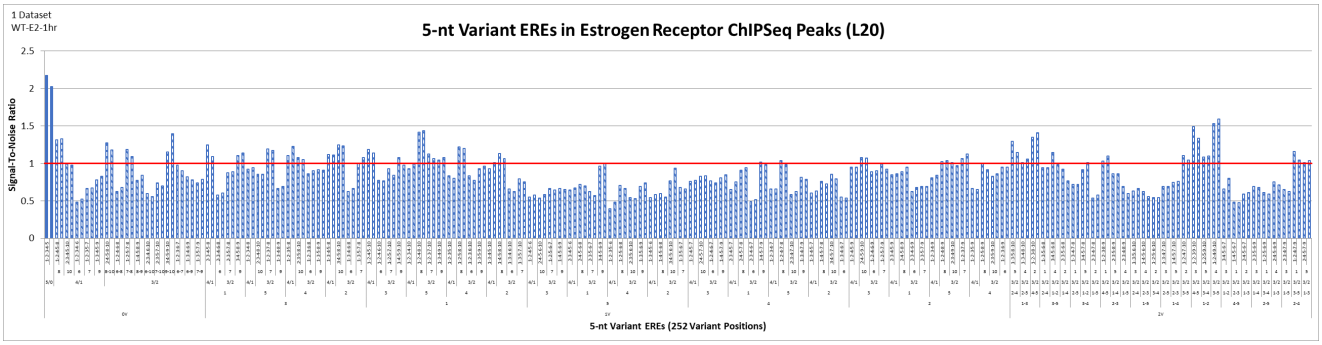
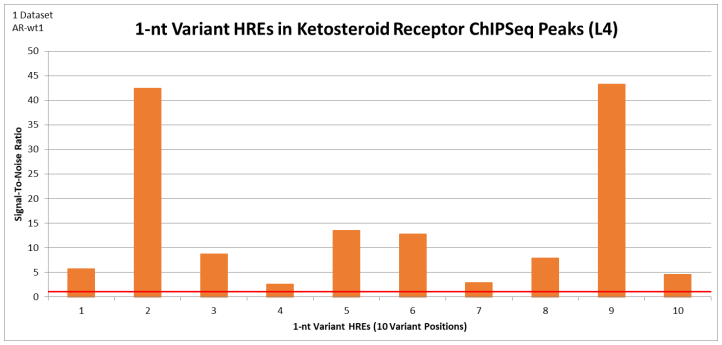
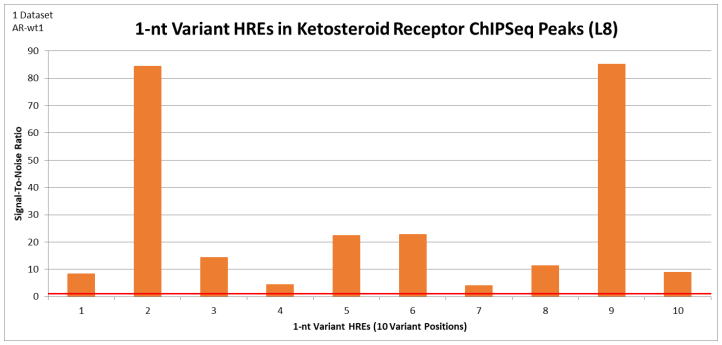
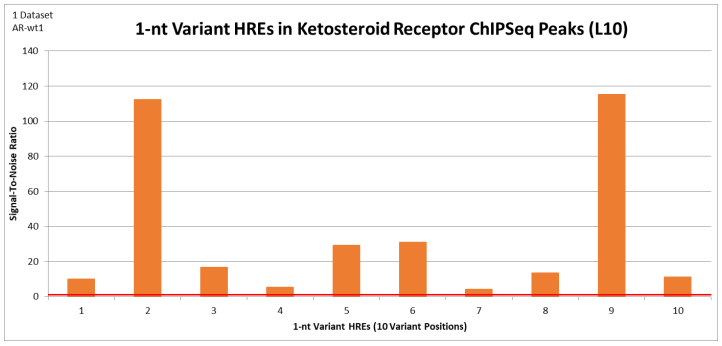
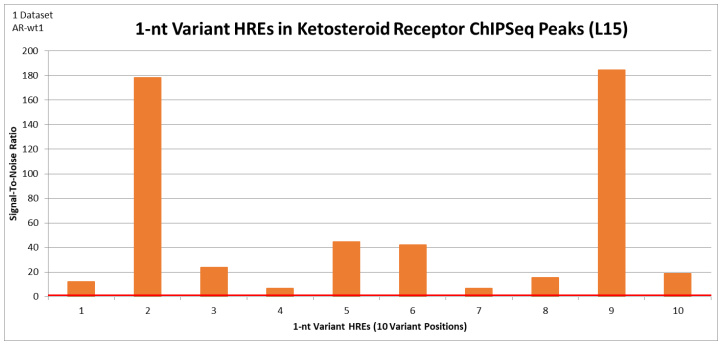


Figure S17. (S/N) analysis of 1-nt Variant HREs in KR ChIPSeq Peaks (Variant Position)









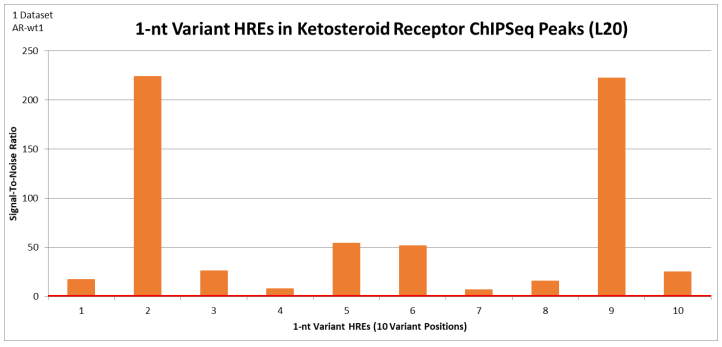
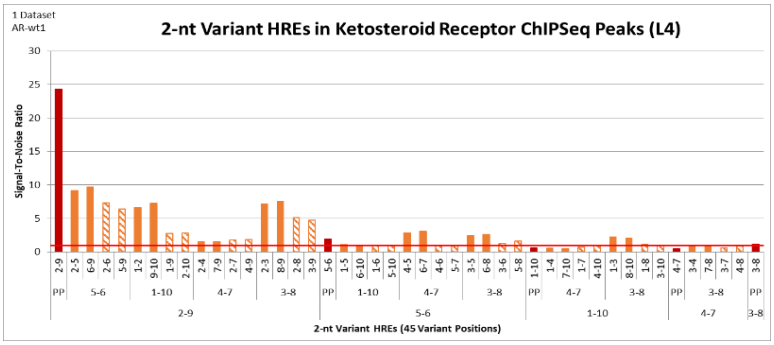
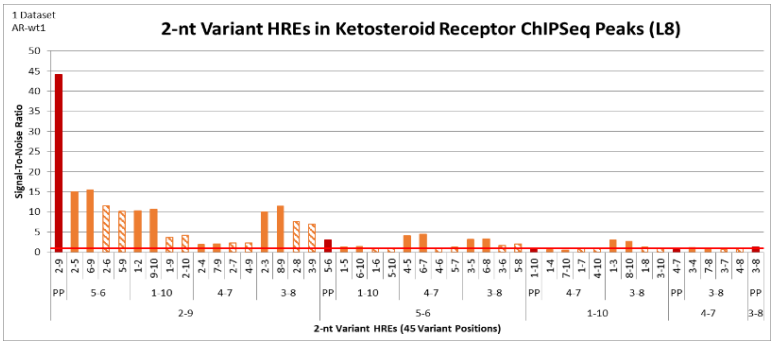
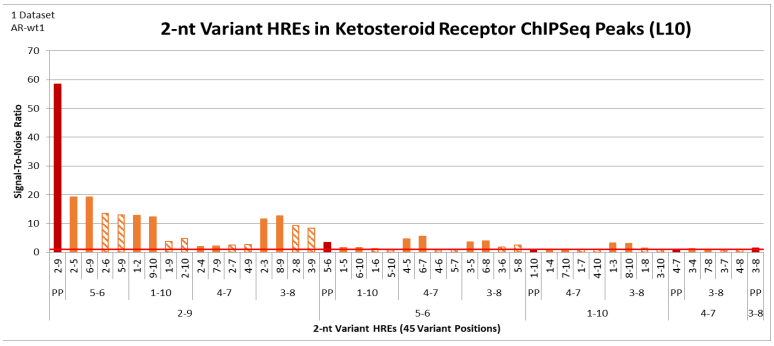
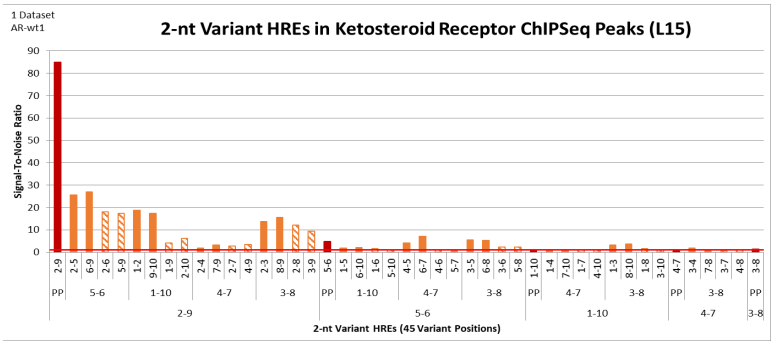


Figure S18. (S/N) analysis of 2-nt Variant HREs in KR ChIPSeq Peaks (Variant Position)









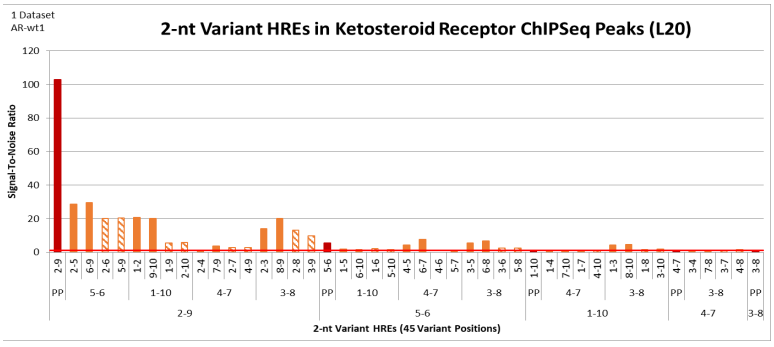
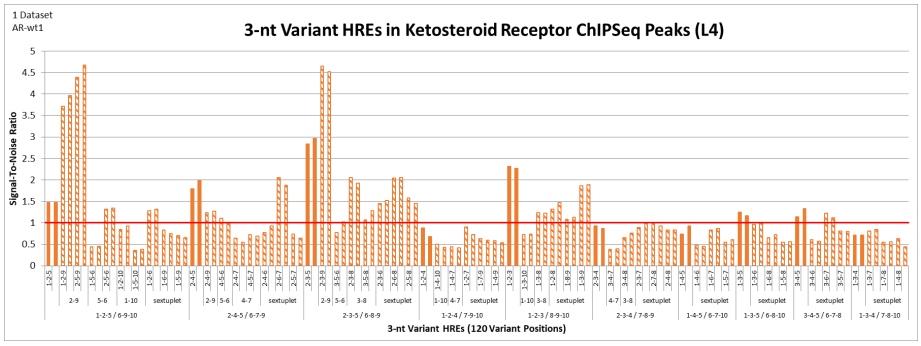
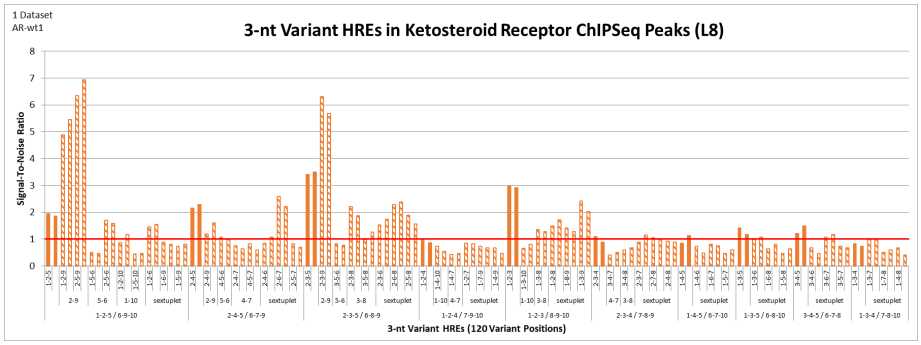
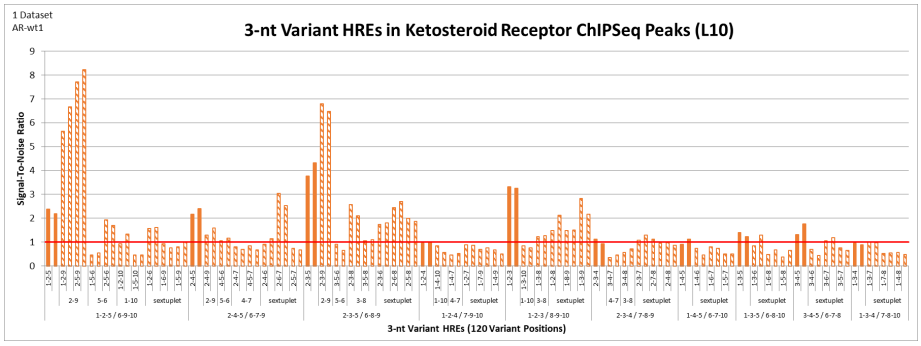
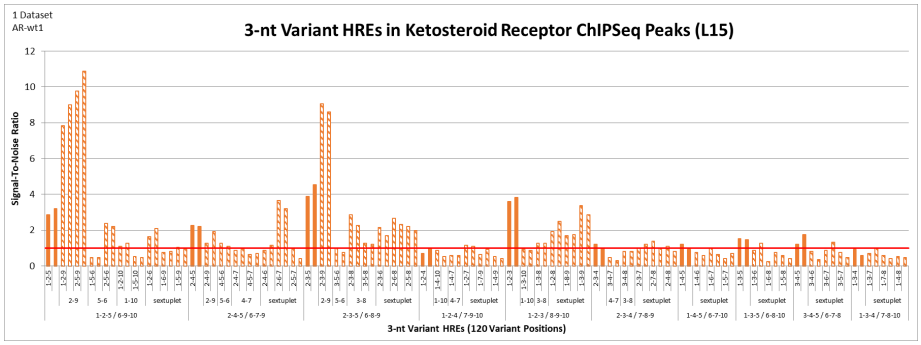


Figure S19. (S/N) analysis of 3-nt Variant HREs in KR ChIPSeq Peaks (Variant Position)









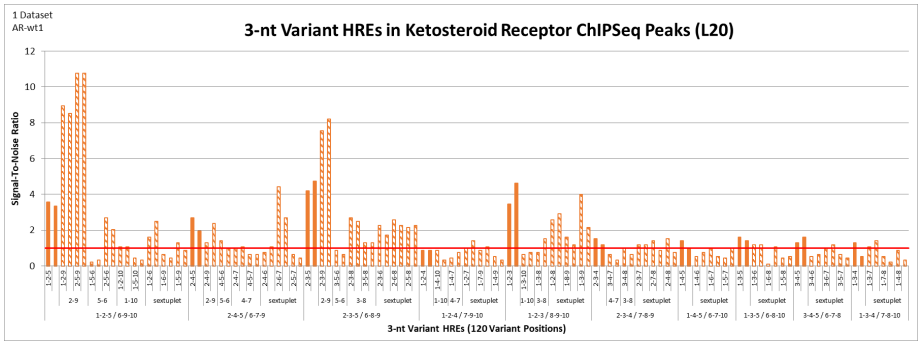
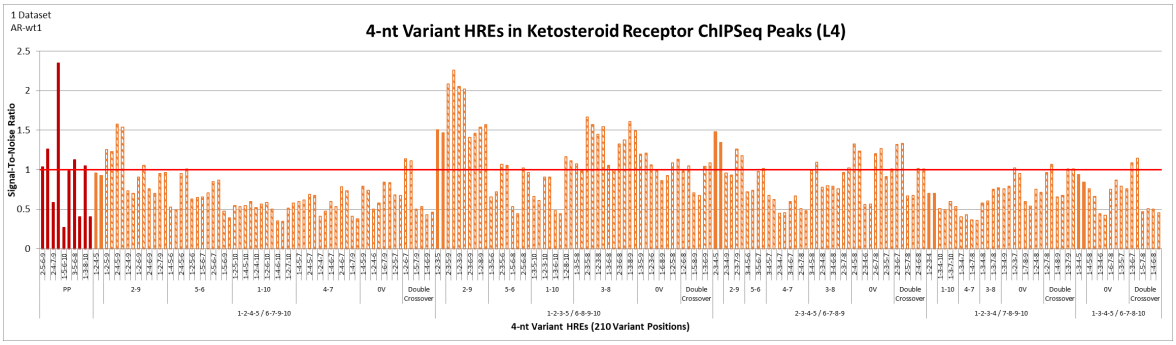
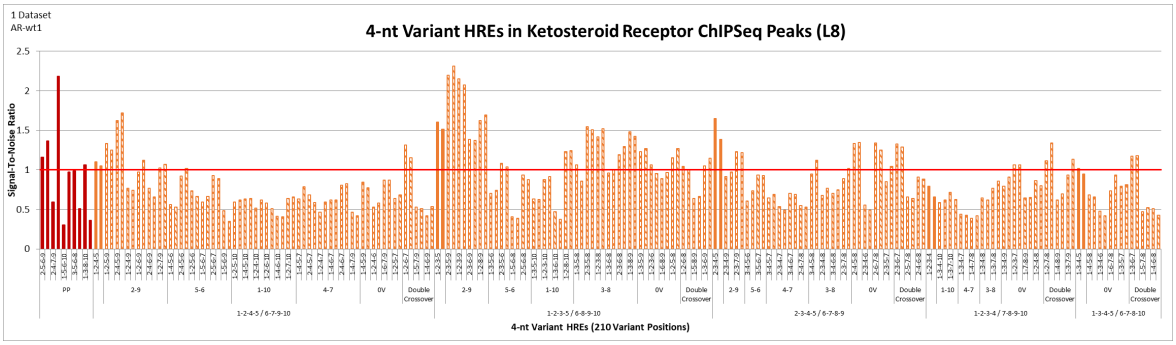
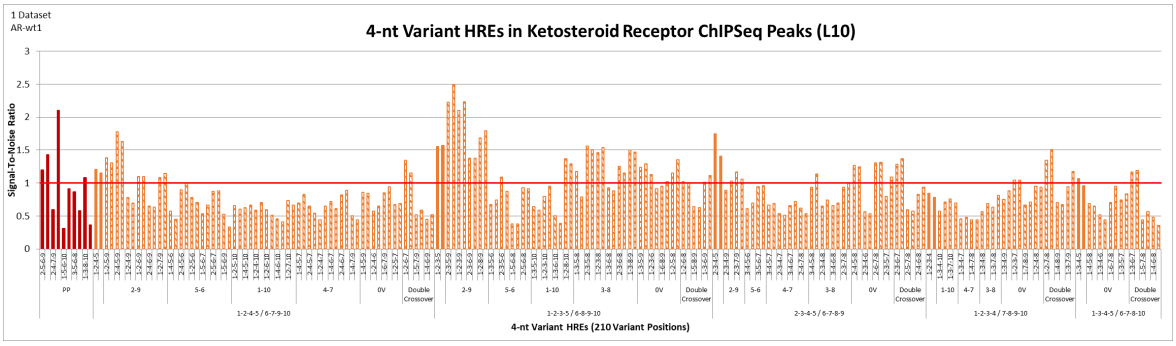
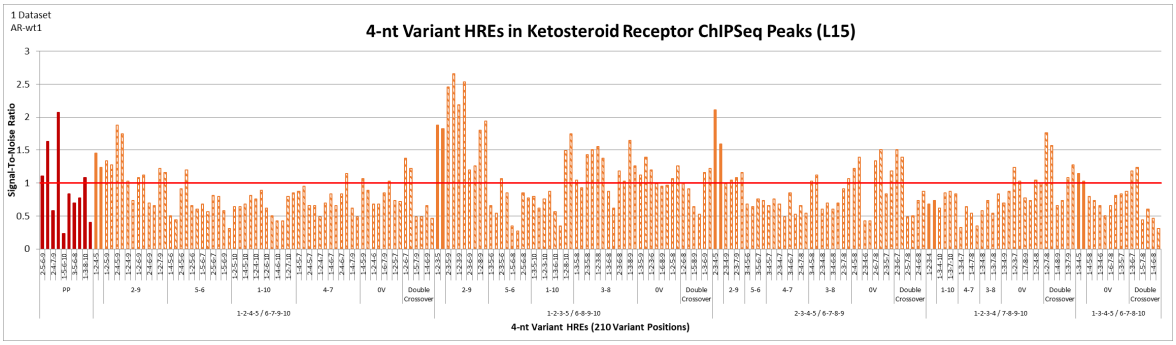


Figure S20. (S/N) analysis of 4-nt Variant HREs in KR ChIPSeq Peaks (Variant Position)









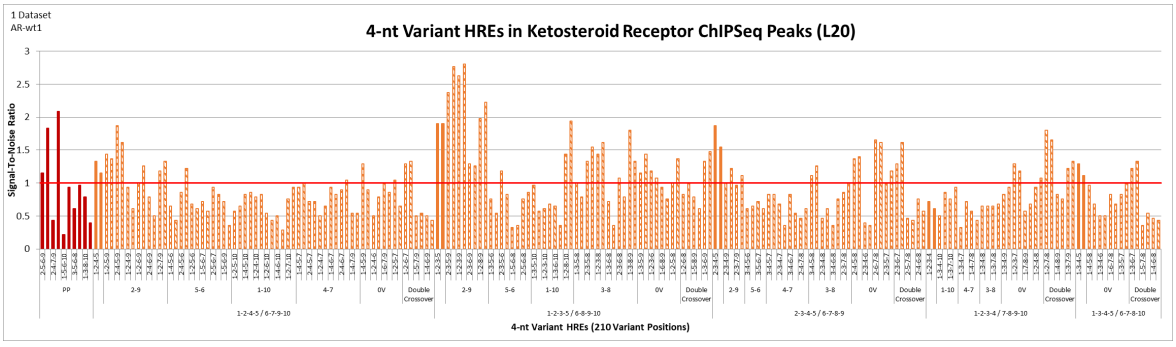
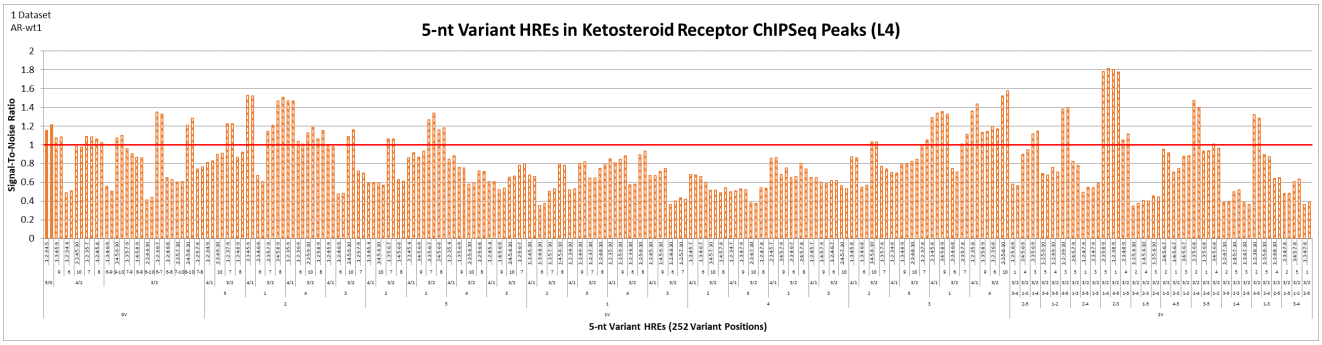
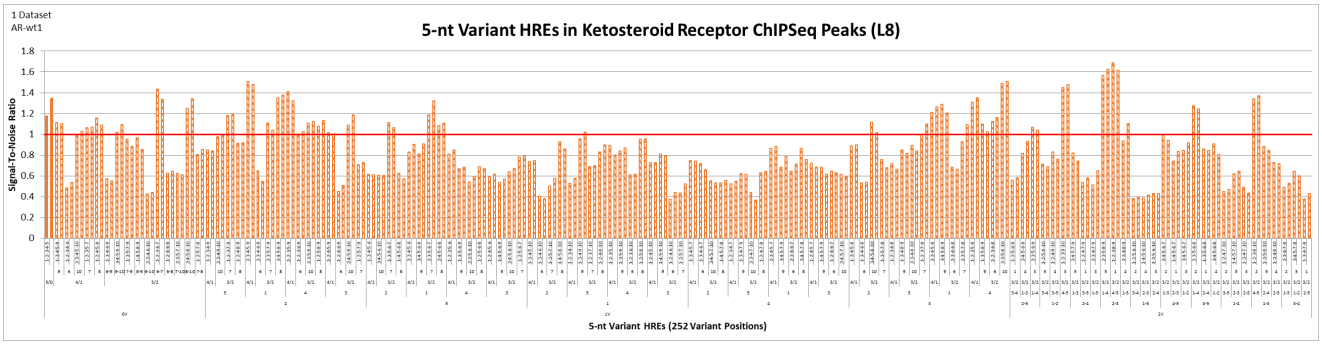
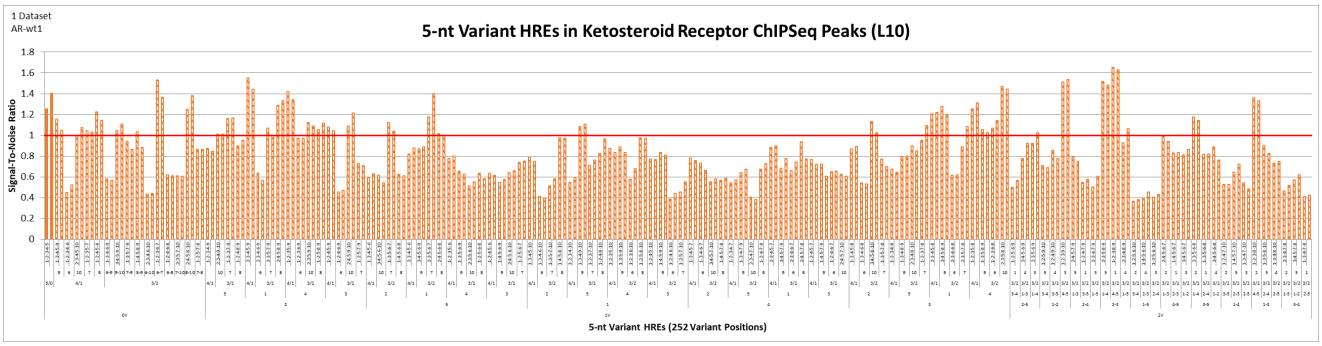
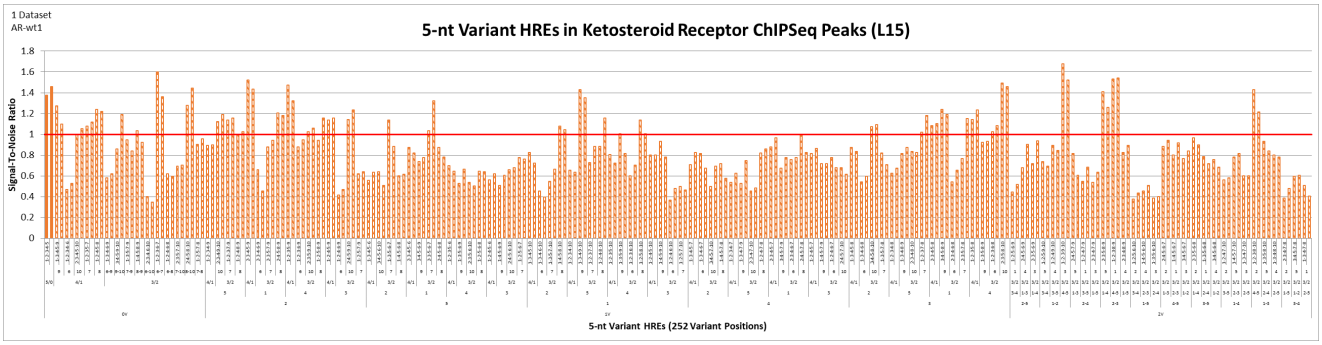


Figure S21. (S/N) analysis of 5-nt Variant HREs in KR ChIPSeq Peaks (Variant Position)









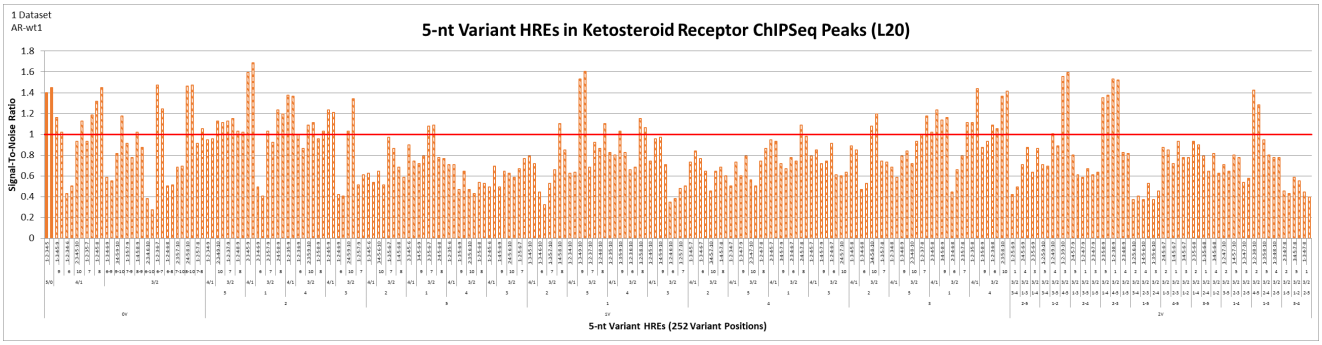
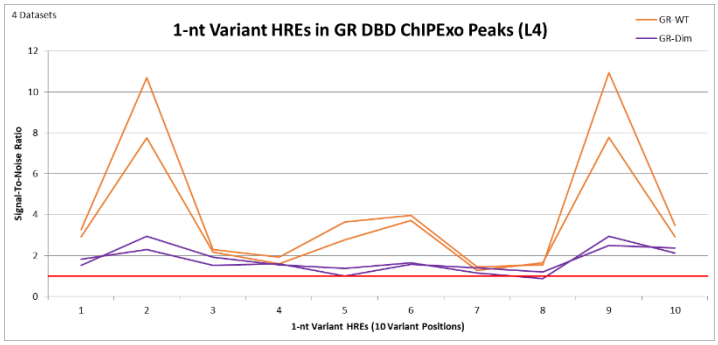


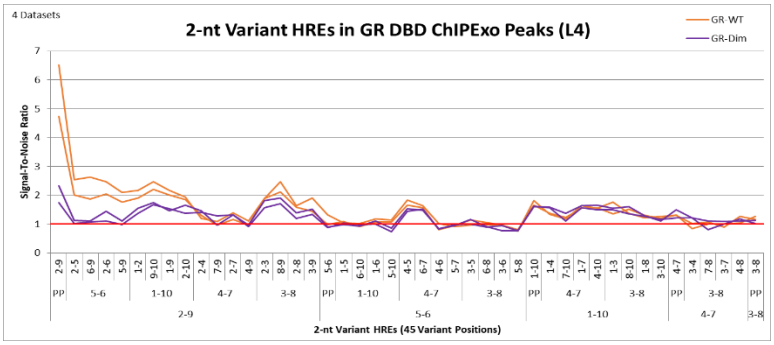
Figure S22. (S/N) analysis of 0-nt to 5-nt Variant HREs in GR and GR-Dim ChIPSeq Peaks #2 (252 Half-Site Groups)

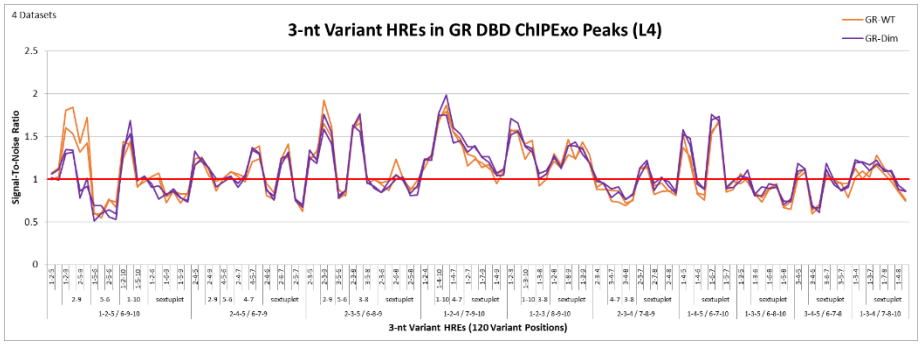
Figure S23. (S/N) analysis of 0-nt to 5-nt Variant HREs in GR and GR-Dim ChIPExo Peaks #1 (252 Half-Site Groups)

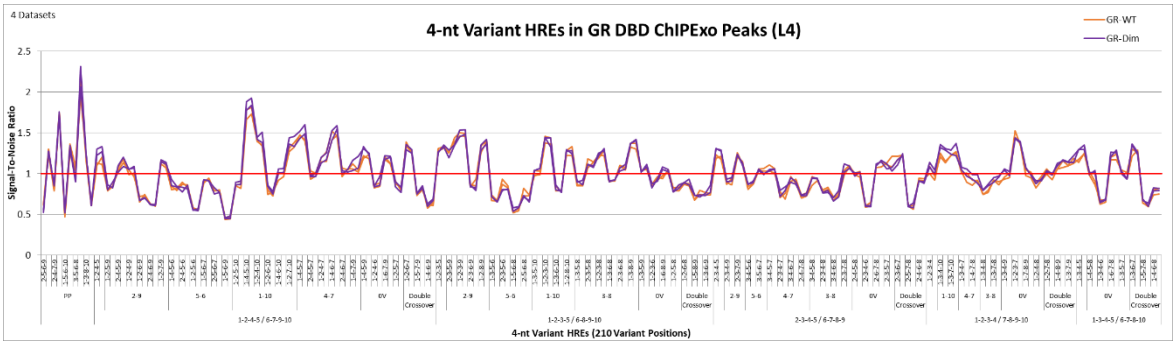
Figure S24. (S/N) analysis of 0-nt to 5-nt Variant HREs in GR and GR-Dim ChIPExo Peaks #2 (252 Half-Site Groups)

Figure S25. (S/N) analysis of 1-nt to 5-nt Variant HREs in GR and GR-Dim ChIPExo Peaks (Variant Position)









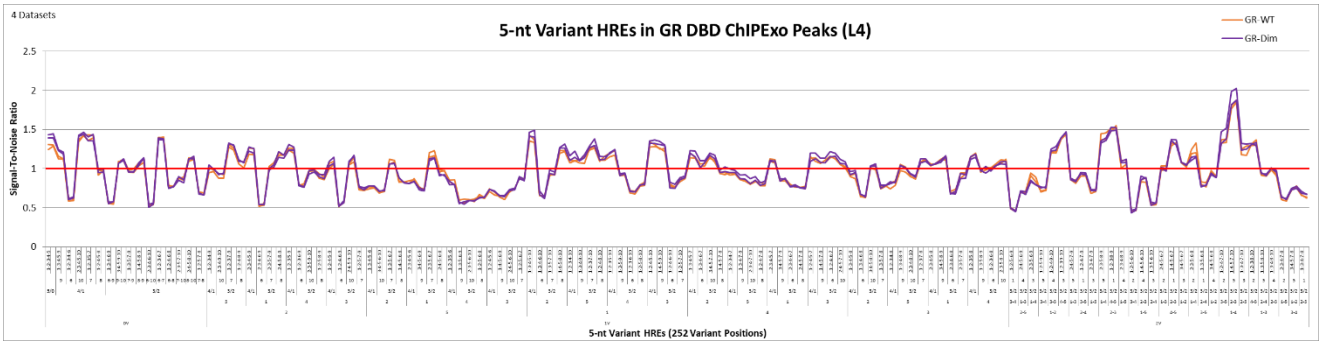


Figure S26-S32 Descriptions: Inversion Symmetry Detection Methodology (Position Weight Matrices)

****All detailed data and statistics associated with every figure are compiled in Data S1****

Scoring DNA sequences against position weight matrices (PWMs) is a widely adopted method to identify TF *cis*-regulatory DNA elements in ChIPSeq or ChIPExo experiments. PWMs are used to scan a DNA sequence for the presence of DNA sequences that are significantly more similar to the PWMs than to the background (1). To assess the resolution capabilities of PWMs-based DNA motif identification, we grouped the ChIPSeq or ChIPExo peaks from each experiment by the number of variants each ERE or HRE DNA element within the peak contained (i.e., peaks that contain a 0-nt variant consensus palindromic DNA element, 1-nt variant DNA element, 2-nt variant DNA element, 3-nt variant DNA element, 4-nt variant DNA element, or 5-nt variant DNA element) and ran these grouped peaks against PWMs. Because peaks can contain multiple ERE or HRE DNA elements, we then assigned the peaks to the ERE or HRE DNA element with the least number of variants relative to the 0-nt consensus palindromic DNA element, thus defining each peak by a single ERE or HRE DNA element (i.e., unique).

Using WT-E2-1hr.L4 [76,163 peaks, 146-nt peak length] as an example experiment, 1,201 peaks contain a 0-nt variant consensus palindromic ERE DNA element, 7,349 peaks contain a 1-nt variant ERE DNA element, 16,368 peaks contain a 2-nt variant ERE DNA element, 34,605 peaks contain a 3-nt variant ERE DNA element, 68,484 peaks contain a 4-nt variant ERE DNA element, and 76,147 peaks contain a 5-nt variant ERE DNA element. PWMs identified the ERE motif (VAGGTCACNSTGACC) in 1,201 (100%) of the peaks that contain 0-nt variant consensus palindromic ERE DNA element, 7,170 (98%) of the peaks that contain a 1-nt variant ERE DNA element, 8,494 (52%) of the peaks that contain a 2-nt variant ERE DNA element, 7,107 (21%) of the peaks that contain a 3-nt variant ERE DNA element, 14,846 (22%) of the peaks that contain a 4-nt variant ERE DNA element, 16,816 (22%) of the peaks that contain a 5-nt variant ERE DNA element, and 16,820 (22%) of all the peaks in the experiment (76,163). We then assigned the peaks to the ERE DNA element with the least number of variants relative to the 0-nt consensus palindromic DNA element, thus defining each peak by a single ERE or HRE DNA element (i.e., unique): 1,201 peaks contain a 0-nt variant consensus palindromic ERE DNA element, 7,332 peaks contain a 1-nt variant ERE DNA element, 15,490 peaks contain a 2-nt variant ERE DNA element, 24,347 peaks contain a 3-nt variant ERE DNA element, 25,194 peaks contain a 4-nt variant ERE DNA element, and 2,599 peaks contains a 5-nt variant ERE DNA element. PWMs identified the ERE motif (VAGGTCACNSTGACC) in 1,201 (100%) of the peaks that contain 0-nt variant consensus palindromic ERE DNA element, 7,153 (98%) of the peaks that contain a 1-nt variant ERE DNA element, 7,631 (49%) of the peaks that contain a 2-nt variant ERE DNA element, 582 (2%) of the peaks that contain a 3-nt variant ERE DNA element, 227 (1%) of the peaks that contain a 4-nt variant ERE DNA element, and 26 (1%) of the peaks that contain a 5-nt variant ERE DNA element. Thus, PWMs identified the ERE motif in the majority of peaks that contain a 0-nt or a 1-nt variant ERE DNA element, about half of the peaks that contain a 2-nt variant ERE DNA element, and did not identify the ERE motif in peaks that contain a 3-nt, 4-nt, or 5-nt variant ERE DNA element. Therefore, PWMs analyses would lead one to conclude that the majority of ER DNA-binding events in the genome are being driven by mechanisms other than the ERE sequence since peaks that contain a 0-nt variant consensus palindromic ERE DNA element or 1-nt variant ERE DNA element constitute 11% of all ER DNA-binding events in the experiment, a proportion observed in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) (2).

Figure S26. PWMs Detection of the ERE DNA Motif in ER ChIPSeq Peaks
(A) Of 157 ER experiments, on average, PWMs identified the ERE motif (VAGGTCACNSTGACC) in 99%-99% (L4-L20) of the peaks that contain 0-nt variant consensus palindromic ERE DNA element, 95%-97% (L4-L20) of the peaks that contain a 1-nt variant ERE DNA element, 50%-58% (L4-L20) of the peaks that contain a 2-nt variant ERE DNA element, 26%-38% (L4-L20) of the peaks that contain a 3-nt variant ERE DNA element, 27%-41% (L4-L20) of the peaks that contain a 4-nt variant ERE DNA element, and 26%-41% (L4-L20) of the peaks that contain a 5-nt variant ERE DNA element. **(B)** Because peaks can contain multiple ERE DNA elements, we then assigned peaks to the ERE DNA element with the least number of variants relative to the 0-nt consensus palindromic ERE, thus defining each peak by a single ERE DNA element (i.e., unique). Now, PWMs identified the ERE motif in 99%-99% (L4-L20) of the peaks that contain 0-nt variant consensus palindromic ERE DNA element, 95%-97% (L4-L20) of the peaks that contain a 1-nt variant ERE DNA element, 46%-52% (L4-L20) of the peaks that contain a 2-nt variant ERE DNA element, 3%-4% (L4-L20) of the peaks that contain a 3-nt variant ERE DNA element, 1%-1% (L4-L20) of the peaks that contain a 4-nt variant ERE DNA element, and 1%-1% (L4-L20) of the peaks that contain a 5-nt variant ERE DNA element.

Figure S27. PWMs Detection of the ARE DNA Motif in KR ChIPSeq Peaks
(A) Of 194 KR experiments, on average, PWMs identified the ARE motif (RGRACASNSTGTYCYB) in 96%-96% (L4-L20) of the peaks that contain 0-nt variant consensus palindromic HRE DNA element, 76%-81% (L4-L20) of the peaks that contain a 1-nt variant HRE DNA element, 41%-50% (L4-L20) of the peaks that contain a 2-nt variant HRE DNA element, 19%-29% (L4-L20) of the peaks that contain a 3-nt variant HRE DNA element, 18%-29% (L4-L20) of the peaks that contain a 4-nt variant HRE DNA element, and 18%-30% (L4-L20) of the peaks that contain a 5-nt variant HRE DNA element. **(B)** Because peaks can contain multiple HRE DNA elements, we then assigned peaks to the HRE DNA element with the least number of variants relative to the 0-nt consensus palindromic HRE, thus defining each peak by a single HRE DNA element (i.e., unique). Now, PWMs identified the ARE motif in 96%-96% (L4-L20) of the peaks that contain 0-nt variant consensus palindromic HRE DNA element, 76%-81% (L4-L20) of the peaks that contain a 1-nt variant HRE DNA element, 39%-47% (L4-L20) of the peaks that contain a 2-nt variant HRE DNA element, 9%-

14% (L4-L20) of the peaks that contain a 3-nt variant HRE DNA element, 1%-2% (L4-L20) of the peaks that contain a 4-nt variant HRE DNA element, and 1%-2% (L4-L20) of the peaks that contain a 5-nt variant HRE DNA element.

Figure S28. PWMs Detection of the GRE DNA Motif in KR ChIPSeq Peaks
(A) Of 194 KR experiments, on average, PWMs identified the GRE motif (NRGVACABNVTGTYCY) in 96%-97% (L4-L20) of the peaks that contain 0-nt variant consensus palindromic HRE DNA element, 64%-72% (L4-L20) of the peaks that contain a 1-nt variant HRE DNA element, 31%-39% (L4-L20) of the peaks that contain a 2-nt variant HRE DNA element, 14%-24% (L4-L20) of the peaks that contain a 3-nt variant HRE DNA element, 14%-24% (L4-L20) of the peaks that contain a 4-nt variant HRE DNA element, and 14%-24% (L4-L20) of the peaks that contain a 5-nt variant HRE DNA element. **(B)** Because peaks can contain multiple HRE DNA elements, we then assigned peaks to the HRE DNA element with the least number of variants relative to the 0-nt consensus palindromic HRE, thus defining each peak by a single HRE DNA element (i.e., unique). Now, PWMs identified the GRE motif in 96%-97% (L4-L20) of the peaks that contain 0-nt variant consensus palindromic HRE DNA element, 64%-71% (L4-L20) of the peaks that contain a 1-nt variant HRE DNA element, 29%-37% (L4-L20) of the peaks that contain a 2-nt variant HRE DNA element, 6%-10% (L4-L20) of the peaks that contain a 3-nt variant HRE DNA element, 1%-2% (L4-L20) of the peaks that contain a 4-nt variant HRE DNA element, and 1%-1% (L4-L20) of the peaks that contain a 5-nt variant HRE DNA element.

Figure S29. PWMs Detection of the GRE 2 DNA Motif in KR ChIPSeq Peaks
(A) Of 194 KR experiments, on average, PWMs identified the GRE 2 motif (VAGRACAKWCTGTYC) in 99%-100% (L4-L20) of the peaks that contain 0-nt variant consensus palindromic HRE DNA element, 84%-88% (L4-L20) of the peaks that contain a 1-nt variant HRE DNA element, 43%-53% (L4-L20) of the peaks that contain a 2-nt variant HRE DNA element, 19%-31% (L4-L20) of the peaks that contain a 3-nt variant HRE DNA element, 19%-31% (L4-L20) of the peaks that contain a 4-nt variant HRE DNA element, and 19%-31% (L4-L20) of the peaks that contain a 5-nt variant HRE DNA element. **(B)** Because peaks can contain multiple HRE DNA elements, we then assigned peaks to the HRE DNA element with the least number of variants relative to the 0-nt consensus palindromic HRE, thus defining each peak by a single HRE DNA element (i.e., unique). Now, PWMs identified the GRE 2 motif in 99%-100% (L4-L20) of the peaks that contain 0-nt variant consensus palindromic HRE DNA element, 84%-88% (L4-L20) of the peaks that contain a 1-nt variant HRE DNA element, 42%-51% (L4-L20) of the peaks that contain a 2-nt variant HRE DNA element, 8%-13% (L4-L20) of the peaks that contain a 3-nt variant HRE DNA element, 1%-2% (L4-L20) of the peaks that contain a 4-nt variant HRE DNA element, and 1%-1% (L4-L20) of the peaks that contain a 5-nt variant HRE DNA element.

Figure S30. PWMs Detection of the PRE DNA Motif in KR ChIPSeq Peaks
(A) Of 194 KR experiments, on average, PWMs identified the PRE motif (VAGRACAKNCTGTBC) in 99%-100% (L4-L20) of the peaks that contain 0-nt variant consensus palindromic HRE DNA element, 99%-99% (L4-L20) of the peaks that contain a 1-nt variant HRE DNA element, 86%-90% (L4-L20) of the peaks that contain a 2-nt variant HRE DNA element, 59%-69% (L4-L20) of the peaks that contain a 3-nt variant HRE DNA element, 50%-64% (L4-L20) of the peaks that contain a 4-nt variant HRE DNA element, and 50%-63% (L4-L20) of the peaks that contain a 5-nt variant HRE DNA element. **(B)** Because peaks can contain multiple HRE DNA elements, we then assigned peaks to the HRE DNA element with the least number of variants relative to the 0-nt consensus palindromic HRE, thus defining each peak by a single HRE DNA element (i.e., unique). Now, PWMs identified the PRE motif in 99%-100% (L4-L20) of the peaks that contain 0-nt variant consensus palindromic HRE DNA element, 99%-99% (L4-L20) of the peaks that contain a 1-nt variant HRE DNA element, 85%-89% (L4-L20) of the peaks that contain a 2-nt variant HRE DNA element, 49%-57% (L4-L20) of the peaks that contain a 3-nt variant HRE DNA element, 19%-25% (L4-L20) of the peaks that contain a 4-nt variant HRE DNA element, and 6%-8% (L4-L20) of the peaks that contain a 5-nt variant HRE DNA element.

Figure S31. PWMs Detection of the ARE Palindrome Half-Site DNA Motif in KR ChIPSeq Peaks
(A) Of 194 KR experiments, on average, PWMs identified the ARE Palindrome Half-Site motif (CCAGGAACAG) in 88%-89% (L4-L20) of the peaks that contain 0-nt variant consensus palindromic HRE DNA element, 77%-80% (L4-L20) of the peaks that contain a 1-nt variant HRE DNA element, 69%-73% (L4-L20) of the peaks that contain a 2-nt variant HRE DNA element, 60%-64% (L4-L20) of the peaks that contain a 3-nt variant HRE DNA element, 57%-63% (L4-L20) of the peaks that contain a 4-nt variant HRE DNA element, and 56%-63% (L4-L20) of the peaks that contain a 5-nt variant HRE DNA element. **(B)** Because peaks can contain multiple HRE DNA elements, we then assigned peaks to the HRE DNA element with the least number of variants relative to the 0-nt consensus palindromic HRE, thus defining each peak by a single HRE DNA element (i.e., unique). Now, PWMs identified the ARE Palindrome Half-Site motif in 88%-89% (L4-L20) of the peaks that contain 0-nt variant consensus palindromic HRE DNA element, 77%-80% (L4-L20) of the peaks that contain a 1-nt variant HRE DNA element, 68%-72% (L4-L20) of the peaks that contain a 2-nt variant HRE DNA element, 56%-59% (L4-L20) of the peaks that contain a 3-nt variant HRE DNA element, 46%-50% (L4-L20) of the peaks that contain a 4-nt variant HRE DNA element, and 33%-39% (L4-L20) of the peaks that contain a 5-nt variant HRE DNA element.

Figure S32. PWMs Detection of the Different HRE DNA Motifs (ARE, GRE, GRE 2, PRE, ARE Half) in KR ChIPSeq Peaks
Displaying **Figure S27-S31** side-by-side for comparison purposes. Of 194 KR experiments, there was a wide distribution in the number of peaks identified by PWMs as containing the ARE motif (RGRACASNSTGTYCYB), GRE motif (NRGVACABNVTGTYCY), GRE 2 motif (VAGRACAKWCTGTYC), PRE motif (VAGRACAKNCTGTBC), or ARE palindrome half-site motif (CCAGGAACAG). For example, of the peaks that contain a 1-nt variant HRE DNA element, on average, 64%-72% (L4-L20) of the peaks were identified as containing the GRE motif, whereas 99%-99% (L4-L20) of those exact same peaks were identified as containing the PRE motif. Or, of the peaks that contain a 2-nt variant HRE DNA element, 31%-39% (L4-L20) of the

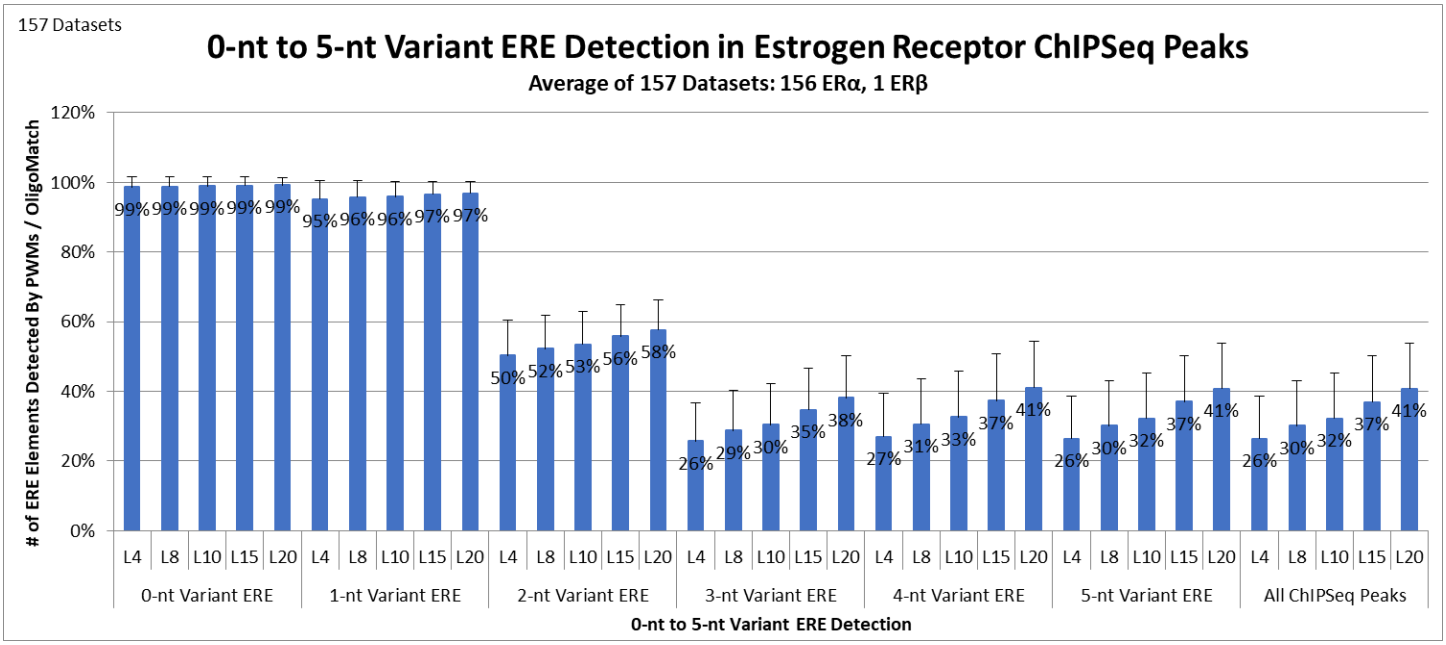
peaks were identified as containing the GRE motif, whereas 86%-90% (L4-L20) of those exact same peaks were identified as containing the PRE motif. Thus, it is apparent why the use of PWMs for DNA motif identification have led investigators to delineate different *cis*-regulatory DNA elements for the different KRs. However, analysis of KR DNA-binding at 81,922 0-nt to 5-nt variant HRE elements in the genome quantitively demonstrates that all members of the KR family follow the same DNA-binding rules, and thus bind the same DNA elements **(Table S8, Table S16-S20)**. These DNA-binding rules at 0-nt to 5-nt variant HRE DNA elements in the genome was observed in 194 KR ChIPSeq experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S8, Table S16-S20)**.

Figure S26. PWMs Detection of the ERE DNA Motif in ER ChIPSeq Peaks

ERE(NR), IR3/MCF7-ERa-ChIPSeq(Unpublished)/Homer

VAGGTCACNSTGACC

(A) PWMs detection of the ERE DNA motif in ER ChIPSeq Peaks



(B) PWMs detection of the ERE DNA motif in ER ChIPSeq Peaks assigned to the ERE DNA element with the least number of variants, defining each peak by a single ERE DNA element (i.e., Unique)

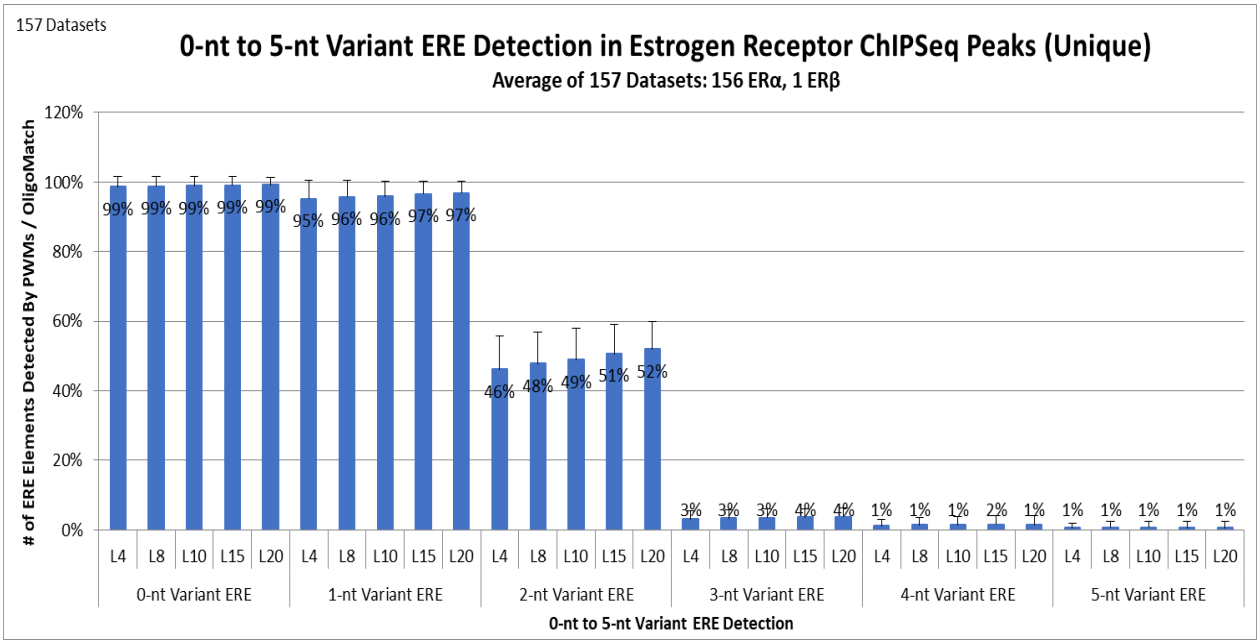
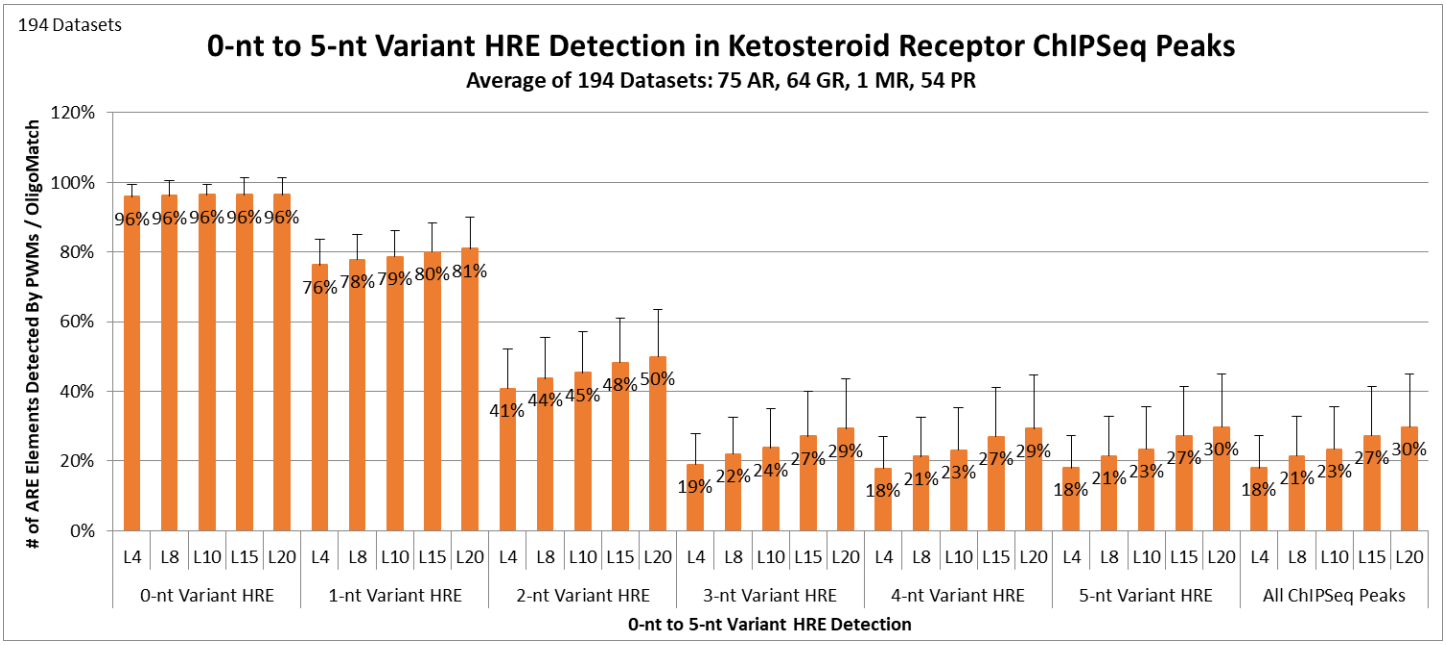


Figure S27. PWMs Detection of the ARE DNA Motif in KR ChIPSeq Peaks

ARE(NR)/LNCAP-AR-ChIPSeq(GSE27824)/Homer

RGRACASNSTGTYCYB

(A) PWMs detection of the ARE DNA motif in KR ChIPSeq Peaks



(B) PWMs detection of the ARE DNA motif in KR ChIPSeq Peaks assigned to the HRE DNA element with the least number of variants, defining each peak by a single HRE DNA element (i.e., Unique)

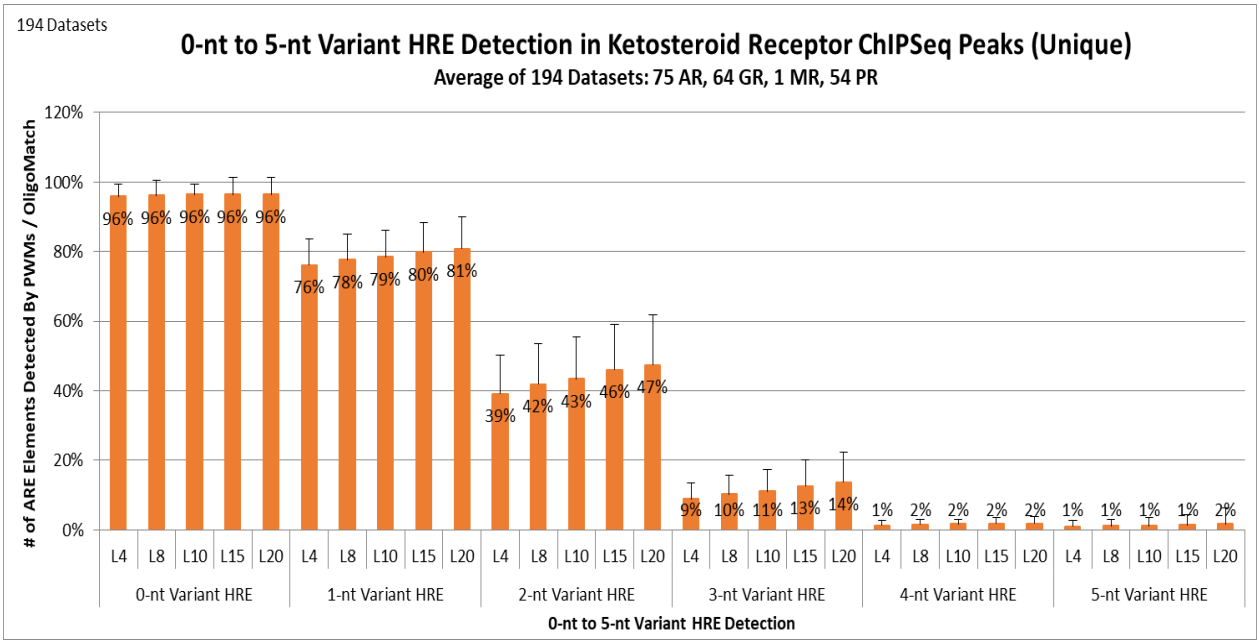
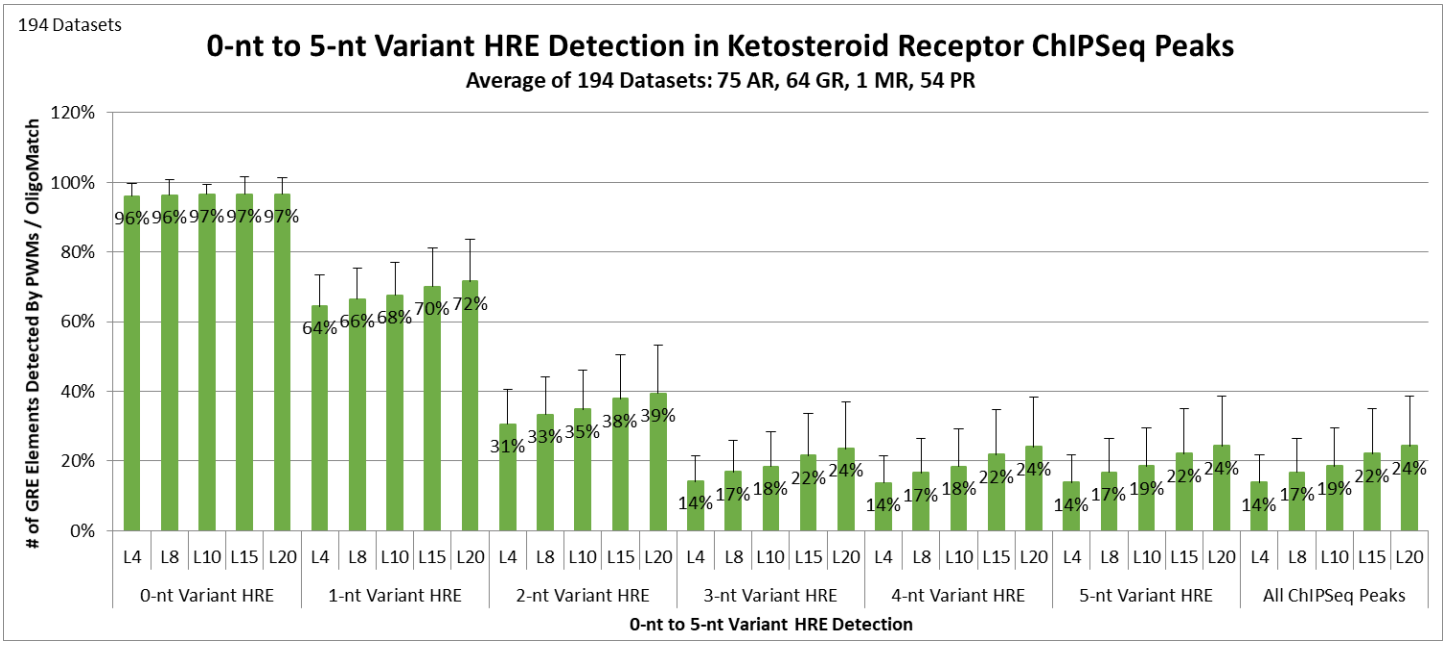


Figure S28. PWMs Detection of the GRE DNA Motif in KR ChIPSeq Peaks

GRE(NR), IR3/A549-GR-ChIPSeq(GSE32465)/Homer

NRGVACABNVTGTYCY

(A) PWMs detection of the GRE DNA motif in KR ChIPSeq Peaks



(B) PWMs detection of the GRE DNA motif in KR ChIPSeq Peaks assigned to the HRE DNA element with the least number of variants, defining each peak by a single HRE DNA element (i.e., Unique)

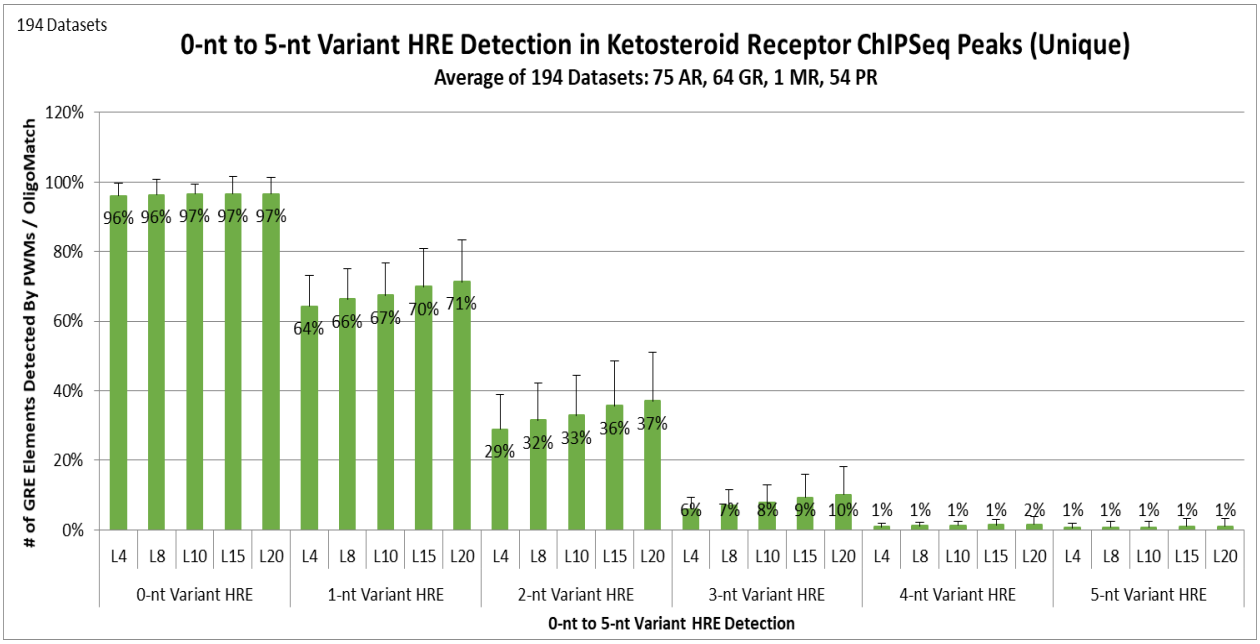
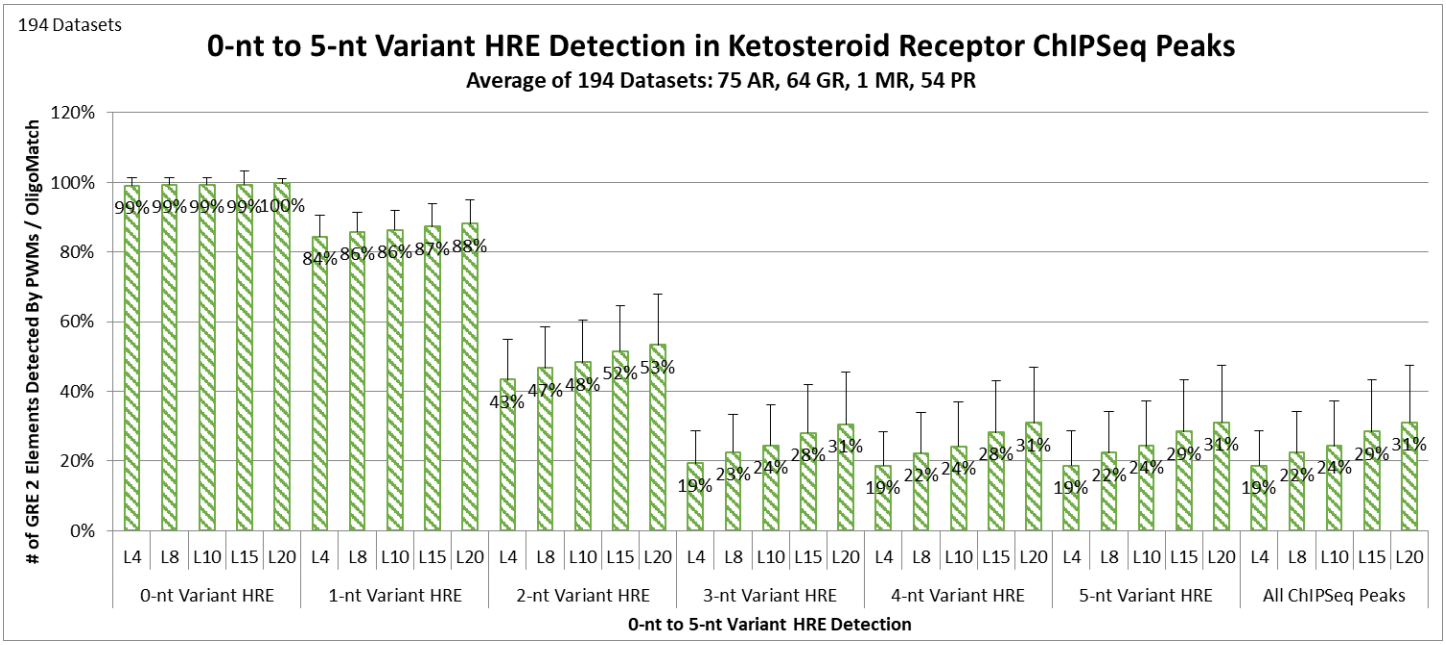


Figure S29. PWMs Detection of the GRE 2 DNA Motif in KR ChIPSeq Peaks

GRE(NR), IR3/RAW264.7-GRE-ChIPSeq(Unpublished)/Homer

VAGRACAKWCTGTYC

(A) PWMs detection of the GRE 2 DNA motif in KR ChIPSeq Peaks



(B) PWMs detection of the GRE 2 DNA motif in KR ChIPSeq Peaks assigned to the HRE DNA element with the least number of variants, defining each peak by a single HRE DNA element (i.e., Unique)

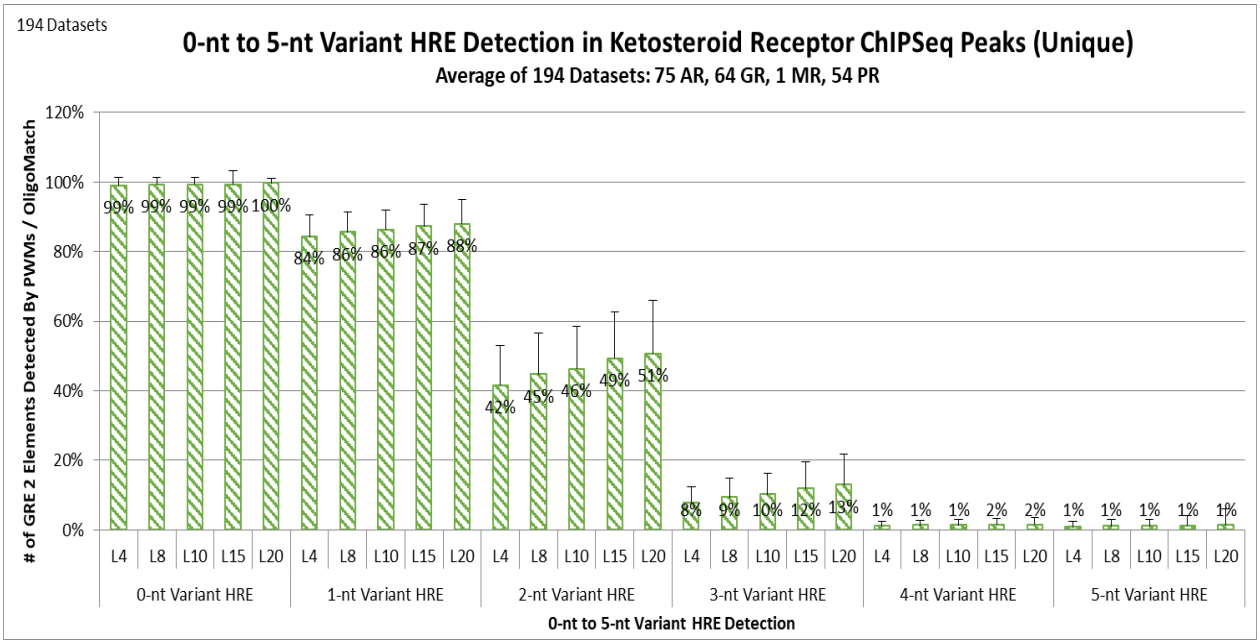
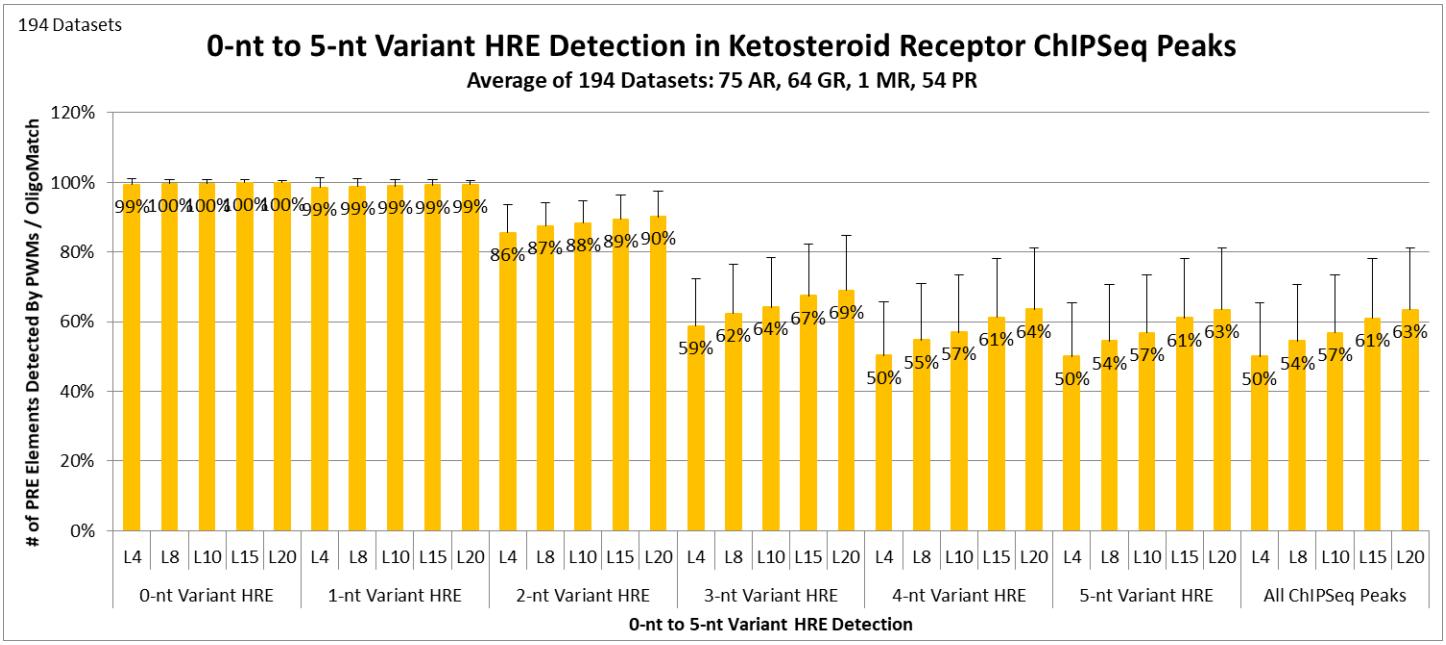


Figure S30. PWMs Detection of the PRE DNA Motif in KR ChIPSeq Peaks

PR(NR)/T47D-PR-ChIPSeq(GSE31130)/Homer

VAGRACAKNCTGTBC

(A) PWMs detection of the PRE DNA motif in KR ChIPSeq Peaks



(B) PWMs detection of the PRE DNA motif in KR ChIPSeq Peaks assigned to the HRE DNA element with the least number of variants, defining each peak by a single HRE DNA element (i.e., Unique)

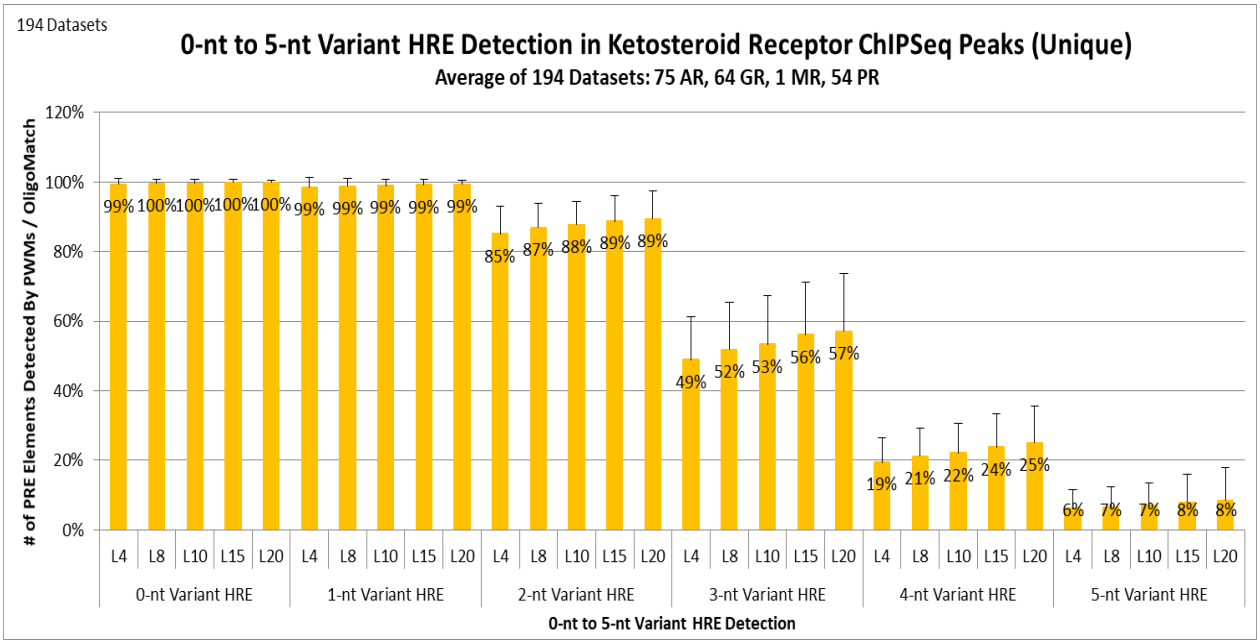
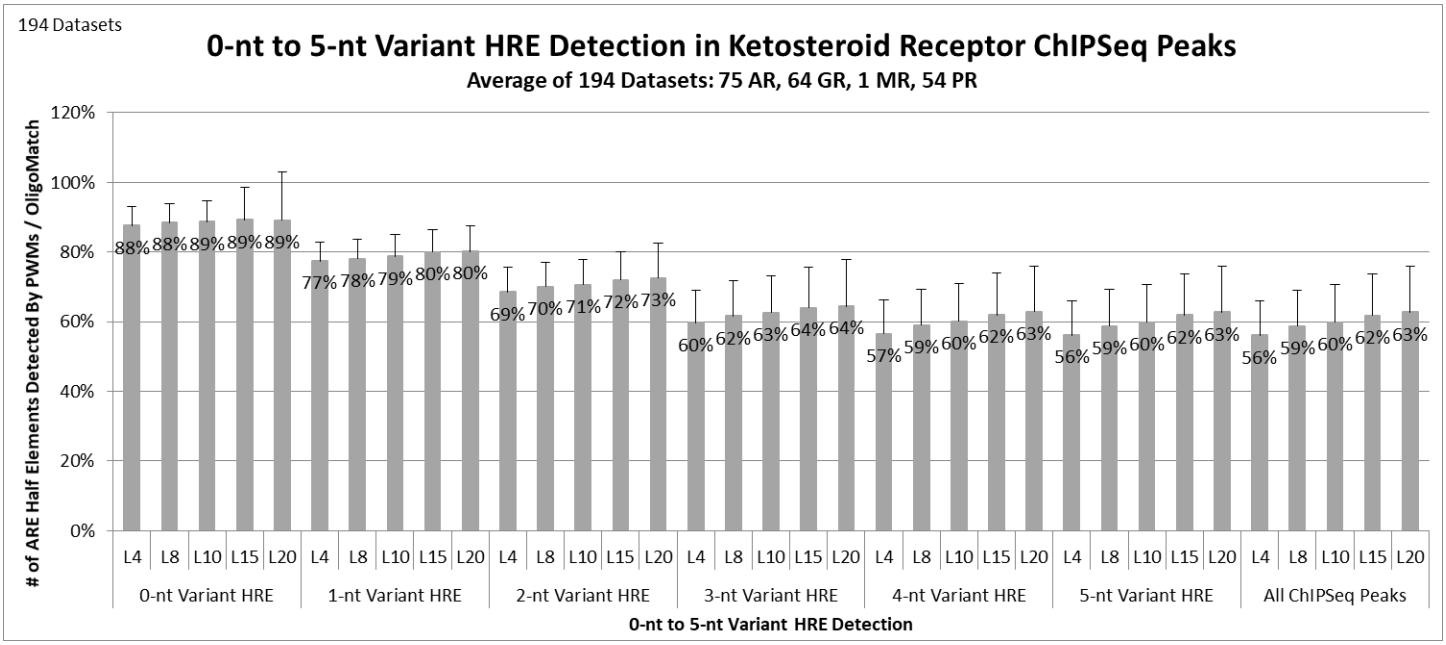


Figure S31. PWMs Detection of the ARE Palindrome Half-Site DNA Motif in KR ChIPSeq Peaks

AR-halfsite(NR)/LNCaP-AR-ChIPSeq(GSE27824)/Homer

CCAGGAACAG

(A) PWMs detection of the ARE Palindrome Half-Site DNA motif in KR ChIPSeq Peaks



(B) PWMs detection of the ARE Palindrome Half-Site DNA motif in KR ChIPSeq Peaks assigned to the HRE DNA element with the least number of variants, defining each peak by a single HRE DNA element (i.e., Unique)

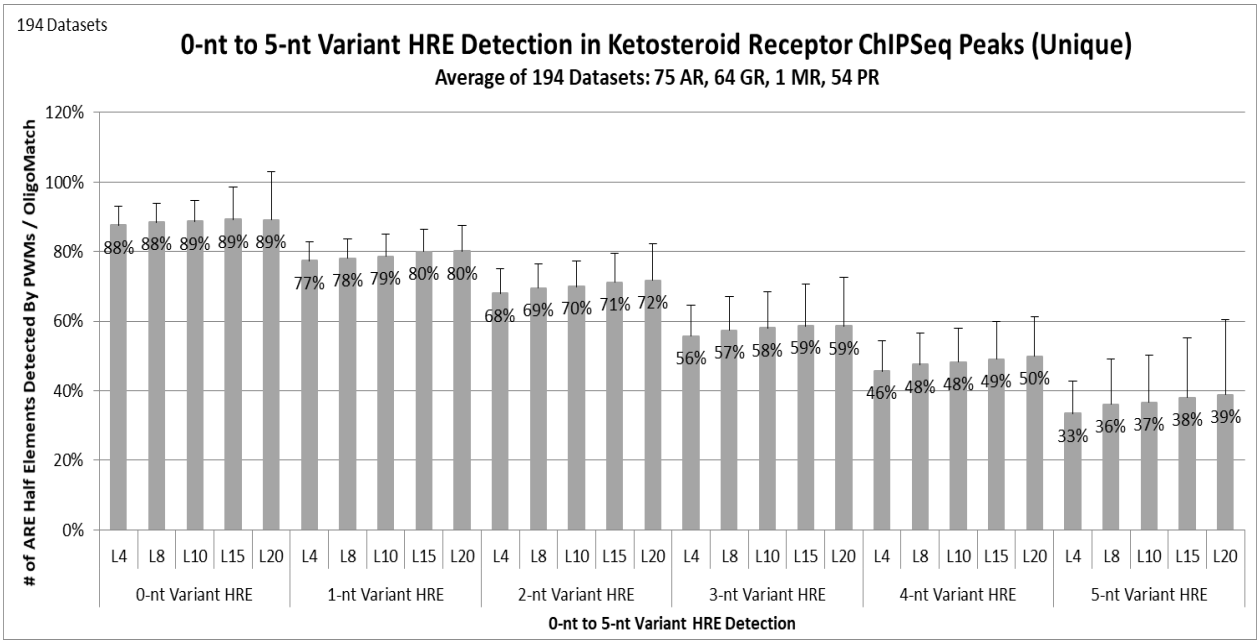
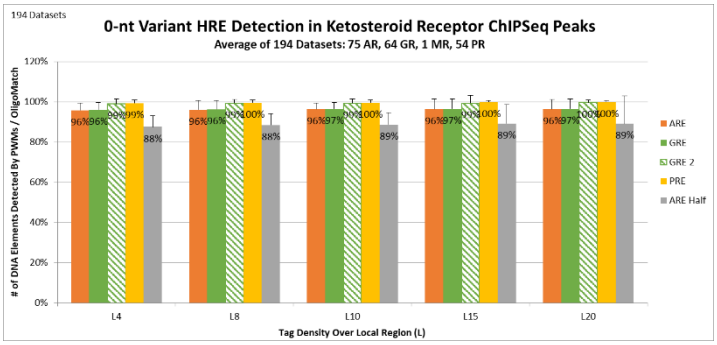
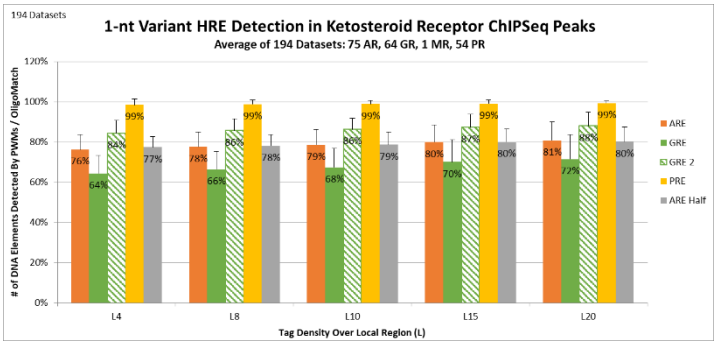
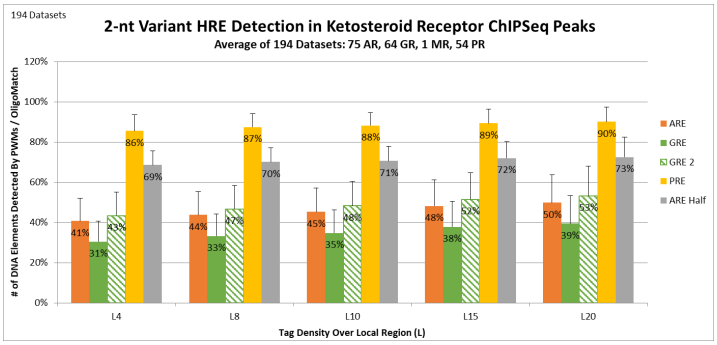


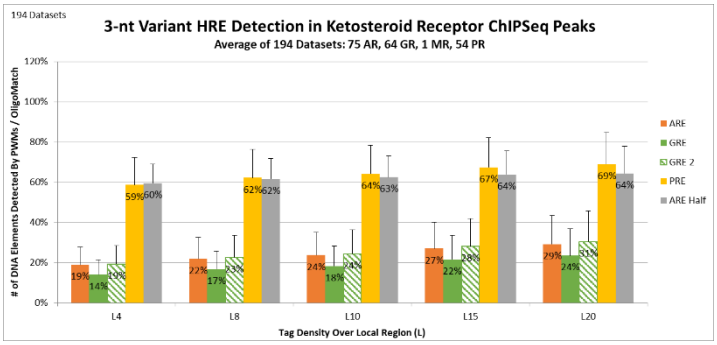
Figure S32. PWMs Detection of the Different HRE Motifs (ARE, GRE, GRE 2, PRE, ARE Half) in KR ChIPSeq Peaks

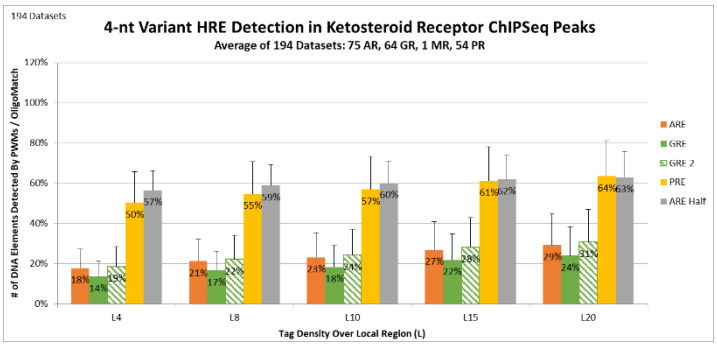
All Peaks



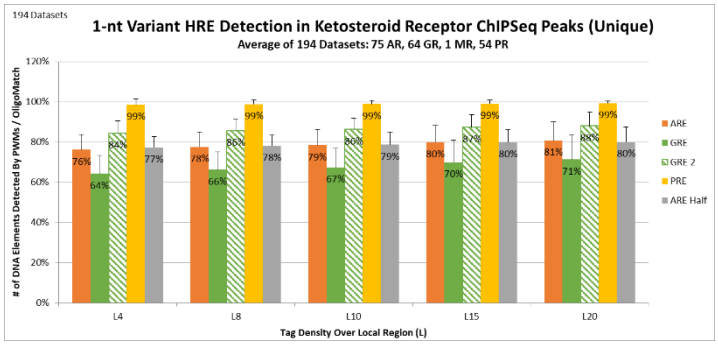


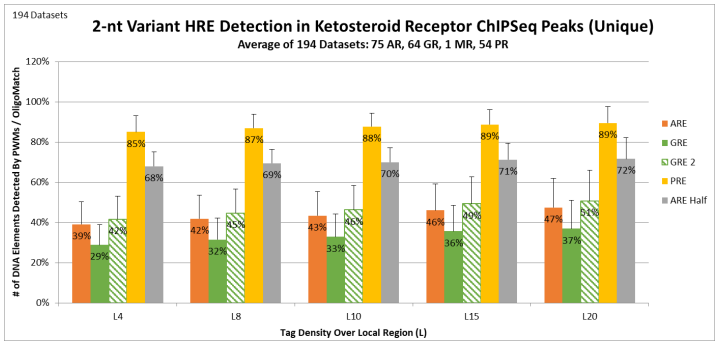


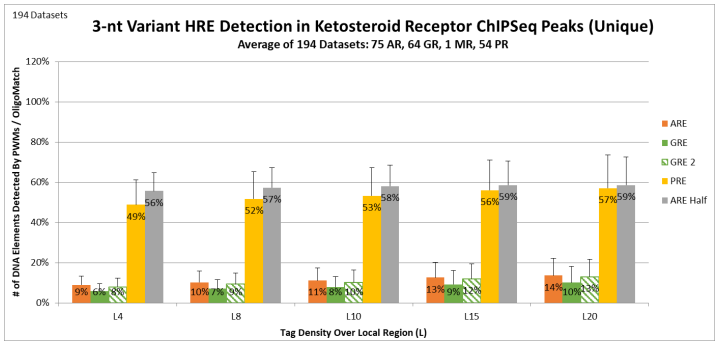


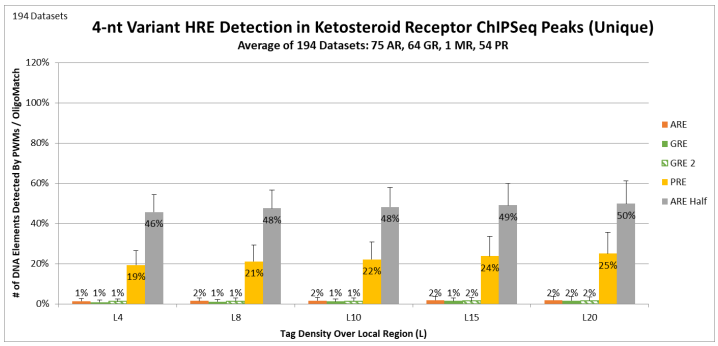


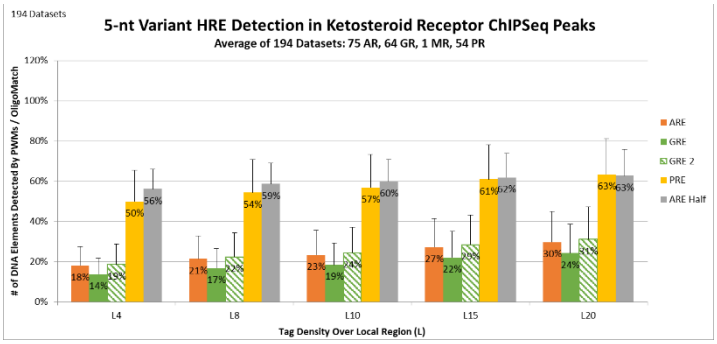
Unique Peaks










194 Datasets
5-nt Variant HRE Detection in Ketosteroid Receptor ChIPSeq Peaks
Average of 194 Datasets: 75 AR, 64 GR, 1 MR, 54 PR
# of DNA Elements Detected By PWMs / OligoMatch
120%
100%
80%
60%
40%
20%
0%
18%
14%
19%
50%
56%
21%
17%
22%
54%
59%
23%
19%
24%
57%
60%
27%
22%
29%
61%
62%
30%
24%
31%
63%
63%
L4
L8
L10
L15
L20
Tag Density Over Local Region (L)
ARE
GRE
GRE 2
PRE
ARE Half

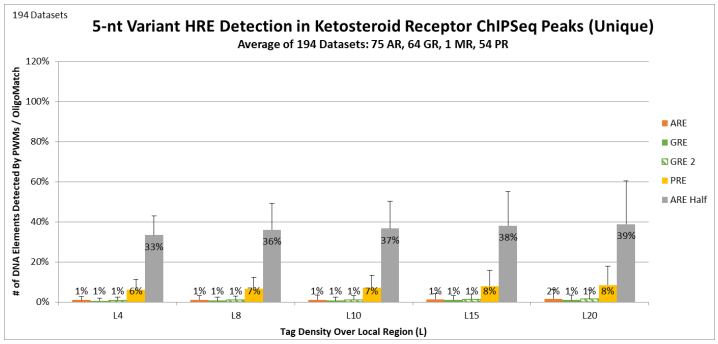

194 Datasets
5-nt Variant HRE Detection in Ketosteroid Receptor ChIPSeq Peaks (Unique)
Average of 194 Datasets: 75 AR, 64 GR, 1 MR, 54 PR
# of DNA Elements Detected By PWMs / OligoMatch
120%
100%
80%
60%
40%
20%
0%
1%
1%
1%
6%
33%
1%
1%
1%
7%
36%
1%
1%
1%
7%
37%
1%
1%
1%
8%
38%
2%
1%
1%
8%
39%
L4
L8
L10
L15
L20
Tag Density Over Local Region (L)
ARE
GRE
GRE 2
PRE
ARE Half

Figure S33-S61 Descriptions: 15-nt ERE and HRE DNA Element Analysis

****All detailed data and statistics associated with every figure are compiled in Data S1****

The analyses in this section were completed by overlapping the location coordinates of each 0-nt to 6-nt variant 15-nt ERE or HRE DNA element in the genome and the location coordinates of the ChIPSeq or ChIPExo peaks in an experiment. The 0-nt variant consensus palindromic 15-nt ERE DNA element (5′-AGGTCAnnnTGACCT-3′) and 0-nt variant consensus palindromic 15-nt HRE DNA element (5′-AGAACAnnnTGTTCT-3′) have twelve (12) primary positions. The 0-nt to 6-nt variant 15-nt ERE or HRE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 36 1-nt variant DNA elements (12 variant positions), 594 2-nt variant DNA elements (66 variant positions), 5,940 3-nt variant DNA elements (220 variant positions), 40,095 4-nt variant DNA elements (495 variant positions), 192,456 5-nt variant DNA elements (792 variant positions), and 673,596 6-nt variant DNA elements (924 variant positions), for a total of 912,718 DNA elements.

Figure S33. 0-nt Variant 15-nt ERE and HRE Identification in ER and KR ChIPSeq Peaks (16 possibilities)
(S/N) analysis of sNR DNA-binding at the 16 possible 15-nt ERE and HRE DNA elements in the genome (i.e., 4-nt possibilities in position 1 and 4-nt possibilities in position 12). sNR DNA-binding was most predominant when an adenine (A) was in position 1 and a thymine (T) was in position 12 of the 15-nt ERE DNA element (5′-AGGTCAnnnTGACCT-3′) for ERs and the 15-nt HRE DNA element (5′-AGAACAnnnTGTTCT-3′) for KRs. This was observed in 157 ER experiments and 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20). Thus, these 15-nt ERE and HRE DNA elements (i.e., adenine (A) in position 1 and thymine (T) in position 12) represent the 0-nt variant consensus palindromic 15-nt ERE and HRE DNA elements for ERs and KRs.

Figure S34. Number of 0-nt to 6-nt Variant 15-nt ERE and HRE DNA Elements in the Mouse and Human Genome
(Top Table) The number of 0-nt to 5-nt variants of the 13-nt ERE and HRE consensus palindromic DNA element on the positive/sense strand in the mouse (mm10) and human (hg19) genome. Here the 13-nt ERE is the estrogen response element (5′-GGTCAnnnTGACC-3′) and the 13-nt HRE is the hormone response element (5′-GAACAnnnTGTTC-3′). The 0-nt to 5-nt variant 13-nt ERE and HRE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements (10 variant positions), 405 2-nt variant DNA elements (45 variant positions), 3,240 3-nt variant DNA elements (120 variant positions), 17,010 4-nt variant DNA elements (210 variant positions), and 61,236 5-nt variant DNA elements (252 variant positions), for a total of 81,922 DNA elements. The population count of each of the 81,922 0-nt to 5-nt variant 13-nt ERE and HRE DNA elements in the mouse (mm10) and human (hg19) genome can be found in **Table S55-S58**.

(Bottom Table) The number of 0-nt to 6-nt variants of the 15-nt ERE and HRE consensus palindromic DNA element on the positive/sense strand in the mouse (mm10) and human (hg19) genome. Here the 15-nt ERE is the estrogen response element (5′-AGGTCAnnnTGACCT-3′) and the 15-nt HRE is the hormone response element (5′-AGAACAnnnTGTTCT-3′). The 0-nt to 6-nt variant 15-nt ERE and HRE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 36 1-nt variant DNA elements (12 variant positions), 594 2-nt variant DNA elements (66 variant positions), 5,940 3-nt variant DNA elements (220 variant positions), 40,095 4-nt variant DNA elements (495 variant positions), 192,456 5-nt variant DNA elements (792 variant positions), and 673,596 6-nt variant DNA elements (924 variant positions), for a total of 912,718 DNA elements. The population count of each of the 912,718 0-nt to 6-nt variant 15-nt ERE and HRE DNA elements in the mouse (mm10) and human (hg19) genome can be found in **Table S59-S62**.

*****

Figure S35. Inversion Symmetry of sNR DNA-Binding at 1-nt Variant EREs and HREs in the Genome (%)
(A) Average distribution of 1-nt variant EREs in ER experiments (156 ERα and 1 ERβ) at the thirty-six (36) 1-nt variant EREs in the genome (the total number of 1-nt variant EREs contained within a ChIPSeq experiment =100%) **(B)** Average distribution of 1-nt variant HREs in 193 KR experiments (75 AR, 64 GR, 54 PR) at the thirty-six (36) 1-nt variant HREs in the genome (the total number of 1-nt variant HREs contained within a ChIPSeq experiment =100%). These thirty-six (36) 1-nt variant EREs or HREs are defined by twelve (12) variant positions (position 1 through position 6, followed by their reverse-complements).

Figure S36. Inversion Symmetry of sNR DNA-Binding at 2-nt Variant EREs and HREs in the Genome (%)
(A) Average distribution of 2-nt variant EREs in ER experiments (156 ERα and 1 ERβ) at the five hundred and ninety-four (594) 2-nt variant EREs in the genome (the total number of 2-nt variant EREs contained within a ChIPSeq experiment =100%) **(B)** Average distribution of 2-nt variant HREs in 193 KR experiments (75 AR, 64 GR, 54 PR) at the five hundred and ninety-four (594) 2-nt variant HREs in the genome (the total number of 2-nt variant HREs contained within a ChIPSeq experiment =100%). These five hundred and ninety-four (594) 2-nt variant EREs or HREs are defined by sixty-six (66) variant positions. Six (6) of those are when the variants are in the palindromic position pairs (PP): 1-12, 2-11, 3-10, 4-9, 5-8, 6-7. The remaining sixty (60) are symmetrically split into 30 groups and their reverse-complements.

*****

Figure S37. (S/N) analysis of 0-nt to 6-nt Variant EREs in ER ChIPSeq Peaks (displayed by number of variants)
(S/N) analysis of ER DNA-binding at 0-nt to 6-nt variant EREs in 157 ER experiments (156 ERα and 1 ERβ). Experiments are ordered by the experiment with highest (S/N) value at the 0-nt variant consensus palindromic ERE DNA element in the genome. The signal-to-noise ratio (S/N) is the absolute number of times a 0-nt to 6-nt variant DNA element occurs within an experiment [defined by the total number of peaks in the experiment and the peak length] compared to the random frequency expectation of that DNA element occurring in the genome (i.e., the probability that any 12-nt DNA element that has a maximum possibility of 4 nucleotides at each of the 12 primary positions will occur is once every 16,777,216 nucleotides (4^{12}) at random occurrence). Thus, the expected noise is calculated by multiplying the total number of peaks in the experiment by the length of the peaks over 4^{12}. For (S/N) analysis of 0-nt to 6-nt variant DNA elements (displayed by the number of variants in the DNA element), the expected noise is multiplied by the number of DNA elements in each group: 1 0-nt variant consensus palindromic DNA element, 36 1-nt variant DNA elements, 594 2-nt variant DNA elements, 5,940 3-nt variant DNA elements, 40,095 4-nt variant DNA elements, 192,456 5-nt variant DNA elements, and 673,596 6-nt variant DNA elements, for a total of 912,718 0-nt to 6-nt variant DNA elements. **Note**: the relative enhancement values of the sNR DNA-binding signals are scale invariant (i.e., the relative ratios between the 0-nt to 6-nt variant groups are constant). Y-axis =the natural log of the (S/N) value (i.e., random =0 and 0 =log(1)).

Figure S38. (S/N) analysis of 0-nt to 6-nt Variant HREs in KR ChIPSeq Peaks (displayed by number of variants)
(S/N) analysis of ER DNA-binding at 0-nt to 6-nt variant HREs in 194 KR experiments (75 AR, 64 GR, 1 MR, 54 PR). Experiments are grouped by KR, and then ordered by the experiment with highest (S/N) value at the 0-nt variant consensus palindromic HRE DNA element in the genome. The signal-to-noise ratio (S/N) is the absolute number of times a 0-nt to 6-nt variant DNA element occurs within an experiment [defined by the total number of peaks in the experiment and the peak length] compared to the random frequency expectation of that DNA element occurring in the genome (i.e., the probability that any 12-nt DNA element that has a maximum possibility of 4 nucleotides at each of the 12 primary positions will occur is once every 16,777,216 nucleotides (4^{12}) at random occurrence). Thus, the expected noise is calculated by multiplying the total number of peaks in the experiment by the length of the peaks over 4^{12}. For (S/N) analysis of 0-nt to 6-nt variant DNA elements (displayed by the number of variants in the DNA element), the expected noise is multiplied by the number of DNA elements in each group: 1 0-nt variant consensus palindromic DNA element, 36 1-nt variant DNA elements, 594 2-nt variant DNA elements, 5,940 3-nt variant DNA elements, 40,095 4-nt variant DNA elements, 192,456 5-nt variant DNA elements, and 673,596 6-nt variant DNA elements, for a total of 912,718 0-nt to 6-nt variant DNA elements. **Note**: the relative enhancement values of the sNR DNA-binding signals are scale invariant (i.e., the relative ratios between the 0-nt to 6-nt variant groups are constant). Y-axis =the natural log of the (S/N) value (i.e., random =0 and 0 =log(1)).

*****

Figure S39. 924 Half-Site Groups Symmetrically Split Into 452 Groups and Their Reverse-Complements (and 20 Palindromes)
(Column A) Categorizing the 673,596 6-nt variant DNA elements into 924 half-site groups, defined by the six (6) positions that are fixed/not varied. The 924 half-site groups are symmetrically split into a set of 452 groups and their 452 reverse-complements, plus 20 innate palindromes. **(Column B)** These 452 groups further split into three distinct sub-groups (zero vacancies, one vacancy, two vacancies) plus the 20 innate palindromes (three vacancies) depending on how many reverse-complement vacancies are in the DNA element (i.e., the information for which nucleotide occupies each of the twelve (12) primary positions of the 0-nt variant consensus palindromic DNA element is missing/replaced by variants/vacant in the position and its reverse-complement position (its palindromic position pair)). Of these 472 groups, thirty-two (32) have zero vacancies, two hundred and forty (240) have one vacancy, one hundred and eighty (180) have two vacancies, and twenty (20) have three vacancies. **(Column C)** The reverse-complement vacancy position ID indicates which of the six (6) palindromic position pairs (i.e., 1-12, 2-11, 3-10, 4-9, 5-8, 6-7) are missing/replaced by variants/vacant. **(Column B)** The number of reverse-complement vacancies equals the number of reverse-complement double occupants in the DNA element (i.e., the information for which nucleotide occupies each of the twelve (12) primary positions of the 0-nt variant consensus palindromic DNA element is occupied in the position and its reverse-complement position (its palindromic position pair)). **(Column D)** The reverse-complement double occupant position ID indicates which of the six (6) palindromic position pairs (i.e., 1-12, 2-11, 3-10, 4-9, 5-8, 6-7) are doubly occupied. **(Column E)** The positions that are replaced by variants/not fixed for 6-nt variant DNA elements, and may be replaced by variants for 0-nt to 5-nt variant DNA elements.

Figure S40. (S/N) analysis of 0-nt to 6-nt Variant EREs in ER ChIPSeq Peaks (displayed by 924 half-site groups)
(S/N) analysis of 0-nt to 6-nt variant EREs (displayed by the 924 half-site groups) in ER (WT-E2-1hr) [76,163 peaks, 146-nt peak length] ChIPSeq peaks. X-axis order =Reverse-Complement Vacancy Position ID 4-9 > 1-12 > 2-11 > 6-7 > 5-8 > 3-10. See **Table S25** for x-axis details. This DNA-binding profile was observed in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S27)**. See **Table S68-S71** for step-by-step instructions of the data analysis from peak selection to (S/N) analysis of ER DNA-binding at 912,718 0-nt to 6-nt variant 15-nt ERE DNA elements in the genome, displayed by the 924 half-site groups.

Figure S41. (S/N) analysis of 0-nt to 6-nt Variant HREs in AR ChIPSeq Peaks (displayed by 924 half-site groups)
(S/N) analysis of 0-nt to 6-nt variant HREs (displayed by the 924 half-site groups) in AR (AR-wt1) [49,859 peaks, 136-nt peak length] ChIPSeq peaks. X-axis order =Reverse-Complement Vacancy Position ID 3-10 > 1-12 > 6-7 > 2-11 > 5-8 > 4-9. See **Table S26** for x-axis details. This DNA-binding profile was observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S28)**.

Figure S42. (S/N) analysis of 0-nt to 6-nt Variant HREs in GR ChIPSeq Peaks (displayed by 924 half-site groups)
(S/N) analysis of 0-nt to 6-nt variant HREs (displayed by the 924 half-site groups) in GR (Liver-GR-WT-pred-6am-2) [23,742 peaks, 142-nt peak length] ChIPSeq peaks. X-axis order =Reverse-Complement Vacancy Position ID 3-10 > 1-12 > 6-7 > 2-11 > 5-8 > 4-9. See **Table S26** for x-axis details. This DNA-binding profile was observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S28)**.

Figure S43. (S/N) analysis of 0-nt to 6-nt Variant HREs in PR ChIPSeq Peaks (displayed by 924 half-site groups)
(S/N) analysis of 0-nt to 6-nt variant HREs (displayed by the 924 half-site groups) in PR (Uterus-PGR-P4) [66,545 peaks, 154-nt peak length] ChIPSeq peaks. X-axis order =Reverse-Complement Vacancy Position ID 3-10 > 1-12 > 6-7 > 2-11 > 5-8 > 4-9. See **Table S26** for x-axis details. This DNA-binding profile was observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S28)**.

*****

Figure S44. (S/N) analysis of 1-nt Variant EREs in ER ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 1-nt variant EREs in ER (WT-E2-1hr) ChIPSeq peaks [76,163 peaks, 146-nt peak length].
Displayed by the 12 variant positions of the 36 1-nt variant EREs. This DNA-binding profile was observed in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S29)**.

Figure S45. (S/N) analysis of 2-nt Variant EREs in ER ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 2-nt variant EREs in ER (WT-E2-1hr) ChIPSeq peaks [76,163 peaks, 146-nt peak length].
Displayed by the 66 variant positions of the 594 2-nt variant EREs. 6 of the 66 variant positions are the six (6) palindromic position pairs: 1-12, 2-11, 3-10, 4-9, 5-8, 6-7 (PP, dark red solid bars). 30 of the 66 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (blue solid bars). 30 of the 66 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (blue diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome). This DNA-binding profile was observed in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S30)**.

Figure S46. (S/N) analysis of 3-nt Variant EREs in ER ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 3-nt variant EREs in ER (WT-E2-1hr) ChIPSeq peaks [76,163 peaks, 146-nt peak length].
Displayed by the 220 variant positions of the 5,940 3-nt variant EREs. 40 of the 220 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (blue solid bars). 180 of the 220 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (blue diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome). This DNA-binding profile was observed in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S31)**.

Figure S47. (S/N) analysis of 4-nt Variant EREs in ER ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 4-nt variant EREs in ER (WT-E2-1hr) ChIPSeq peaks [76,163 peaks, 146-nt peak length].
Displayed by the 495 variant positions of the 40,095 4-nt variant EREs. 15 of the 495 variant positions are the five (5) palindromic position pairs: 2-4-9-11, 4-6-7-9, 4-5-8-9, 3-4-9-10, 2-6-7-11, 2-5-8-11, 2-3-10-11, 5-6-7-8, 3-6-7-10, 3-5-8-10, 1-4-9-12, 1-2-11-12, 1-6-7-12, 1-5-8-12, 1-3-10-12 (PP, dark red solid bars). 30 of the 495 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (blue solid bars). 450 of the 495 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (blue diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome). This DNA-binding profile was observed in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S32)**.

Figure S48. (S/N) analysis of 5-nt Variant EREs in ER ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 5-nt variant EREs in ER (WT-E2-1hr) ChIPSeq peaks [76,163 peaks, 146-nt peak length].
Displayed by the 792 variant positions of the 192,456 5-nt variant EREs. 12 of the 792 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (blue solid bars). 780 of the 792 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (blue diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the

number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome). This DNA-binding profile was observed in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S33)**.

Figure S49. (S/N) analysis of 6-nt Variant EREs in ER ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 6-nt variant EREs in ER (WT-E2-1hr) ChIPSeq peaks [76,163 peaks, 146-nt peak length].
Displayed by the 924 variant positions of the 673,596 6-nt variant EREs. 20 of the 924 variant positions are the five (5) palindromic position pairs: 1-4-6-7-9-12, 2-3-4-9-10-11, 1-3-4-9-10-12, 1-2-4-9-11-12, 2-4-6-7-9-11, 1-4-5-8-9-12, 2-4-5-8-9-11, 3-4-6-7-9-10, 1-2-6-7-11-12, 4-5-6-7-8-9, 1-2-5-8-11-12, 1-3-6-7-10-12, 1-2-3-10-11-12, 3-4-5-8-9-10, 1-5-6-7-8-12, 2-3-5-8-10-11, 2-3-6-7-10-11, 1-3-5-8-10-12, 3-5-6-7-8-10, 2-5-6-7-8-11 (PP, dark red solid bars). 2 of the 924 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (blue solid bars). 922 of the 924 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (blue diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome). This DNA-binding profile was observed in 157 ER experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S34)**.

*****

Figure S50. (S/N) analysis of 1-nt Variant HREs in KR ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 1-nt variant HREs in AR (AR-wt1) ChIPSeq peaks [49,859 peaks, 136-nt peak length].
Displayed by the 12 variant positions of the 36 1-nt variant HREs. This DNA-binding profile was observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S35)**.

Figure S51. (S/N) analysis of 2-nt Variant HREs in KR ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 2-nt variant HREs in AR (AR-wt1) ChIPSeq peaks [49,859 peaks, 136-nt peak length].
Displayed by the 66 variant positions of the 594 2-nt variant HREs. 6 of the 66 variant positions are the six (6) palindromic position pairs: 1-12, 2-11, 3-10, 4-9, 5-8, 6-7 (PP, dark red solid bars). 30 of the 66 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (orange solid bars). 30 of the 66 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (orange diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome). This DNA-binding profile was observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S36)**.

Figure S52. (S/N) analysis of 3-nt Variant HREs in KR ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 3-nt variant HREs in AR (AR-wt1) ChIPSeq peaks [49,859 peaks, 136-nt peak length].
Displayed by the 220 variant positions of the 5,940 3-nt variant HREs. 40 of the 220 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (orange solid bars). 180 of the 220 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (orange diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome). This DNA-binding profile was observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S37)**.

Figure S53. (S/N) analysis of 4-nt Variant HREs in KR ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 4-nt variant HREs in AR (AR-wt1) ChIPSeq peaks [49,859 peaks, 136-nt peak length].
Displayed by the 495 variant positions of the 40,095 4-nt variant HREs. 15 of the 495 variant positions are the five (5) palindromic position pairs: 3-6-7-10, 2-3-10-11, 3-5-8-10, 3-4-9-10, 2-6-7-11, 5-6-7-8, 4-6-7-9, 2-5-8-11, 2-4-9-11, 4-5-8-9, 1-3-10-12, 1-6-7-12, 1-2-11-12, 1-5-8-12, 1-4-9-12 (PP, dark red solid bars). 30 of the 495 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (orange solid bars). 450 of the 495 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (orange diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome). This DNA-binding profile was observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S38)**.

Figure S54. (S/N) analysis of 5-nt Variant HREs in KR ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 5-nt variant HREs in AR (AR-wt1) ChIPSeq peaks [49,859 peaks, 136-nt peak length].
Displayed by the 792 variant positions of the 192,456 5-nt variant HREs. 12 of the 792 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (orange solid bars). 780 of the 792 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (orange diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome). This DNA-binding profile was observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S39)**.

Figure S55. (S/N) analysis of 6-nt Variant HREs in KR ChIPSeq Peaks (displayed by variant position)
(S/N) analysis of 6-nt variant HREs in AR (AR-wt1) ChIPSeq peaks [49,859 peaks, 136-nt peak length].
Displayed by the 924 variant positions of the 673,596 6-nt variant HREs. 20 of the 924 variant positions are the five (5) palindromic position pairs: 1-2-3-10-11-12, 3-4-5-8-9-10, 1-3-4-9-10-12, 1-2-4-9-11-12, 2-3-6-7-10-11, 1-3-5-8-10-12, 3-5-6-7-8-10, 2-5-6-7-8-11, 1-2-6-7-11-12, 4-5-6-7-8-9, 1-5-6-7-8-12, 2-3-5-8-10-11, 1-4-6-7-9-12, 2-3-4-9-10-11, 1-2-5-8-11-12, 1-3-6-7-10-12, 2-4-6-7-9-11, 1-4-5-8-9-12, 2-4-5-8-9-11, 3-4-6-7-9-10 (PP, dark red solid bars). 2 of the 924 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (orange solid bars). 922 of the 924 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (orange diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome). This DNA-binding profile was observed in 194 KR experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20) **(Table S40)**.

*****

Here we show replicate experiments for GR and GR-Dim to illustrate that the slight decrease in DNA-binding at a few specific 3-nt variant HRE DNA elements by GR-Dim is made apparent by the DNA-binding enhancement of a particular experiment at 0-nt to 2-nt variant HRE DNA elements as the relative (S/N) values of sNR DNA-binding signals are scale invariant (i.e., the relative ratios between the 0-nt to 5-nt variant groups are constant) **(Figure S37-S38)**. Thus, the reduced amplitude in DNA-binding at a few specific 3-nt variant HRE DNA elements only occurs in experiments with the highest (S/N) values at 0-nt to 2-nt variant HRE DNA elements.
Note: see section *"Transform the 13-nt DNA Element Analysis to the 15-nt DNA Element Analysis"* to understand how the 13-nt ERE and HRE DNA elements split into the 15-nt ERE and HRE DNA elements (e.g., explaining why some DNA elements are a 2-nt variant DNA element in the 13-nt analysis and a 4-nt variant DNA element in the 15-nt analysis).

Figure S56. (S/N) analysis of 0-nt to 6-nt Variant HREs in GR and GR-Dim ChIPSeq Peaks #1 (displayed by 924 half-site groups)
(S/N) analysis of 0-nt to 6-nt variant HREs (displayed by the 924 half-site groups) in GR (Liver-GR-WT-pred-6am-2) [23,742 peaks, 142-nt peak length] ChIPSeq peaks and GR-Dim (Liver-GR-Dim-pred-6am-2) [34,966 peaks, 148-nt peak length] ChIPSeq peaks. X-axis order =Reverse-Complement Vacancy Position ID 3-10 > 1-12 > 6-7 > 2-11 > 5-8 > 4-9. See **Table S26** for x-axis details.

Figure S57. (S/N) analysis of 0-nt to 6-nt variant HREs in GR and GR-Dim ChIPSeq Peaks #2 (displayed by 924 half-site groups)
(S/N) analysis of 0-nt to 6-nt variant HREs (displayed by the 924 half-site groups) in GR (Liver-GR-WT-pred-6am-1) [34,758 peaks, 151-nt peak length] ChIPSeq peaks and GR-Dim (Liver-GR-Dim-pred-6am-1) [22,130 peaks, 163-nt peak length] ChIPSeq peaks. X-axis order =Reverse-Complement Vacancy Position ID 3-10 > 1-12 > 6-7 > 2-11 > 5-8 > 4-9. See **Table S26** for x-axis details.

Figure S58. (S/N) analysis of 0-nt to 6-nt variant HREs in GR and GR-Dim ChIPExo Peaks #1 (displayed by 924 half-site groups)
(S/N) analysis of 0-nt to 6-nt variant HREs (displayed by the 924 half-site groups) in GR (Liver-GR-WT-pred-6am-1) [20,966 peaks, 445-nt peak length] ChIPExo peaks and GR-Dim (Liver-GR-Dim-pred-6am-1) [19,480 peaks, 391-nt peak length] ChIPExo peaks. X-axis order =Reverse-Complement Vacancy Position ID 3-10 > 1-12 > 6-7 > 2-11 > 5-8 > 4-9. See **Table S26** for x-axis details.

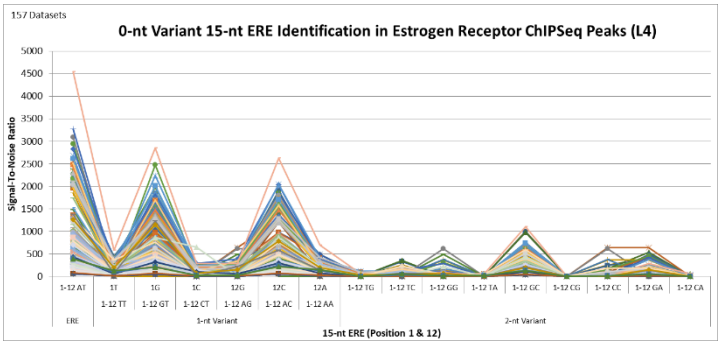
Figure S59. (S/N) analysis of 0-nt to 6-nt variant HREs in GR and GR-Dim ChIPExo Peaks #2 (displayed by 924 half-site groups)
(S/N) analysis of 0-nt to 6-nt variant HREs (displayed by the 924 half-site groups) in GR (Liver-GR-WT-pred-6am-2) [22,129 peaks, 338-nt peak length] ChIPExo peaks and GR-Dim (Liver-GR-Dim-pred-6am-2) [25,670 peaks, 408-nt peak length] ChIPExo peaks. X-axis order =Reverse-Complement Vacancy Position ID 3-10 > 1-12 > 6-7 > 2-11 > 5-8 > 4-9. See **Table S26** for x-axis details.

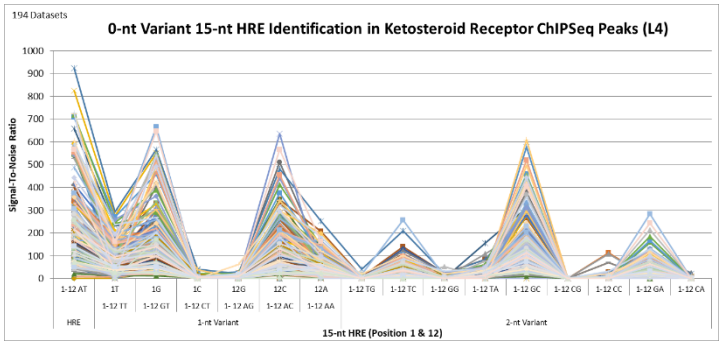
Figure S60. (S/N) analysis of 1-nt to 6-nt variant HREs in GR and GR-Dim ChIPSeq Peaks (displayed by variant position)

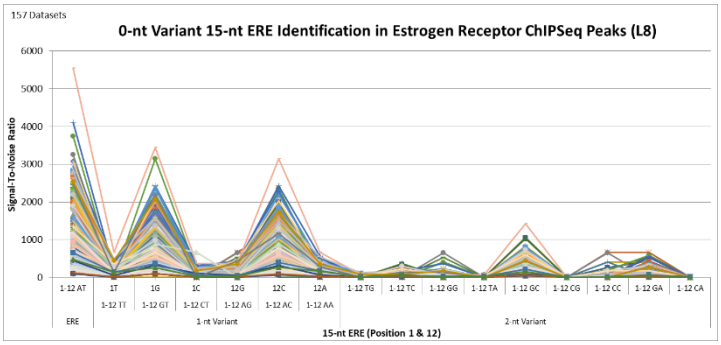
(S/N) analysis of 1-nt to 6-nt variant HREs (displayed by variant position) in GR (Liver-GR-WT-pred-6am-1) [23,742 peaks, 142-nt peak length] (Liver-GR-WT-pred-6am-2) [34,758 peaks, 151-nt peak length] ChIPSeq peaks and GR-Dim (Liver-GR-Dim-pred-6am-1) [34,966 peaks, 148-nt peak length] (Liver-GR-Dim-pred-6am-2) [22,130 peaks, 163-nt peak length] ChIPSeq peaks.

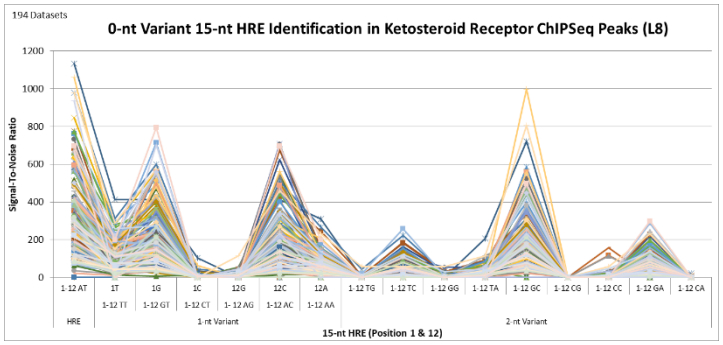
Figure S61. (S/N) analysis of 1-nt to 6-nt variant HREs in GR and GR-Dim ChIPExo Peaks (displayed by variant position) (S/N) analysis of 1-nt to 6-nt variant HREs (displayed by variant position) in GR (Liver-GR-WT-pred-6am-1) [20,966 peaks, 445-nt peak length] (Liver-GR-WT-pred-6am-2) [22,129 peaks, 338-nt peak length] ChIPExo peaks and GR-Dim (Liver-GR-Dim-pred-6am-1) [19,480 peaks, 391-nt peak length] (Liver-GR-Dim-pred-6am-2) [25,670 peaks, 408-nt peak length] ChIPExo peaks.

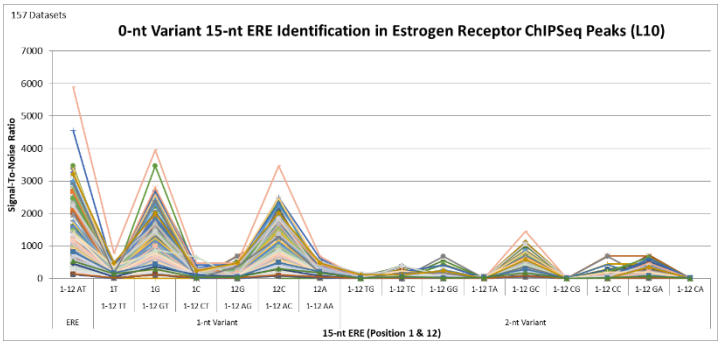
Figure S33. 0-nt Variant 15-nt ERE and HRE Identification in ER and KR ChIPSeq Peaks (16 possibilities)

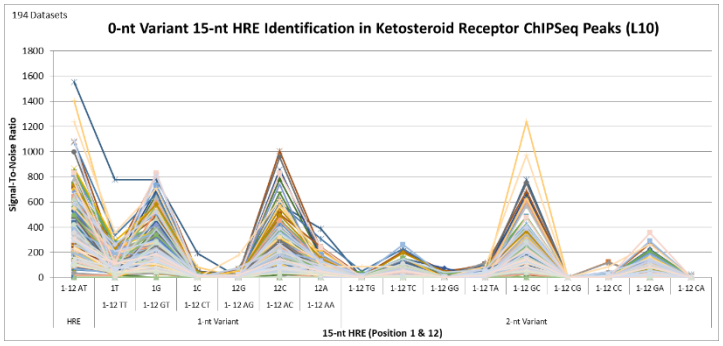


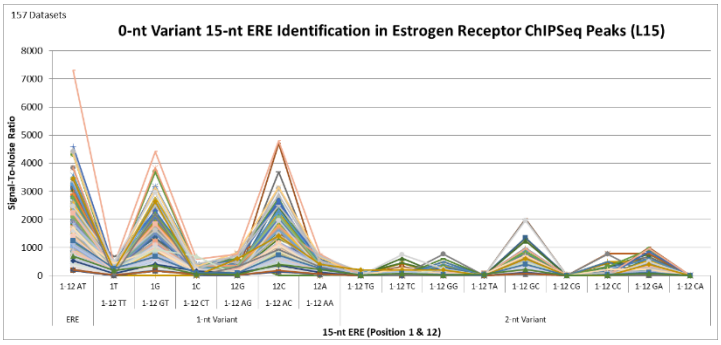


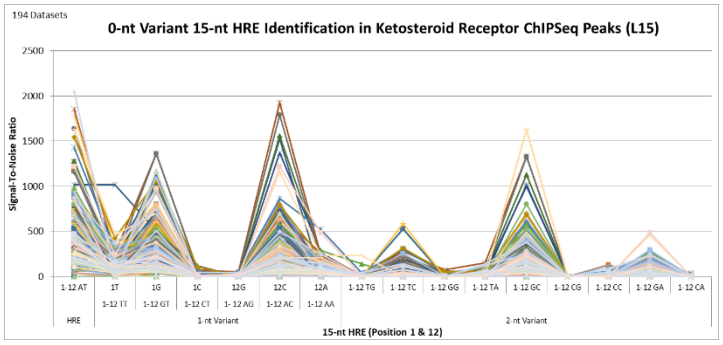


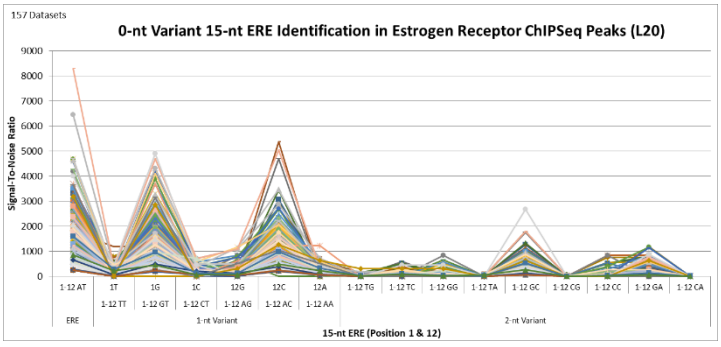












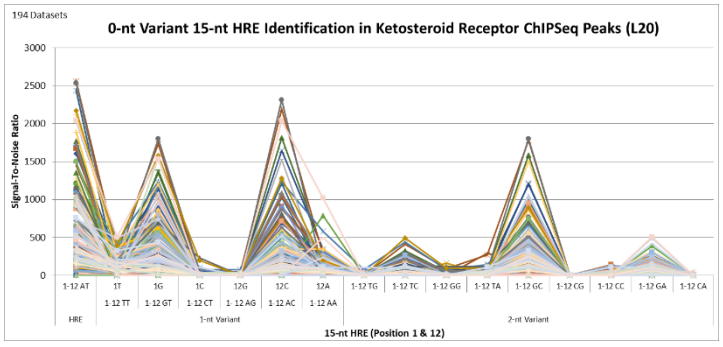


Figure S34. Number of 0-nt to 6-nt Variant 15-nt ERE and HRE DNA Elements in the Mouse and Human Genome

13-nt ERE and HRE DNA Elements (*n*=10)

	DNA Element	Combinatorial Counts				DNA Element Frequency in Genome			
A	B	C	D	E	F	Mouse Genome (mm10)		Human Genome (hg19)	
k	**13-nt ERE** GGTCAnnnTGACC	Combinations $C = \binom{n}{k}$	4-nts $D = C \times 4^k$	Unique $E = C \times 3^k$	Total Unique $F = \Sigma E$	Unique	Total Unique	Unique	Total Unique
0	0-nt Variant ERE	1	1	1	1	2,367		2,194	
1	1-nt Variant ERE	10	40	30	31	71,428	73,795	60,313	62,507
2	2-nt Variant ERE	45	720	405	436	898,155	971,950	914,726	977,233
3	3-nt Variant ERE	120	7,680	3,240	3,676	6,750,607	7,722,557	7,516,184	8,493,417
4	4-nt Variant ERE	210	53,760	17,010	20,686	35,508,190	43,230,747	38,222,674	46,716,091
5	5-nt Variant ERE	252	258,048	61,236	81,922	134,732,965	177,963,712	141,280,093	187,996,184

A	B	C	D	E	F	Mouse Genome (mm10)		Human Genome (hg19)	
k	**13-nt HRE** GAACAnnnTGTTC	Combinations $C = \binom{n}{k}$	4-nts $D = C \times 4^k$	Unique $E = C \times 3^k$	Total Unique $F = \Sigma E$	Unique	Total Unique	Unique	Total Unique
0	0-nt Variant HRE	1	1	1	1	3,444		3,535	
1	1-nt Variant HRE	10	40	30	31	97,039	100,483	104,767	108,302
2	2-nt Variant HRE	45	720	405	436	1,337,516	1,437,999	1,339,543	1,447,845
3	3-nt Variant HRE	120	7,680	3,240	3,676	10,461,197	11,899,196	10,771,159	12,219,004
4	4-nt Variant HRE	210	53,760	17,010	20,686	49,391,434	61,290,630	54,136,564	66,355,568
5	5-nt Variant HRE	252	258,048	61,236	81,922	169,440,498	230,731,128	187,469,898	253,825,466

A	B	C	D	E	F
k		Combinations $C = \binom{n}{k}$	4-nts $D = C \times 4^k$	Unique $E = C \times 3^k$	Total Unique $F = \Sigma E$
0	0-nt Variant	10!/(0! x 10!)	1 x 1	1 x 1	1 + 0
1	1-nt Variant	10!/(1! x 9!)	10 x 4	10 x 3	30 + 1
2	2-nt Variant	10!/(2! x 8!)	45 x 16	45 x 9	405 + 31
3	3-nt Variant	10!/(3! x 7!)	120 x 64	120 x 27	3,240 + 436
4	4-nt Variant	10!/(4! x 6!)	210 x 256	210 x 81	17,010 +3,676
5	5-nt Variant	10!/(5! x 5!)	252 x 1,024	252 x 243	61,236 + 20,686

15-nt ERE and HRE DNA Elements (*n*=12)

	DNA Element	Combinatorial Counts				DNA Element Frequency in Genome			
A	B	C	D	E	F	Mouse Genome (mm10)		Human Genome (hg19)	
k	**15-nt ERE** AGGTCAnnnTGACCT	Combinations $C = \binom{n}{k}$	4-nts $D = C \times 4^k$	Unique $E = C \times 3^k$	Total Unique $F = \Sigma E$	Unique	Total Unique	Unique	Total Unique
0	0-nt Variant ERE	1	1	1	1	404		342	
1	1-nt Variant ERE	12	48	36	37	10,964	11,368	9,223	9,565
2	2-nt Variant ERE	66	1056	594	631	152,140	163,508	118,354	127,919
3	3-nt Variant ERE	220	14,080	5,940	6,571	1,256,647	1,420,155	1,213,963	1,341,882
4	4-nt Variant ERE	495	126,720	40,095	46,666	6,916,002	8,336,157	7,349,740	8,691,622
5	5-nt Variant ERE	792	811,008	192,456	239,122	30,575,777	38,911,934	33,052,166	41,743,788
6	6-nt Variant ERE	924	3,784,704	673,596	912,718	104,588,671	143,500,605	113,051,144	154,794,932

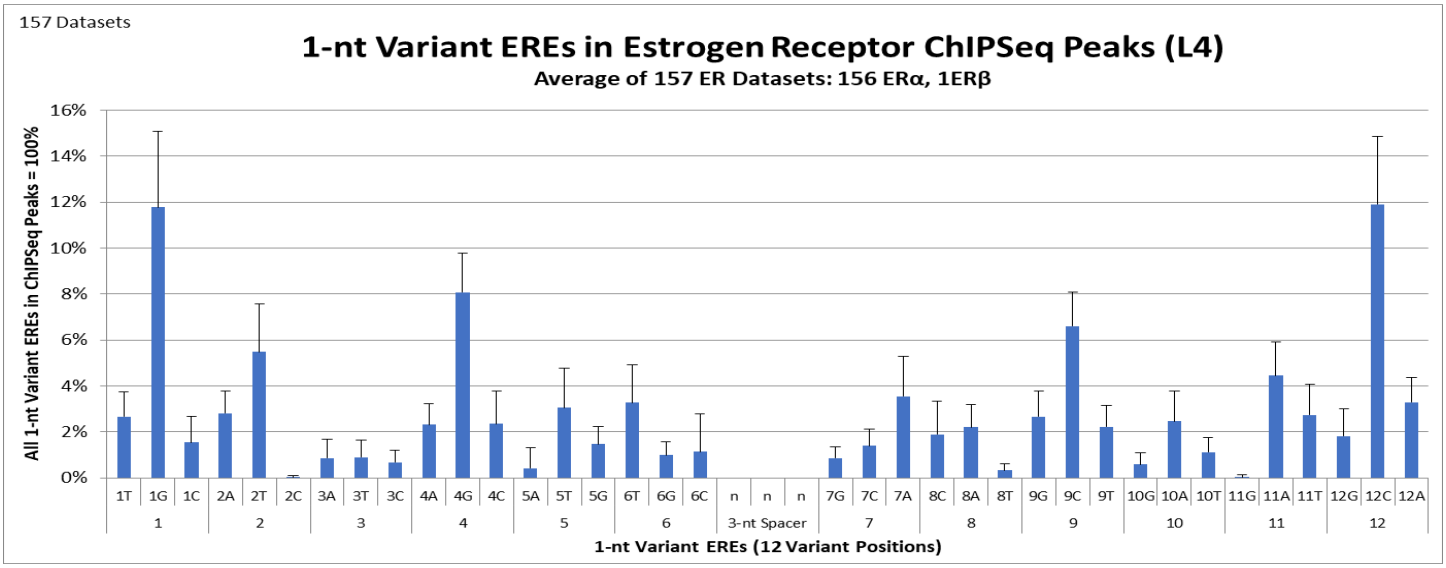
A	B	C	D	E	F	Mouse Genome (mm10)		Human Genome (hg19)	
k	**15-nt HRE** AGAACAnnnTGTTCT	Combinations $C = \binom{n}{k}$	4-nts $D = C \times 4^k$	Unique $E = C \times 3^k$	Total Unique $F = \Sigma E$	Unique	Total Unique	Unique	Total Unique
0	0-nt Variant HRE	1	1	1	1	639		565	
1	1-nt Variant HRE	12	48	36	37	17,847	18,486	14,639	15,204
2	2-nt Variant HRE	66	1056	594	631	256,540	275,026	206,983	222,187
3	3-nt Variant HRE	220	14,080	5,940	6,571	2,188,277	2,463,303	1,978,470	2,200,657
4	4-nt Variant HRE	495	126,720	40,095	46,666	11,128,133	13,591,436	11,661,975	13,862,632
5	5-nt Variant HRE	792	811,008	192,456	239,122	43,400,607	56,992,043	48,017,520	61,880,152
6	6-nt Variant HRE	924	3,784,704	673,596	912,718	134,064,927	191,056,970	151,083,618	212,963,770

A	B	C	D	E	F
k		Combinations $C = \binom{n}{k}$	4-nts $D = C \times 4^k$	Unique $E = C \times 3^k$	Total Unique $F = \Sigma E$
0	0-nt Variant	12!/(0! x 12!)	1 x 1	1 x 1	1 + 0
1	1-nt Variant	12!/(1! x 11!)	12 x 4	12 x 3	36 + 1
2	2-nt Variant	12!/(2! x 10!)	66 x 16	66 x 9	594 + 37
3	3-nt Variant	12!/(3! x 9!)	220 x 64	220 x 27	5,940 + 631
4	4-nt Variant	12!/(4! x 8!)	495 x 256	495 x 81	40,095 + 6,571
5	5-nt Variant	12!/(5! x 7!)	792 x 1,024	792 x 243	192,456 + 46,666
6	6-nt Variant	12!/(6! x 6!)	924 x 4,096	924 x 729	673,596 + 239,122

Figure S35. Inversion Symmetry of sNR DNA-Binding at 1-nt Variant EREs and HREs in the Genome (%)

(A) 1-nt Variant EREs in ER ChIPSeq Peaks (157 Experiments)

ERE = AGGTCAnnnTGACCT
1 2 3 4 5 6 7 8 9 10 11 12



(B) 1-nt Variant HREs in KR ChIPSeq Peaks (193 Experiments)

HRE = AGAACAnnnTGTTCT
1 2 3 4 5 6 7 8 9 10 11 12

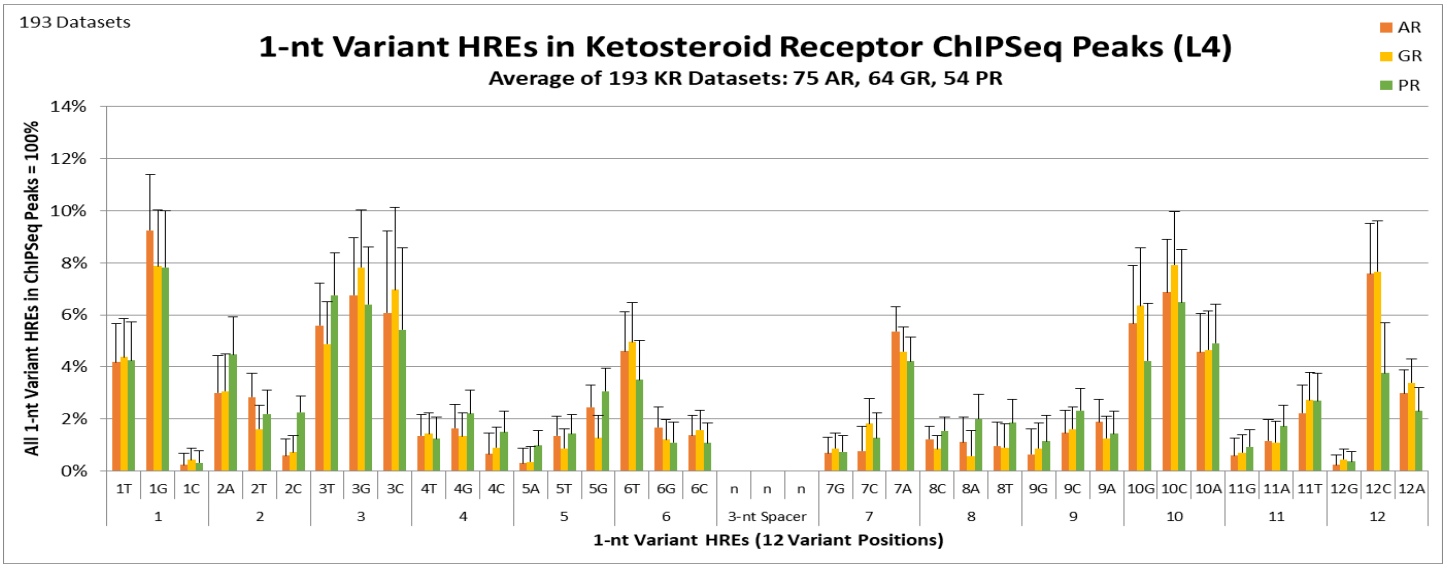
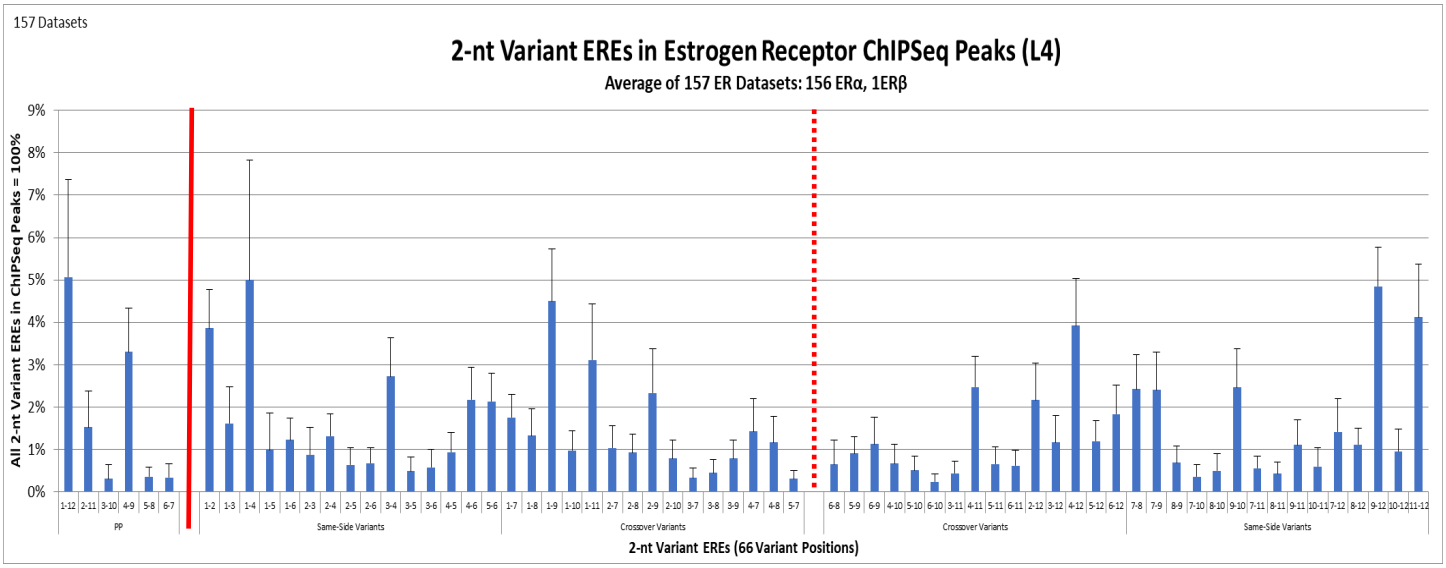


Figure S36. Inversion Symmetry of sNR DNA-Binding at 2-nt Variant EREs and HREs in the Genome (%)

(A) 2-nt Variant EREs in ER ChIPSeq Peaks (157 Experiments)

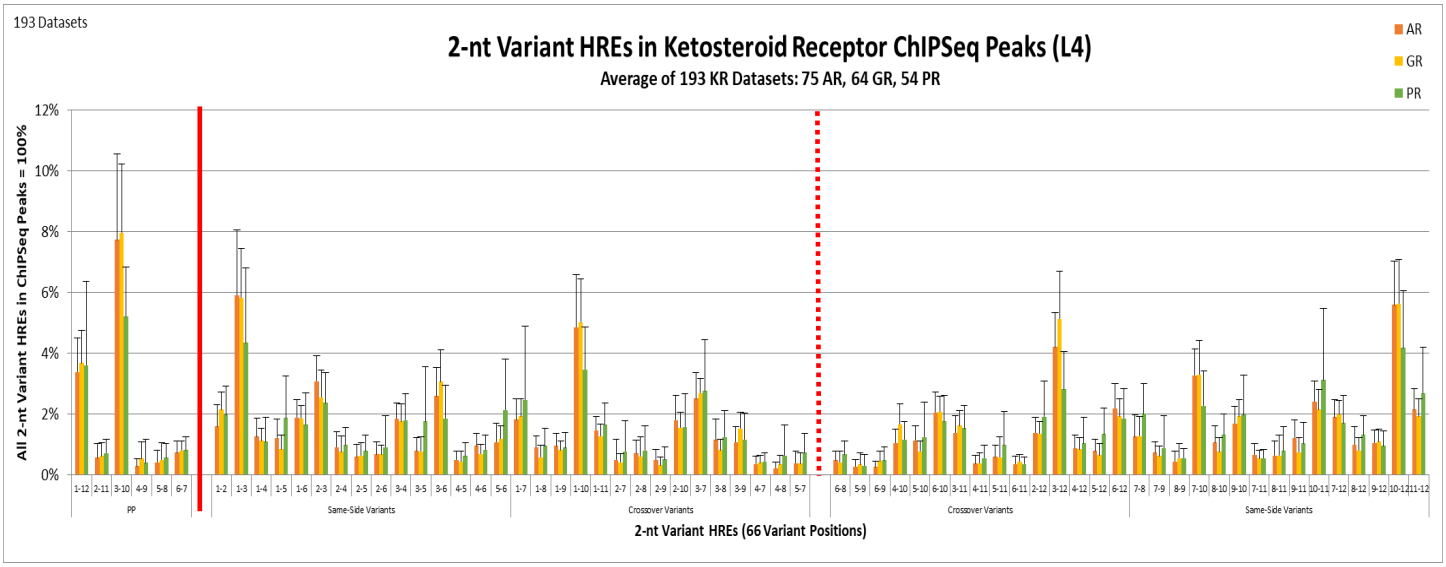
ERE = AGGTCAnnnTGACCT
1 2 3 4 5 6 7 8 9 10 11 12



---- **Symmetry Divider**

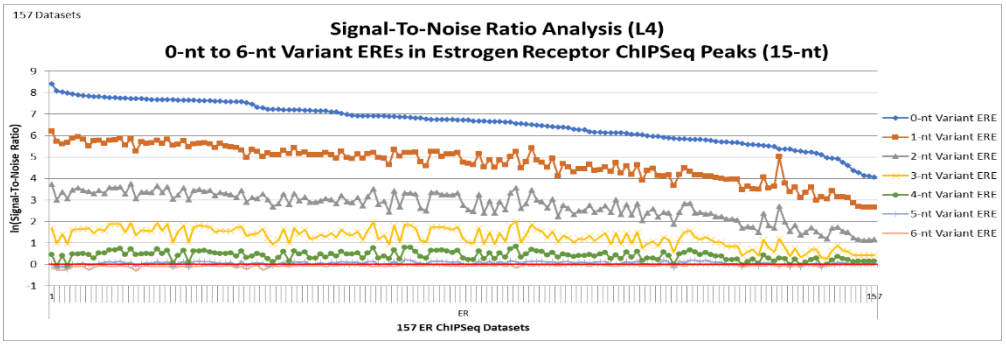
(B) 2-nt Variant HREs in KR ChIPSeq Peaks (193 Experiments)

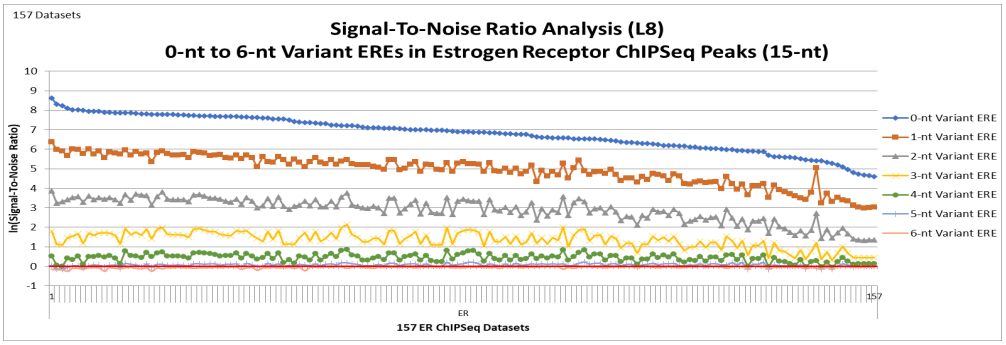
HRE = AGAACAnnnTGTTCT
1 2 3 4 5 6 7 8 9 10 11 12

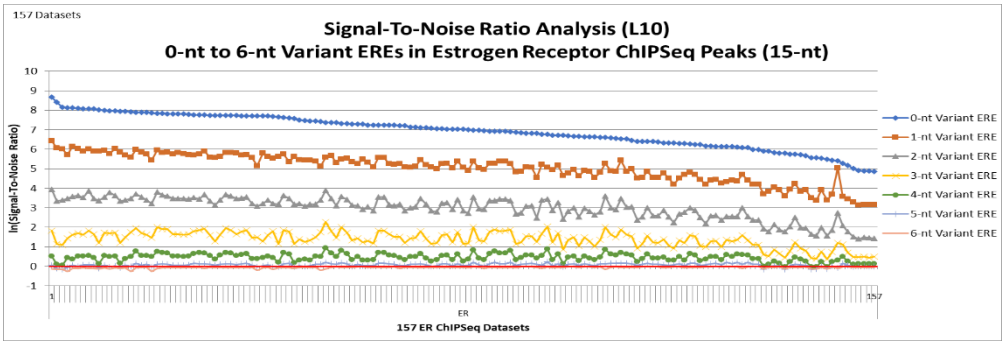


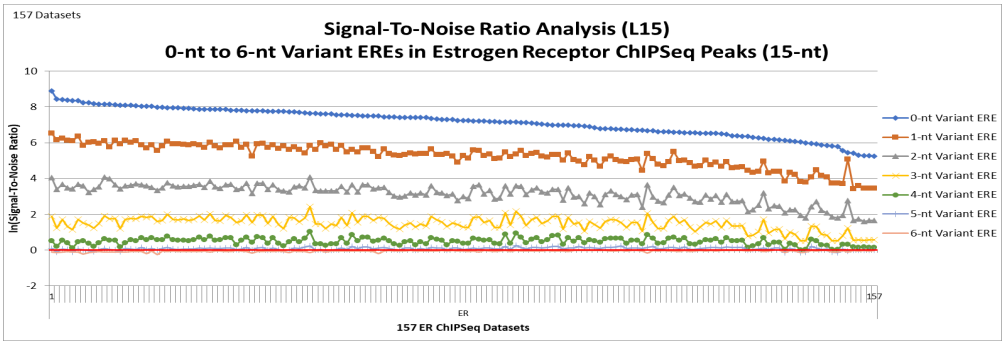
---- **Symmetry Divider**

Figure S37. (S/N) analysis of 0-nt to 6-nt Variant EREs in ER ChIPSeq Peaks (# of Variants)









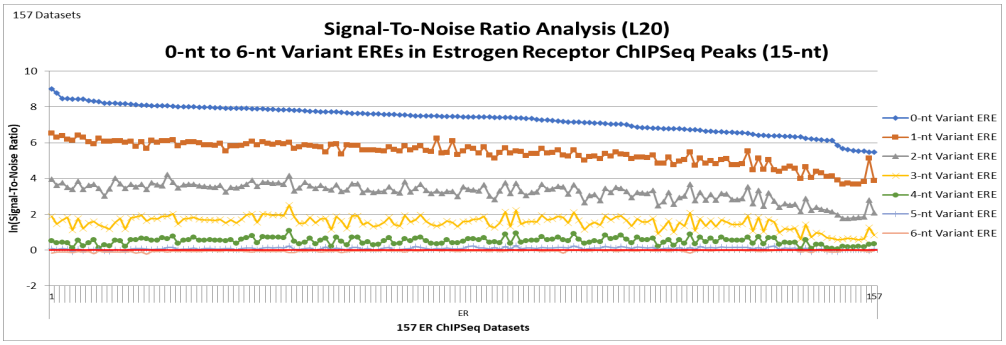
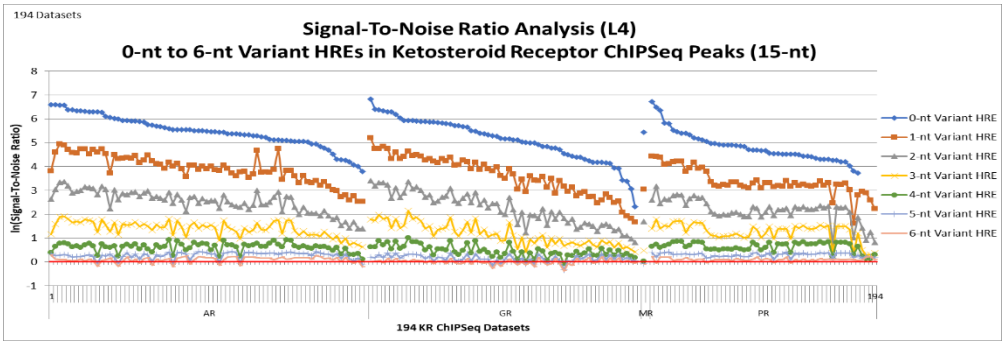
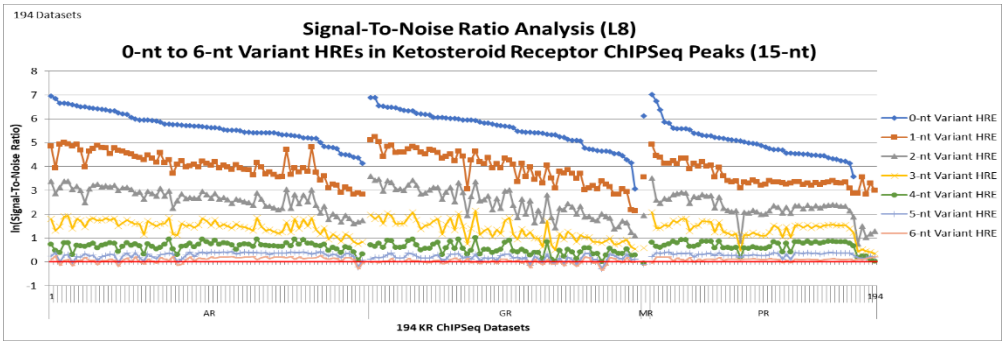
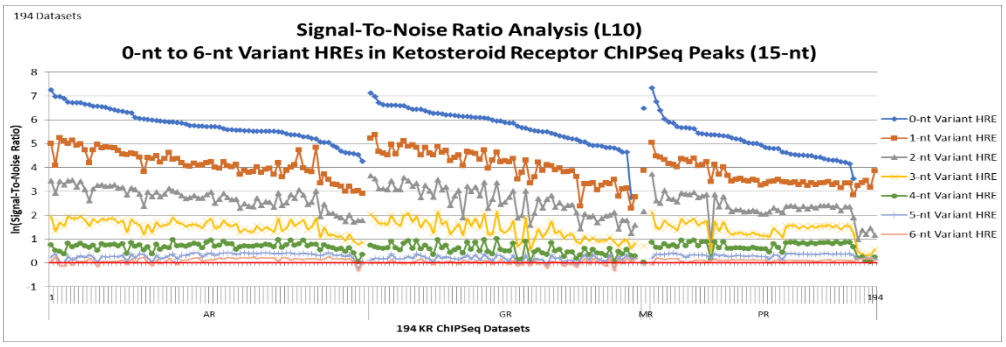
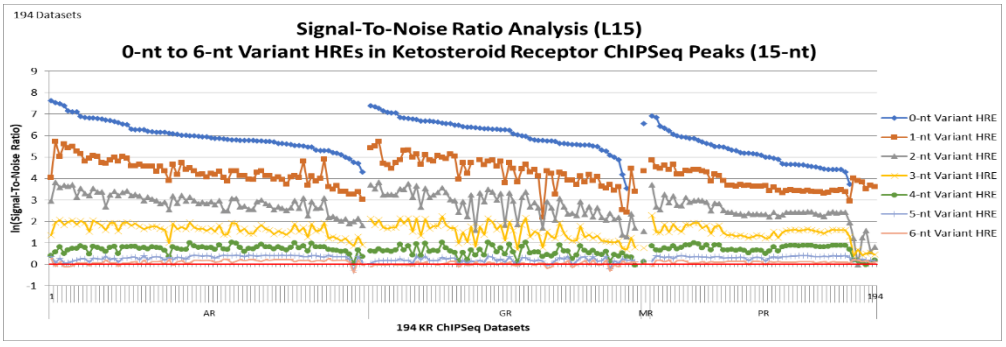


Figure S38. (S/N) analysis of 0-nt to 6-nt Variant HREs in KR ChIPSeq Peaks (# of Variants)









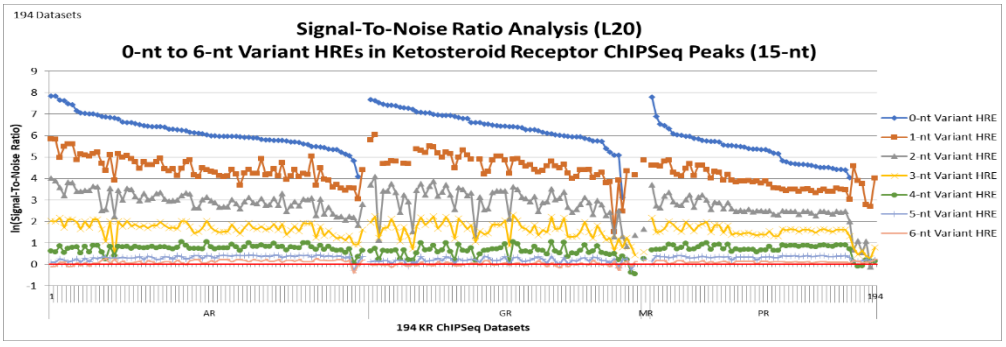


Figure S39. 924 Half-Site Groups

	A	B	C	D	E
	Fixed Half-Site Positions	Reverse-Complement Vacancies & Double Occupants	Reverse-Complement Vacancy Position ID	Reverse-Complement Double Occupant Position ID	Positions That Can Be Varied
			Zero Vacancies		
1	1-2-3-4-5-6------	0V			------7-8-9-10-11-12
2	1-2-3--5-6---9---	0V			---4---7-8--10-11-12
3	-2-3-4-5-6------12	0V			1------7-8-9-10-11-
4	1--3-4-5-6-----11-	0V			-2-----7-8-9-10--12
5	1-2-3-4-5--7-----	0V			-----6--8-9-10-11-12
6	1-2-3-4--6--8----	0V			----5--7--9-10-11-12
7	1-2--4-5-6----10--	0V			--3----7-8-9--11-12
8	-2-3--5-6---9---12	0V			1---4---7-8--10-11-
9	1--3--5-6---9--11-	0V			-2--4---7-8--10--12
10	1-2-3--5--7--9---	0V			---4--6--8--10-11-12
11	1-2-3---6--8-9---	0V			---4-5--7---10-11-12
12	1-2---5-6---9-10--	0V			--3-4---7-8---11-12
13	--3-4-5-6-----11-12	0V			1-2-----7-8-9-10--
14	-2-3-4-5--7-----12	0V			1-----6--8-9-10-11-
15	-2-3-4--6--8----12	0V			1----5--7--9-10-11-
16	-2--4-5-6----10--12	0V			1--3----7-8-9--11-
17	1--3-4-5--7----11-	0V			-2----6--8-9-10--12
18	1--3-4--6--8---11-	0V			-2---5--7--9-10--12
19	1---4-5-6----10-11-	0V			-2-3----7-8-9---12
20	1-2-3-4---7-8----	0V			----5-6---9-10-11-12
21	1-2--4-5--7---10--	0V			--3---6--8-9--11-12
22	1-2---4--6--8--10--	0V			--3--5--7--9--11-12
23	-2-3--5--7--9---12	0V			1---4--6--8--10-11-
24	1--3--5--7--9--11-	0V			-2--4--6--8--10--12
25	1-2-3----7-8-9---	0V			---4-5-6----10-11-12
26	1-2---5--7--9-10--	0V			--3-4--6--8---11-12
27	--3-4-5--7----11-12	0V			1-2----6--8-9-10--
28	-2-3-4---7-8----12	0V			1----5-6---9-10-11-
29	-2--4-5--7---10--12	0V			1--3---6--8-9--11-
30	1--3-4---7-8---11-	0V			-2---5-6---9-10--12
31	1---4-5--7---10-11-	0V			-2-3---6--8-9---12
32	1-2--4---7-8--10--	0V			--3--5-6---9--11-12
			One Vacancy		
33	1-2-3--5--7-----12	1V	4	1	---4--6--8-9-10-11-
34	1--3--5--7----11-12	1V	4	1	-2--4--6--8-9-10--
35	1-2-3----7-8----12	1V	4	1	---4-5-6---9-10-11-
36	1-2---5--7---10--12	1V	4	1	--3-4--6--8-9--11-
37	1--3----7-8---11-12	1V	4	1	-2--4-5-6---9-10--
38	1----5--7---10-11-12	1V	4	1	-2-3-4--6--8-9---
39	1-2-----7-8--10--12	1V	4	1	--3-4-5-6---9--11-
40	1------7-8--10-11-12	1V	4	1	-2-3-4-5-6---9---
41	1-2-3--5--7----11-	1V	4	2	---4--6--8-9-10--12
42	-2-3--5--7----11-12	1V	4	2	1---4--6--8-9-10--
43	1-2-3----7-8---11-	1V	4	2	---4-5-6---9-10--12
44	1-2---5--7---10-11-	1V	4	2	--3-4--6--8-9---12
45	-2-3----7-8---11-12	1V	4	2	1---4-5-6---9-10--
46	-2---5--7---10-11-12	1V	4	2	1--3-4--6--8-9---
47	1-2-----7-8--10-11-	1V	4	2	--3-4-5-6---9---12
48	-2-----7-8--10-11-12	1V	4	2	1--3-4-5-6---9---
49	1-2-3---6-7-8----	1V	4	6	---4-5----9-10-11-12
50	-2-3---6-7-8----12	1V	4	6	1---4-5----9-10-11-
51	1--3---6-7-8---11-	1V	4	6	-2--4-5----9-10--12
52	1-2----6-7-8--10--	1V	4	6	--3-4-5----9--11-12
53	--3---6-7-8---11-12	1V	4	6	1-2--4-5----9-10--
54	-2----6-7-8--10--12	1V	4	6	1--3-4-5----9--11-
55	1-----6-7-8--10-11-	1V	4	6	-2-3-4-5----9---12
56	-----6-7-8--10-11-12	1V	4	6	1-2-3-4-5----9---
57	1-2-3--5--7-8----	1V	4	5	---4--6---9-10-11-12
58	-2-3--5--7-8----12	1V	4	5	1---4--6---9-10-11-
59	1--3--5--7-8---11-	1V	4	5	-2--4--6---9-10--12
60	1-2---5--7-8--10--	1V	4	5	--3-4--6---9--11-12
61	--3--5--7-8---11-12	1V	4	5	1-2--4--6---9-10--
62	-2---5--7-8--10--12	1V	4	5	1--3-4--6---9--11-
63	1----5--7-8--10-11-	1V	4	5	-2-3-4--6---9---12
64	----5--7-8--10-11-12	1V	4	5	1-2-3-4--6---9---
65	1-2-3--5--7---10--	1V	4	3	---4--6--8-9--11-12
66	-2-3--5--7---10--12	1V	4	3	1---4--6--8-9--11-
67	1--3--5--7---10-11-	1V	4	3	-2--4--6--8-9---12
68	1-2-3----7-8--10--	1V	4	3	---4-5-6---9--11-12
69	--3--5--7---10-11-12	1V	4	3	1-2--4--6--8-9---
70	-2-3----7-8--10--12	1V	4	3	1---4-5-6---9--11-
71	1--3----7-8--10-11-	1V	4	3	-2--4-5-6---9---12
72	--3----7-8--10-11-12	1V	4	3	1-2--4-5-6---9---
73	-2-3-4-5--7--9---	1V	1	4	1-----6--8--10-11-12

74	--3-4-5--7--9--11-	1V	1	4	1-2----6--8--10--12
75	-2-3-4---7-8-9---	1V	1	4	1----5-6----10-11-12
76	-2--4-5--7--9-10--	1V	1	4	1--3---6--8---11-12
77	--3-4---7-8-9--11-	1V	1	4	1-2---5-6----10--12
78	---4-5--7--9-10-11-	1V	1	4	1-2-3---6--8----12
79	-2--4---7-8-9-10--	1V	1	4	1--3--5-6-----11-12
80	---4---7-8-9-10-11-	1V	1	4	1-2-3--5-6------12
81	-2-3-4-5--7----11-	1V	1	2	1-----6--8-9-10--12
82	-2-3--5--7--9--11-	1V	1	2	1---4--6--8--10--12
83	-2-3-4---7-8---11-	1V	1	2	1----5-6---9-10--12
84	-2--4-5--7---10-11-	1V	1	2	1--3---6--8-9---12
85	-2-3----7-8-9--11-	1V	1	2	1---4-5-6----10--12
86	-2---5--7--9-10-11-	1V	1	2	1--3-4--6--8----12
87	-2--4---7-8--10-11-	1V	1	2	1--3--5-6---9---12
88	-2-----7-8-9-10-11-	1V	1	2	1--3-4-5-6------12
89	-2-3-4--6-7-8----	1V	1	6	1----5----9-10-11-12
90	-2-3---6-7-8-9---	1V	1	6	1---4-5-----10-11-12
91	--3-4--6-7-8---11-	1V	1	6	1-2---5----9-10--12
92	-2--4--6-7-8--10--	1V	1	6	1--3--5----9--11-12
93	--3---6-7-8-9--11-	1V	1	6	1-2--4-5-----10--12
94	-2----6-7-8-9-10--	1V	1	6	1--3-4-5------11-12
95	---4--6-7-8--10-11-	1V	1	6	1-2-3--5----9---12
96	-----6-7-8-9-10-11-	1V	1	6	1-2-3-4-5-------12
97	-2-3-4-5--7-8----	1V	1	5	1-----6---9-10-11-12
98	-2-3--5--7-8-9---	1V	1	5	1---4--6----10-11-12
99	--3-4-5--7-8---11-	1V	1	5	1-2----6---9-10--12
100	-2--4-5--7-8--10--	1V	1	5	1--3---6---9--11-12
101	--3--5--7-8-9--11-	1V	1	5	1-2--4--6----10--12
102	-2---5--7-8-9-10--	1V	1	5	1--3-4--6-----11-12
103	---4-5--7-8--10-11-	1V	1	5	1-2-3---6---9---12
104	----5--7-8-9-10-11-	1V	1	5	1-2-3-4--6------12
105	-2-3-4-5--7---10--	1V	1	3	1-----6--8-9--11-12
106	-2-3--5--7--9-10--	1V	1	3	1---4--6--8---11-12
107	--3-4-5--7---10-11-	1V	1	3	1-2----6--8-9---12
108	-2-3-4---7-8--10--	1V	1	3	1----5-6---9--11-12
109	--3--5--7--9-10-11-	1V	1	3	1-2--4--6--8----12
110	-2-3----7-8-9-10--	1V	1	3	1---4-5-6-----11-12
111	--3-4---7-8--10-11-	1V	1	3	1-2---5-6---9---12
112	--3----7-8-9-10-11-	1V	1	3	1-2--4-5-6------12
113	1--3-4-5--7--9---	1V	2	4	-2----6--8--10-11-12
114	--3-4-5--7--9---12	1V	2	4	1-2----6--8--10-11-
115	1--3-4---7-8-9---	1V	2	4	-2---5-6----10-11-12
116	1---4-5--7--9-10--	1V	2	4	-2-3---6--8---11-12
117	--3-4---7-8-9---12	1V	2	4	1-2---5-6----10-11-
118	---4-5--7--9-10--12	1V	2	4	1-2-3---6--8---11-
119	1---4---7-8-9-10--	1V	2	4	-2-3--5-6-----11-12
120	---4---7-8-9-10--12	1V	2	4	1-2-3--5-6-----11-
121	1--3-4-5--7-----12	1V	2	1	-2----6--8-9-10-11-
122	1--3--5--7--9---12	1V	2	1	-2--4--6--8--10-11-
123	1--3-4---7-8----12	1V	2	1	-2---5-6---9-10-11-
124	1---4-5--7---10--12	1V	2	1	-2-3---6--8-9--11-
125	1--3----7-8-9---12	1V	2	1	-2--4-5-6----10-11-
126	1----5--7--9-10--12	1V	2	1	-2-3-4--6--8---11-
127	1---4---7-8--10--12	1V	2	1	-2-3--5-6---9--11-
128	1------7-8-9-10--12	1V	2	1	-2-3-4-5-6-----11-
129	1--3-4--6-7-8----	1V	2	6	-2---5----9-10-11-12
130	1--3---6-7-8-9---	1V	2	6	-2--4-5-----10-11-12
131	--3-4--6-7-8----12	1V	2	6	1-2---5----9-10-11-
132	1---4--6-7-8--10--	1V	2	6	-2-3--5----9--11-12
133	--3---6-7-8-9---12	1V	2	6	1-2--4-5-----10-11-
134	1-----6-7-8-9-10--	1V	2	6	-2-3-4-5------11-12
135	---4--6-7-8--10--12	1V	2	6	1-2-3--5----9--11-
136	-----6-7-8-9-10--12	1V	2	6	1-2-3-4-5------11-
137	1--3-4-5--7-8----	1V	2	5	-2----6---9-10-11-12
138	1--3--5--7-8-9---	1V	2	5	-2--4--6----10-11-12
139	--3-4-5--7-8----12	1V	2	5	1-2----6---9-10-11-
140	1---4-5--7-8--10--	1V	2	5	-2-3---6---9--11-12
141	--3--5--7-8-9---12	1V	2	5	1-2--4--6----10-11-
142	1----5--7-8-9-10--	1V	2	5	-2-3-4--6-----11-12
143	---4-5--7-8--10--12	1V	2	5	1-2-3---6---9--11-
144	----5--7-8-9-10--12	1V	2	5	1-2-3-4--6-----11-
145	1--3-4-5--7---10--	1V	2	3	-2----6--8-9--11-12
146	1--3--5--7--9-10--	1V	2	3	-2--4--6--8---11-12
147	--3-4-5--7---10--12	1V	2	3	1-2----6--8-9--11-
148	1--3-4---7-8--10--	1V	2	3	-2---5-6---9--11-12
149	--3--5--7--9-10--12	1V	2	3	1-2--4--6--8---11-
150	1--3----7-8-9-10--	1V	2	3	-2--4-5-6-----11-12
151	--3-4---7-8--10--12	1V	2	3	1-2---5-6---9--11-
152	--3----7-8-9-10--12	1V	2	3	1-2--4-5-6-----11-
153	1-2-3-4----8-9---	1V	6	4	----5-6-7---10-11-12
154	-2-3-4----8-9---12	1V	6	4	1----5-6-7---10-11-
155	1--3-4----8-9--11-	1V	6	4	-2---5-6-7---10--12
156	1-2--4----8-9-10--	1V	6	4	--3--5-6-7----11-12

157	--3-4----8-9--11-12	1V	6	4	1-2---5-6-7---10--
158	-2--4----8-9-10--12	1V	6	4	1--3--5-6-7----11-
159	1---4----8-9-10-11-	1V	6	4	-2-3--5-6-7-----12
160	---4----8-9-10-11-12	1V	6	4	1-2-3--5-6-7-----
161	1-2-3-4----8----12	1V	6	1	----5-6-7--9-10-11-
162	1-2-3-----8-9---12	1V	6	1	---4-5-6-7---10-11-
163	1--3-4----8---11-12	1V	6	1	-2---5-6-7--9-10--
164	1-2--4----8--10--12	1V	6	1	--3--5-6-7--9--11-
165	1--3-----8-9--11-12	1V	6	1	-2--4-5-6-7---10--
166	1-2------8-9-10--12	1V	6	1	--3-4-5-6-7----11-
167	1---4----8--10-11-12	1V	6	1	-2-3--5-6-7--9---
168	1-------8-9-10-11-12	1V	6	1	-2-3-4-5-6-7-----
169	1-2-3-4----8---11-	1V	6	2	----5-6-7--9-10--12
170	1-2-3-----8-9--11-	1V	6	2	---4-5-6-7---10--12
171	-2-3-4----8---11-12	1V	6	2	1----5-6-7--9-10--
172	1-2--4----8--10-11-	1V	6	2	--3--5-6-7--9---12
173	-2-3-----8-9--11-12	1V	6	2	1---4-5-6-7---10--
174	1-2------8-9-10-11-	1V	6	2	--3-4-5-6-7-----12
175	-2--4----8--10-11-12	1V	6	2	1--3--5-6-7--9---
176	-2------8-9-10-11-12	1V	6	2	1--3-4-5-6-7-----
177	1-2-3--5---8-9---	1V	6	5	---4--6-7---10-11-12
178	-2-3--5---8-9---12	1V	6	5	1---4--6-7---10-11-
179	1--3--5---8-9--11-	1V	6	5	-2--4--6-7---10--12
180	1-2---5---8-9-10--	1V	6	5	--3-4--6-7----11-12
181	--3--5---8-9--11-12	1V	6	5	1-2--4--6-7---10--
182	-2---5---8-9-10--12	1V	6	5	1--3-4--6-7----11-
183	1----5---8-9-10-11-	1V	6	5	-2-3-4--6-7-----12
184	----5---8-9-10-11-12	1V	6	5	1-2-3-4--6-7-----
185	1-2-3-4----8--10--	1V	6	3	----5-6-7--9--11-12
186	1-2-3-----8-9-10--	1V	6	3	---4-5-6-7----11-12
187	-2-3-4----8--10--12	1V	6	3	1----5-6-7--9--11-
188	1--3-4----8--10-11-	1V	6	3	-2---5-6-7--9---12
189	-2-3-----8-9-10--12	1V	6	3	1---4-5-6-7----11-
190	1--3-----8-9-10-11-	1V	6	3	-2--4-5-6-7-----12
191	--3-4----8--10-11-12	1V	6	3	1-2---5-6-7--9---
192	--3-----8-9-10-11-12	1V	6	3	1-2--4-5-6-7-----
193	1-2-3-4---7--9---	1V	5	4	----5-6--8--10-11-12
194	-2-3-4---7--9---12	1V	5	4	1----5-6--8--10-11-
195	1--3-4---7--9--11-	1V	5	4	-2---5-6--8--10--12
196	1-2--4---7--9-10--	1V	5	4	--3--5-6--8---11-12
197	--3-4---7--9--11-12	1V	5	4	1-2---5-6--8--10--
198	-2--4---7--9-10--12	1V	5	4	1--3--5-6--8---11-
199	1---4---7--9-10-11-	1V	5	4	-2-3--5-6--8----12
200	---4---7--9-10-11-12	1V	5	4	1-2-3--5-6--8----
201	1-2-3-4---7-----12	1V	5	1	----5-6--8-9-10-11-
202	1-2-3----7--9---12	1V	5	1	---4-5-6--8--10-11-
203	1--3-4---7----11-12	1V	5	1	-2---5-6--8-9-10--
204	1-2--4---7---10--12	1V	5	1	--3--5-6--8-9--11-
205	1--3----7--9--11-12	1V	5	1	-2--4-5-6--8--10--
206	1-2-----7--9-10--12	1V	5	1	--3-4-5-6--8---11-
207	1---4---7---10-11-12	1V	5	1	-2-3--5-6--8-9---
208	1------7--9-10-11-12	1V	5	1	-2-3-4-5-6--8----
209	1-2-3-4---7----11-	1V	5	2	----5-6--8-9-10--12
210	1-2-3----7--9--11-	1V	5	2	---4-5-6--8--10--12
211	-2-3-4---7----11-12	1V	5	2	1----5-6--8-9-10--
212	1-2--4---7---10-11-	1V	5	2	--3--5-6--8-9---12
213	-2-3----7--9--11-12	1V	5	2	1---4-5-6--8--10--
214	1-2-----7--9-10-11-	1V	5	2	--3-4-5-6--8----12
215	-2--4---7---10-11-12	1V	5	2	1--3--5-6--8-9---
216	-2-----7--9-10-11-12	1V	5	2	1--3-4-5-6--8----
217	1-2-3---6-7--9---	1V	5	6	---4-5---8--10-11-12
218	-2-3---6-7--9---12	1V	5	6	1---4-5---8--10-11-
219	1--3---6-7--9--11-	1V	5	6	-2--4-5---8--10--12
220	1-2----6-7--9-10--	1V	5	6	--3-4-5---8---11-12
221	--3---6-7--9--11-12	1V	5	6	1-2--4-5---8--10--
222	-2----6-7--9-10--12	1V	5	6	1--3-4-5---8---11-
223	1-----6-7--9-10-11-	1V	5	6	-2-3-4-5---8----12
224	-----6-7--9-10-11-12	1V	5	6	1-2-3-4-5---8----
225	1-2-3-4---7---10--	1V	5	3	----5-6--8-9--11-12
226	1-2-3----7--9-10--	1V	5	3	---4-5-6--8---11-12
227	-2-3-4---7---10--12	1V	5	3	1----5-6--8-9--11-
228	1--3-4---7---10-11-	1V	5	3	-2---5-6--8-9---12
229	-2-3----7--9-10--12	1V	5	3	1---4-5-6--8---11-
230	1--3----7--9-10-11-	1V	5	3	-2--4-5-6--8----12
231	--3-4---7---10-11-12	1V	5	3	1-2---5-6--8-9---
232	--3----7--9-10-11-12	1V	5	3	1-2--4-5-6--8----
233	1-2--4-5--7--9---	1V	3	4	--3---6--8--10-11-12
234	-2--4-5--7--9---12	1V	3	4	1--3---6--8--10-11-
235	1---4-5--7--9--11-	1V	3	4	-2-3---6--8--10--12
236	1-2--4---7-8-9---	1V	3	4	--3--5-6----10-11-12
237	---4-5--7--9--11-12	1V	3	4	1-2-3---6--8--10--
238	-2--4---7-8-9---12	1V	3	4	1--3--5-6----10-11-
239	1---4---7-8-9--11-	1V	3	4	-2-3--5-6----10--12

240	---4---7-8-9--11-12	1V	3	4	1-2-3--5-6----10--
241	1-2--4-5--7-----12	1V	3	1	--3---6--8-9-10-11-
242	1-2---5--7--9---12	1V	3	1	--3-4--6--8--10-11-
243	1---4-5--7----11-12	1V	3	1	-2-3---6--8-9-10--
244	1-2--4---7-8----12	1V	3	1	--3--5-6---9-10-11-
245	1----5--7--9--11-12	1V	3	1	-2-3-4--6--8--10--
246	1-2-----7-8-9---12	1V	3	1	--3-4-5-6----10-11-
247	1---4---7-8---11-12	1V	3	1	-2-3--5-6---9-10--
248	1------7-8-9--11-12	1V	3	1	-2-3-4-5-6----10--
249	1-2--4-5--7----11-	1V	3	2	--3---6--8-9-10--12
250	1-2---5--7--9--11-	1V	3	2	--3-4--6--8--10--12
251	-2--4-5--7----11-12	1V	3	2	1--3---6--8-9-10--
252	1-2--4---7-8---11-	1V	3	2	--3--5-6---9-10--12
253	-2---5--7--9--11-12	1V	3	2	1--3-4--6--8--10--
254	1-2-----7-8-9--11-	1V	3	2	--3-4-5-6----10--12
255	-2--4---7-8---11-12	1V	3	2	1--3--5-6---9-10--
256	-2-----7-8-9--11-12	1V	3	2	1--3-4-5-6----10--
257	1-2--4--6-7-8----	1V	3	6	--3--5----9-10-11-12
258	1-2----6-7-8-9---	1V	3	6	--3-4-5-----10-11-12
259	-2--4--6-7-8----12	1V	3	6	1--3--5----9-10-11-
260	1---4--6-7-8---11-	1V	3	6	-2-3--5----9-10--12
261	-2----6-7-8-9---12	1V	3	6	1--3-4-5-----10-11-
262	1-----6-7-8-9--11-	1V	3	6	-2-3-4-5-----10--12
263	---4--6-7-8---11-12	1V	3	6	1-2-3--5----9-10--
264	-----6-7-8-9--11-12	1V	3	6	1-2-3-4-5-----10--
265	1-2--4-5--7-8----	1V	3	5	--3---6---9-10-11-12
266	1-2---5--7-8-9---	1V	3	5	--3-4--6----10-11-12
267	-2--4-5--7-8----12	1V	3	5	1--3---6---9-10-11-
268	1---4-5--7-8---11-	1V	3	5	-2-3---6---9-10--12
269	-2---5--7-8-9---12	1V	3	5	1--3-4--6----10-11-
270	1----5--7-8-9--11-	1V	3	5	-2-3-4--6----10--12
271	---4-5--7-8---11-12	1V	3	5	1-2-3---6---9-10--
272	----5--7-8-9--11-12	1V	3	5	1-2-3-4--6----10--
			Two Vacancies		
273	-2-3---6-7-8---11-	2V	1-4	2-6	1---4-5----9-10--12
274	-2----6-7-8--10-11-	2V	1-4	2-6	1--3-4-5----9---12
275	-2-3--5--7-8---11-	2V	1-4	2-5	1---4--6---9-10--12
276	-2---5--7-8--10-11-	2V	1-4	2-5	1--3-4--6---9---12
277	-2-3--5--7---10-11-	2V	1-4	2-3	1---4--6--8-9---12
278	-2-3----7-8--10-11-	2V	1-4	2-3	1---4-5-6---9---12
279	-2-3--5-6-7-8----	2V	1-4	5-6	1---4-----9-10-11-12
280	--3--5-6-7-8---11-	2V	1-4	5-6	1-2--4-----9-10--12
281	-2-3---6-7-8--10--	2V	1-4	3-6	1---4-5----9--11-12
282	--3---6-7-8--10-11-	2V	1-4	3-6	1-2--4-5----9---12
283	-2-3--5--7-8--10--	2V	1-4	3-5	1---4--6---9--11-12
284	--3--5--7-8--10-11-	2V	1-4	3-5	1-2--4--6---9---12
285	1--3---6-7-8----12	2V	2-4	1-6	-2--4-5----9-10-11-
286	1-----6-7-8--10--12	2V	2-4	1-6	-2-3-4-5----9--11-
287	1--3---5--7-8----12	2V	2-4	1-5	-2--4--6---9-10-11-
288	1----5--7-8--10--12	2V	2-4	1-5	-2-3-4--6---9--11-
289	1--3--5--7---10--12	2V	2-4	1-3	-2--4--6--8-9--11-
290	1--3----7-8--10--12	2V	2-4	1-3	-2--4-5-6---9--11-
291	1--3--5-6-7-8----	2V	2-4	5-6	-2--4-----9-10-11-12
292	--3--5-6-7-8----12	2V	2-4	5-6	1-2--4-----9-10-11-
293	1--3---6-7-8--10--	2V	2-4	3-6	-2--4-5----9--11-12
294	--3---6-7-8--10--12	2V	2-4	3-6	1-2--4-5----9--11-
295	1--3--5--7-8--10--	2V	2-4	3-5	-2--4--6---9--11-12
296	--3--5--7-8--10--12	2V	2-4	3-5	1-2--4--6---9--11-
297	1-2-3-----8---11-12	2V	4-6	1-2	---4-5-6-7--9-10--
298	1-2------8--10-11-12	2V	4-6	1-2	--3-4-5-6-7--9---
299	1-2-3--5---8----12	2V	4-6	1-5	---4--6-7--9-10-11-
300	1--3--5---8---11-12	2V	4-6	1-5	-2--4--6-7--9-10--
301	1-2-3-----8--10--12	2V	4-6	1-3	---4-5-6-7--9--11-
302	1--3-----8--10-11-12	2V	4-6	1-3	-2--4-5-6-7--9---
303	1-2-3--5---8---11-	2V	4-6	2-5	---4--6-7--9-10--12
304	-2-3--5---8---11-12	2V	4-6	2-5	1---4--6-7--9-10--
305	1-2-3-----8--10-11-	2V	4-6	2-3	---4-5-6-7--9---12
306	-2-3-----8--10-11-12	2V	4-6	2-3	1---4-5-6-7--9---
307	1-2-3--5---8--10--	2V	4-6	3-5	---4--6-7--9--11-12
308	-2-3--5---8--10--12	2V	4-6	3-5	1---4--6-7--9--11-
309	1-2-3----7----11-12	2V	4-5	1-2	---4-5-6--8-9-10--
310	1-2-----7---10-11-12	2V	4-5	1-2	--3-4-5-6--8-9---
311	1-2-3---6-7-----12	2V	4-5	1-6	---4-5---8-9-10-11-
312	1--3---6-7----11-12	2V	4-5	1-6	-2--4-5---8-9-10--
313	1-2-3----7---10--12	2V	4-5	1-3	---4-5-6--8-9--11-
314	1--3----7---10-11-12	2V	4-5	1-3	-2--4-5-6--8-9---
315	1-2-3---6-7----11-	2V	4-5	2-6	---4-5---8-9-10--12
316	-2-3---6-7----11-12	2V	4-5	2-6	1---4-5---8-9-10--
317	1-2-3----7---10-11-	2V	4-5	2-3	---4-5-6--8-9---12
318	-2-3----7---10-11-12	2V	4-5	2-3	1---4-5-6--8-9---
319	1-2-3---6-7---10--	2V	4-5	3-6	---4-5---8-9--11-12
320	-2-3---6-7---10--12	2V	4-5	3-6	1---4-5---8-9--11-
321	1-2---5--7----11-12	2V	3-4	1-2	--3-4--6--8-9-10--

322	1-2-----7-8---11-12	2V	3-4	1-2	--3-4-5-6---9-10--
323	1-2----6-7-8----12	2V	3-4	1-6	--3-4-5----9-10-11-
324	1-----6-7-8---11-12	2V	3-4	1-6	-2-3-4-5----9-10--
325	1-2---5--7-8----12	2V	3-4	1-5	--3-4--6---9-10-11-
326	1----5--7-8---11-12	2V	3-4	1-5	-2-3-4--6---9-10--
327	1-2----6-7-8---11-	2V	3-4	2-6	--3-4-5----9-10--12
328	-2----6-7-8---11-12	2V	3-4	2-6	1--3-4-5----9-10--
329	1-2---5--7-8---11-	2V	3-4	2-5	--3-4--6---9-10--12
330	-2---5--7-8---11-12	2V	3-4	2-5	1--3-4--6---9-10--
331	1-2---5-6-7-8----	2V	3-4	5-6	--3-4-----9-10-11-12
332	-2---5-6-7-8----12	2V	3-4	5-6	1--3-4-----9-10-11-
333	--3-4--6-7-8-9---	2V	1-2	4-6	1-2---5-----10-11-12
334	---4--6-7-8-9-10--	2V	1-2	4-6	1-2-3--5------11-12
335	--3-4-5--7-8-9---	2V	1-2	4-5	1-2----6----10-11-12
336	---4-5--7-8-9-10--	2V	1-2	4-5	1-2-3---6-----11-12
337	--3-4-5--7--9-10--	2V	1-2	3-4	1-2----6--8---11-12
338	--3-4---7-8-9-10--	2V	1-2	3-4	1-2---5-6-----11-12
339	--3--5-6-7-8-9---	2V	1-2	5-6	1-2--4------10-11-12
340	----5-6-7-8-9-10--	2V	1-2	5-6	1-2-3-4-------11-12
341	--3-4--6-7-8--10--	2V	1-2	3-6	1-2---5----9--11-12
342	--3---6-7-8-9-10--	2V	1-2	3-6	1-2--4-5------11-12
343	--3-4-5--7-8--10--	2V	1-2	3-5	1-2----6---9--11-12
344	--3--5--7-8-9-10--	2V	1-2	3-5	1-2--4--6-----11-12
345	-2-3-4----8-9--11-	2V	1-6	2-4	1----5-6-7---10--12
346	-2--4----8-9-10-11-	2V	1-6	2-4	1--3--5-6-7-----12
347	-2-3-4-5---8-9---	2V	1-6	4-5	1-----6-7---10-11-12
348	--3-4-5---8-9--11-	2V	1-6	4-5	1-2----6-7---10--12
349	-2-3-4----8-9-10--	2V	1-6	3-4	1----5-6-7----11-12
350	--3-4----8-9-10-11-	2V	1-6	3-4	1-2---5-6-7-----12
351	-2-3--5---8-9--11-	2V	1-6	2-5	1---4--6-7---10--12
352	-2---5---8-9-10-11-	2V	1-6	2-5	1--3-4--6-7-----12
353	-2-3-4----8--10-11-	2V	1-6	2-3	1----5-6-7--9---12
354	-2-3-----8-9-10-11-	2V	1-6	2-3	1---4-5-6-7-----12
355	-2-3--5---8-9-10--	2V	1-6	3-5	1---4--6-7----11-12
356	--3--5---8-9-10-11-	2V	1-6	3-5	1-2--4--6-7-----12
357	-2-3-4---7--9--11-	2V	1-5	2-4	1----5-6--8--10--12
358	-2--4---7--9-10-11-	2V	1-5	2-4	1--3--5-6--8----12
359	-2-3-4--6-7--9---	2V	1-5	4-6	1----5---8--10-11-12
360	--3-4--6-7--9--11-	2V	1-5	4-6	1-2---5---8--10--12
361	-2-3-4---7--9-10--	2V	1-5	3-4	1----5-6--8---11-12
362	--3-4---7--9-10-11-	2V	1-5	3-4	1-2---5-6--8----12
363	-2-3---6-7--9--11-	2V	1-5	2-6	1---4-5---8--10--12
364	-2----6-7--9-10-11-	2V	1-5	2-6	1--3-4-5---8----12
365	-2-3-4---7---10-11-	2V	1-5	2-3	1----5-6--8-9---12
366	-2-3----7--9-10-11-	2V	1-5	2-3	1---4-5-6--8----12
367	-2-3---6-7--9-10--	2V	1-5	3-6	1---4-5---8---11-12
368	--3---6-7--9-10-11-	2V	1-5	3-6	1-2--4-5---8----12
369	-2--4-5--7--9--11-	2V	1-3	2-4	1--3---6--8--10--12
370	-2--4---7-8-9--11-	2V	1-3	2-4	1--3--5-6----10--12
371	-2--4--6-7-8-9---	2V	1-3	4-6	1--3--5-----10-11-12
372	---4--6-7-8-9--11-	2V	1-3	4-6	1-2-3--5-----10--12
373	-2--4-5--7-8-9---	2V	1-3	4-5	1--3---6----10-11-12
374	---4-5--7-8-9--11-	2V	1-3	4-5	1-2-3---6----10--12
375	-2--4--6-7-8---11-	2V	1-3	2-6	1--3--5----9-10--12
376	-2----6-7-8-9--11-	2V	1-3	2-6	1--3-4-5-----10--12
377	-2--4-5--7-8---11-	2V	1-3	2-5	1--3---6---9-10--12
378	-2---5--7-8-9--11-	2V	1-3	2-5	1--3-4--6----10--12
379	-2---5-6-7-8-9---	2V	1-3	5-6	1--3-4------10-11-12
380	----5-6-7-8-9--11-	2V	1-3	5-6	1-2-3-4------10--12
381	1--3-4----8-9---12	2V	2-6	1-4	-2---5-6-7---10-11-
382	1---4----8-9-10--12	2V	2-6	1-4	-2-3--5-6-7----11-
383	1--3-4-5---8-9---	2V	2-6	4-5	-2----6-7---10-11-12
384	--3-4-5---8-9---12	2V	2-6	4-5	1-2----6-7---10-11-
385	1--3-4----8-9-10--	2V	2-6	3-4	-2---5-6-7----11-12
386	--3-4----8-9-10--12	2V	2-6	3-4	1-2---5-6-7----11-
387	1--3--5---8-9---12	2V	2-6	1-5	-2--4--6-7---10-11-
388	1----5---8-9-10--12	2V	2-6	1-5	-2-3-4--6-7----11-
389	1--3-4----8--10--12	2V	2-6	1-3	-2---5-6-7--9--11-
390	1--3-----8-9-10--12	2V	2-6	1-3	-2--4-5-6-7----11-
391	1--3--5---8-9-10--	2V	2-6	3-5	-2--4--6-7----11-12
392	--3--5---8-9-10--12	2V	2-6	3-5	1-2--4--6-7----11-
393	1--3-4---7--9---12	2V	2-5	1-4	-2---5-6--8--10-11-
394	1---4---7--9-10--12	2V	2-5	1-4	-2-3--5-6--8---11-
395	1--3-4--6-7--9---	2V	2-5	4-6	-2---5---8--10-11-12
396	--3-4--6-7--9---12	2V	2-5	4-6	1-2---5---8--10-11-
397	1--3-4---7--9-10--	2V	2-5	3-4	-2---5-6--8---11-12
398	--3-4---7--9-10--12	2V	2-5	3-4	1-2---5-6--8---11-
399	1--3---6-7--9---12	2V	2-5	1-6	-2--4-5---8--10-11-
400	1-----6-7--9-10--12	2V	2-5	1-6	-2-3-4-5---8---11-
401	1--3-4---7---10--12	2V	2-5	1-3	-2---5-6--8-9--11-
402	1--3----7--9-10--12	2V	2-5	1-3	-2--4-5-6--8---11-
403	1--3---6-7--9-10--	2V	2-5	3-6	-2--4-5---8---11-12
404	--3---6-7--9-10--12	2V	2-5	3-6	1-2--4-5---8---11-

405	1---4-5--7--9---12	2V	2-3	1-4	-2-3---6--8--10-11-
406	1---4---7-8-9---12	2V	2-3	1-4	-2-3--5-6----10-11-
407	1---4--6-7-8-9---	2V	2-3	4-6	-2-3--5-----10-11-12
408	---4--6-7-8-9---12	2V	2-3	4-6	1-2-3--5-----10-11-
409	1---4-5--7-8-9---	2V	2-3	4-5	-2-3---6----10-11-12
410	---4-5--7-8-9---12	2V	2-3	4-5	1-2-3---6----10-11-
411	1---4--6-7-8----12	2V	2-3	1-6	-2-3--5----9-10-11-
412	1-----6-7-8-9---12	2V	2-3	1-6	-2-3-4-5-----10-11-
413	1---4-5--7-8----12	2V	2-3	1-5	-2-3---6---9-10-11-
414	1----5--7-8-9---12	2V	2-3	1-5	-2-3-4--6----10-11-
415	1----5-6-7-8-9---	2V	2-3	5-6	-2-3-4------10-11-12
416	----5-6-7-8-9---12	2V	2-3	5-6	1-2-3-4------10-11-
417	1-2-3-4-----9---12	2V	5-6	1-4	----5-6-7-8--10-11-
418	1--3-4-----9--11-12	2V	5-6	1-4	-2---5-6-7-8--10--
419	1-2-3-4-----9--11-	2V	5-6	2-4	----5-6-7-8--10--12
420	-2-3-4-----9--11-12	2V	5-6	2-4	1----5-6-7-8--10--
421	1-2-3-4-----9-10--	2V	5-6	3-4	----5-6-7-8---11-12
422	-2-3-4-----9-10--12	2V	5-6	3-4	1----5-6-7-8---11-
423	1-2-3------9--11-12	2V	5-6	1-2	---4-5-6-7-8--10--
424	1-2-------9-10-11-12	2V	5-6	1-2	--3-4-5-6-7-8----
425	1-2-3------9-10--12	2V	5-6	1-3	---4-5-6-7-8---11-
426	1--3------9-10-11-12	2V	5-6	1-3	-2--4-5-6-7-8----
427	1-2-3------9-10-11-	2V	5-6	2-3	---4-5-6-7-8----12
428	-2-3------9-10-11-12	2V	5-6	2-3	1---4-5-6-7-8----
429	1-2--4----8-9---12	2V	3-6	1-4	--3--5-6-7---10-11-
430	1---4----8-9--11-12	2V	3-6	1-4	-2-3--5-6-7---10--
431	1-2--4----8-9--11-	2V	3-6	2-4	--3--5-6-7---10--12
432	-2--4----8-9--11-12	2V	3-6	2-4	1--3--5-6-7---10--
433	1-2--4-5---8-9---	2V	3-6	4-5	--3---6-7---10-11-12
434	-2--4-5---8-9---12	2V	3-6	4-5	1--3---6-7---10-11-
435	1-2--4----8---11-12	2V	3-6	1-2	--3--5-6-7--9-10--
436	1-2------8-9--11-12	2V	3-6	1-2	--3-4-5-6-7---10--
437	1-2---5---8-9---12	2V	3-6	1-5	--3-4--6-7---10-11-
438	1----5---8-9--11-12	2V	3-6	1-5	-2-3-4--6-7---10--
439	1-2---5---8-9--11-	2V	3-6	2-5	--3-4--6-7---10--12
440	-2---5---8-9--11-12	2V	3-6	2-5	1--3-4--6-7---10--
441	1-2--4---7--9---12	2V	3-5	1-4	--3--5-6--8--10-11-
442	1---4---7--9--11-12	2V	3-5	1-4	-2-3--5-6--8--10--
443	1-2--4---7--9--11-	2V	3-5	2-4	--3--5-6--8--10--12
444	-2--4---7--9--11-12	2V	3-5	2-4	1--3--5-6--8--10--
445	1-2--4--6-7--9---	2V	3-5	4-6	--3--5---8--10-11-12
446	-2--4--6-7--9---12	2V	3-5	4-6	1--3--5---8--10-11-
447	1-2--4---7----11-12	2V	3-5	1-2	--3--5-6--8-9-10--
448	1-2-----7--9--11-12	2V	3-5	1-2	--3-4-5-6--8--10--
449	1-2----6-7--9---12	2V	3-5	1-6	--3-4-5---8--10-11-
450	1-----6-7--9--11-12	2V	3-5	1-6	-2-3-4-5---8--10--
451	1-2----6-7--9--11-	2V	3-5	2-6	--3-4-5---8--10--12
452	-2----6-7--9--11-12	2V	3-5	2-6	1--3-4-5---8--10--
Three Vacancies					
453	1----5-6-7-8----12	3V	2-3-4	1-5-6	-2-3-4-----9-10-11-
454	--3--5-6-7-8--10--	3V	1-2-4	3-5-6	1-2--4-----9--11-12
455	-2-3---6-7---10-11-	3V	1-4-5	2-3-6	1---4-5---8-9---12
456	1-2---5---8---11-12	3V	3-4-6	1-2-5	--3-4--6-7--9-10--
457	1-2-3-------10-11-12	3V	4-5-6	1-2-3	---4-5-6-7-8-9---
458	-2--4-5---8-9--11-	3V	1-3-6	2-4-5	1--3---6-7---10--12
459	1-2----6-7----11-12	3V	3-4-5	1-2-6	--3-4-5---8-9-10--
460	1---4--6-7--9---12	3V	2-3-5	1-4-6	-2-3--5---8--10-11-
461	-2--4--6-7--9--11-	3V	1-3-5	2-4-6	1--3--5---8--10--12
462	1--3-4-----9-10--12	3V	2-5-6	1-3-4	-2---5-6-7-8---11-
463	1-2--4-----9--11-12	3V	3-5-6	1-2-4	--3--5-6-7-8--10--
464	1---4-5---8-9---12	3V	2-3-6	1-4-5	-2-3---6-7---10-11-
465	-2-3-4-----9-10-11-	3V	1-5-6	2-3-4	1----5-6-7-8----12
466	---4-5-6-7-8-9---	3V	1-2-3	4-5-6	1-2-3-------10-11-12
467	--3-4--6-7--9-10--	3V	1-2-5	3-4-6	1-2---5---8---11-12
468	--3-4-5---8-9-10--	3V	1-2-6	3-4-5	1-2----6-7----11-12
469	1--3---6-7---10--12	3V	2-4-5	1-3-6	-2--4-5---8-9--11-
470	1--3--5---8--10--12	3V	2-4-6	1-3-5	-2--4--6-7--9--11-
471	-2---5-6-7-8---11-	3V	1-3-4	2-5-6	1--3-4-----9-10--12
472	-2-3--5---8--10-11-	3V	1-4-6	2-3-5	1---4--6-7--9---12

Figure S40. (S/N) analysis of 0-nt to 6-nt Variant EREs in ER ChIPSeq Peaks (924 Half-Site Groups)

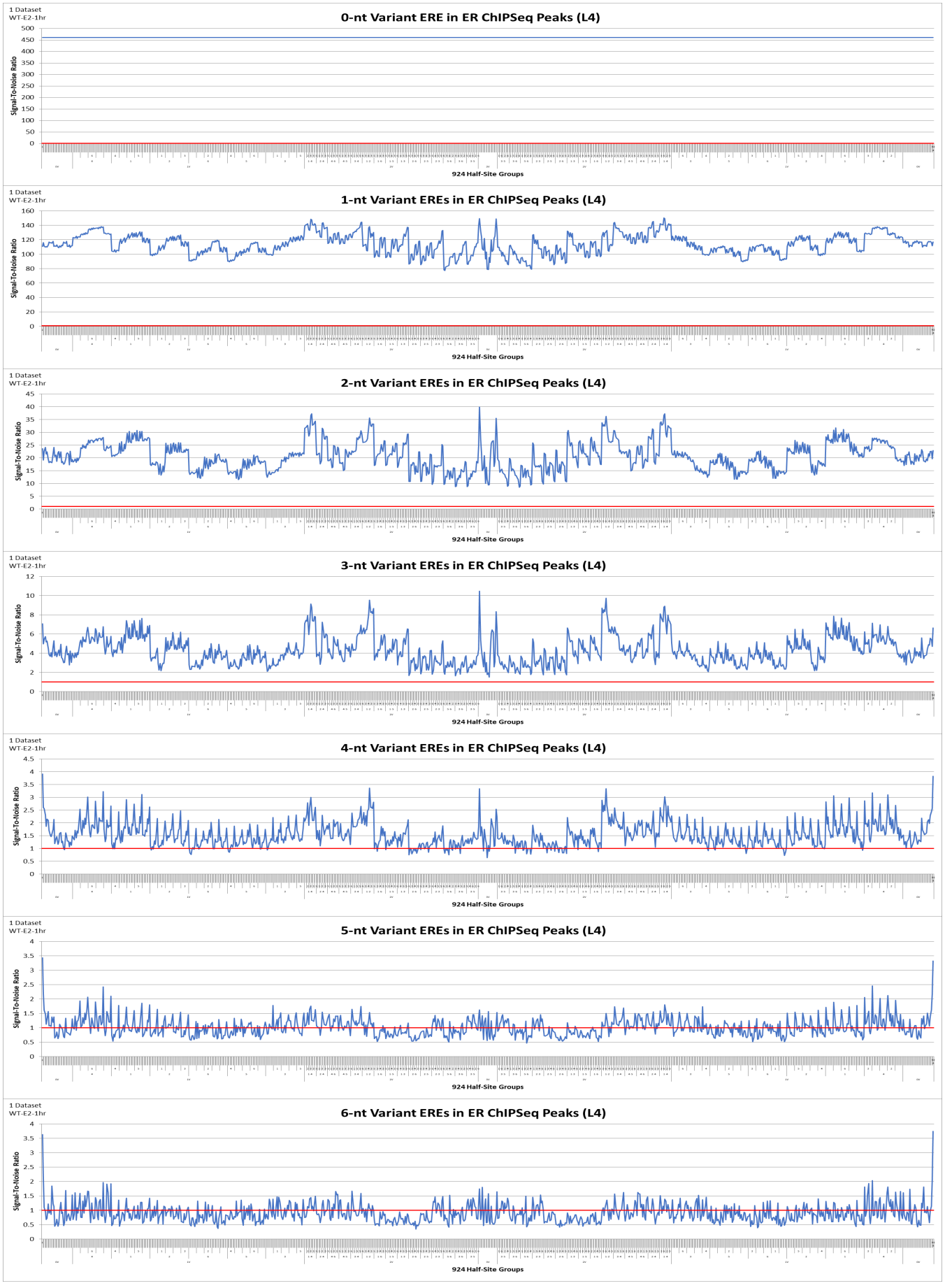


Figure S41. (S/N) analysis of 0-nt to 6-nt Variant HREs in AR ChIPSeq Peaks (924 Half-Site Groups)

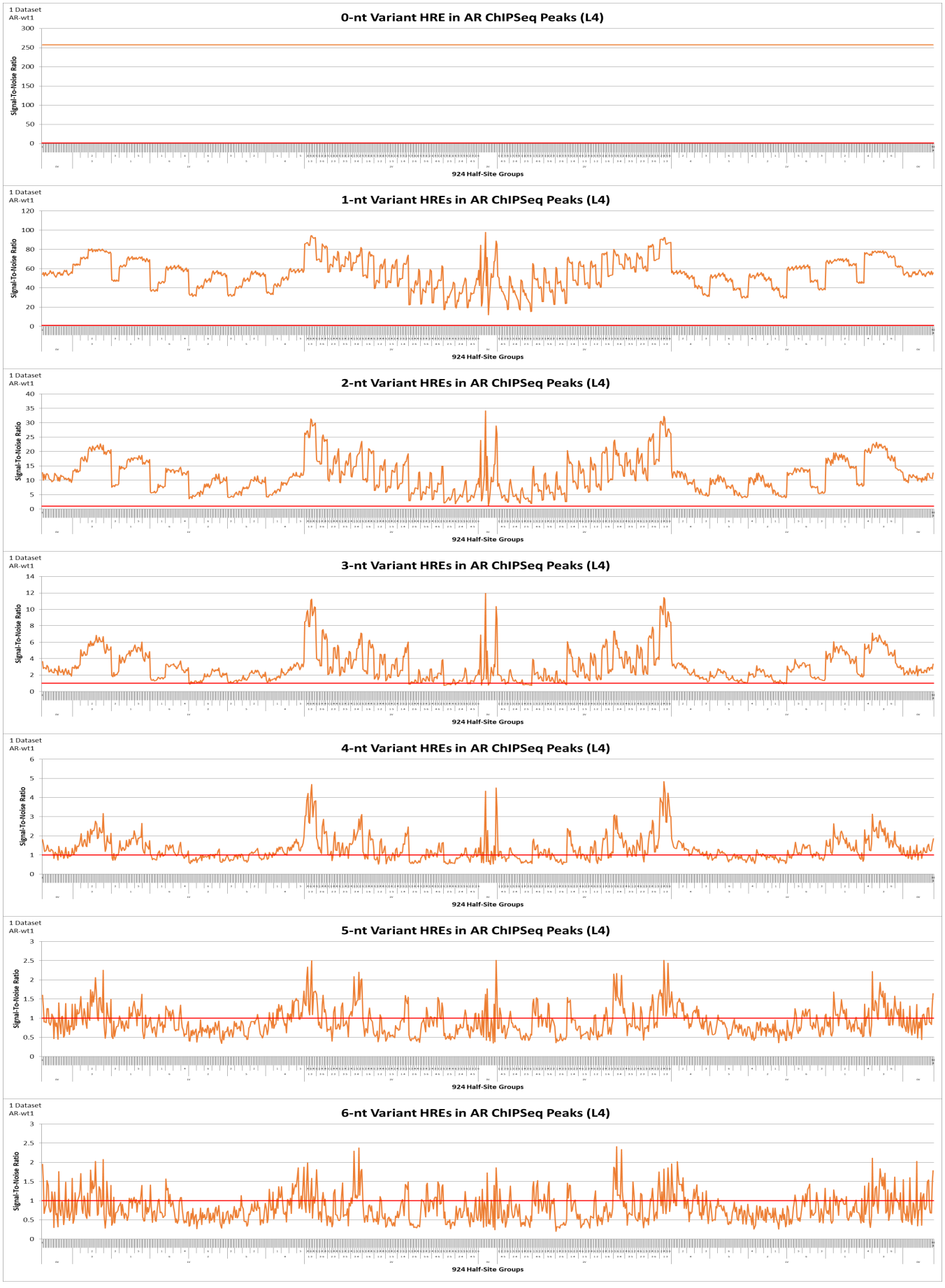
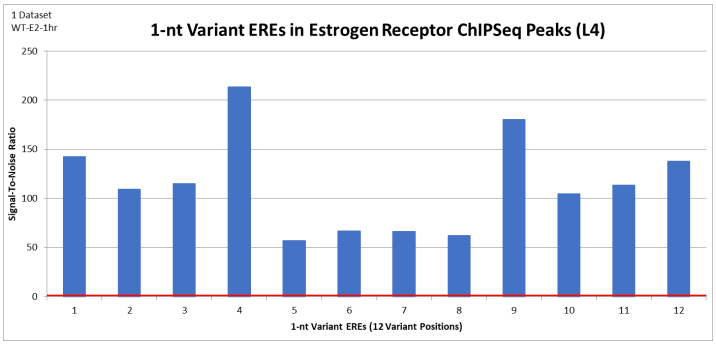
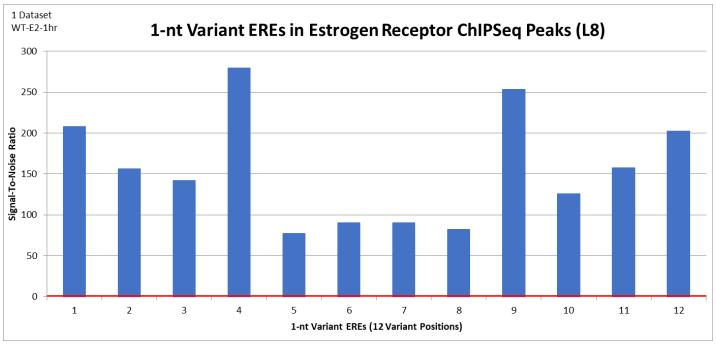


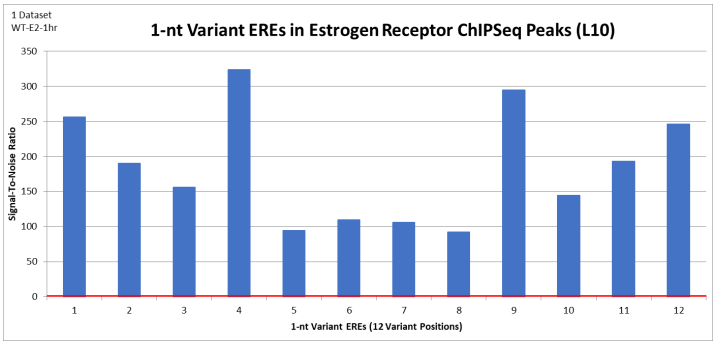
Figure S42. (S/N) analysis of 0-nt to 6-nt Variant HREs in GR ChIPSeq Peaks (924 Half-Site Groups)

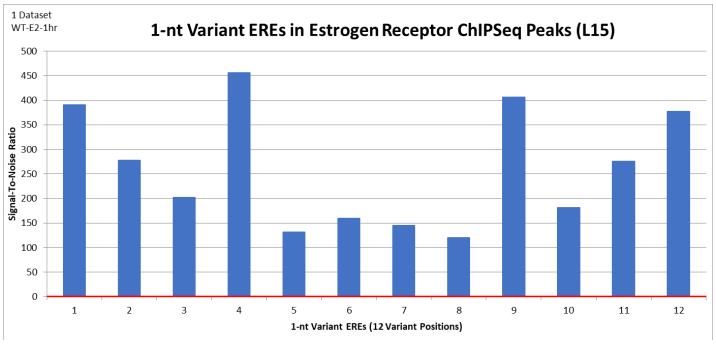
Figure S43. (S/N) analysis of 0-nt to 6-nt Variant HREs in PR ChIPSeq Peaks (924 Half-Site Groups)

Figure S44. (S/N) analysis of 1-nt Variant EREs in ER ChIPSeq Peaks (Variant Position)









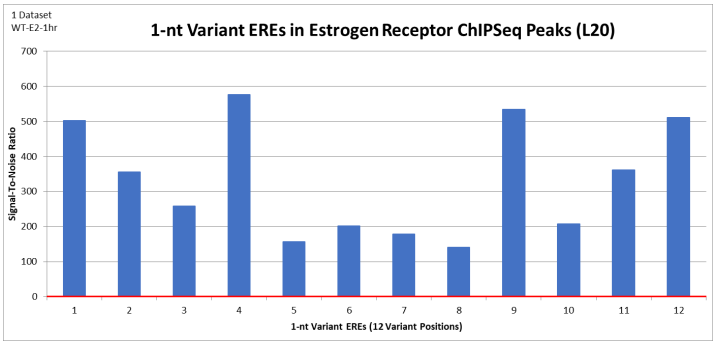
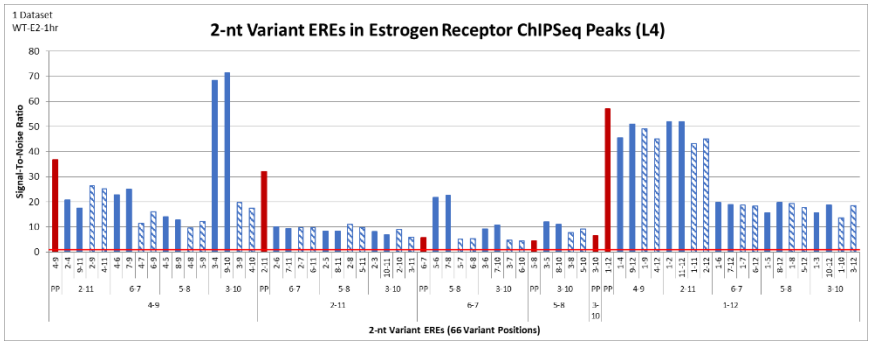
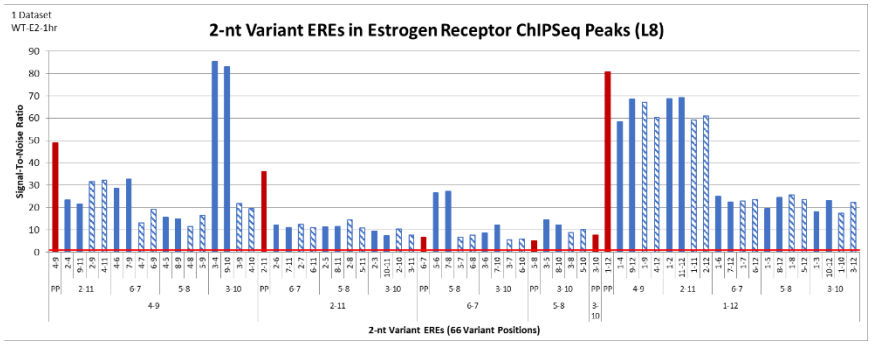
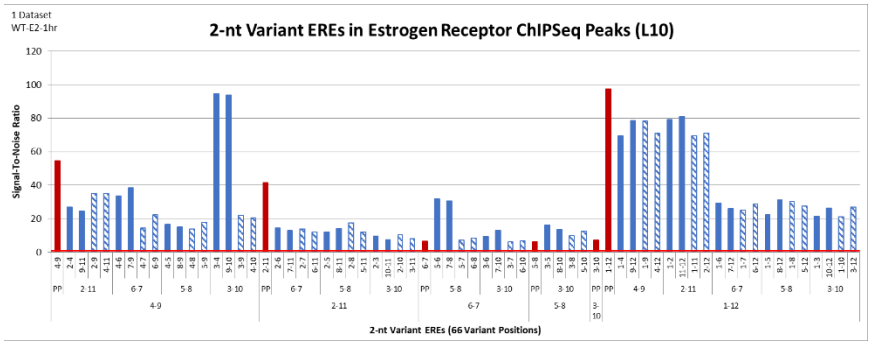
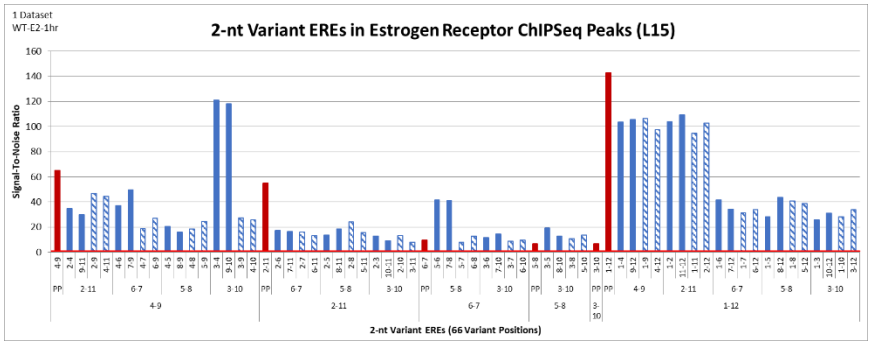


Figure S45. (S/N) analysis of 2-nt Variant EREs in ER ChIPSeq Peaks (Variant Position)









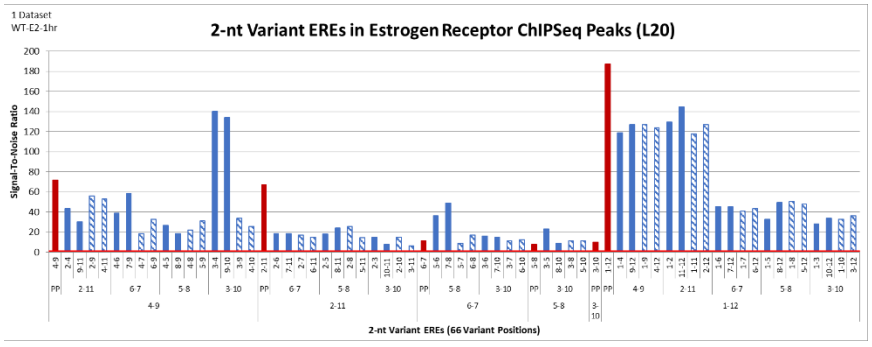
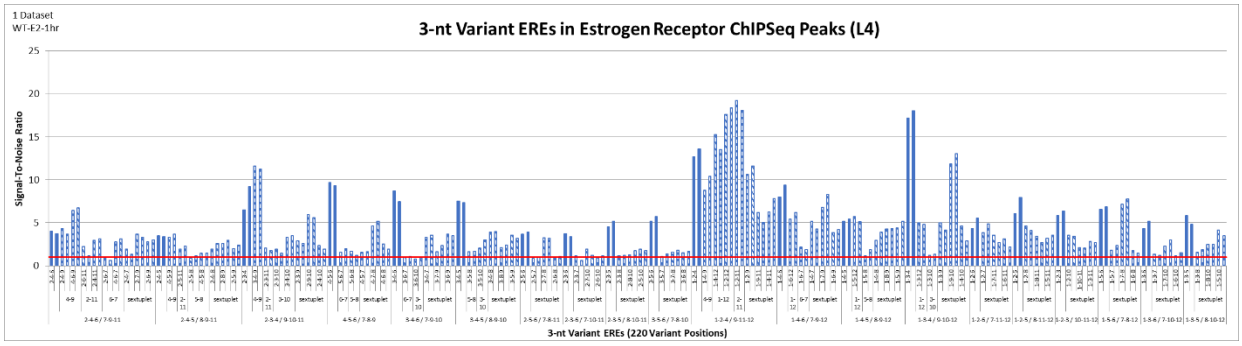
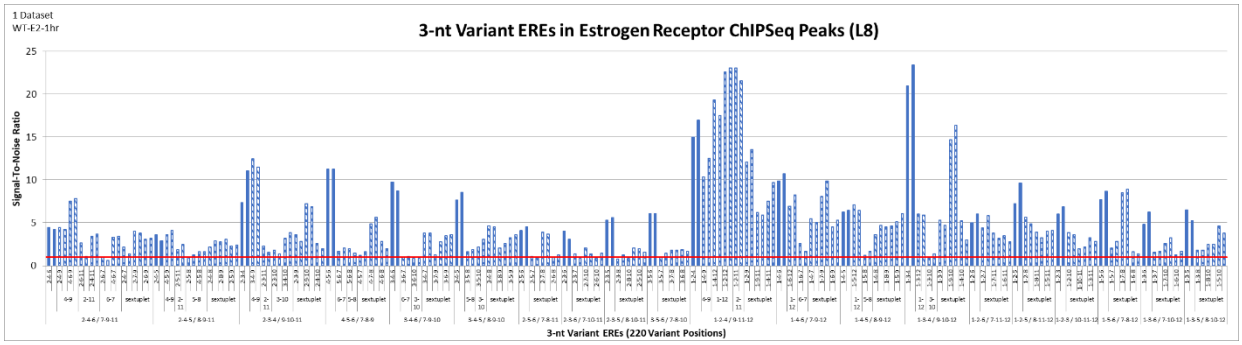
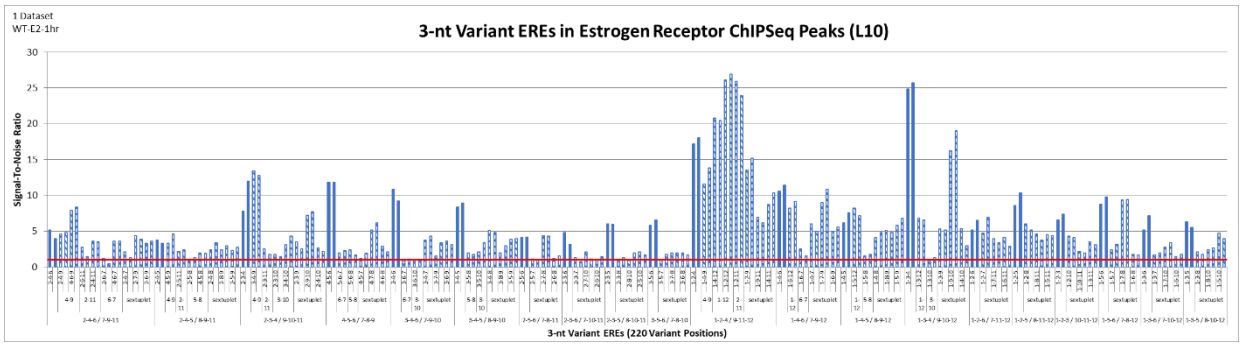
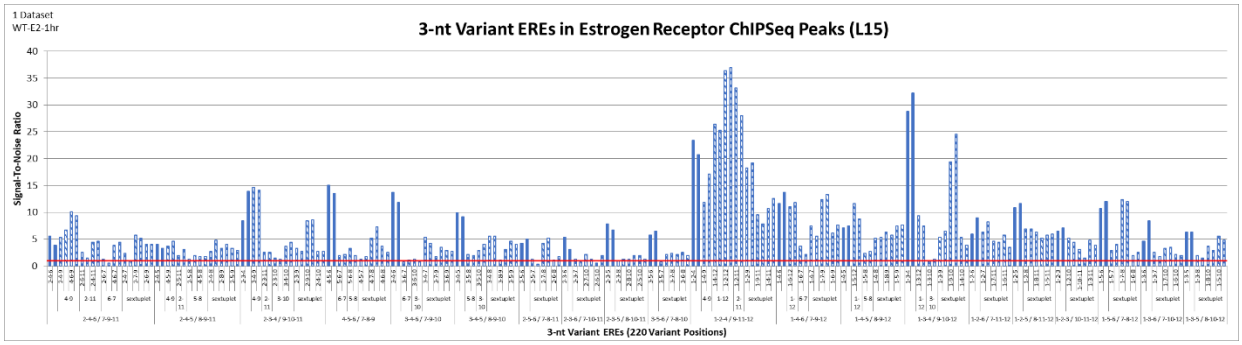


Figure S46. (S/N) analysis of 3-nt Variant EREs in ER ChIPSeq Peaks (Variant Position)









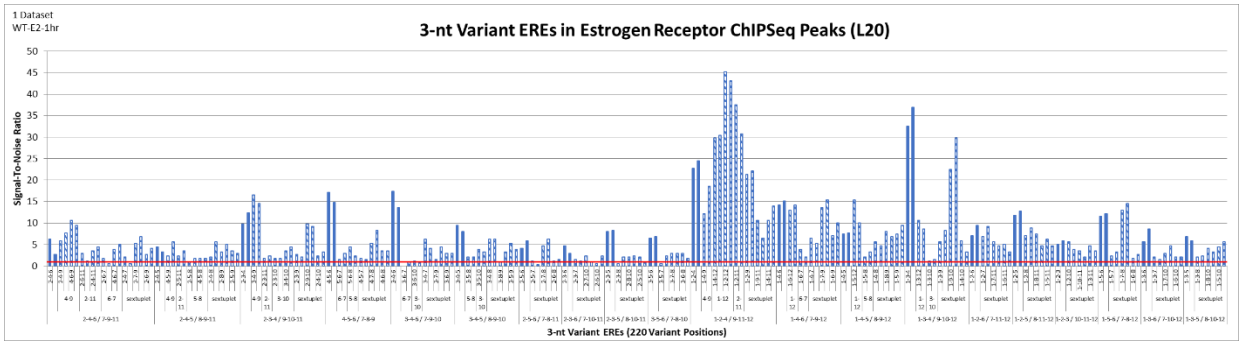
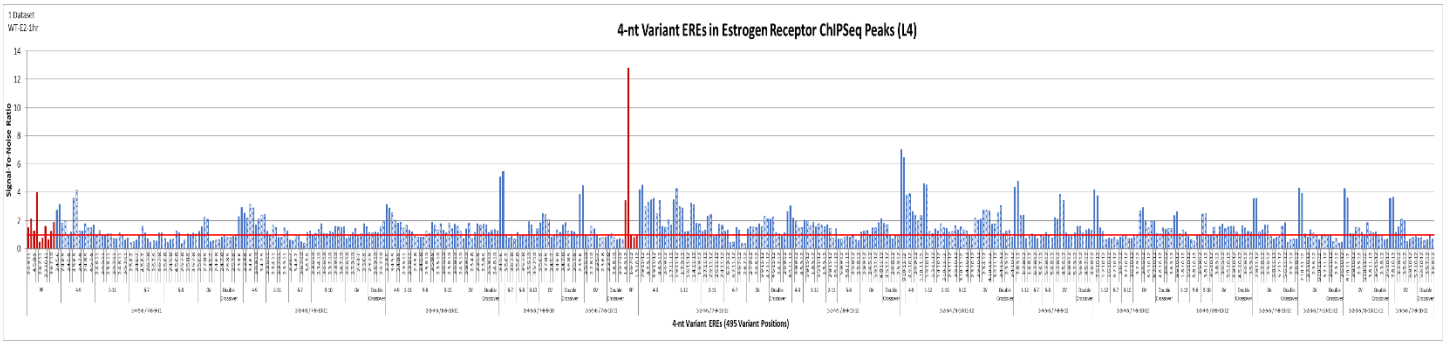
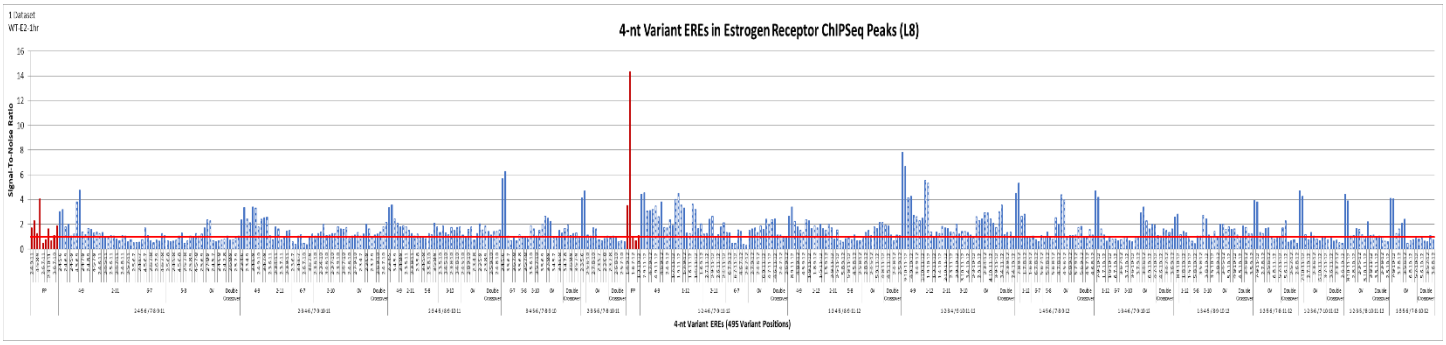
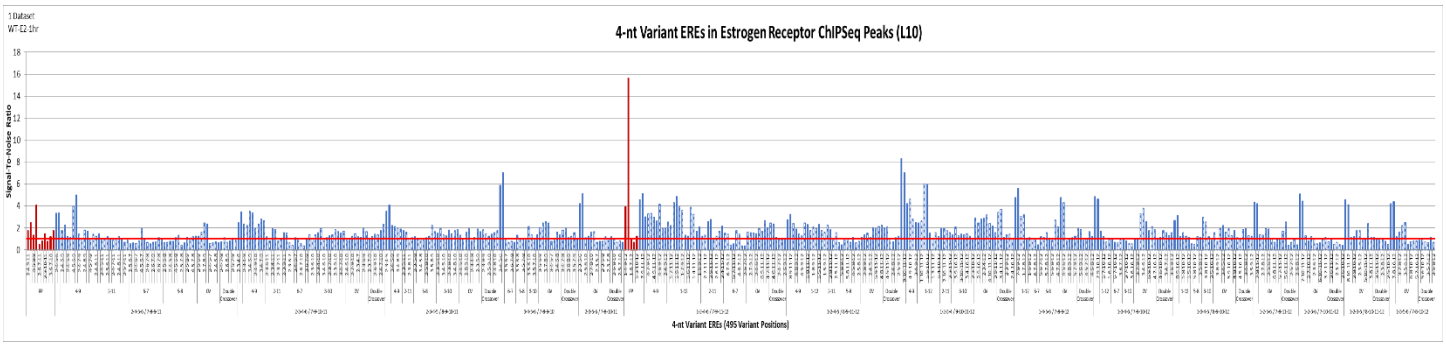
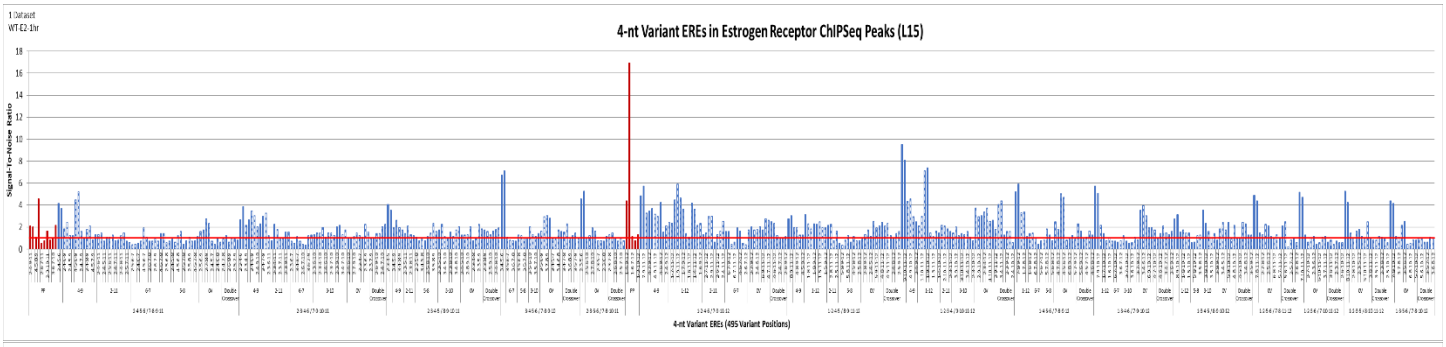


Figure S47. (S/N) analysis of 4-nt Variant EREs in ER ChIPSeq Peaks (Variant Position)









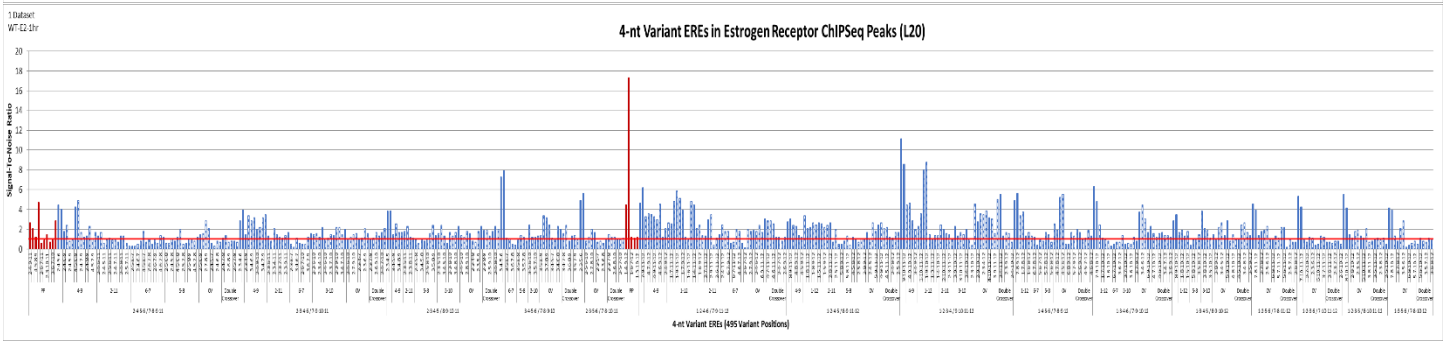


Figure S48. (S/N) analysis of 5-nt Variant EREs in ER ChIPSeq Peaks (Variant Position)

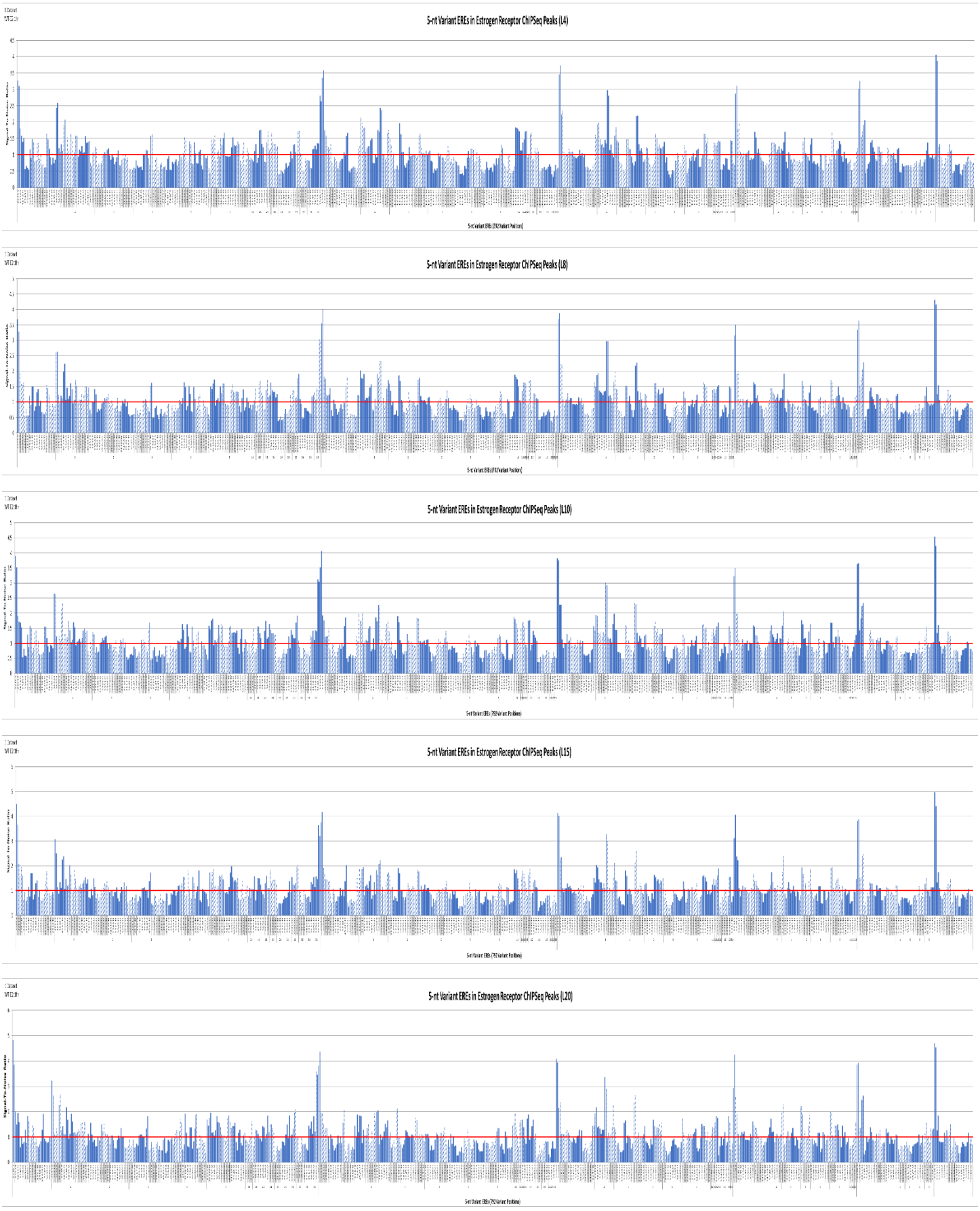
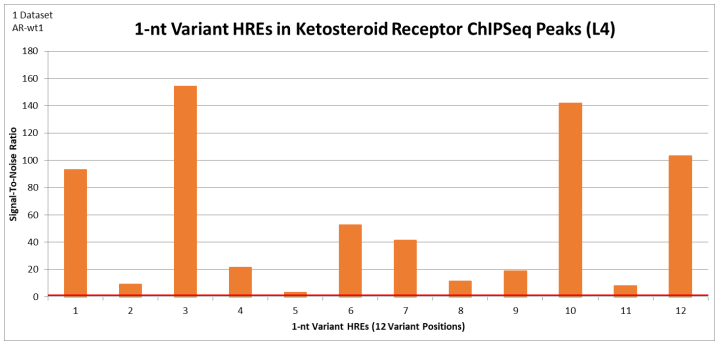
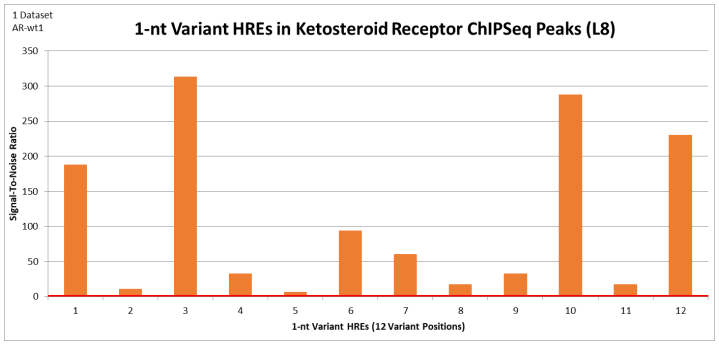
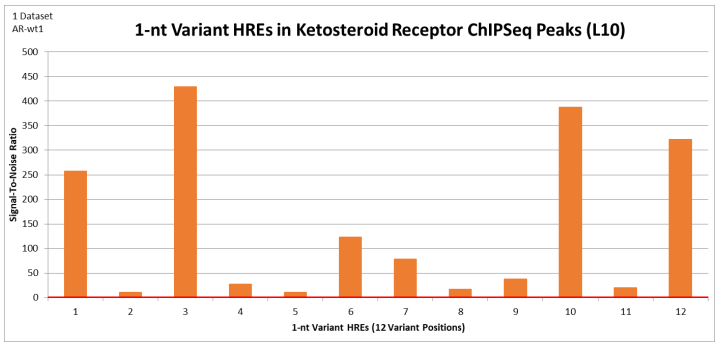


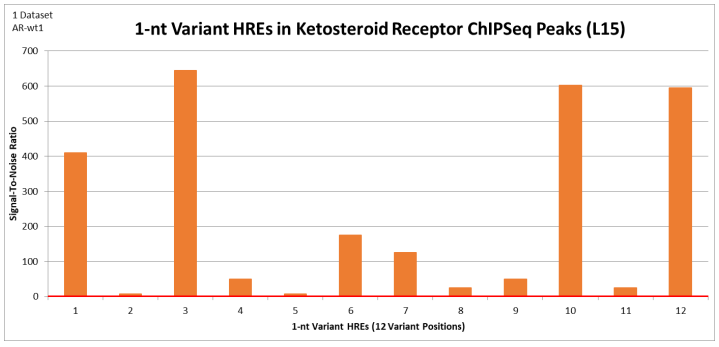
Figure S49. (S/N) analysis of 6-nt Variant EREs in ER ChIPSeq Peaks (Variant Position)

Figure S50. (S/N) analysis of 1-nt Variant HREs in KR ChIPSeq Peaks (Variant Position)









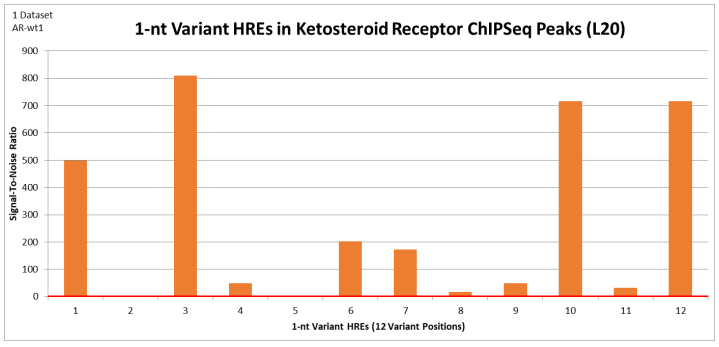
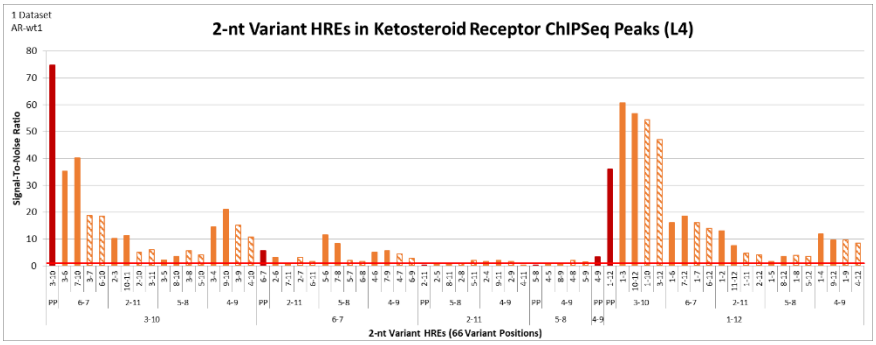
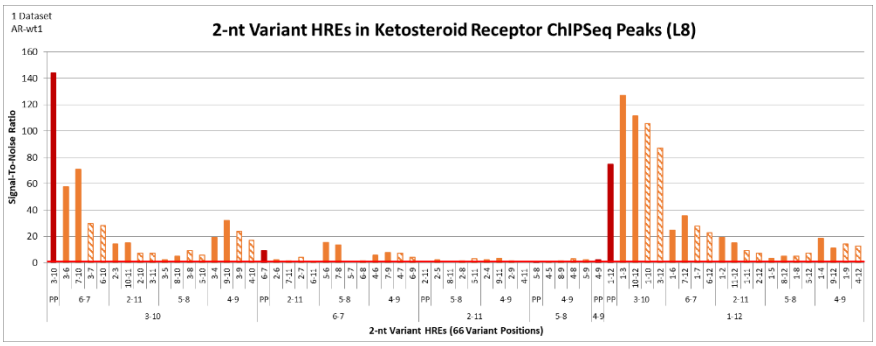
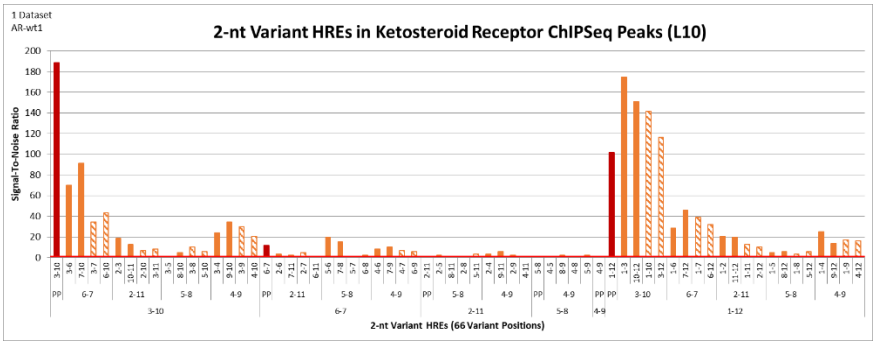
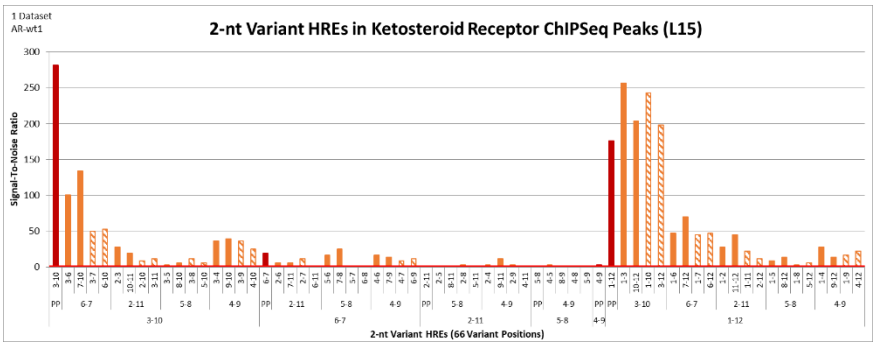


Figure S51. (S/N) analysis of 2-nt Variant HREs in KR ChIPSeq Peaks (Variant Position)









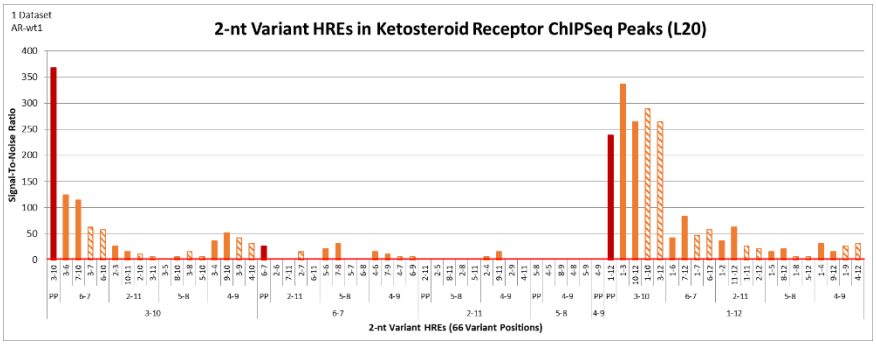
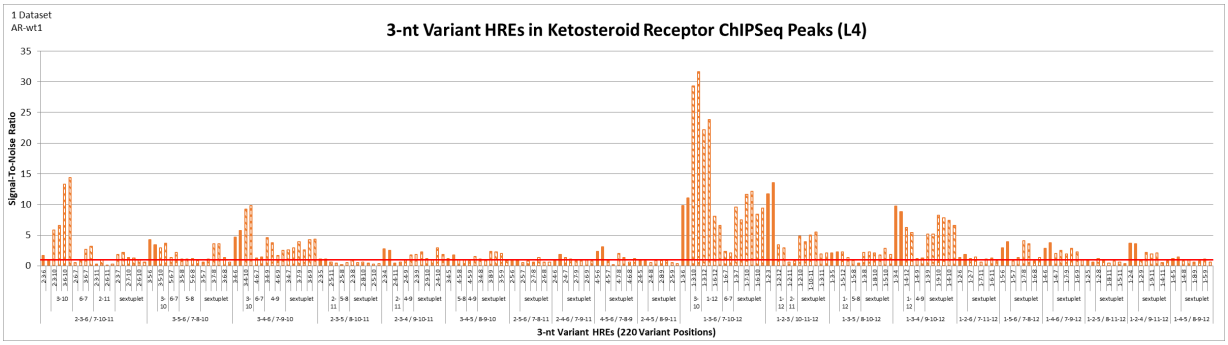
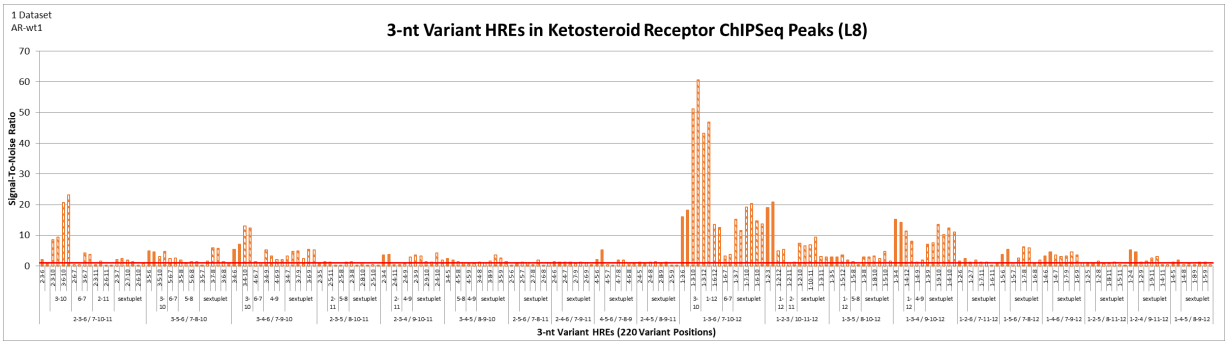
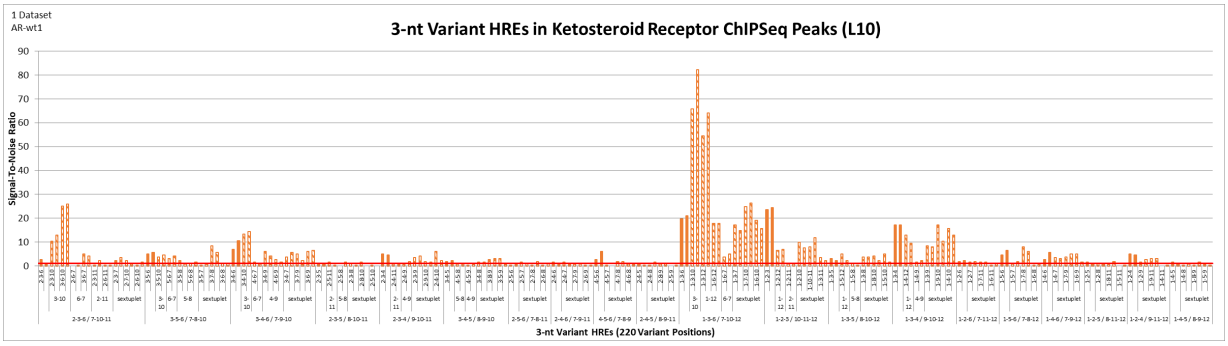
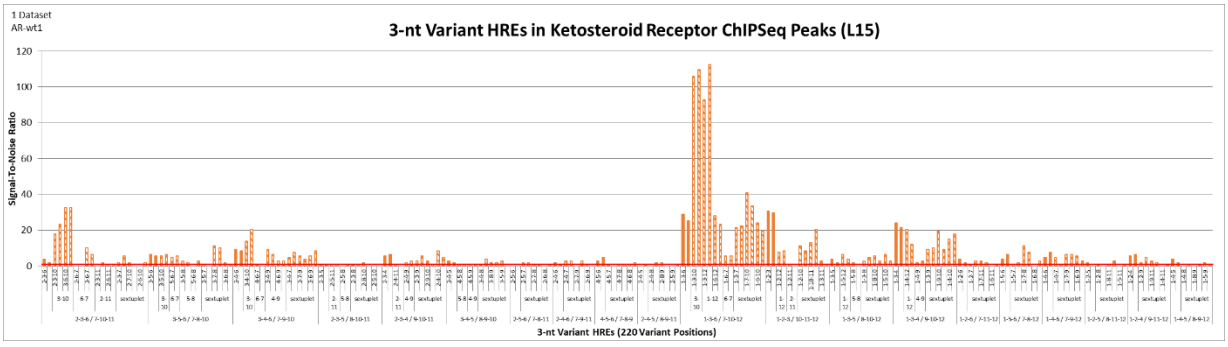


Figure S52. (S/N) analysis of 3-nt Variant HREs in KR ChIPSeq Peaks (Variant Position)









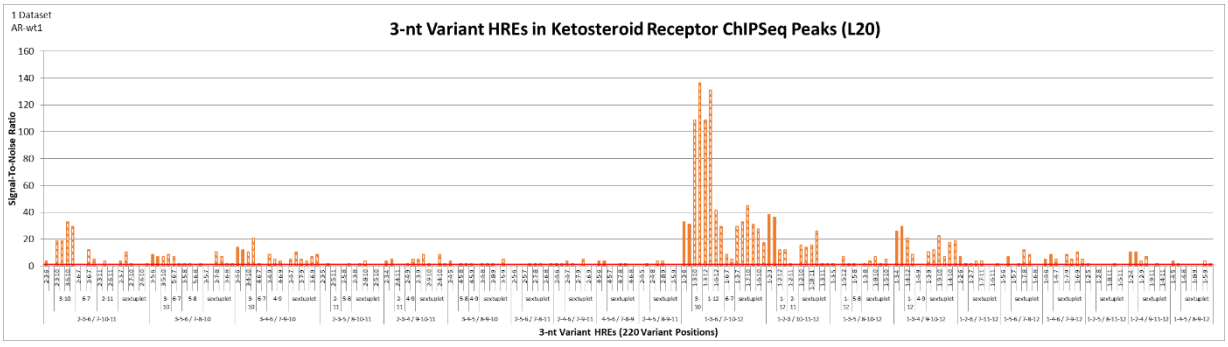
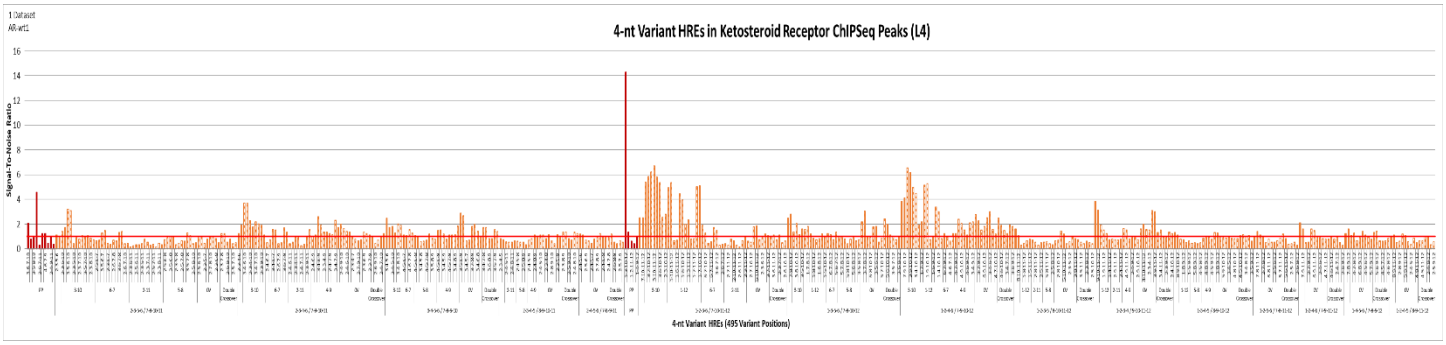
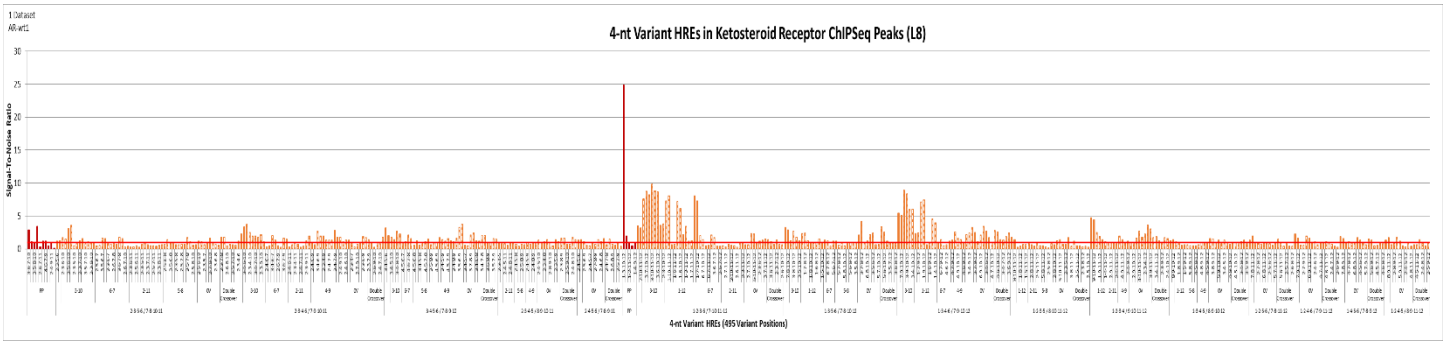
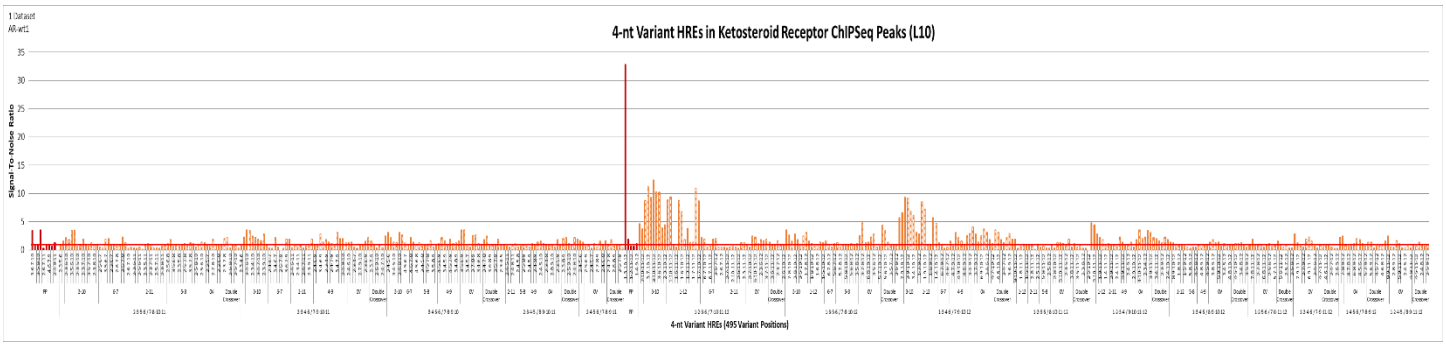
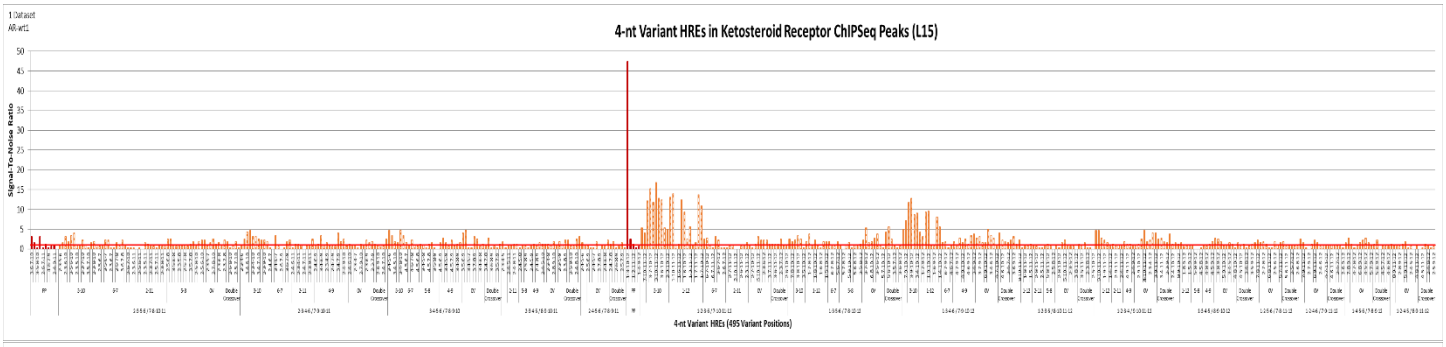


Figure S53. (S/N) analysis of 4-nt Variant HREs in KR ChIPSeq Peaks (Variant Position)









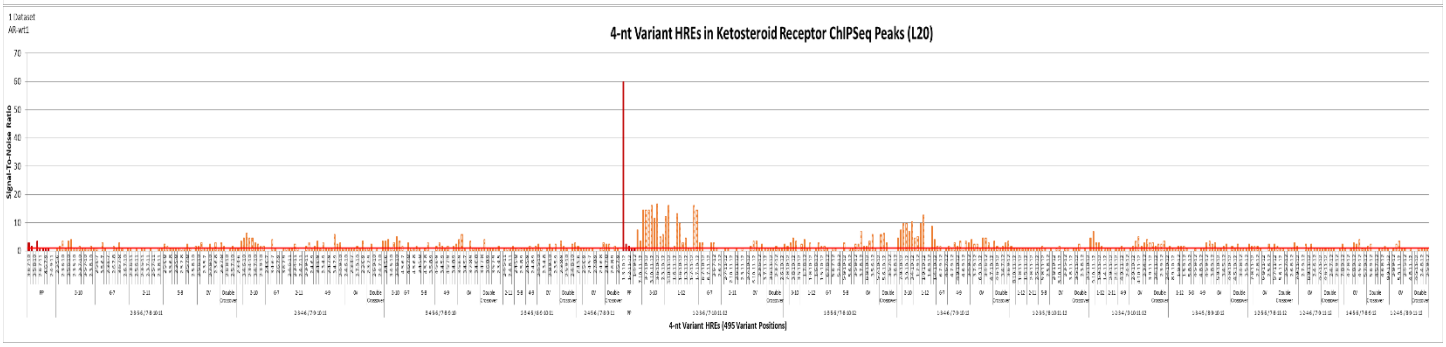


Figure S54. (S/N) analysis of 5-nt Variant HREs in KR ChIPSeq Peaks (Variant Position)

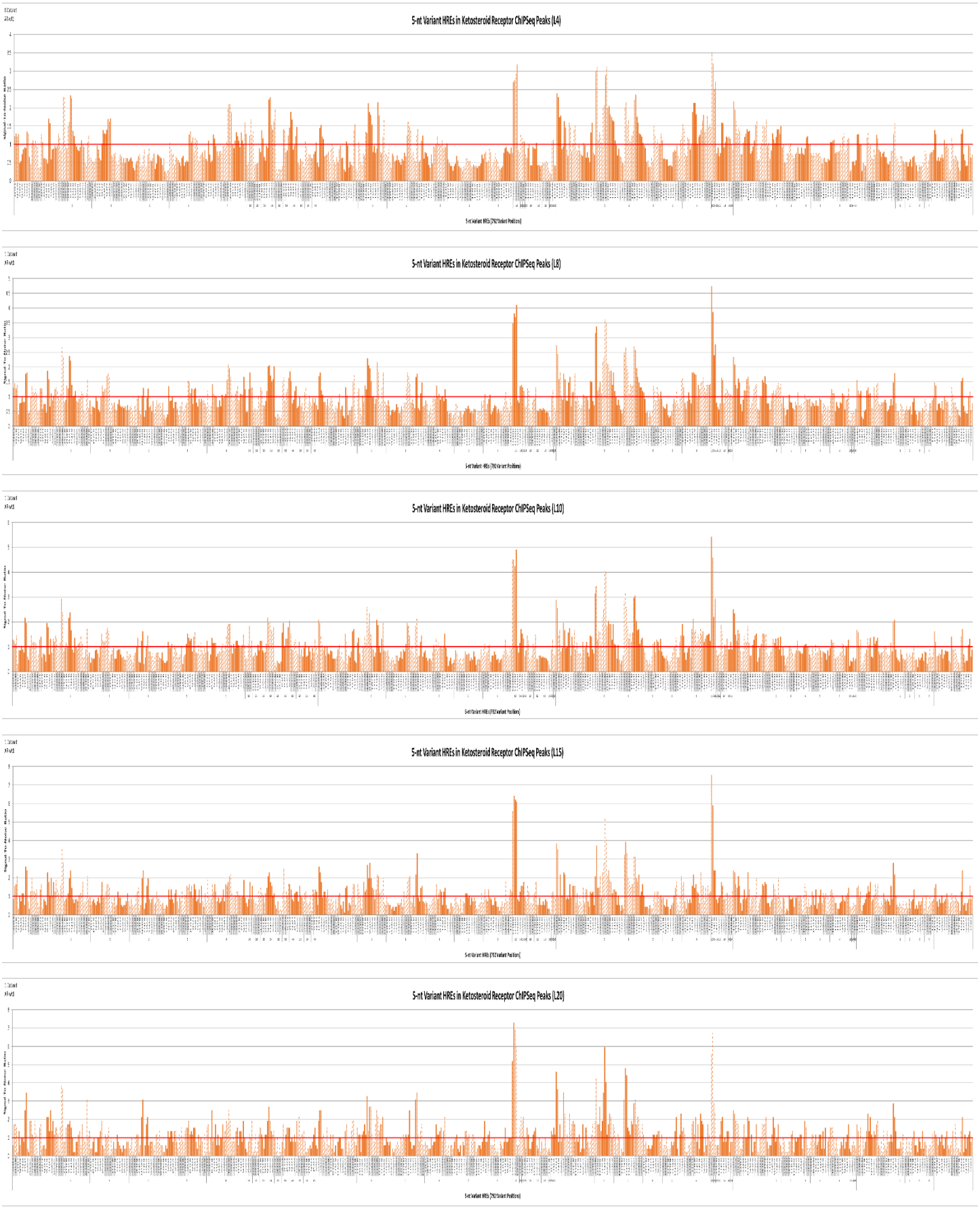


Figure S55. (S/N) analysis of 6-nt Variant HREs in KR ChIPSeq Peaks (Variant Position)

Figure S56. (S/N) analysis of 0-nt to 6-nt Variant HREs in GR and GR-Dim ChIPSeq Peaks #1 (924 Half-Site Groups)

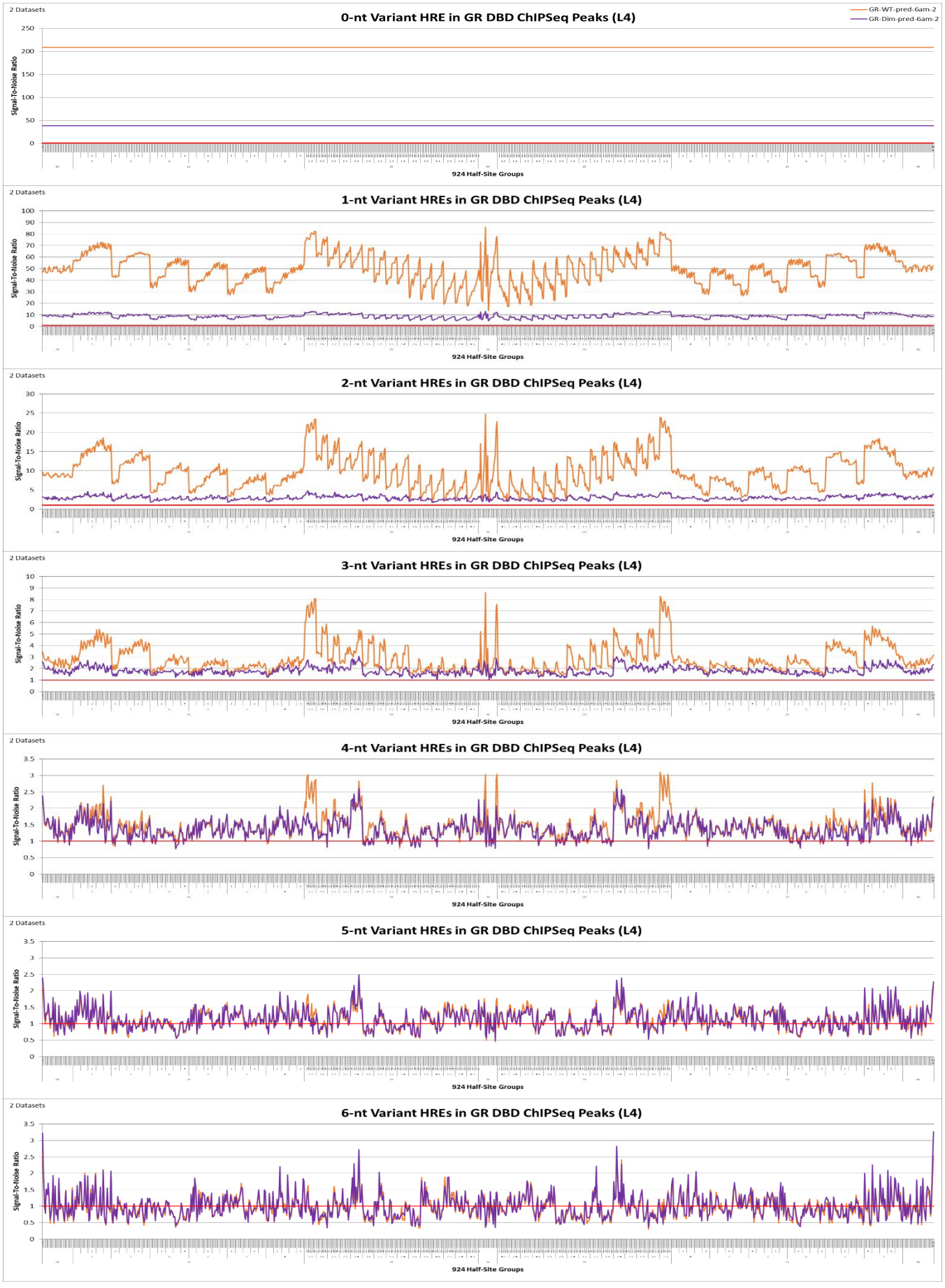


Figure S57. (S/N) analysis of 0-nt to 6-nt Variant HREs in GR and GR-Dim ChIPSeq Peaks #2 (924 Half-Site Groups)

Figure S58. (S/N) analysis of 0-nt to 6-nt Variant HREs in GR and GR-Dim ChIPExo Peaks #1 (924 Half-Site Groups)

Figure S59. (S/N) analysis of 0-nt to 6-nt Variant HREs in GR and GR-Dim ChIPExo Peaks #2 (924 Half-Site Groups)

Figure S60. (S/N) analysis of 1-nt to 6-nt Variant HREs in GR and GR-Dim ChIPSeq Peaks (Variant Position)

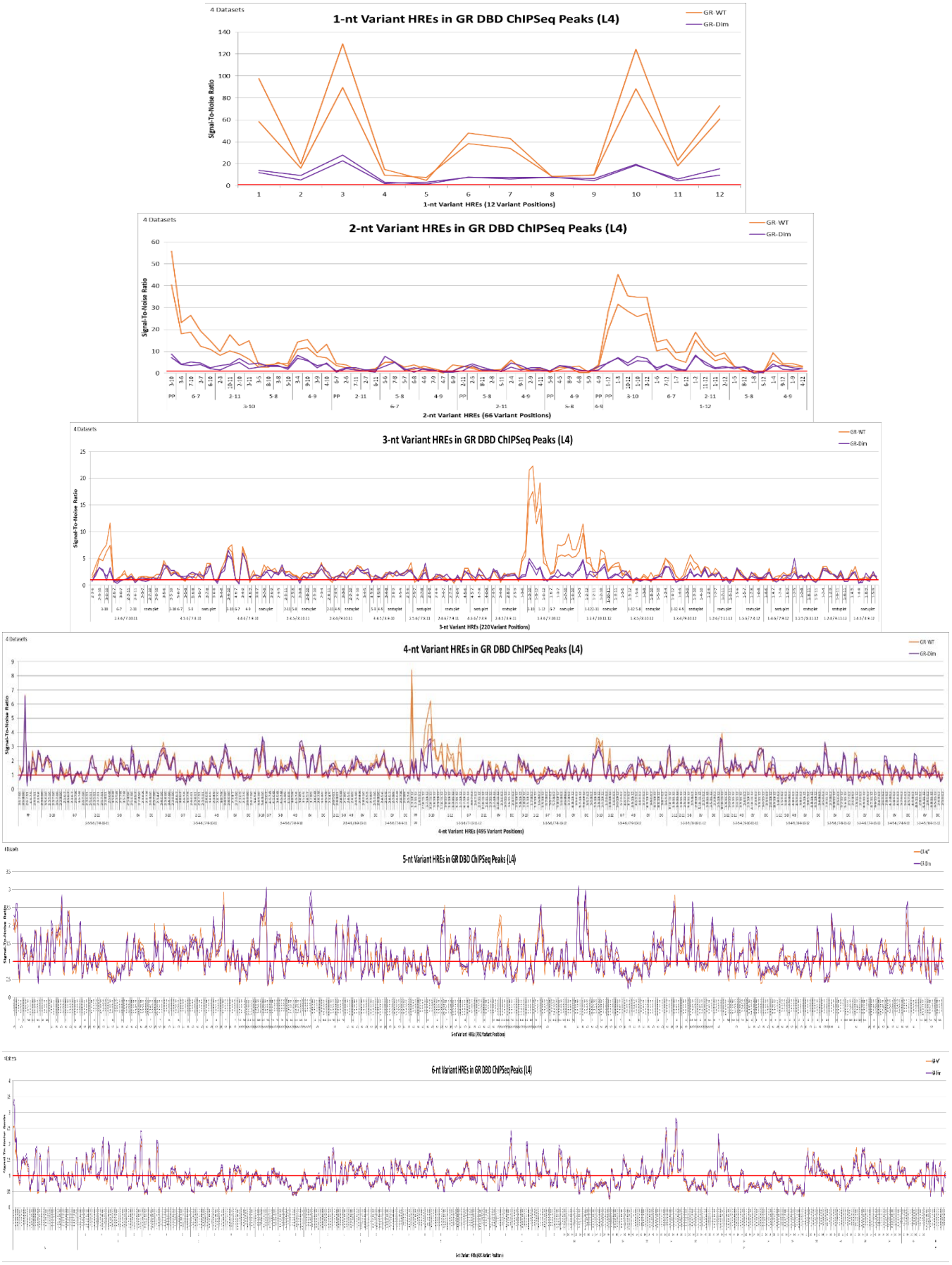


Figure S61. (S/N) analysis of 1-nt to 6-nt Variant HREs in GR and GR-Dim ChIPExo Peaks (Variant Position)

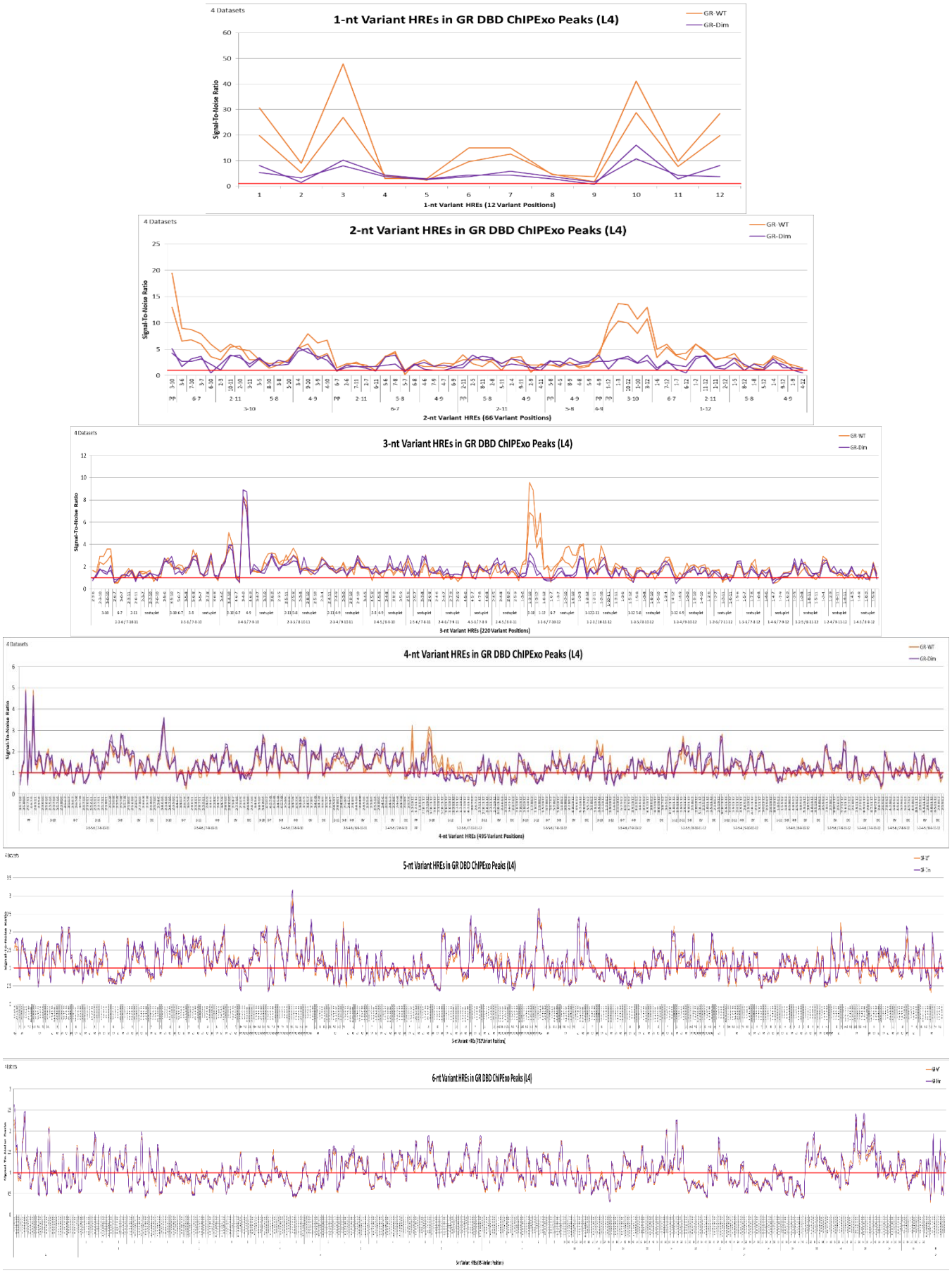


Figure S62-S71 Descriptions: Compare the 13-nt DNA Element Analysis to the 15-nt DNA Element Analysis

Comparative Figure Summary: 13-nt DNA Element Analysis vs. 15-nt DNA Element Analysis

We repeated the same sNR DNA-binding analysis (as was done for sNR DNA-binding at 0-nt to 5-nt variant 13-nt ERE and HRE DNA elements in the genome) for the newly identified 0-nt variant consensus palindromic 15-nt ERE and HRE DNA elements and extending through 6-nt variant DNA elements (i.e., a 6-nt variant =a half-site for a 15-nt DNA element with 12 primary positions) in the genome. Analysis of sNR DNA-binding at 13-nt ERE and HRE DNA elements included the 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements (10 variant positions), 405 2-nt variant DNA elements (45 variant positions), 3,240 3-nt variant DNA elements (120 variant positions), 17,010 4-nt variant DNA elements (210 variant positions), and 61,236 5-nt variant DNA elements (252 variant positions), for a total of 81,922 DNA elements. Analysis of sNR DNA-binding at 15-nt ERE and HRE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 36 1-nt variant DNA elements (12 variant positions), 594 2-nt variant DNA elements (66 variant positions), 5,940 3-nt variant DNA elements (220 variant positions), 40,095 4-nt variant DNA elements (495 variant positions), 192,456 5-nt variant DNA elements (792 variant positions), and 673,596 6-nt variant DNA elements (924 variant positions), for a total of 912,718 DNA elements. All data analyses completed for sNR DNA-binding at 0-nt to 5-nt variant 13-nt ERE and HRE DNA elements in the genome was repeated for sNR DNA-binding at 0-nt to 6-nt variant 15-nt ERE and HRE DNA elements in the genome.

Comparative Figure			
13-nt DNA Element Analysis vs. 15-nt DNA Element Analysis			
	Figure Title	**13-nt DNA Element**	**15-nt DNA Element**
1	Number of ERE and HRE DNA Elements in the Mouse and Human Genome	**Table 1**	Figure S34
2	Inversion Symmetry of sNR DNA-Binding at 1-nt Variant DNA Elements in the Genome (%)	**Figure 1**	Figure S35
3	Inversion Symmetry of sNR DNA-Binding at 2-nt Variant DNA Elements in the Genome (%)	Figure S1	Figure S36
4	(S/N) analysis of 0-nt to 5-nt/6-nt Variant EREs in ER ChIPSeq Peaks (# of Variants)	Figure S2	Figure S37
5	(S/N) analysis of 0-nt to 5-nt/6-nt Variant HREs in KR ChIPSeq Peaks (# of Variants)	Figure S3	Figure S38
6	Half-Site Groups [252] [924]	**Table 2**	Figure S39
7	(S/N) analysis of 0-nt to 5-nt/6-nt Variant EREs in ER ChIPSeq Peaks (Half-Site Groups)	**Figure 2**	Figure S40
8	(S/N) analysis of 0-nt to 5-nt/6-nt Variant HREs in AR ChIPSeq Peaks (Half-Site Groups)	**Figure 3**	Figure S41
9	(S/N) analysis of 0-nt to 5-nt/6-nt Variant HREs in GR ChIPSeq Peaks (Half-Site Groups)	Figure S4	Figure S42
10	(S/N) analysis of 0-nt to 5-nt/6-nt Variant HREs in PR ChIPSeq Peaks (Half-Site Groups)	Figure S5	Figure S43
11	(S/N) analysis of 1-nt Variant EREs in ER ChIPSeq Peaks (Variant Position) [10] [12]	Figure S12	Figure S44
12	(S/N) analysis of 2-nt Variant EREs in ER ChIPSeq Peaks (Variant Position) [45] [66]	Figure S13	Figure S45
13	(S/N) analysis of 3-nt Variant EREs in ER ChIPSeq Peaks (Variant Position) [120] [220]	Figure S14	Figure S46
14	(S/N) analysis of 4-nt Variant EREs in ER ChIPSeq Peaks (Variant Position) [210] [495]	Figure S15	Figure S47
15	(S/N) analysis of 5-nt Variant EREs in ER ChIPSeq Peaks (Variant Position) [252] [792]	Figure S16	Figure S48
16	(S/N) analysis of 1-nt Variant HREs in KR ChIPSeq Peaks (Variant Position) [10] [12]	Figure S17	Figure S50
17	(S/N) analysis of 2-nt Variant HREs in KR ChIPSeq Peaks (Variant Position) [45] [66]	Figure S18	Figure S51
18	(S/N) analysis of 3-nt Variant HREs in KR ChIPSeq Peaks (Variant Position) [120] [220]	Figure S19	Figure S52
19	(S/N) analysis of 4-nt Variant HREs in KR ChIPSeq Peaks (Variant Position) [210] [495]	Figure S20	Figure S53
20	(S/N) analysis of 5-nt Variant HREs in KR ChIPSeq Peaks (Variant Position) [252] [792]	Figure S21	Figure S54
21	(S/N) analysis of 0-nt to 5-nt/6-nt Variant HREs in GR and GR-Dim ChIPSeq Peaks #1 (Half-Site Groups)	**Figure 4**	Figure S56
22	(S/N) analysis of 0-nt to 5-nt/6-nt Variant HREs in GR and GR-Dim ChIPSeq Peaks #2 (Half-Site Groups)	Figure S22	Figure S57
23	(S/N) analysis of 0-nt to 5-nt/6-nt Variant HREs in GR and GR-Dim ChIPExo Peaks #1 (Half-Site Groups)	Figure S23	Figure S58
24	(S/N) analysis of 0-nt to 5-nt/6-nt Variant HREs in GR and GR-Dim ChIPExo Peaks #2 (Half-Site Groups)	Figure S24	Figure S59
25	(S/N) analysis of 0-nt to 5-nt/6-nt Variant HREs in GR and GR-Dim ChIPSeq Peaks (Variant Position)	**Figure 5**	Figure S60
26	(S/N) analysis of 0-nt to 5-nt/6-nt Variant HREs in GR and GR-Dim ChIPExo Peaks (Variant Position)	Figure S25	Figure S61

*****

****All detailed data and statistics associated with every figure are compiled in Data S1****

To allow for a more direct comparison between the sNR DNA-binding profile at the 81,922 0-nt to 5-nt variant 13-nt ERE or HRE DNA elements in the genome versus at the 912,718 0-nt to 6-nt variant 15-nt ERE or HRE DNA elements in the genome, in this section we have removed the bars associated with sNR DNA-binding at DNA elements that have a variant in position 1 or position 12 from the plots of the 15-nt ERE and HRE DNA element analysis. For example, looking at **Figure S62**, the plots on the left-hand side of the page show ER DNA-binding at the 30 1-nt variant 13-nt ERE DNA elements (10 variant positions) in the genome (5′-GGTCAnnnTGACC-3′), while the plots on the right-hand side of the page show ER DNA-binding at the 36 1-nt variant 15-nt ERE DNA elements (12 variant positions) in the genome (5′-AGGTCAnnnTGACCT-3′). We have removed the bars that represent ER DNA-binding at position 1 and position 12 from the plots on the right-hand side of the page (i.e., the 15-nt ERE DNA element analysis), as ER DNA-binding at those sites is expanded information from the 13-nt ERE DNA element analysis. Thus, the 15-nt ERE and HRE DNA element analysis plots (i.e., on the right-hand side of the page) are the same plots from **Figure S44-S55**, except we have removed this expanded information (i.e., bars that represent sNR DNA-binding at DNA elements that have a variant in position 1 or position 12 of the 15-nt ERE and HRE DNA elements). The DNA-binding profile at the 13-nt DNA elements in the genome versus at the 15-nt DNA elements in the genome are almost equivalent, confirming that the DNA-binding rules established by the 13-nt DNA element analysis also apply to the 15-nt DNA element analysis. **Note**: see section *"Transform the 13-nt DNA Element Analysis to the 15-nt DNA Element Analysis"* to understand the minor differences in the relative amplitudes of the sNR DNA-binding profile at the 13-nt DNA elements in the genome versus at the 15-nt DNA elements in the genome.

Figure S62. Compare (S/N) analysis of 1-nt Variant EREs in ER ChIPSeq Peaks (13-nt vs. 15-nt)
(S/N) analysis of 1-nt variant EREs in ER (WT-E2-1hr) ChIPSeq peaks [76,163 peaks, 146-nt peak length].
Displayed by the 10 variant positions of the 30 1-nt variant EREs [13-nt ERE] versus the 10 of 12 variant positions of the 36 1-nt variant EREs [15-nt ERE].

Figure S63. Compare (S/N) analysis of 2-nt Variant EREs in ER ChIPSeq Peaks (13-nt vs. 15-nt)
(S/N) analysis of 2-nt variant EREs in ER (WT-E2-1hr) ChIPSeq peaks [76,163 peaks, 146-nt peak length].
Displayed by the 45 variant positions of the 405 2-nt variant EREs [13-nt ERE] versus the 45 of 66 variant positions of the 594 2-nt variant EREs [15-nt ERE]. 5 of the 45 variant positions are the five (5) palindromic position pairs: 1-10, 2-9, 3-8, 4-7, 5-6 (PP, dark red solid bars). 20 of the 45 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (blue solid bars). 20 of the 45 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (blue diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome).

Figure S64. Compare (S/N) analysis of 3-nt Variant EREs in ER ChIPSeq Peaks (13-nt vs. 15-nt)
(S/N) analysis of 3-nt variant EREs in ER (WT-E2-1hr) ChIPSeq peaks [76,163 peaks, 146-nt peak length].
Displayed by the 120 variant positions of the 3,240 3-nt variant EREs [13-nt ERE] versus the 120 of 220 variant positions of the 5,940 3-nt variant EREs [15-nt ERE]. 20 of the 120 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (blue solid bars). 100 of the 120 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (blue diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome).

Figure S65. Compare (S/N) analysis of 4-nt Variant EREs in ER ChIPSeq Peaks (13-nt vs. 15-nt)
(S/N) analysis of 4-nt variant EREs in ER (WT-E2-1hr) ChIPSeq peaks [76,163 peaks, 146-nt peak length].
Displayed by the 210 variant positions of the 17,010 4-nt variant EREs [13-nt ERE] versus the 210 of 495 variant positions of the 40,095 4-nt variant EREs [15-nt ERE]. 10 of the 210 variant positions are the five (5) palindromic position pairs: 1-3-8-10, 3-5-6-8, 3-4-7-8, 2-3-8-9, 1-5-6-10, 1-4-7-10, 1-2-9-10, 4-5-6-7, 2-5-6-9, 2-4-7-9 (PP, dark red solid bars). 10 of the 210 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (blue solid bars). 190 of the 45 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (blue diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome).

Figure S66. Compare (S/N) analysis of 5-nt Variant EREs in ER ChIPSeq Peaks (13-nt vs. 15-nt)
(S/N) analysis of 5-nt variant EREs in ER (WT-E2-1hr) ChIPSeq peaks [76,163 peaks, 146-nt peak length].
Displayed by the 252 variant positions of the 61,236 5-nt variant EREs [13-nt ERE] versus the 252 of 792 variant positions of the 192,456 5-nt variant EREs [15-nt ERE]. 2 of the 252 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (blue solid bars). 250 of the 252 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (blue diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome).

*****

Figure S67. Compare (S/N) analysis of 1-nt Variant HREs in KR ChIPSeq Peaks (13-nt vs. 15-nt)
(S/N) analysis of 1-nt variant HREs in AR (AR-wt1) ChIPSeq peaks [49,859 peaks, 136-nt peak length].
Displayed by the 10 variant positions of the 30 1-nt variant HREs [13-nt HRE] versus the 10 of 12 variant positions of the 36 1-nt variant HREs [15-nt HRE].

Figure S68. Compare (S/N) analysis of 2-nt Variant HREs in KR ChIPSeq Peaks (13-nt vs. 15-nt)
(S/N) analysis of 2-nt variant HREs in AR (AR-wt1) ChIPSeq peaks [49,859 peaks, 136-nt peak length].
Displayed by the 45 variant positions of the 405 2-nt variant HREs [13-nt HRE] versus the 45 of 66 variant positions of the 594 2-nt variant HREs [15-nt HRE]. 5 of the 45 variant positions are the five (5) palindromic position pairs: 1-10, 2-9, 3-8, 4-7, 5-6 (PP, dark red solid bars). 20 of the 45 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (orange solid bars). 20 of the 45 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (orange diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a

particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome).

Figure S69. Compare (S/N) analysis of 3-nt Variant HREs in KR ChIPSeq Peaks (13-nt vs. 15-nt)
(S/N) analysis of 3-nt variant HREs in AR (AR-wt1) ChIPSeq peaks [49,859 peaks, 136-nt peak length].
Displayed by the 120 variant positions of the 3,240 3-nt variant HREs [13-nt HRE] versus the 120 of 220 variant positions of the 5,940 3-nt variant HREs [15-nt HRE]. 20 of the 120 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (orange solid bars). 100 of the 120 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (orange diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome).

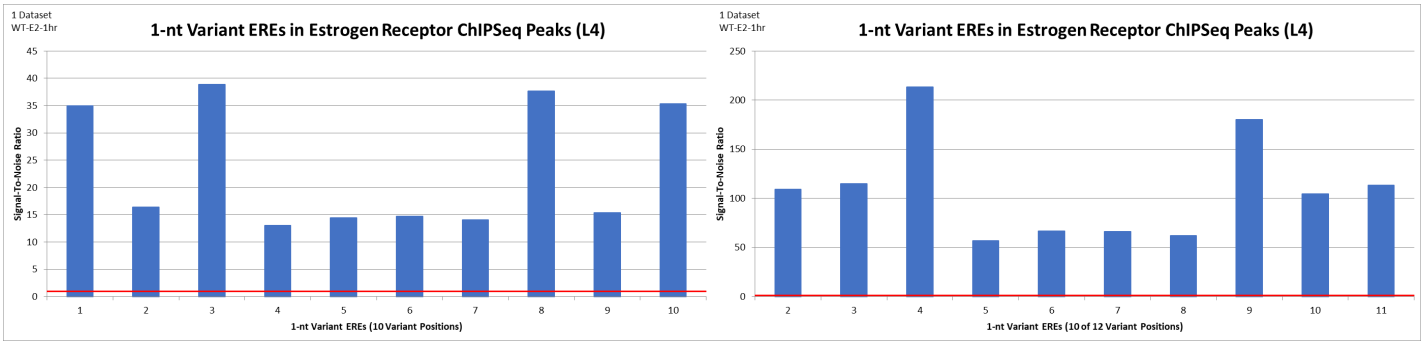
Figure S70. Compare (S/N) analysis of 4-nt Variant HREs in KR ChIPSeq Peaks (13-nt vs. 15-nt)
(S/N) analysis of 4-nt variant HREs in AR (AR-wt1) ChIPSeq peaks [49,859 peaks, 136-nt peak length].
Displayed by the 210 variant positions of the 17,010 4-nt variant HREs [13-nt HRE] versus the 210 of 495 variant positions of the 40,095 4-nt variant HREs [15-nt HRE]. 10 of the 210 variant positions are the five (5) palindromic position pairs: 2-5-6-9, 1-2-9-10, 2-4-7-9, 2-3-8-9, 1-5-6-10, 4-5-6-7, 3-5-6-8, 1-4-7-10, 1-3-8-10, 3-4-7-8 (PP, dark red solid bars). 10 of the 210 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (orange solid bars). 190 of the 45 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (orange diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome).

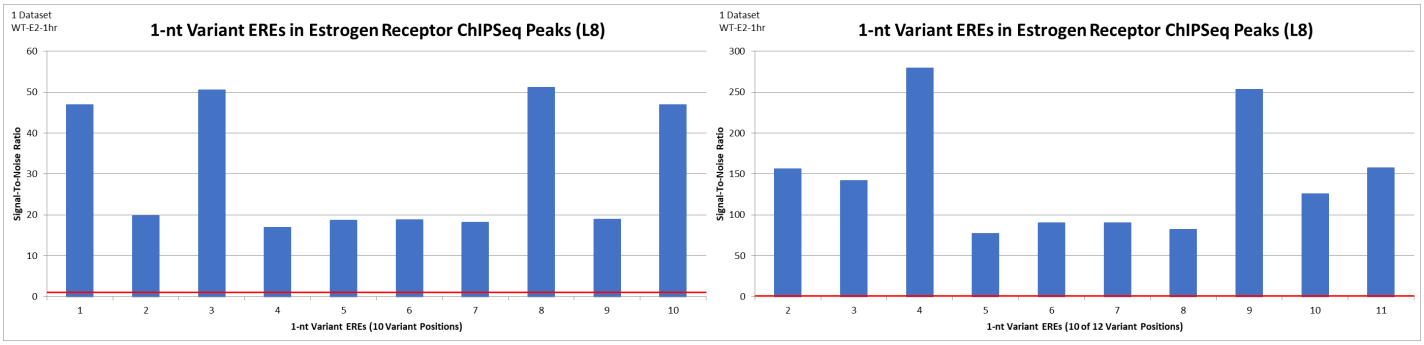
Figure S71. Compare (S/N) analysis of 5-nt Variant HREs in KR ChIPSeq Peaks (13-nt vs. 15-nt)
(S/N) analysis of 5-nt variant HREs in AR (AR-wt1) ChIPSeq peaks [49,859 peaks, 136-nt peak length].
Displayed by the 252 variant positions of the 61,236 5-nt variant HREs [13-nt HRE] versus the 252 of 792 variant positions of the 192,456 5-nt variant HREs [15-nt HRE]. 2 of the 252 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer) (orange solid bars). 250 of the 252 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer) (orange diagonal stripped bars). Each variant position is immediately followed by its reverse-complement variant position on the x-axis, demonstrating that transcription factor binding in the genome follows inversion symmetry (i.e., the number of transcription factor binding events at a particular DNA element in the genome is equivalent to the number of transcription factor binding events at its reverse-complement DNA element in the genome).

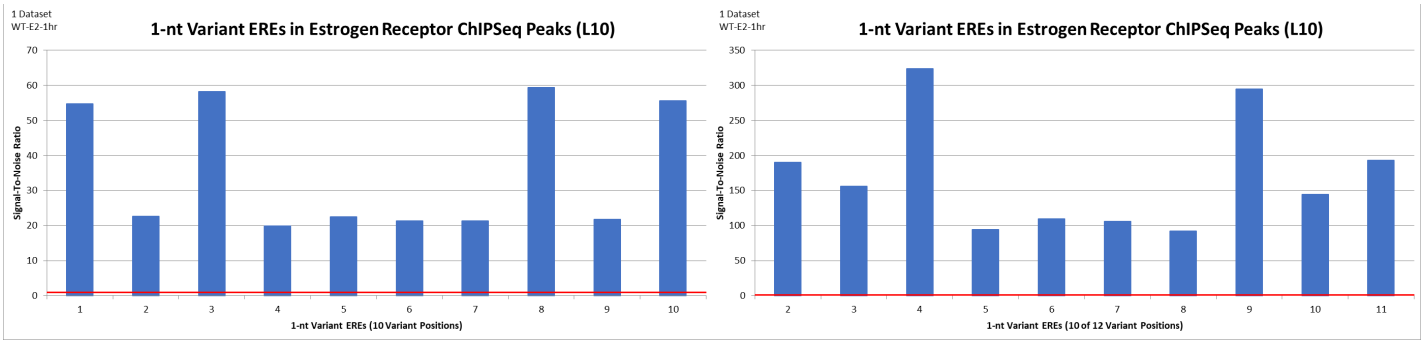
Figure S62. Compare (S/N) analysis of 1-nt Variant EREs in ER ChIPSeq Peaks (13-nt vs. 15-nt)

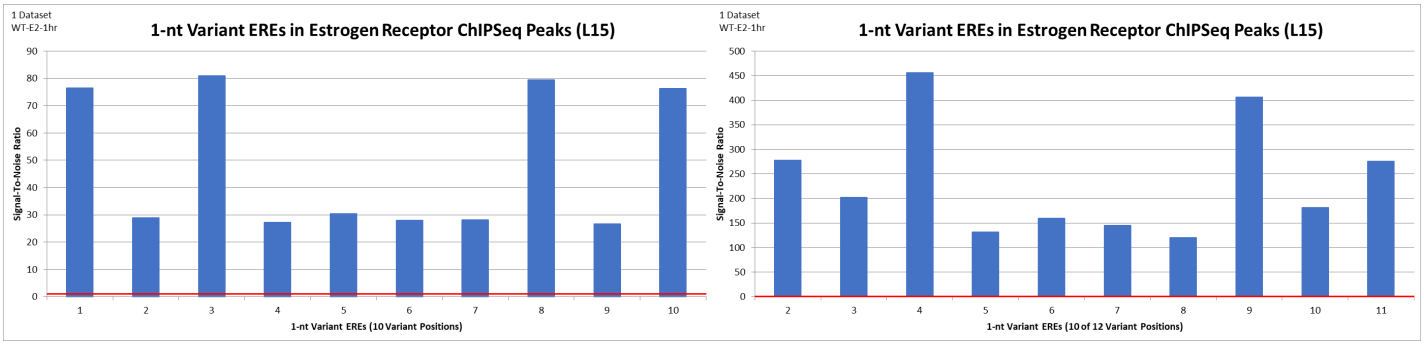
1-nt Variant 13-nt ERE DNA Elements

1-nt Variant 15-nt ERE DNA Elements









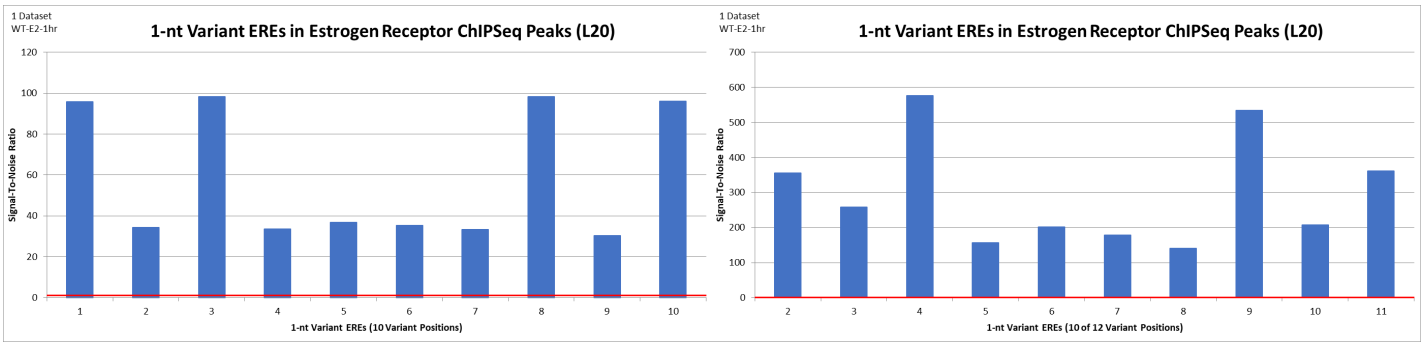
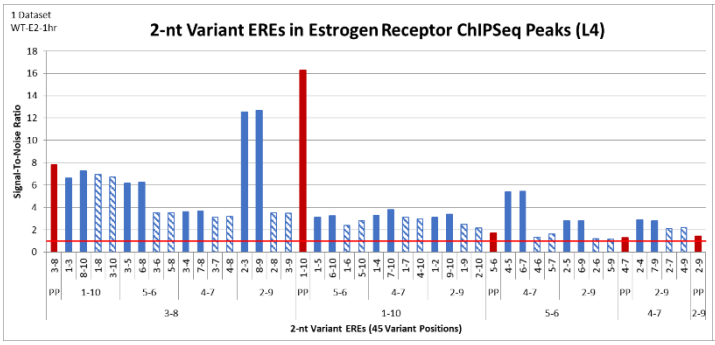
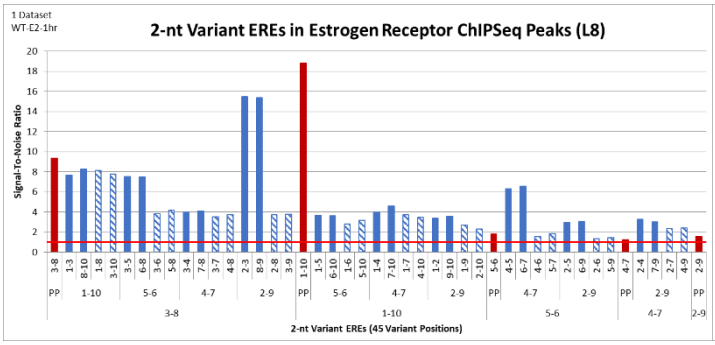
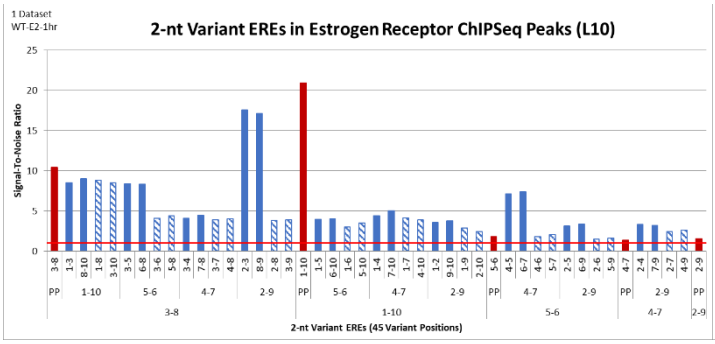


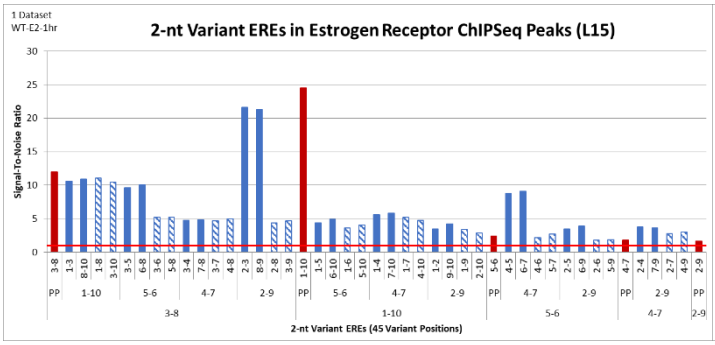
Figure S63. Compare (S/N) analysis of 2-nt Variant EREs in ER ChIPSeq Peaks (13-nt vs. 15-nt)

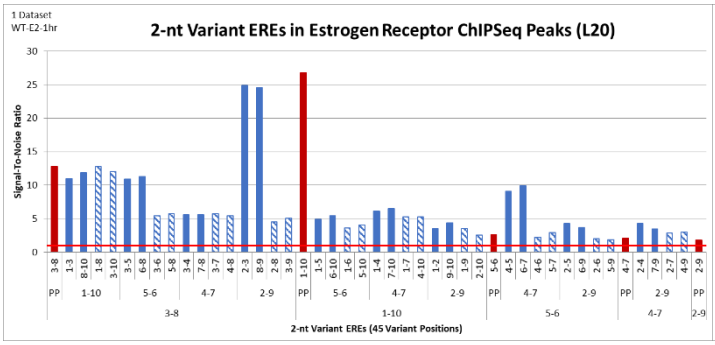
2-nt Variant 13-nt ERE DNA Elements



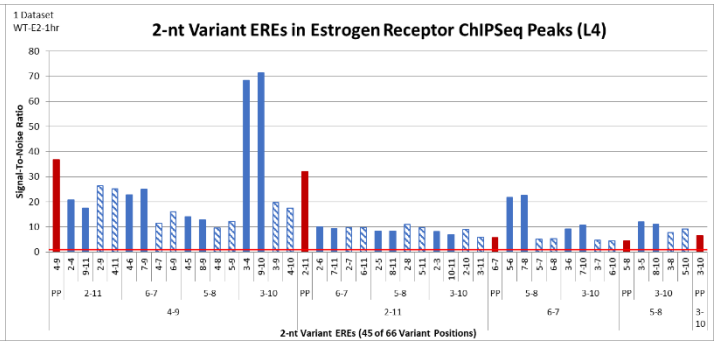


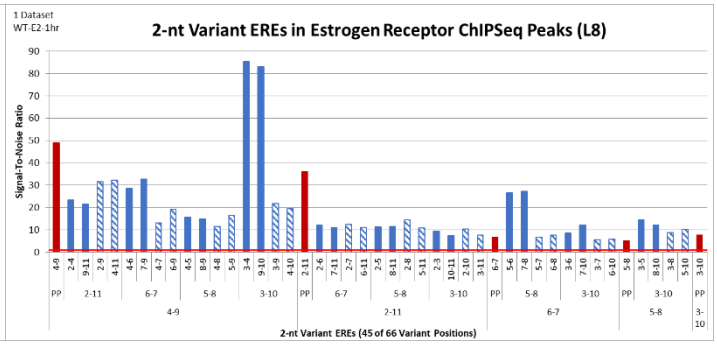


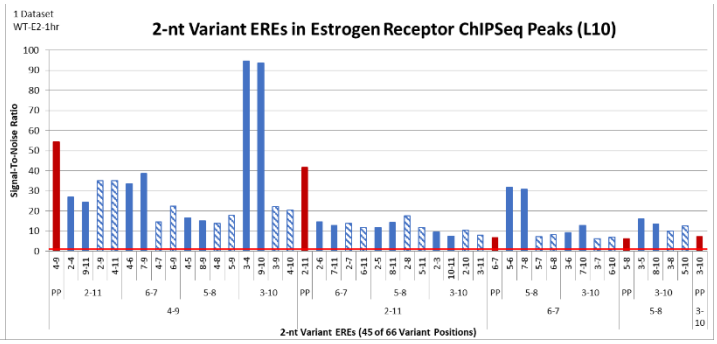


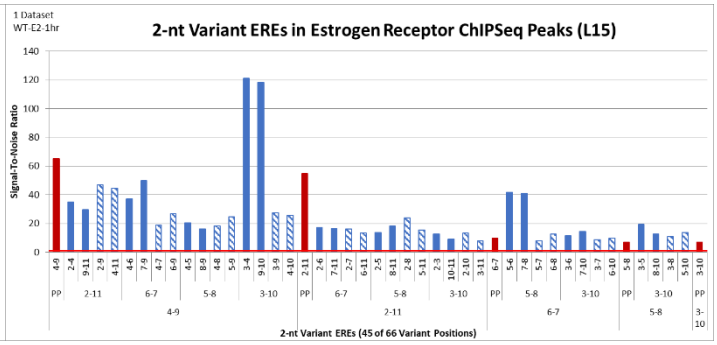


2-nt Variant 15-nt ERE DNA Elements









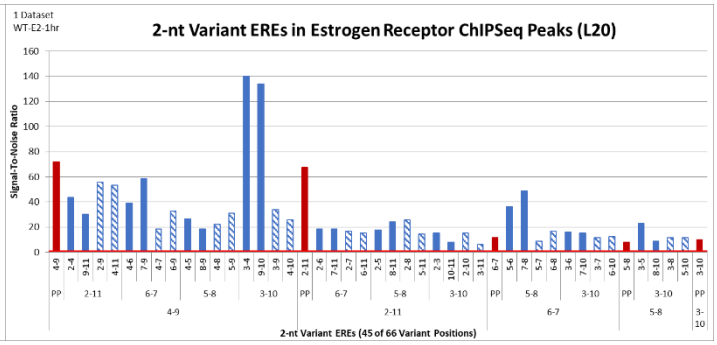
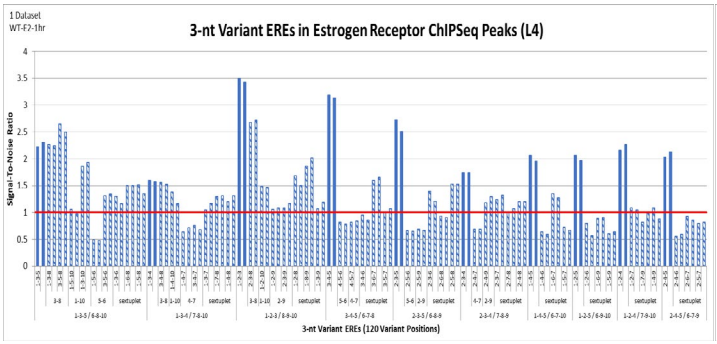
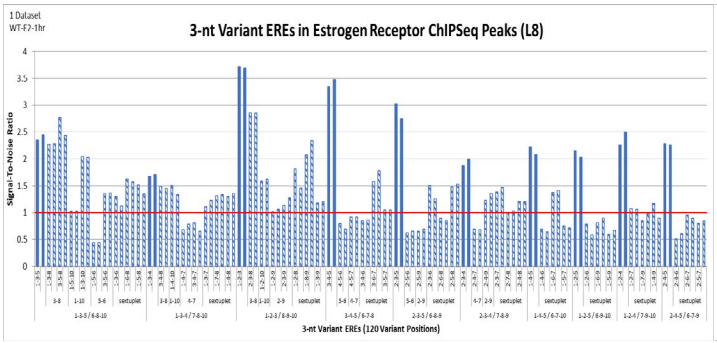
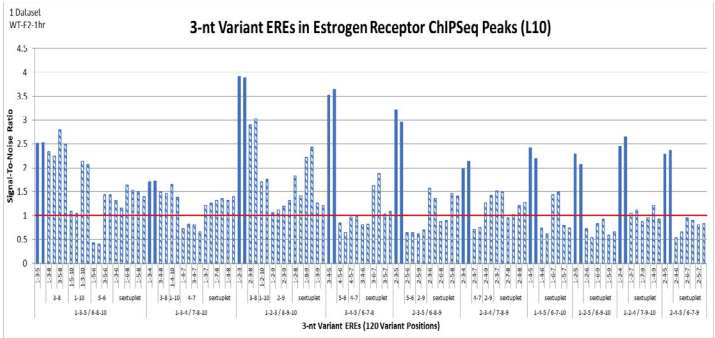


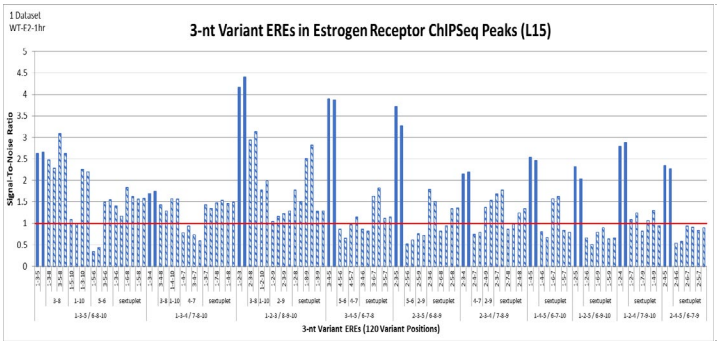
Figure S64. Compare (S/N) analysis of 3-nt Variant EREs in ER ChIPSeq Peaks (13-nt vs. 15-nt)

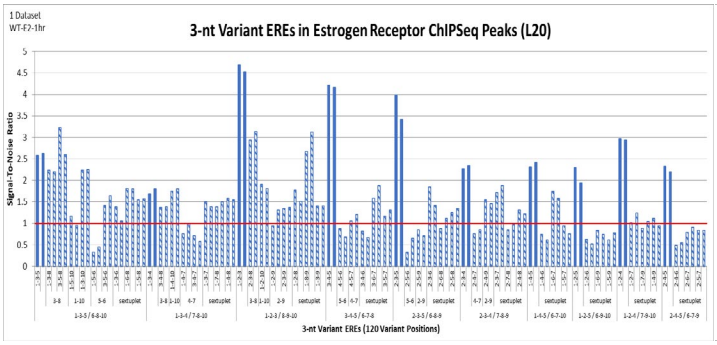
3-nt Variant 13-nt ERE DNA Elements



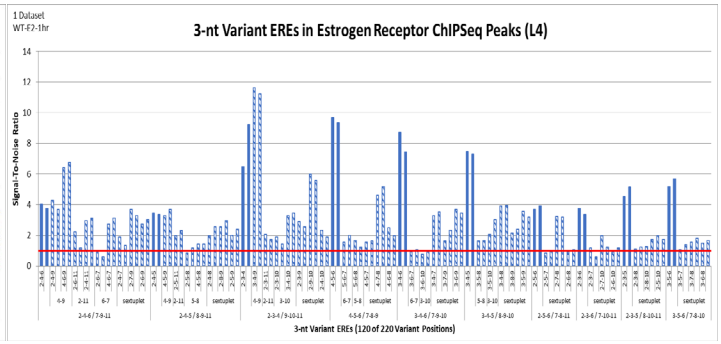


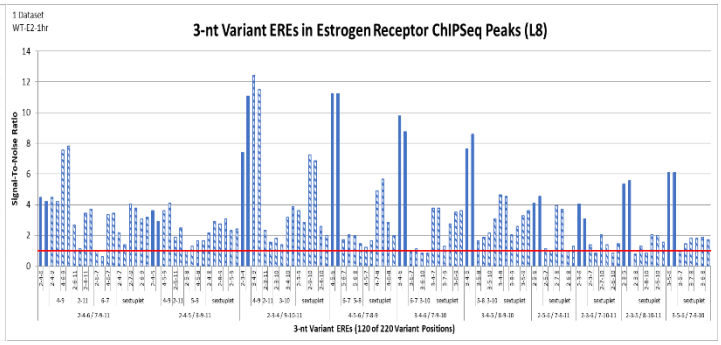


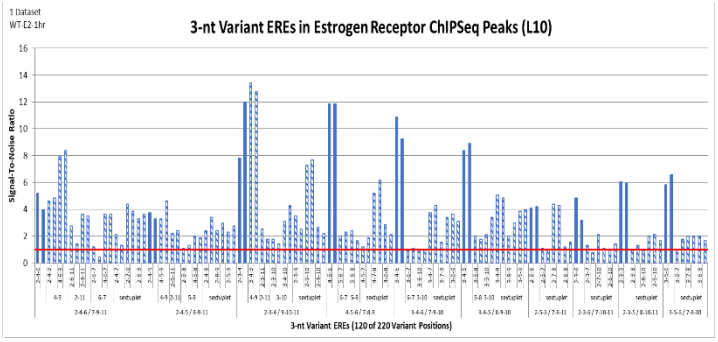


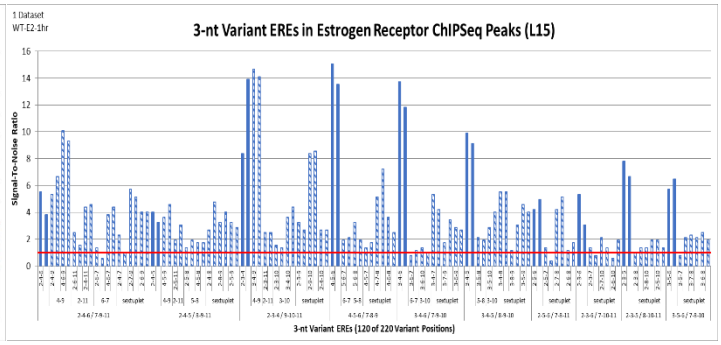


3-nt Variant 15-nt ERE DNA Elements









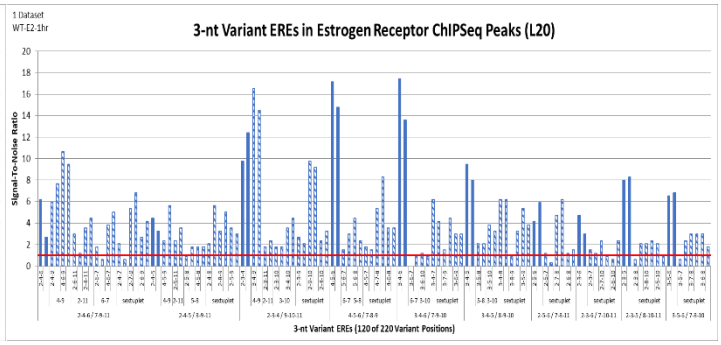
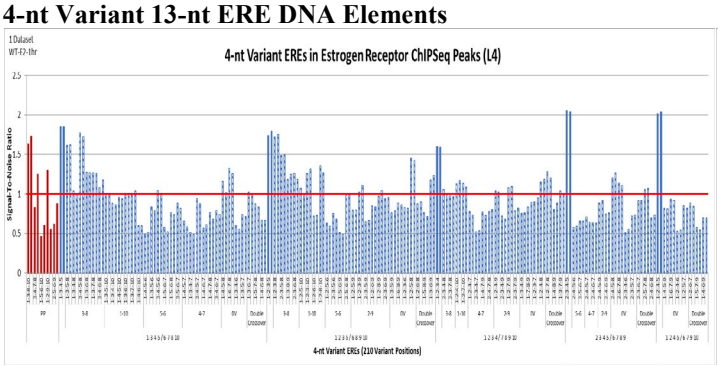
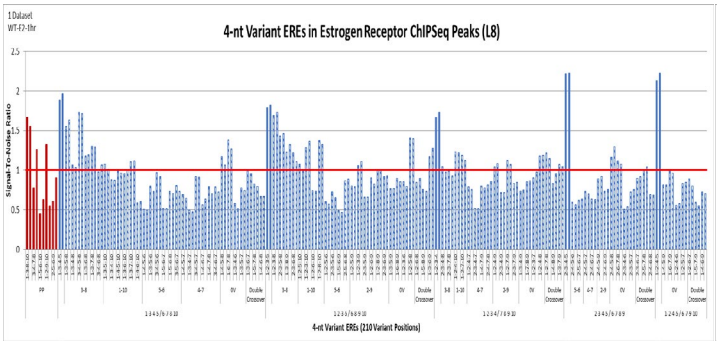
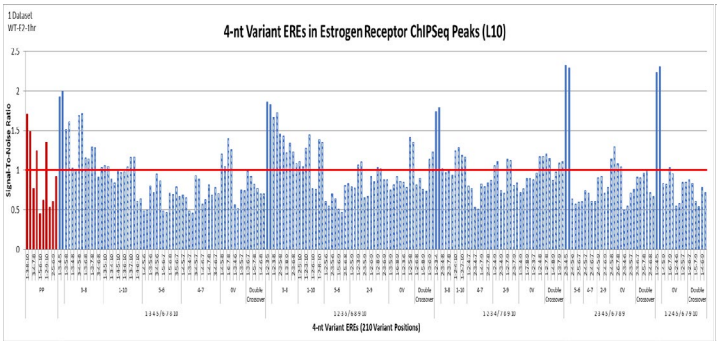


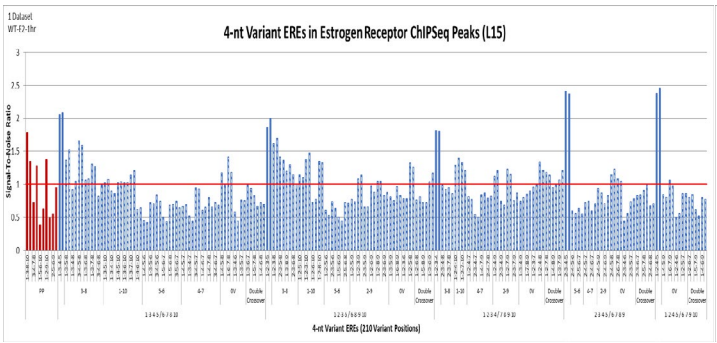
Figure S65. Compare (S/N) analysis of 4-nt Variant EREs in ER ChIPSeq Peaks (13-nt vs. 15-nt)

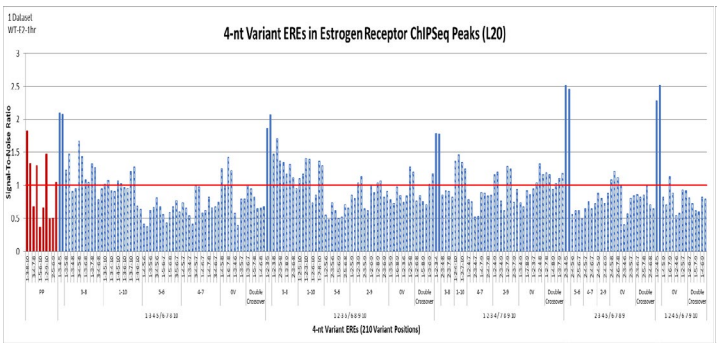
4-nt Variant 13-nt ERE DNA Elements



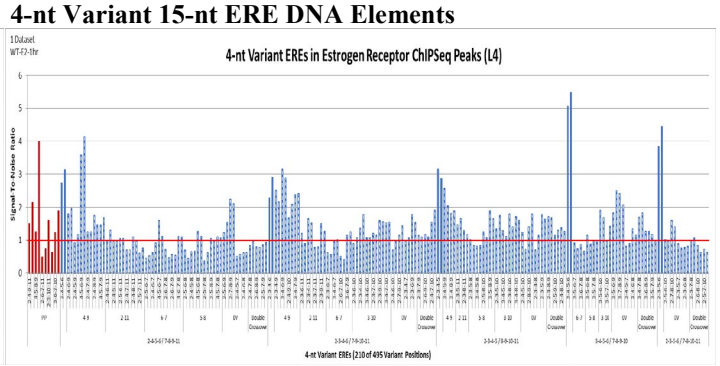


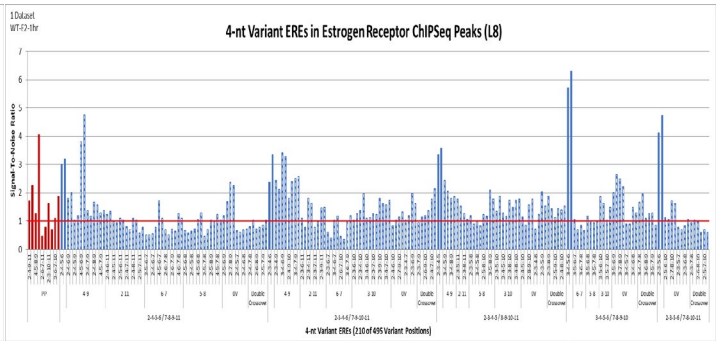


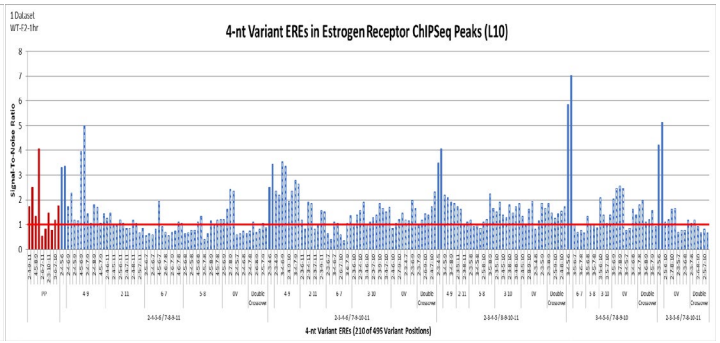


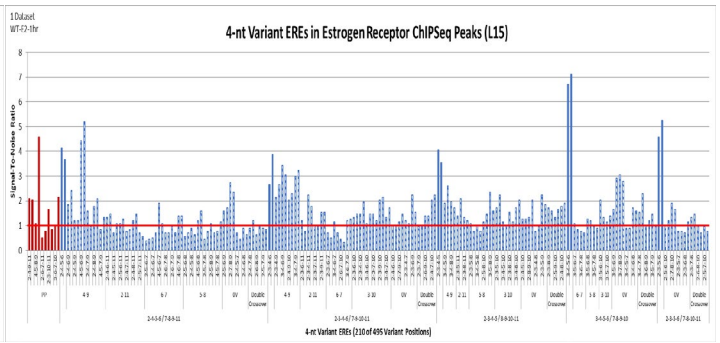


4-nt Variant 15-nt ERE DNA Elements









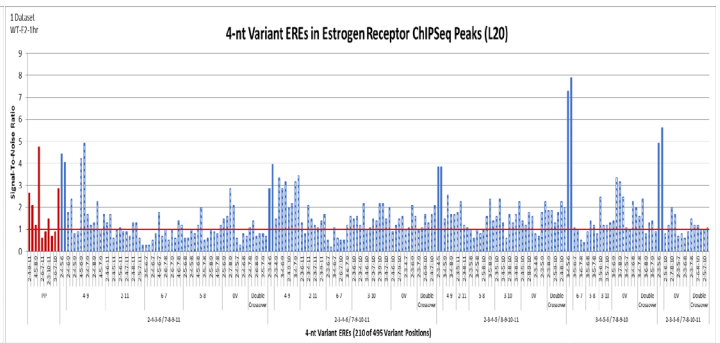
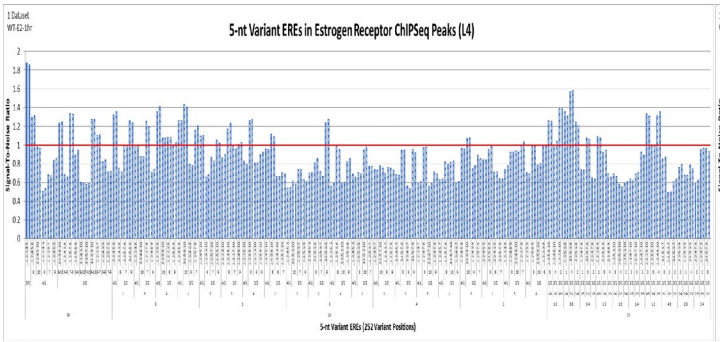
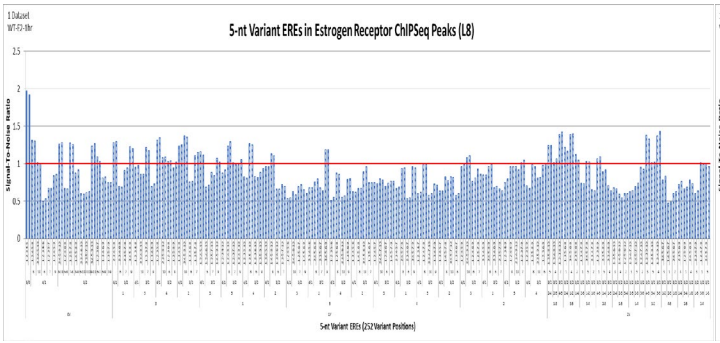
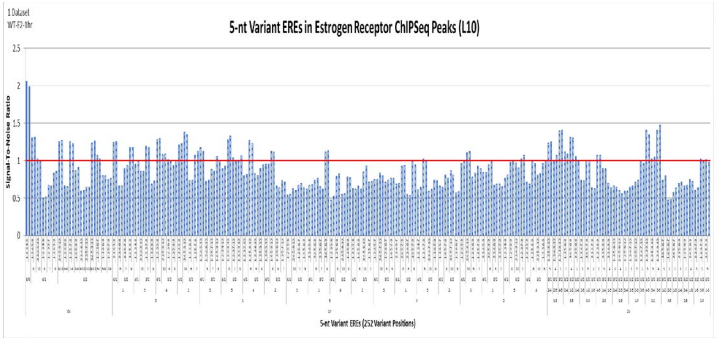


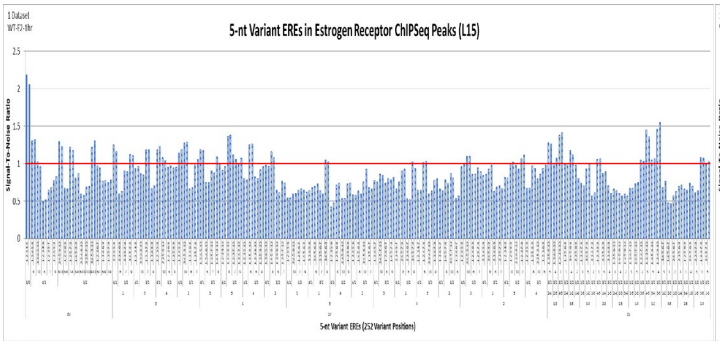
Figure S66. Compare (S/N) analysis of 5-nt Variant EREs in ER ChIPSeq Peaks (13-nt vs. 15-nt)

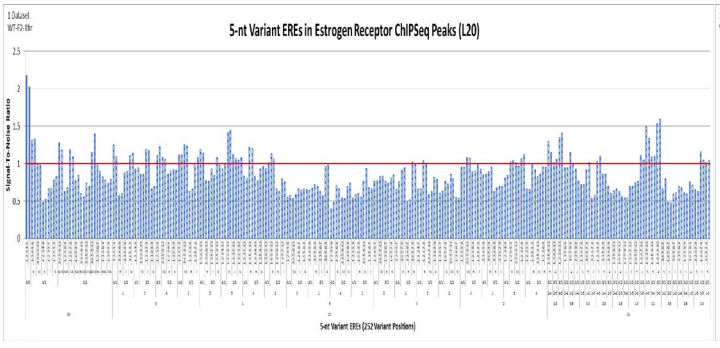
5-nt Variant 13-nt ERE DNA Elements



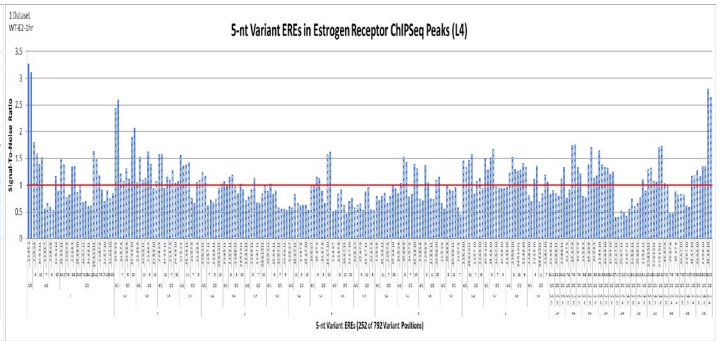


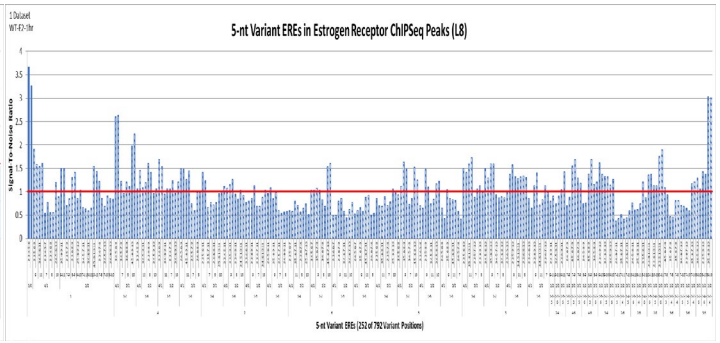


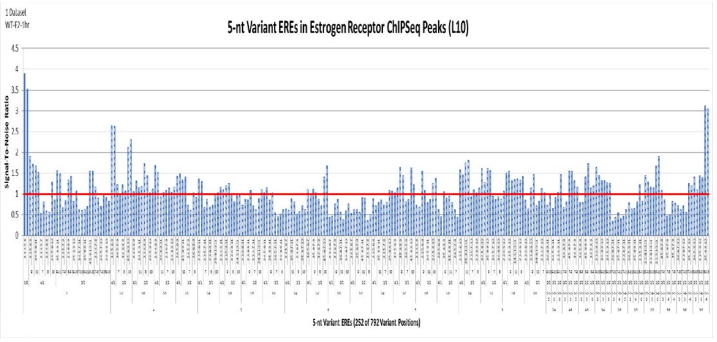


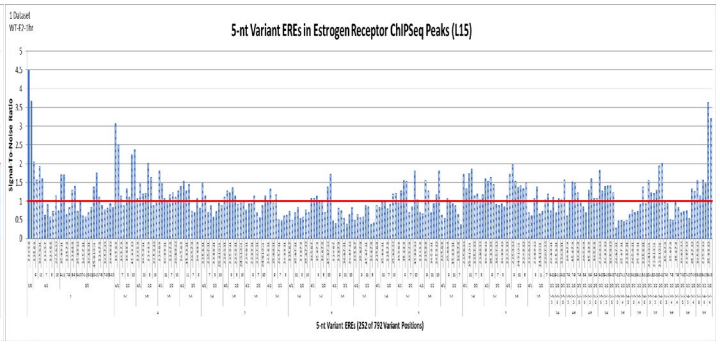


5-nt Variant 15-nt ERE DNA Elements









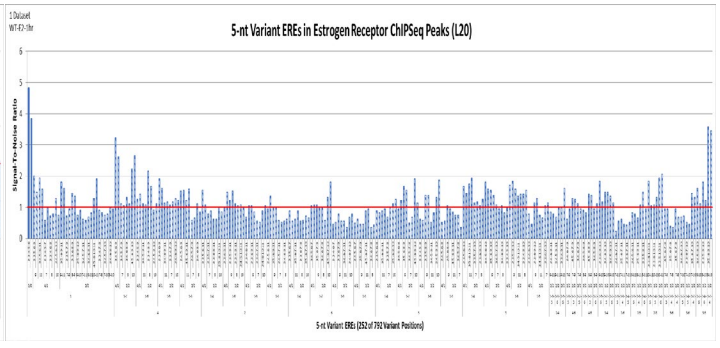
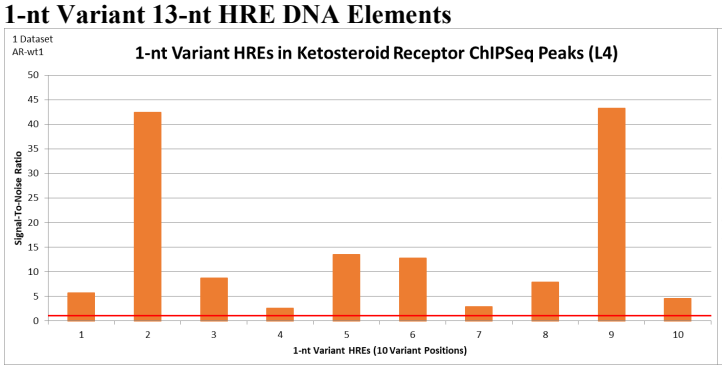
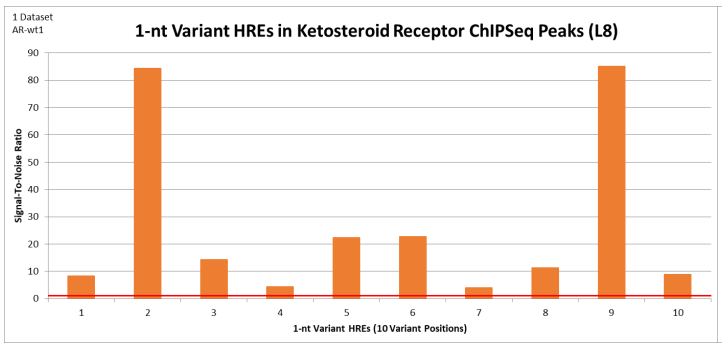
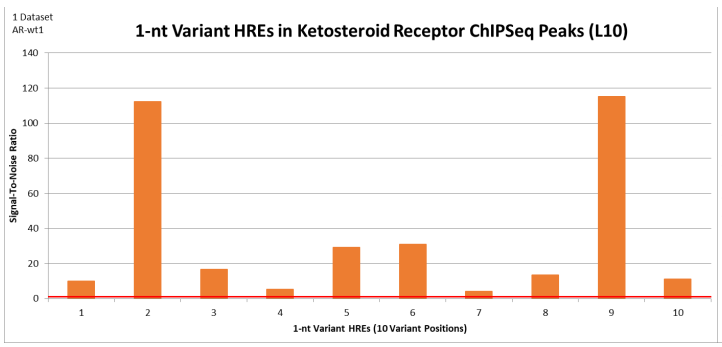


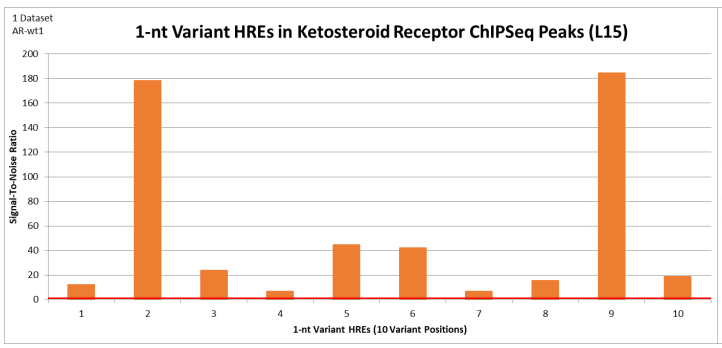
Figure S67. Compare (S/N) analysis of 1-nt Variant HREs in KR ChIPSeq Peaks (13-nt vs. 15-nt)

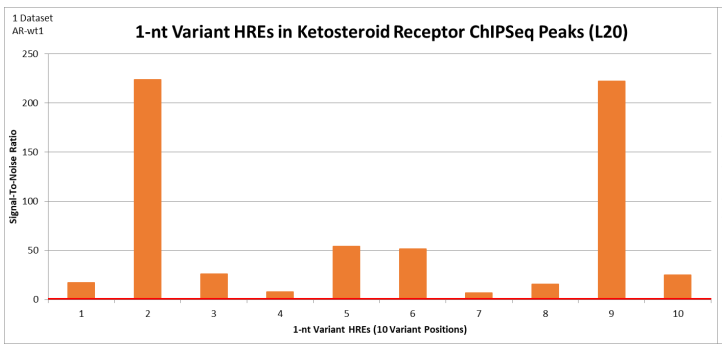
1-nt Variant 13-nt HRE DNA Elements



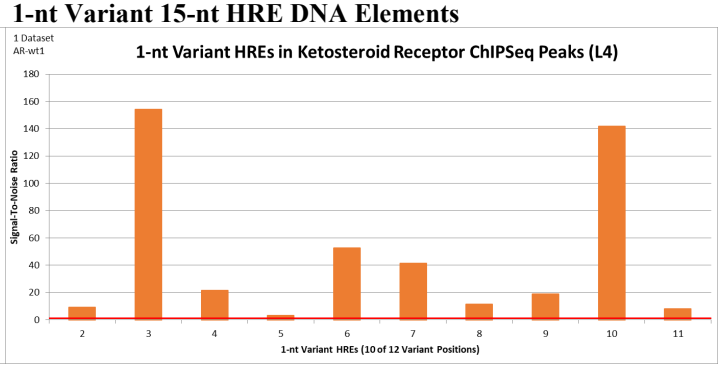


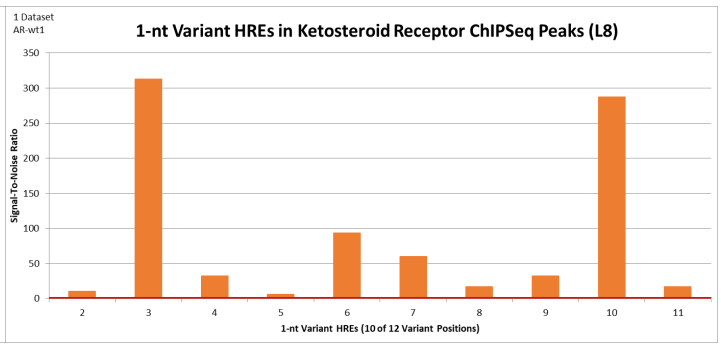


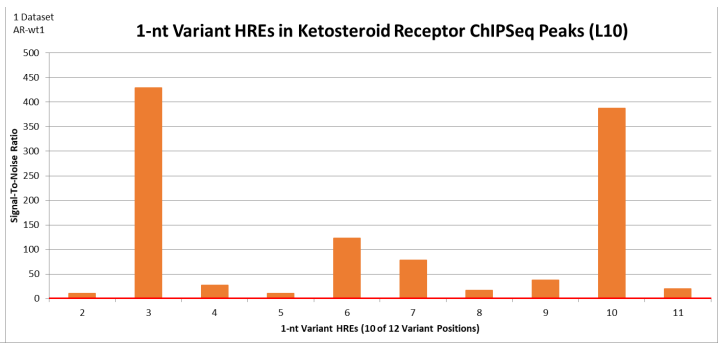


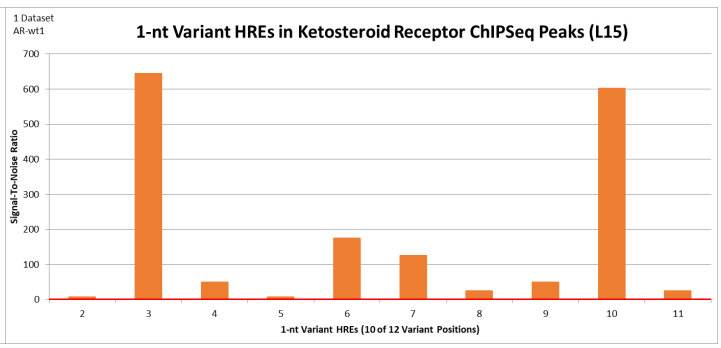


1-nt Variant 15-nt HRE DNA Elements









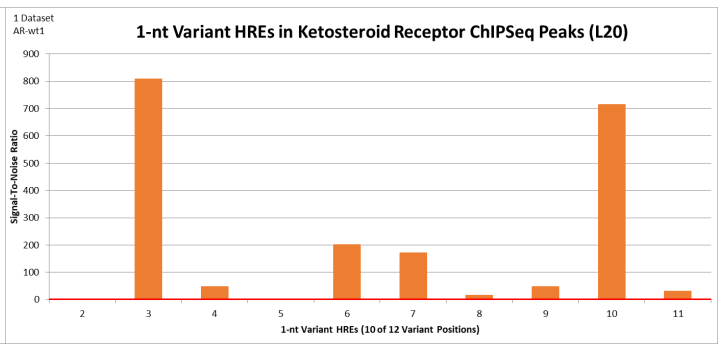
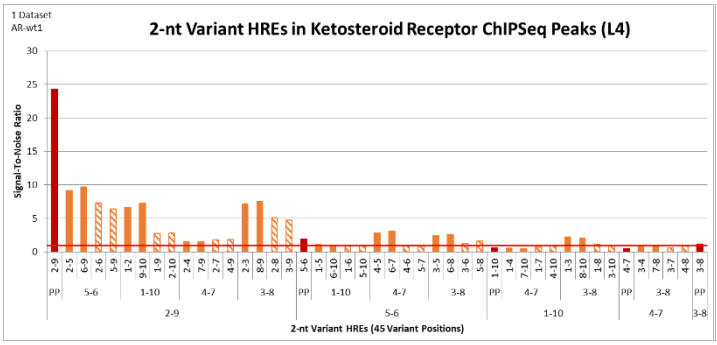
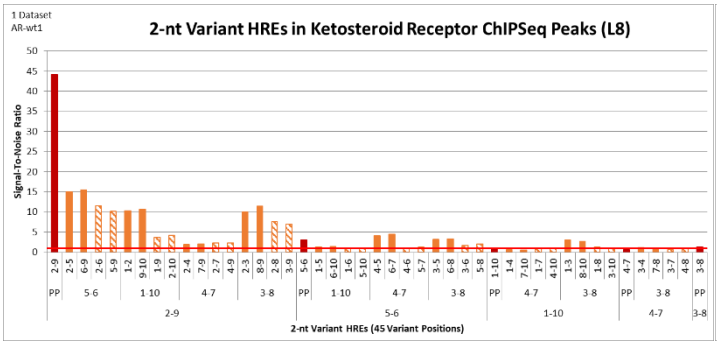
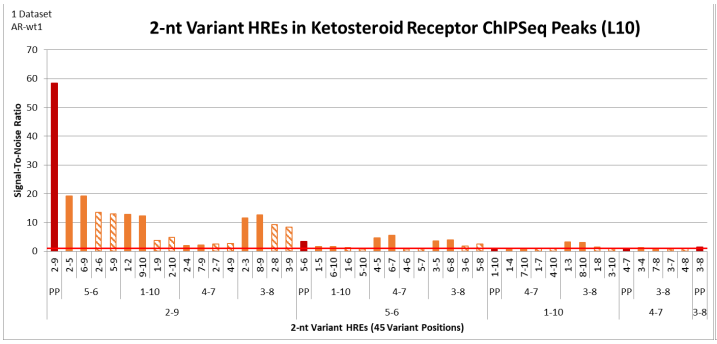


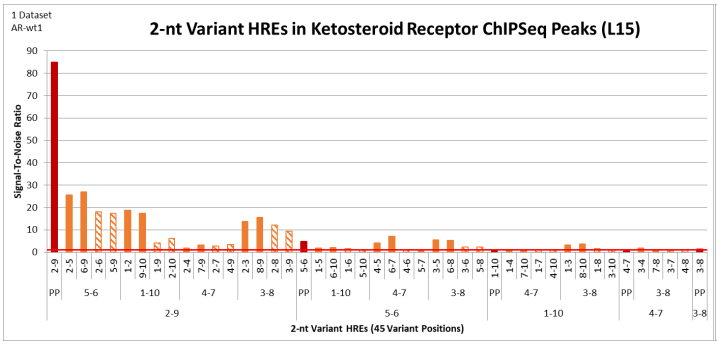
Figure S68. Compare (S/N) analysis of 2-nt Variant HREs in KR ChIPSeq Peaks (13-nt vs. 15-nt)

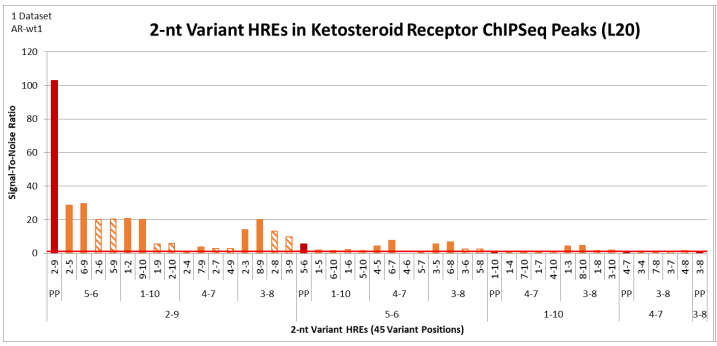
2-nt Variant 13-nt HRE DNA Elements



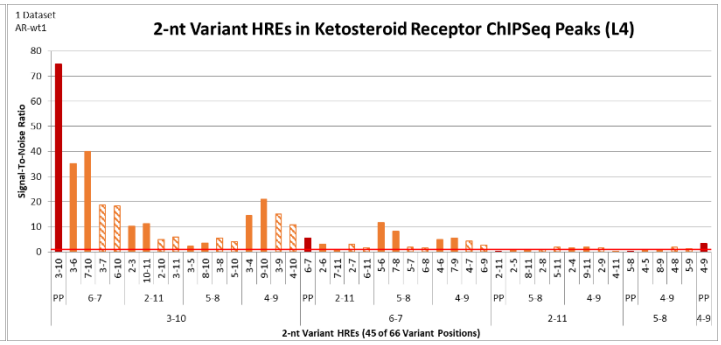


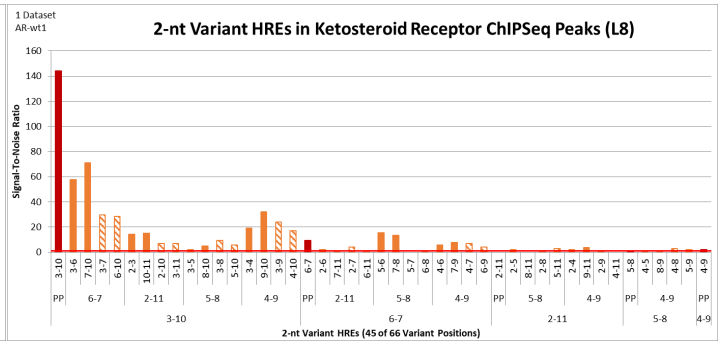


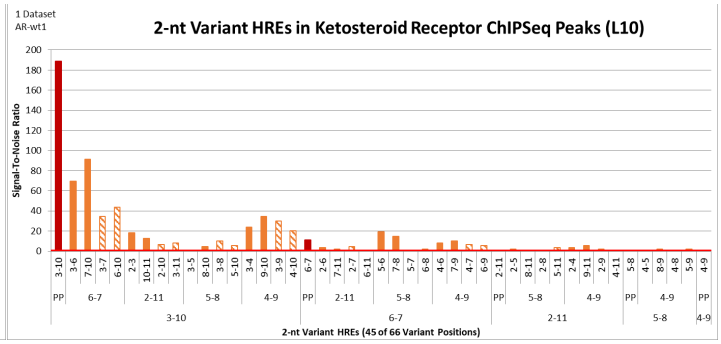


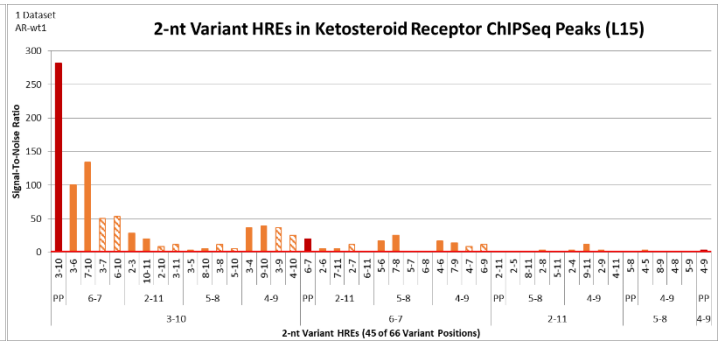


2-nt Variant 15-nt HRE DNA Elements









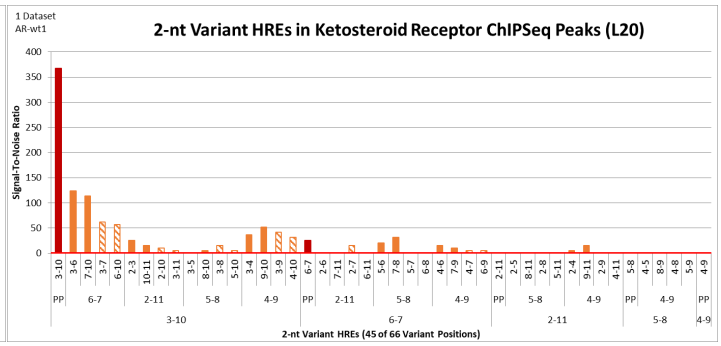
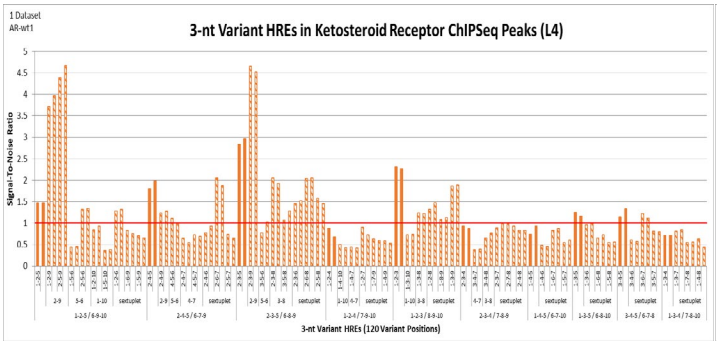
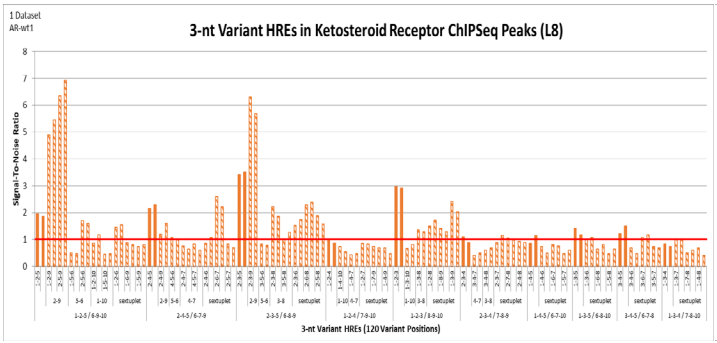
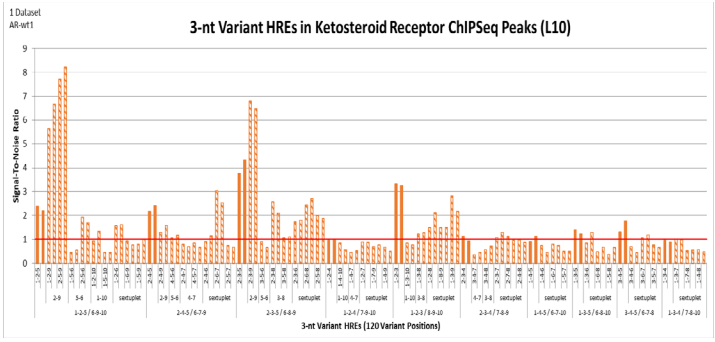


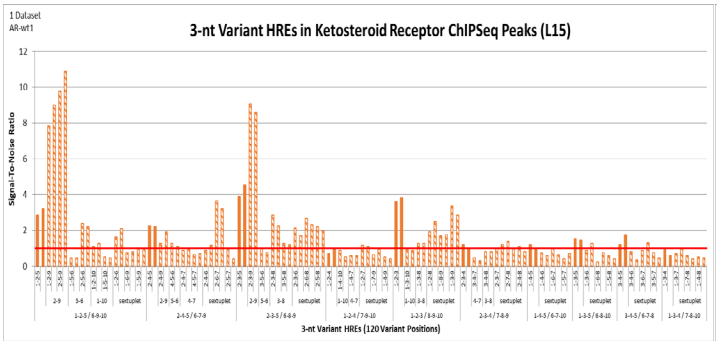
Figure S69. Compare (S/N) analysis of 3-nt Variant HREs in KR ChIPSeq Peaks (13-nt vs. 15-nt)

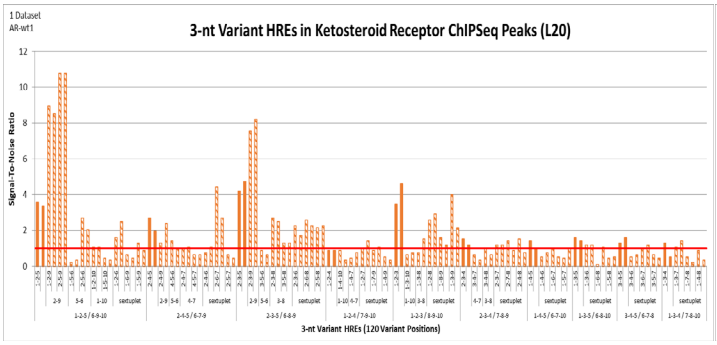
3-nt Variant 13-nt HRE DNA Elements



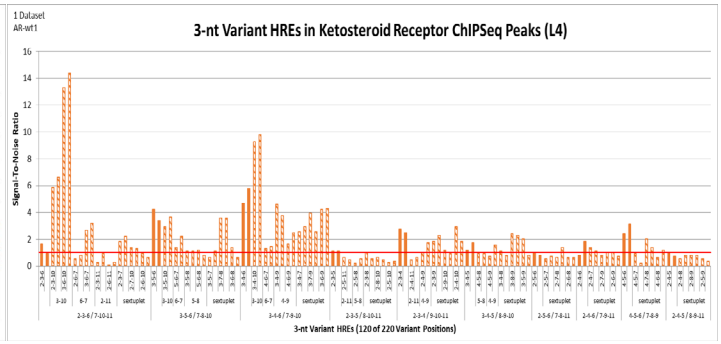


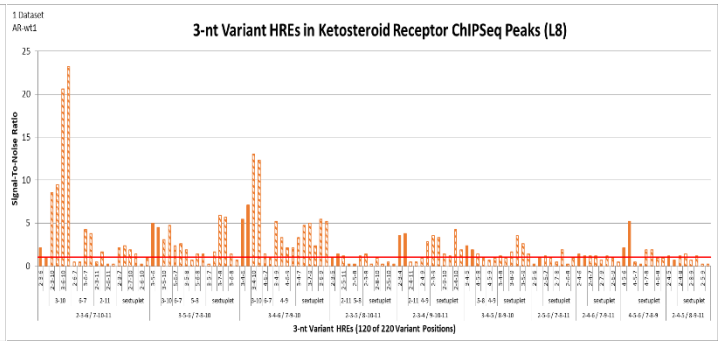


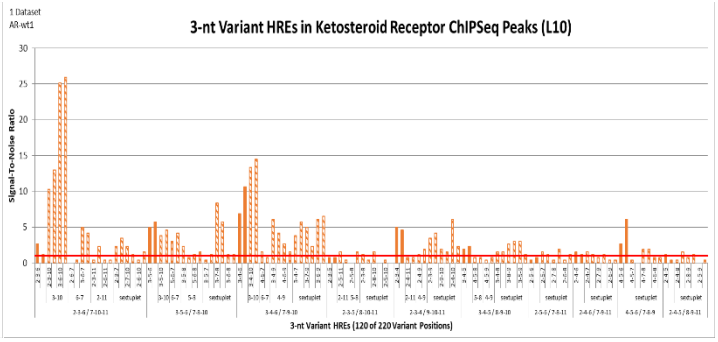


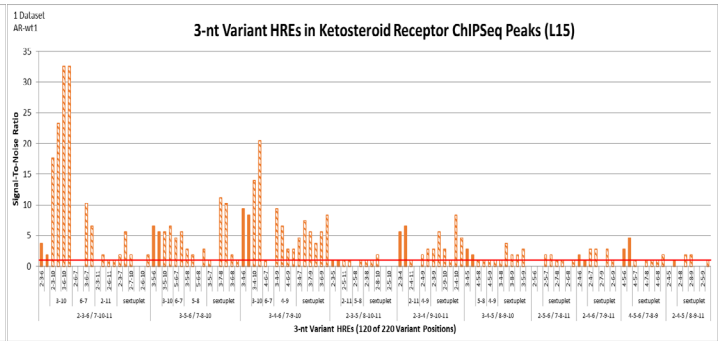


3-nt Variant 15-nt HRE DNA Elements









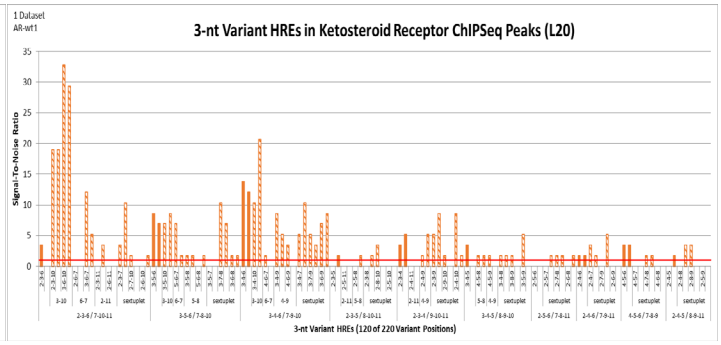
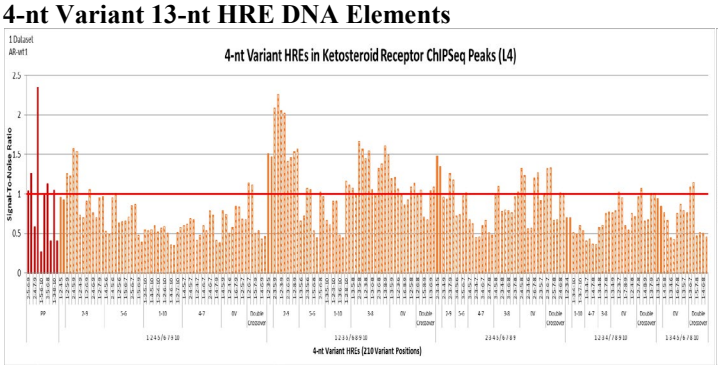
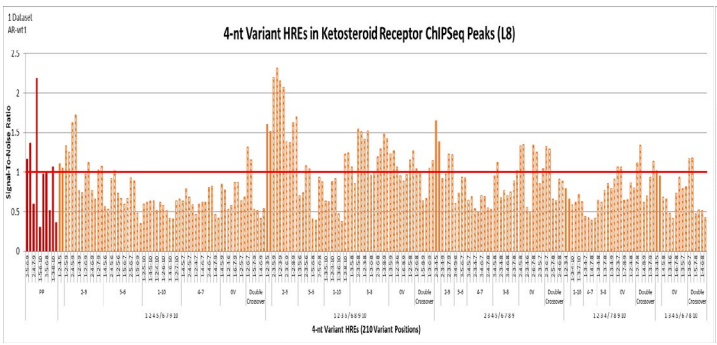
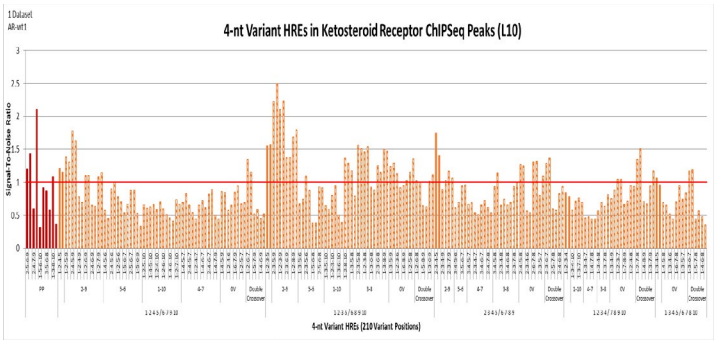


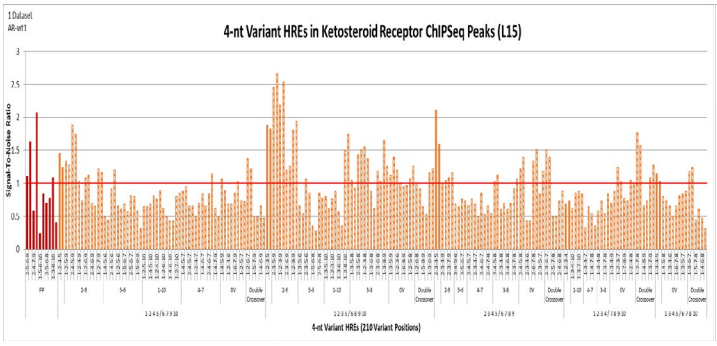
Figure S70. Compare (S/N) analysis of 4-nt Variant HREs in KR ChIPSeq Peaks (13-nt vs. 15-nt)

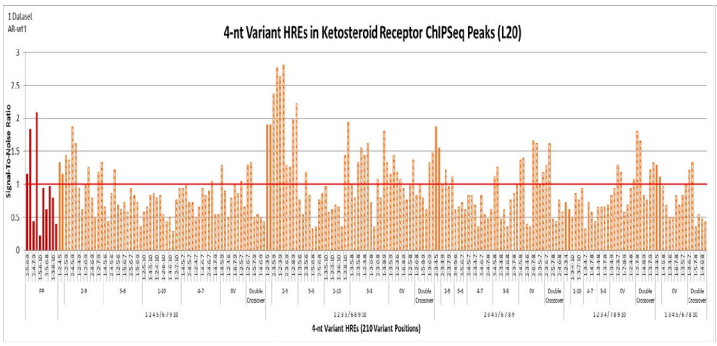
4-nt Variant 13-nt HRE DNA Elements



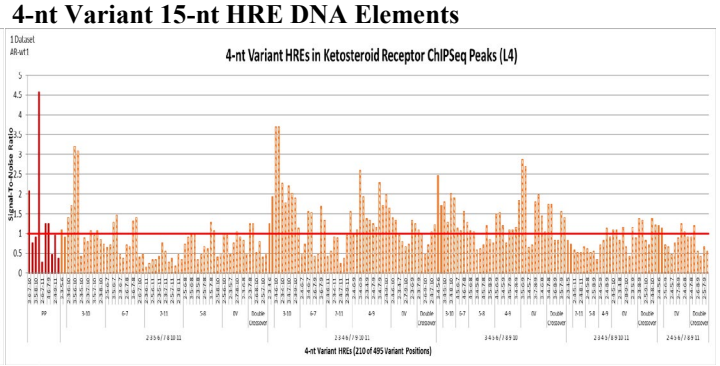


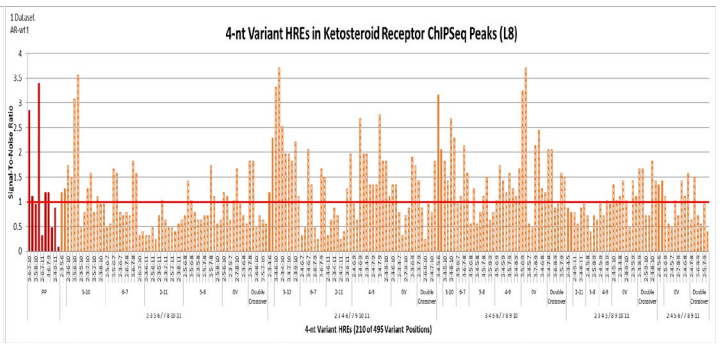


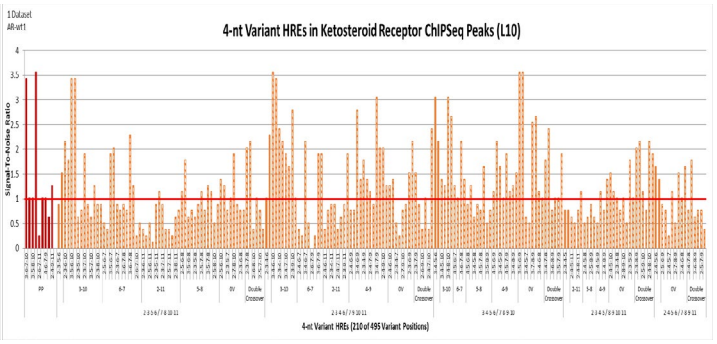


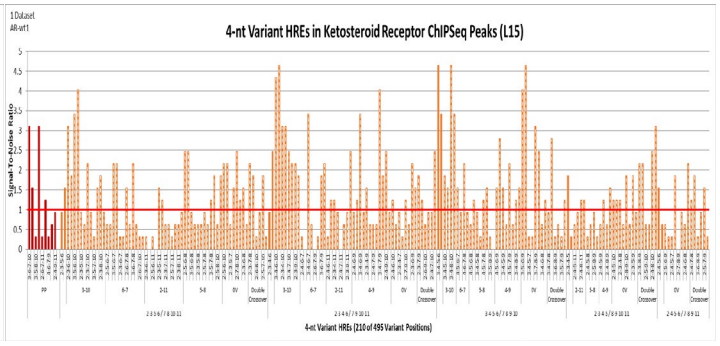


4-nt Variant 15-nt HRE DNA Elements









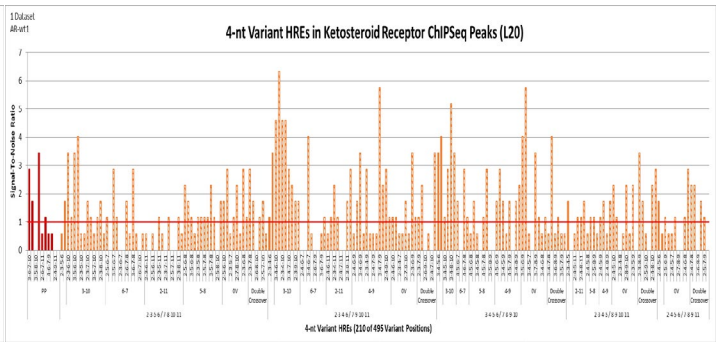
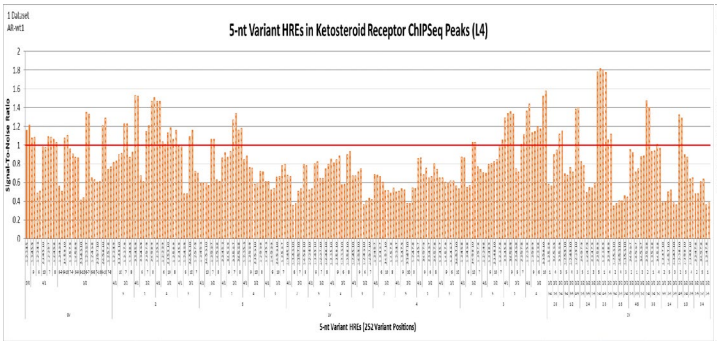
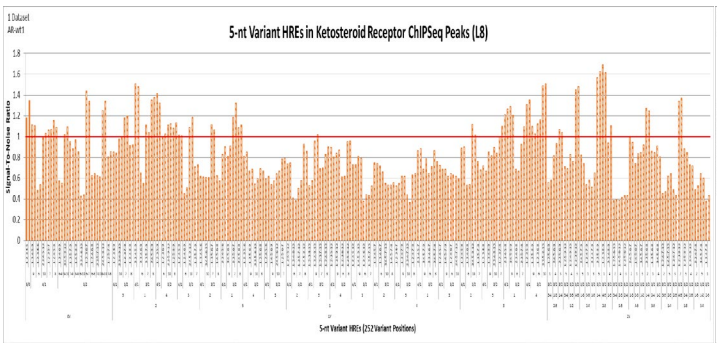
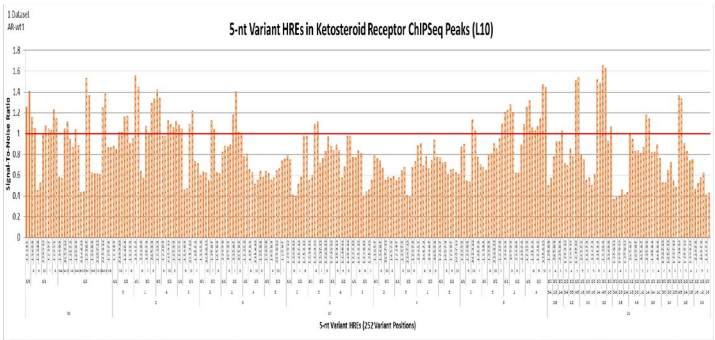


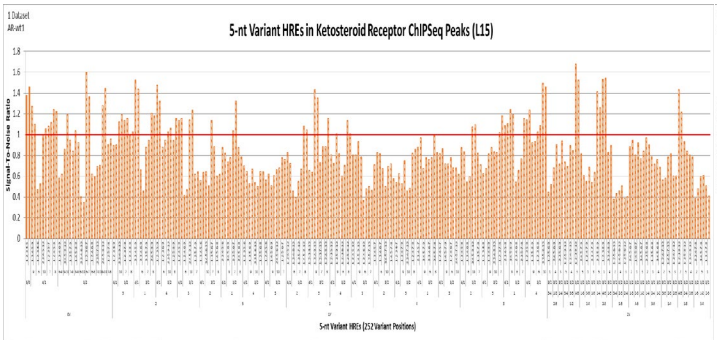
Figure S71. Compare (S/N) analysis of 5-nt Variant HREs in KR ChIPSeq Peaks (13-nt vs. 15-nt)

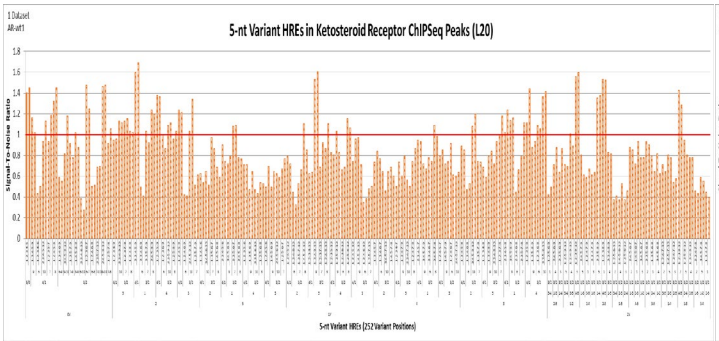
5-nt Variant 13-nt HRE DNA Elements



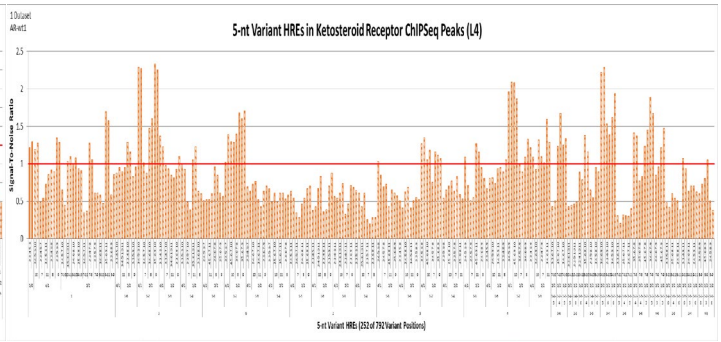


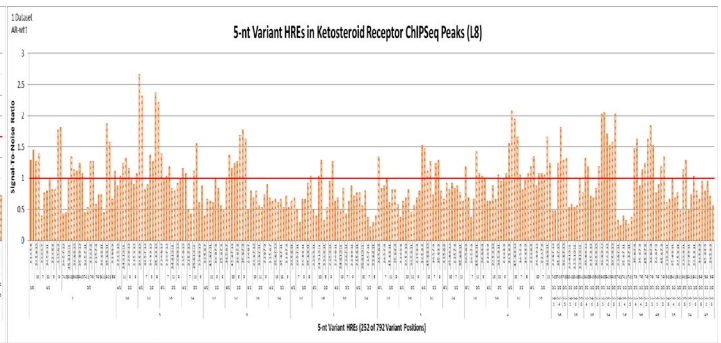


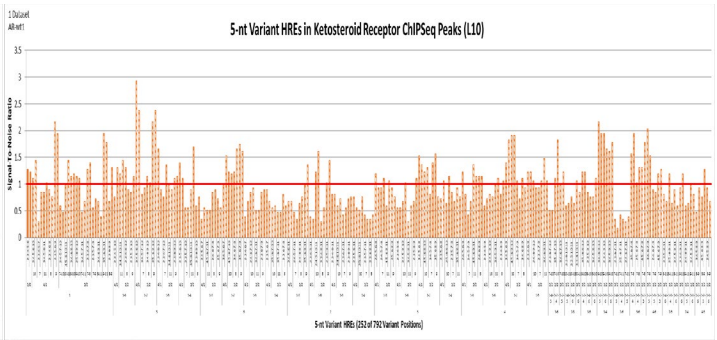


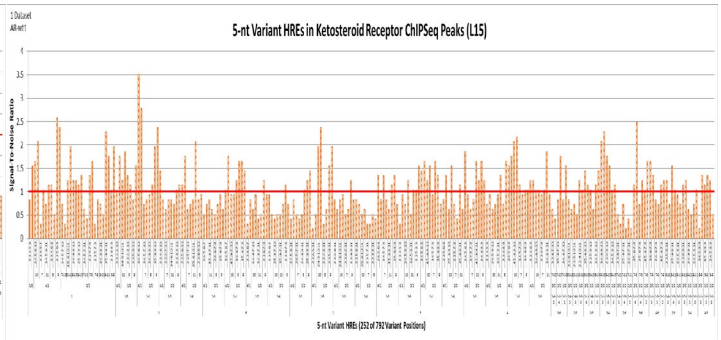


5-nt Variant 15-nt HRE DNA Elements









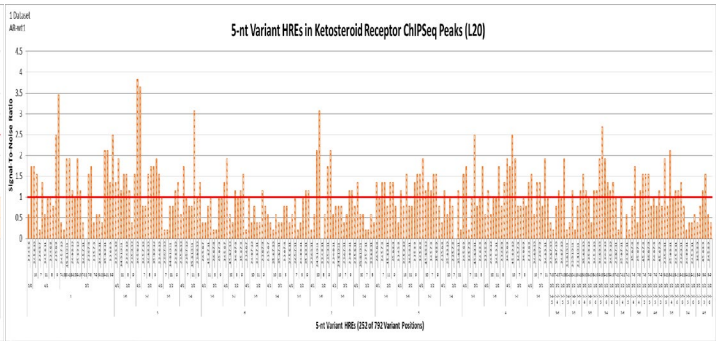


Figure S72-S77 Descriptions: Transform the 13-nt DNA Element Analysis to the 15-nt DNA Element Analysis

****All detailed data and statistics associated with every figure are compiled in Data S1****

Because the analysis of sNR DNA-binding at 13-nt ERE and HRE DNA elements was done by overlapping the location coordinates of each 0-nt to 5-nt variant ERE or HRE DNA element in the genome and the location coordinates of the ChIPSeq or ChIPExo peaks from each experiment, the counts of all sNR DNA-binding events at 15-nt ERE or HRE DNA elements in the genome are included in the 13-nt ERE and HRE DNA element analysis (i.e., every 13-nt ERE and HRE DNA element in the genome is part of a 15-nt ERE and HRE DNA element). Thus, assessing the role of the flanking nucleotides in a 15-nt ERE or HRE DNA element on sNR DNA-binding in the genome requires segregating the counts from the 13-nt ERE and HRE DNA element analysis and displaying them separately. Therefore, the 13-nt ERE or HRE DNA element with 10 primary positions is converted to a 15-nt ERE or HRE DNA element with 12 primary positions.

Analysis of sNR DNA-binding at 13-nt ERE and HRE DNA elements included the 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements (10 variant positions), 405 2-nt variant DNA elements (45 variant positions), 3,240 3-nt variant DNA elements (120 variant positions), 17,010 4-nt variant DNA elements (210 variant positions), and 61,236 5-nt variant DNA elements (252 variant positions), for a total of 81,922 DNA elements. Analysis of sNR DNA-binding at 15-nt ERE and HRE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 36 1-nt variant DNA elements (12 variant positions), 594 2-nt variant DNA elements (66 variant positions), 5,940 3-nt variant DNA elements (220 variant positions), 40,095 4-nt variant DNA elements (495 variant positions), 192,456 5-nt variant DNA elements (792 variant positions), and 673,596 6-nt variant DNA elements (924 variant positions), for a total of 912,718 DNA elements.

Figure S72. Transform the 13-nt DNA Element Analysis to the 15-nt DNA Element Analysis
Each of the 81,922 0-nt to 5-nt variant 13-nt DNA elements splits into 16 categories of the 15-nt DNA elements. When an adenine (A) is in position 1 and a thymine (T) is in position 12 of the 15-nt DNA element, this is the 0-nt variant difference DNA element [=1 DNA element]. When one of the three (3) alternative nucleotide possibilities are in position 1 (C, G, T) or position 12 (A, C, G), these are 1-nt variant difference DNA elements [=6 DNA elements]. When one of the three (3) alternative nucleotide possibilities are in position 1 and position 12, these are 2-nt variant difference DNA elements [=9 DNA elements].

Using WT-E2-1hr.L4 [76,163 peaks, 146-nt peak length] as an example experiment, the 0-nt variant consensus palindromic 13-nt ERE DNA element (5′-GGTCAnnnTGACC-3′) occurred 1,202 times in a particular ER experiment (WT-E2-1hr.L4) (Figure S2). Each of those 1,202 0-nt variant consensus palindromic 13-nt ERE DNA elements are part of a 15-nt ERE DNA element. Thus, analysis of ER DNA-binding at all possible 15-nt ERE DNA elements splits those 1,202 0-nt variant consensus palindromic 13-nt ERE DNA elements into 16 categories (i.e., 4-nt possibilities in position 1 and 4-nt possibilities in position 12): 109 (1T), 155 (1G), 19 (1C), 13 (12G), 151 (12C), 110 (12A), 11 (1-12 TG), 67 (1-12 TC), 17 (1-12 GG), 305 (1-12 AT), 29 (1-12 TA), 107 (1-12 GC), 5 (1-12 CG), 14 (1-12 CC), 88 (1-12 GA), 2 (1-12 CA), totaling 1,202 **(Figure S33).**

However, there are some DNA elements that are unique to each list (i.e., the 81,922 0-nt to 5-nt variant 13-nt DNA elements and the 912,718 0-nt to 6-nt variant 15-nt DNA elements). The 81,922 0-nt to 5-nt variant 13-nt DNA element list includes 551,124 DNA elements that are not on the 912,718 0-nt to 6-nt variant 15-nt DNA element list (these represent a 7-nt variant of the 15-nt DNA elements). The 912,718 0-nt to 6-nt variant 15-nt DNA element list includes 153,090 DNA elements that are not on the 81,922 0-nt to 5-nt variant 13-nt DNA element list (these represent a 6-nt variant of the 13-nt DNA elements). Thus, of the 81,922 0-nt to 5-nt variant 13-nt DNA element list and the 912,718 0-nt to 6-nt variant 15-nt DNA element list, 759,628 DNA elements correspond to each other. The splitting of the 81,922 0-nt to 5-nt variant 13-nt ERE and HRE DNA elements into the 16 categories of the 15-nt ERE and HRE DNA elements can be found in **Table S41-S46**.

Using WT-E2-1hr.L4 [76,163 peaks, 146-nt peak length] as an example experiment, there are 827,919 0-nt to 5-nt variant 13-nt EREs identified by overlapping the location coordinates of the 81,922 0-nt to 5-nt variant 13-nt EREs in the genome and the location coordinates of the ChIPSeq peaks in this experiment. There are 605,166 0-nt to 6-nt variant 15-nt EREs identified by overlapping the location coordinates of the 912,718 0-nt to 6-nt variant 15-nt EREs in the genome and the location coordinates of the ChIPSeq peaks in this experiment. Of the 605,166 0-nt to 6-nt variant 15-nt EREs identified in the experiment, 95,296 of those are the 153,090 DNA elements that are unique to the 15-nt ERE list (i.e., they are 6-nt variant 15-nt variant EREs that categorize to 6-nt variant 13-nt variant EREs) (=509,870). Of the 827,919 0-nt to 5-nt variant 13-nt EREs identified, 314,956 of those are the 551,124 DNA elements that are unique to the 13-nt ERE list (i.e., they are 5-nt variant 13-nt variant EREs that categorize to 7-nt variant 15-nt variant EREs) (=512,963). Of those 512,963 0-nt to 5-nt variant 13-nt EREs, 3093 of them were missing from the 15-nt ERE analysis due to their location being at the ends of the ChIPSeq peaks (i.e., the entire 13-nt DNA element was required to be within the peak boundaries for the 13-nt ERE analysis and the entire 15-nt DNA element was required to be within the peak boundaries for the 15-nt ERE analysis) (=509,870). Thus, in the WT-E2-1hr.L4 experiment, 509,870 are legitimately represented in both the 13-nt DNA element analysis and 15-nt DNA element analysis.

Figure S73. Transform the 13-nt DNA Element Analysis to the 15-nt DNA Element Analysis (1-nt Variant EREs)
Splitting (S/N) analysis of 1-nt variant 13-nt EREs in ER (WT-E2-1hr) [76,163 peaks, 146-nt peak length] ChIPSeq peaks into the 16 categories of the 15-nt ERE DNA elements (i.e., 1 stays in the 0-nt variant difference category, 6 move to the 1-nt variant difference category, 9 move to the 2-nt variant difference category). The distribution ratio into each of the three categories should stay constant if the two flanking nucleotides of the 15-nt DNA element are not playing any role. Therefore, although the DNA-binding profile at the 13-nt DNA elements in the genome versus at the 15-nt DNA elements in the genome are almost equivalent (confirming that the DNA-binding rules established by the 13-nt DNA element analysis also apply to the 15-nt DNA element analysis), there are minor effects. For example, by fixing the flanking nucleotides of the 15-nt ERE (adenine (A) is in position 1 and thymine (T) is in position 12), ER DNA-binding is reduced at 15-nt EREs that have variants in position 2 or position 11.

Figure S74. Transform the 13-nt DNA Element Analysis to the 15-nt DNA Element Analysis (2-nt Variant EREs)
Splitting (S/N) analysis of 2-nt variant 13-nt EREs in ER (WT-E2-1hr) [76,163 peaks, 146-nt peak length] ChIPSeq peaks into the 16 categories of the 15-nt ERE DNA elements (i.e., 1 stays in the 0-nt variant difference category, 6 move to the 1-nt variant difference category, 9 move to the 2-nt variant difference category). The distribution ratio into each of the three categories should stay constant if the two flanking nucleotides of the 15-nt DNA element are not playing any role. Therefore, although the DNA-binding profile at the 13-nt DNA elements in the genome versus at the 15-nt DNA elements in the genome are almost equivalent (confirming that the DNA-binding rules established by the 13-nt DNA element analysis also apply to the 15-nt DNA element analysis), there are minor effects.

Figure S75. Transform the 13-nt DNA Element Analysis to the 15-nt DNA Element Analysis (3-nt Variant EREs)
Splitting (S/N) analysis of 3-nt variant 13-nt EREs in ER (WT-E2-1hr) [76,163 peaks, 146-nt peak length] ChIPSeq peaks into the 16 categories of the 15-nt ERE DNA elements (i.e., 1 stays in the 0-nt variant difference category, 6 move to the 1-nt variant difference category, 9 move to the 2-nt variant difference category). The distribution ratio into each of the three categories should stay constant if the two flanking nucleotides of the 15-nt DNA element are not playing any role. Therefore, although the DNA-binding profile at the 13-nt DNA elements in the genome versus at the 15-nt DNA elements in the genome are almost equivalent (confirming that the DNA-binding rules established by the 13-nt DNA element analysis also apply to the 15-nt DNA element analysis), there are minor effects.

Figure S76. Transform the 13-nt DNA Element Analysis to the 15-nt DNA Element Analysis (4-nt Variant EREs)
Splitting (S/N) analysis of 4-nt variant 13-nt EREs in ER (WT-E2-1hr) [76,163 peaks, 146-nt peak length] ChIPSeq peaks into the 16 categories of the 15-nt ERE DNA elements (i.e., 1 stays in the 0-nt variant difference category, 6 move to the 1-nt variant difference category, 9 move to the 2-nt variant difference category). The distribution ratio into each of the three categories should stay constant if the two flanking nucleotides of the 15-nt DNA element are not playing any role. Therefore, although the DNA-binding profile at the 13-nt DNA elements in the genome versus at the 15-nt DNA elements in the genome are almost equivalent (confirming that the DNA-binding rules established by the 13-nt DNA element analysis also apply to the 15-nt DNA element analysis), there are minor effects.

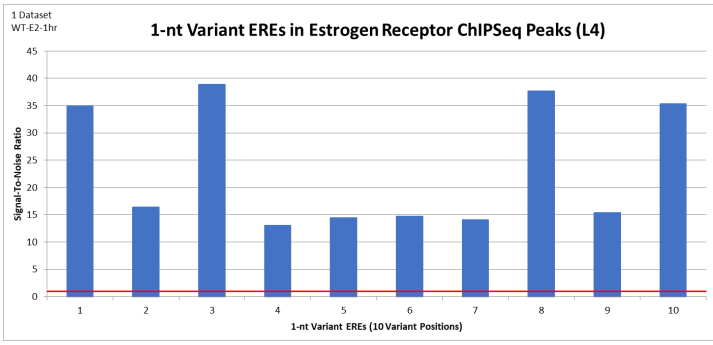
Figure S77. Transform the 13-nt DNA Element Analysis to the 15-nt DNA Element Analysis (5-nt Variant EREs)
Splitting (S/N) analysis of 5-nt variant 13-nt EREs in ER (WT-E2-1hr) [76,163 peaks, 146-nt peak length] ChIPSeq peaks into the 16 categories of the 15-nt ERE DNA elements (i.e., 1 stays in the 0-nt variant difference category, 6 move to the 1-nt variant difference category, 9 move to the 2-nt variant difference category). The distribution ratio into each of the three categories should stay constant if the two flanking nucleotides of the 15-nt DNA element are not playing any role. Therefore, although the DNA-binding profile at the 13-nt DNA elements in the genome versus at the 15-nt DNA elements in the genome are almost equivalent (confirming that the DNA-binding rules established by the 13-nt DNA element analysis also apply to the 15-nt DNA element analysis), there are minor effects. The green stripped bars represent the 314,956 counts that are 5-nt variant 13-nt variant EREs that represent 7-nt variant 15-nt variant EREs (i.e., the 551,124 DNA elements that are unique to the 13-nt ERE list).

Figure S72. Transform the 13-nt DNA Element Analysis to the 15-nt DNA Element Analysis

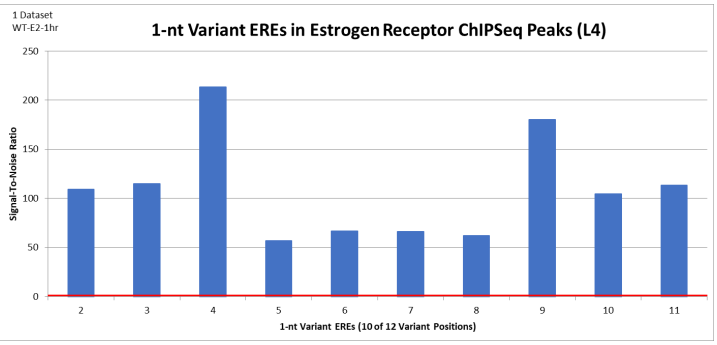
			15-nt DNA Element Analysis								
			1	**36**	**594**	**5,940**	**40,095**	**192,456**	**673,596**	**1,732,104**	
			0-nt Variant	**1-nt Variant**	**2-nt Variant**	**3-nt Variant**	**4-nt Variant**	**5-nt Variant**	**6-nt Variant**	***7-nt Variant***	
13-nt DNA Element Analysis	**1**	**0-nt Variant**	1	6	9						**16**
	30	**1-nt Variant**		30	180	270					**480**
	405	**2-nt Variant**			405	2,430	3,645				**6,480**
	3,240	**3-nt Variant**				3,240	19,440	29,160			**51,840**
	17,010	**4-nt Variant**					17,010	102,060	153,090		**272,160**
	61,236	**5-nt Variant**						61,236	367,416	551,124	**428,652**
	153,090	***6-nt Variant***							153,090		
											759,628

Figure S73. Transform the 13-nt DNA Element Analysis to the 15-nt DNA Element Analysis (1-nt Variant EREs)

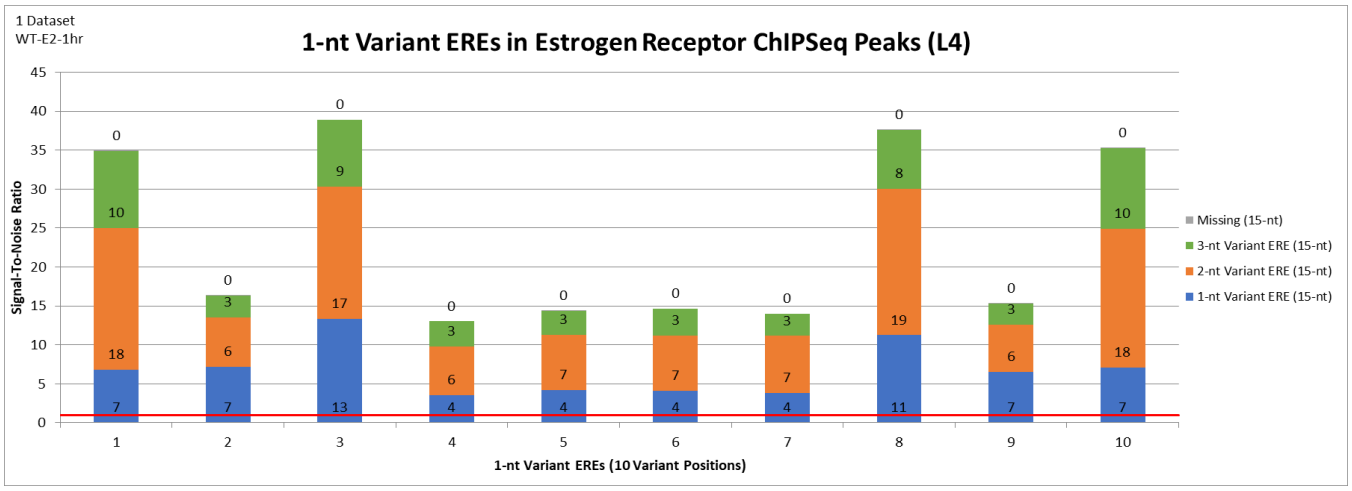
1-nt Variant 13-nt ERE DNA Elements



1-nt Variant 15-nt ERE DNA Elements



Categorize (S/N) Analysis of ER DNA-Binding at 1-nt Variant 13-nt EREs Into Three Categories of 15-nt EREs



Categorize (S/N) Analysis of ER DNA-Binding at 1-nt Variant 13-nt EREs Into Three Categories of 15-nt EREs (% of (S/N))

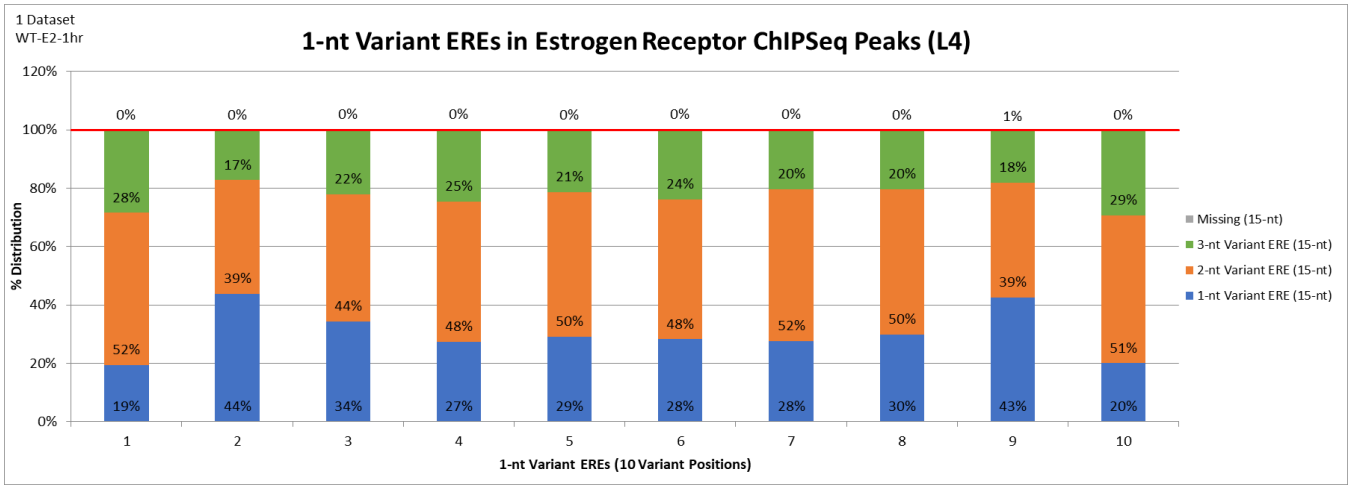
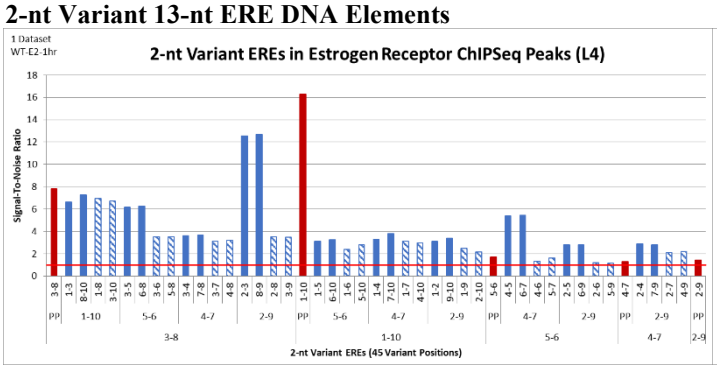
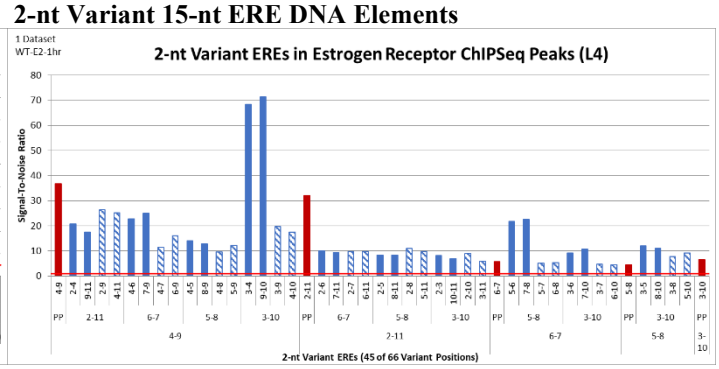


Figure S74. Transform the 13-nt DNA Element Analysis to the 15-nt DNA Element Analysis (2-nt Variant EREs)

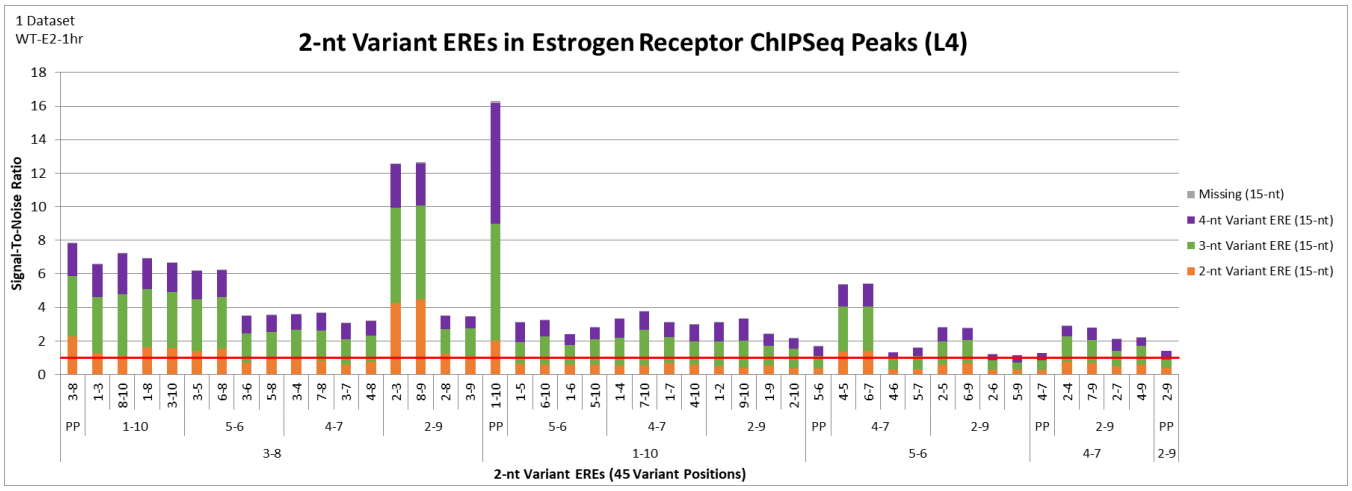
2-nt Variant 13-nt ERE DNA Elements



2-nt Variant 15-nt ERE DNA Elements



Categorize (S/N) Analysis of ER DNA-Binding at 2-nt Variant 13-nt EREs Into Three Categories of 15-nt EREs



Categorize (S/N) Analysis of ER DNA-Binding at 2-nt Variant 13-nt EREs Into Three Categories of 15-nt EREs (% of (S/N))

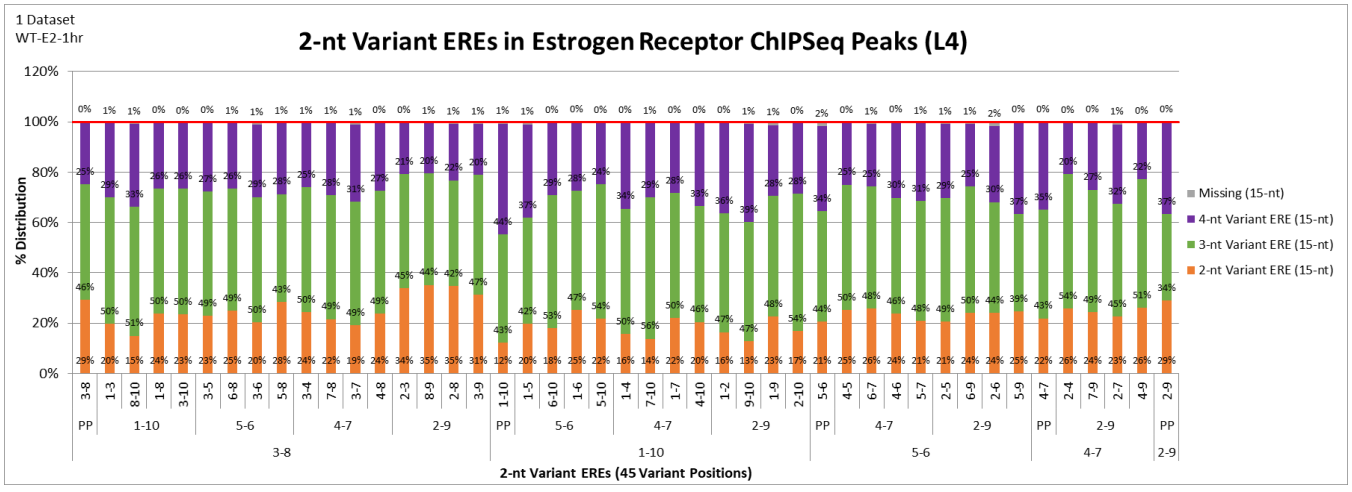
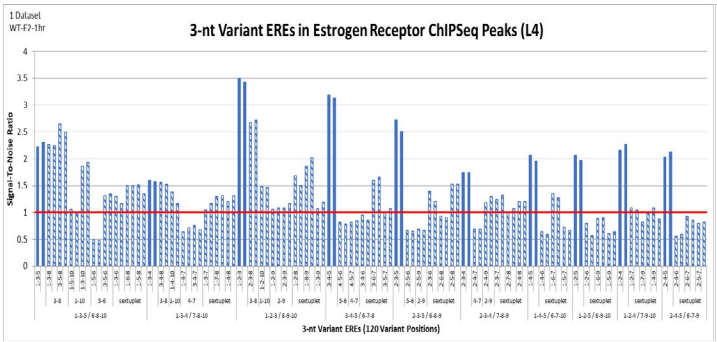
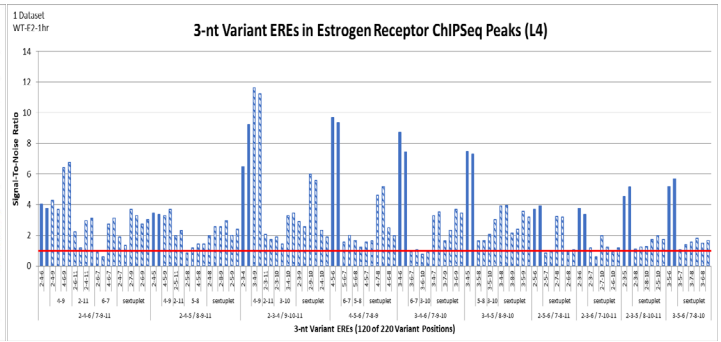


Figure S75. Transform the 13-nt DNA Element Analysis to the 15-nt DNA Element Analysis (3-nt Variant EREs)

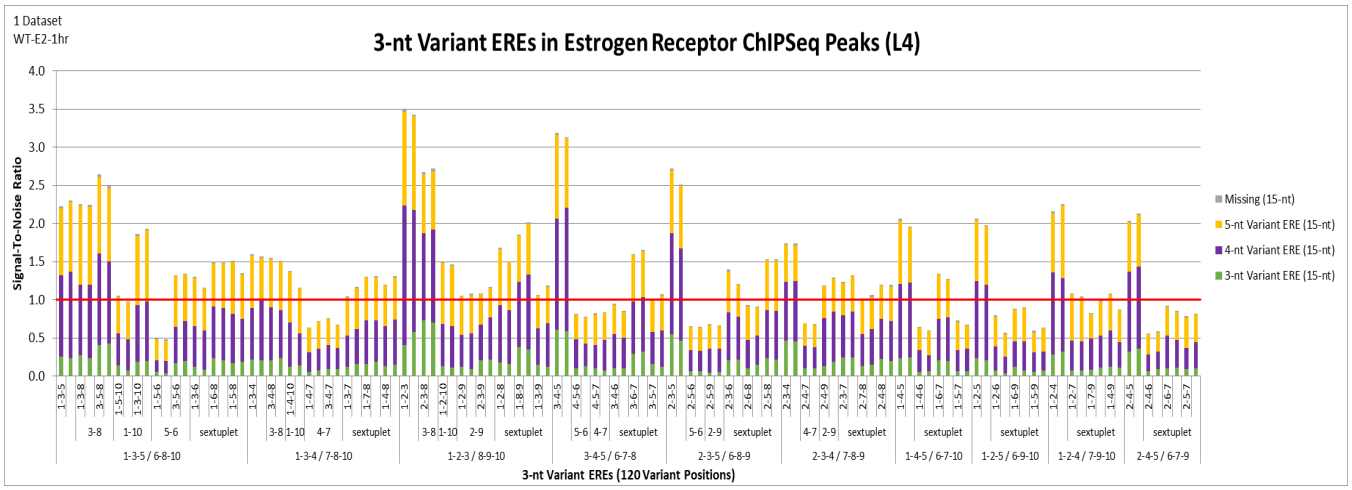
3-nt Variant 13-nt ERE DNA Elements



3-nt Variant 15-nt ERE DNA Elements



Categorize (S/N) Analysis of ER DNA-Binding at 3-nt Variant 13-nt EREs Into Three Categories of 15-nt EREs



Categorize (S/N) Analysis of ER DNA-Binding at 3-nt Variant 13-nt EREs Into Three Categories of 15-nt EREs (% of (S/N))

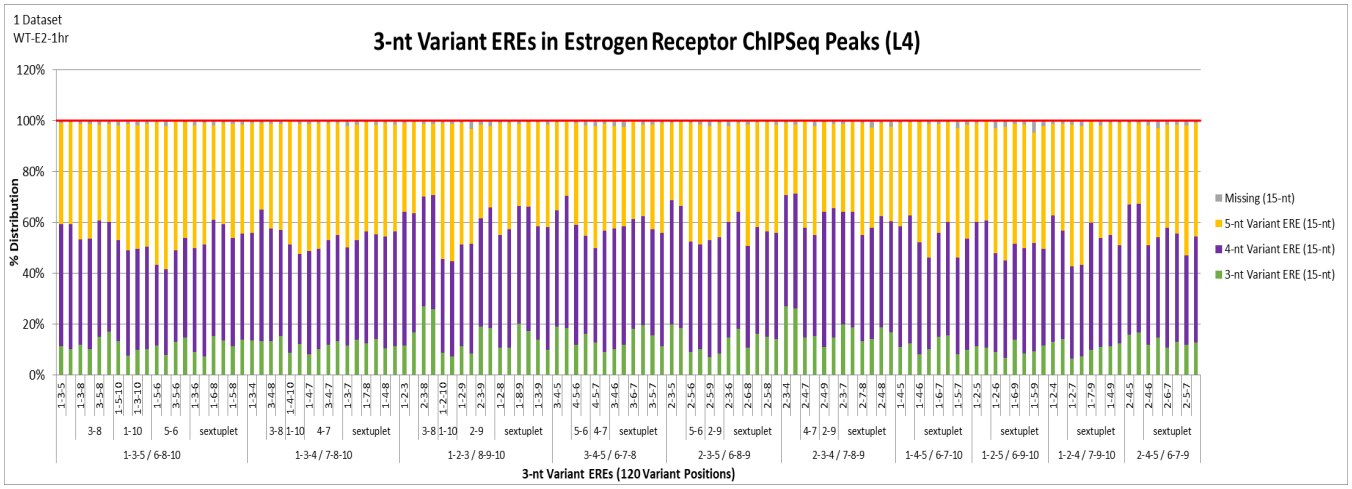
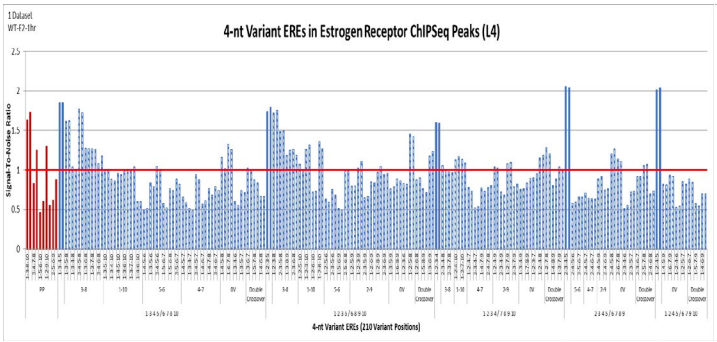
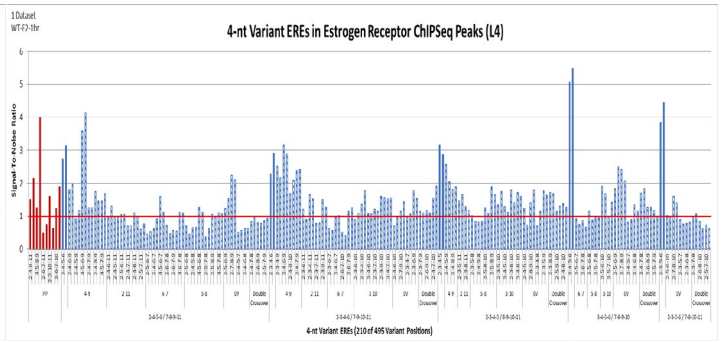


Figure S76. Transform the 13-nt DNA Element Analysis to the 15-nt DNA Element Analysis (4-nt Variant EREs)

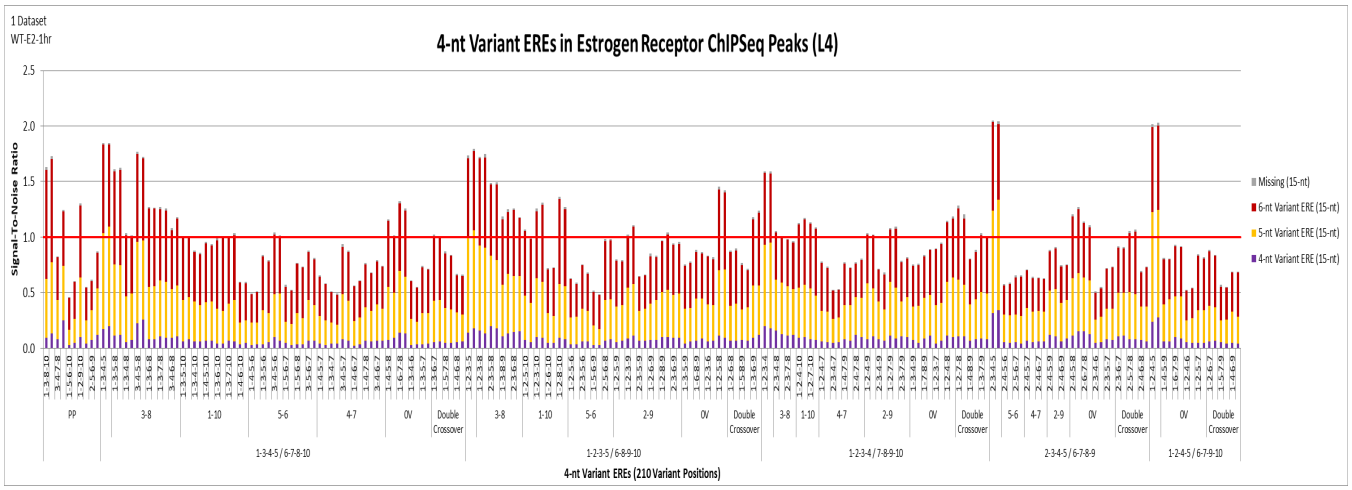
4-nt Variant 13-nt ERE DNA Elements



4-nt Variant 15-nt ERE DNA Elements



Categorize (S/N) Analysis of ER DNA-Binding at 4-nt Variant 13-nt EREs Into Three Categories of 15-nt EREs



Categorize (S/N) Analysis of ER DNA-Binding at 4-nt Variant 13-nt EREs Into Three Categories of 15-nt EREs (% of (S/N))

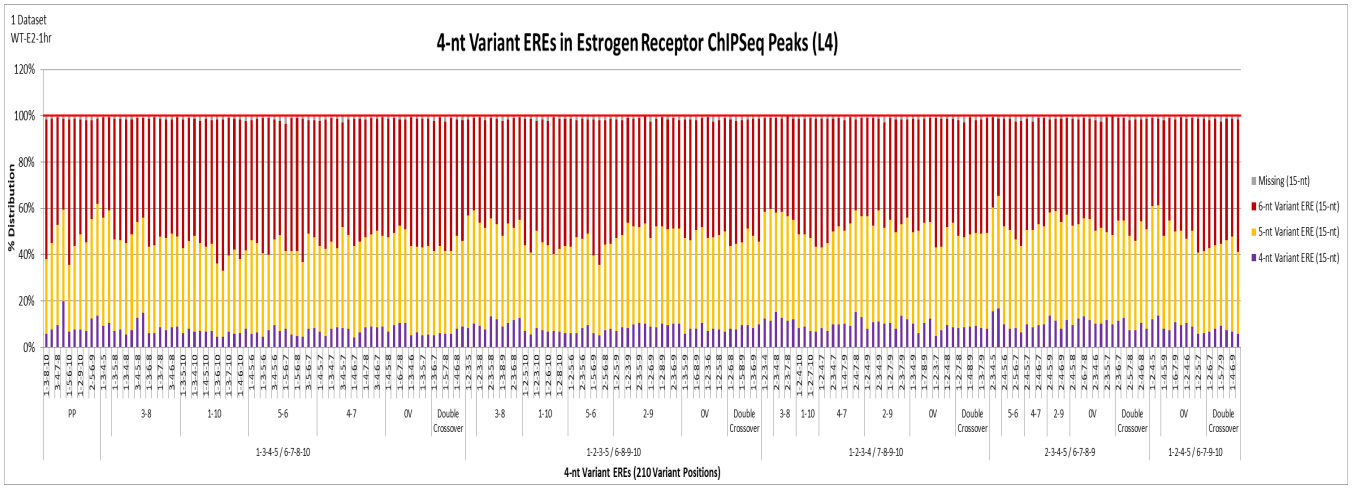
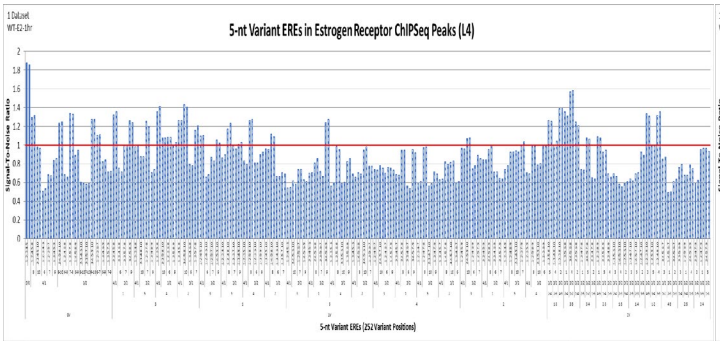
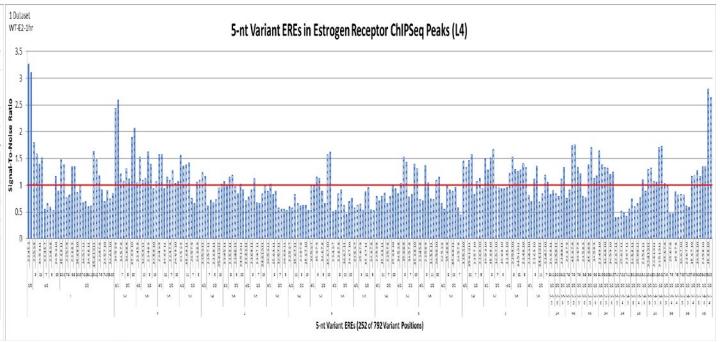


Figure S77. Transform the 13-nt DNA Element Analysis to the 15-nt DNA Element Analysis (5-nt Variant EREs)

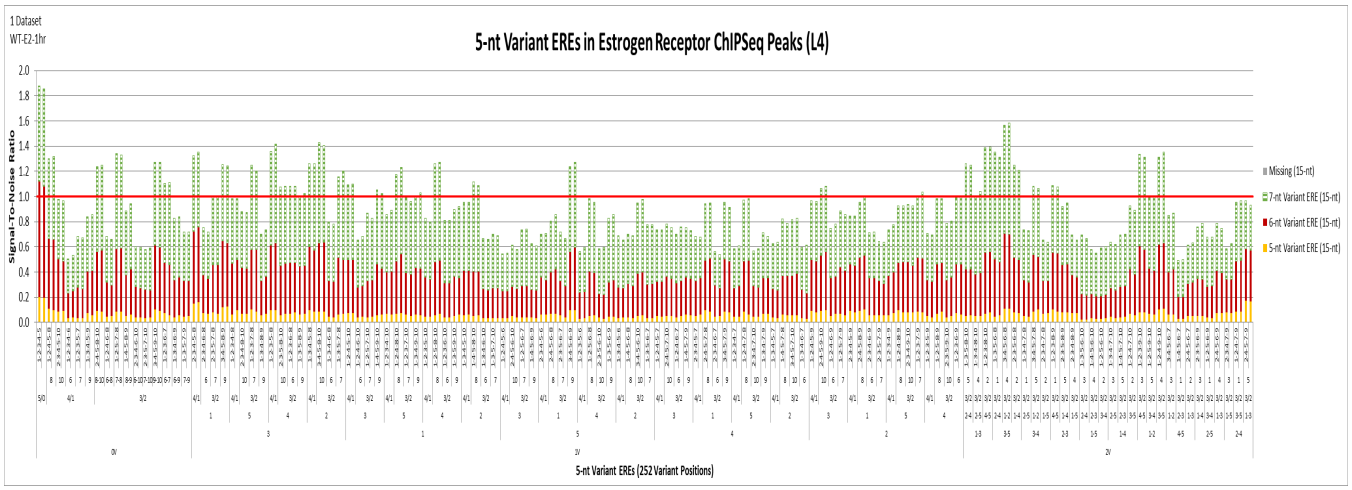
5-nt Variant 13-nt ERE DNA Elements



5-nt Variant 15-nt ERE DNA Elements



Categorize (S/N) Analysis of ER DNA-Binding at 5-nt Variant 13-nt EREs Into Three Categories of 15-nt EREs



Categorize (S/N) Analysis of ER DNA-Binding at 5-nt Variant 13-nt EREs Into Three Categories of 15-nt EREs (% of (S/N))

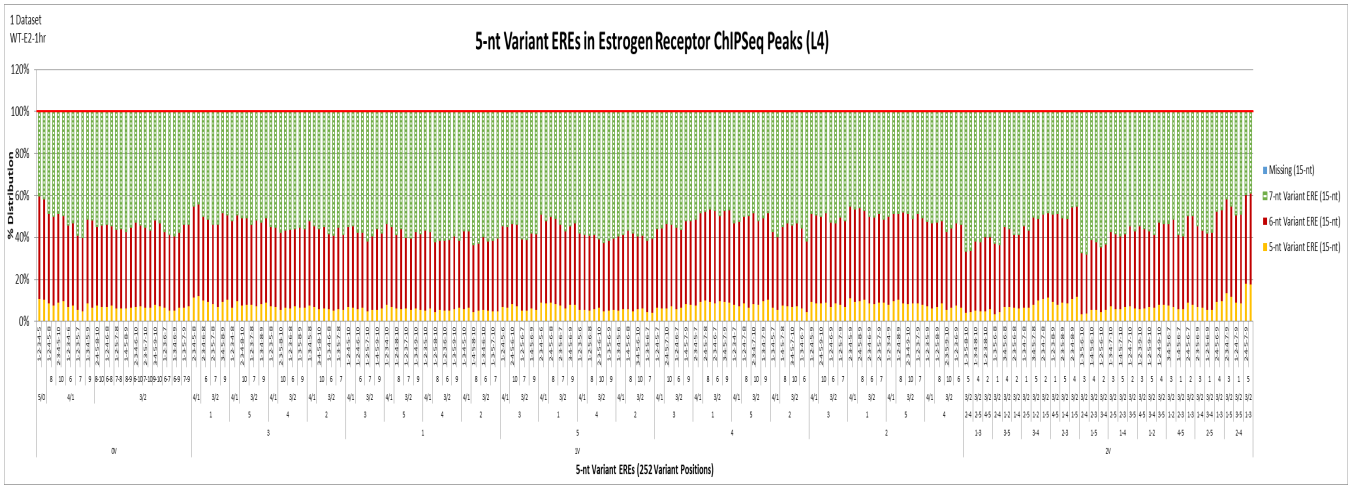


Figure S78-S79 Descriptions: sNR DNA-Binding at 13-nt and 15-nt DNA Elements in the Genome (Time-Course)

****All detailed data and statistics associated with every figure are compiled in Data S1****

Figure S78. (S/N) analysis of 0-nt to 5-nt Variant 13-nt EREs in ER (Time-Course) ChIPSeq Peaks (displayed by 252 half-site groups)
(S/N) analysis of 0-nt to 5-nt variant EREs (displayed by the 252 half-site groups) in ER (MCF7-ERa-E2-2min) [8,476 peaks, 157-nt peak length] (MCF7-ERa-E2-5min) [8,967 peaks, 157-nt peak length] (MCF7-ERa-E2-10min-A) [13,125 peaks, 163-nt peak length] (MCF7-ERa-E2-10min-B) [15,386 peaks, 168-nt peak length] (MCF7-ERa-E2-40min-A) [11,408 peaks, 176-nt peak length] (MCF7-ERa-E2-40min-B) [7,855 peaks, 208-nt peak length] (MCF7-ERa-E2-160min) [3,159 peaks, 164-nt peak length] ChIPSeq peaks. X-axis order =Reverse-Complement Vacancy Position ID 3-8 > 1-10 > 5-6 > 4-7 > 2-9. See **Table S5** for x-axis details.

Figure S79. (S/N) analysis of 0-nt to 6-nt Variant 15-nt EREs in ER (Time-Course) ChIPSeq Peaks (displayed by 924 half-site groups)
(S/N) analysis of 0-nt to 6-nt variant EREs (displayed by the 924 half-site groups) in ER (MCF7-ERa-E2-2min) [8,476 peaks, 157-nt peak length] (MCF7-ERa-E2-5min) [8,967 peaks, 157-nt peak length] (MCF7-ERa-E2-10min-A) [13,125 peaks, 163-nt peak length] (MCF7-ERa-E2-10min-B) [15,386 peaks, 168-nt peak length] (MCF7-ERa-E2-40min-A) [11,408 peaks, 176-nt peak length] (MCF7-ERa-E2-40min-B) [7,855 peaks, 208-nt peak length] (MCF7-ERa-E2-160min) [3,159 peaks, 164-nt peak length] ChIPSeq peaks. X-axis order =Reverse-Complement Vacancy Position ID 4-9 > 1-12 > 2-11 > 6-7 > 5-8 > 3-10. See **Table S25** for x-axis details.

Figure S78. (S/N) analysis of 0-nt to 5-nt Variant EREs in ER (Time-Course) ChIPSeq Peaks (252 Half-Site Groups)

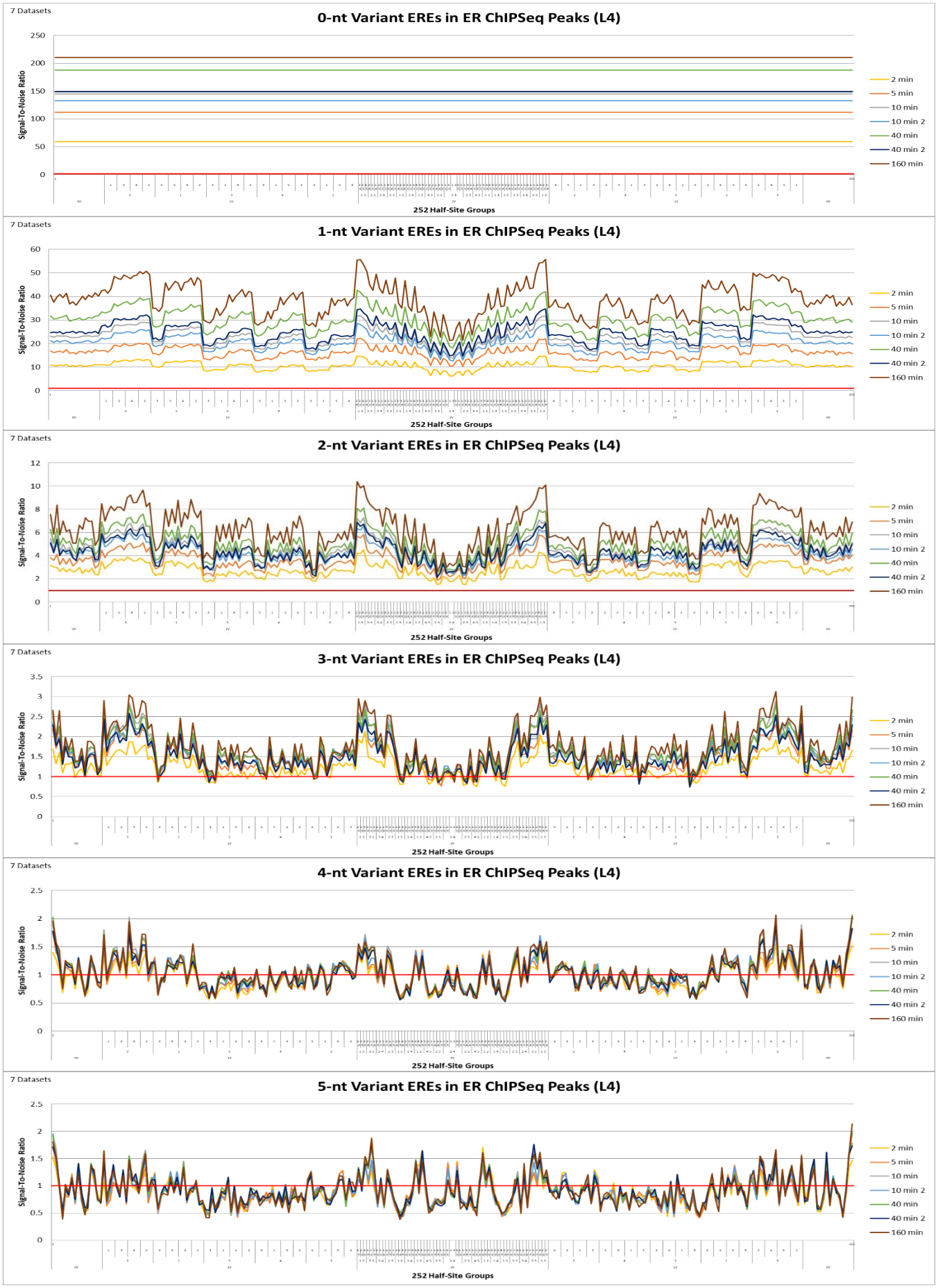


Figure S79. (S/N) analysis of 0-nt to 6-nt Variant EREs in ER (Time-Course) ChIPSeq Peaks (924 Half-Site Groups)

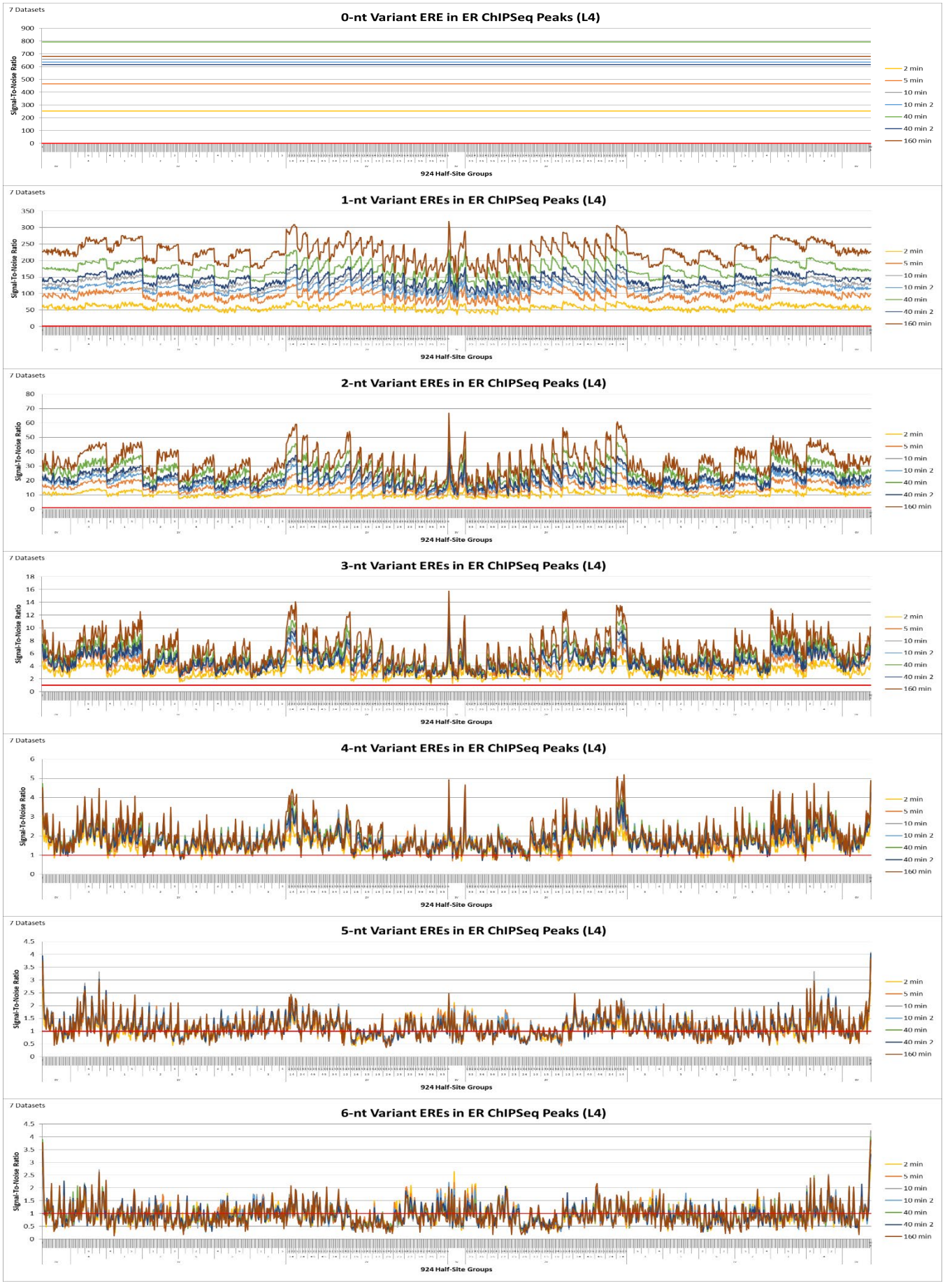


Figure S80-S81 Descriptions: p53 DNA Element Analysis

****All detailed data and statistics associated with every figure are compiled in Data S1****

To further illustrate that the structural mechanisms (by which inversion symmetry ascribes TF DNA-binding and functionality) are universally applicable, we repeated the same DNA-binding analysis as was done with the six (6) sNRs (at ERE and HRE DNA elements) on the unrelated TF p53 (at p53RE DNA elements) in the genome. The tumor suppressor p53 is a TF that controls cellular stress responses (3). Once activated, p53 binds DNA and regulates gene expression programs that contribute to apoptosis, senescence or cell cycle arrest, preventing the dissemination of damaged cells (3). These processes are involved in tumor suppression, setting the selective pressure for p53 inactivation in tumors (3).

Figure S80. Number of 0-nt to 5-nt Variant p53RE DNA Elements in the Mouse Genome
The number of 0-nt to 5-nt variants of the 10-nt p53RE consensus palindromic DNA element on the positive/sense strand in the mouse (mm10) genome. Here the 10-nt p53RE is the p53 response element (5′-TGCCCGGGCA-3′). The 0-nt to 5-nt variant p53RE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements (10 variant positions), 405 2-nt variant DNA elements (45 variant positions), 3,240 3-nt variant DNA elements (120 variant positions), 17,010 4-nt variant DNA elements (210 variant positions), and 61,236 5-nt variant DNA elements (252 variant positions), for a total of 81,922 DNA elements. The population count of each of the 81,922 0-nt to 5-nt variant 10-nt p53RE DNA elements in the mouse (mm10) genome can be found in **Table S63**.

Figure S81. (S/N) analysis of 0-nt to 5-nt Variant p53REs in p53 ChIPSeq Peaks (displayed by 252 half-site groups)
(S/N) analysis of 0-nt to 5-nt variant p53REs (displayed by the 252 half-site groups) in p53 (p53-T5-3hrs-DMSO) [4,489 peaks, 307-nt peak length] ChIPSeq peaks. X-axis order =Reverse-Complement Vacancy Position ID 5-6 > 1-10 > 3-8 > 4-7 > 2-9. See **Table S53** for x-axis details. This DNA-binding profile was observed in 22 p53 experiments, and across multiple peak selection criteria (L4-L20) **(Table S54)**.

Figure S80. Number of 0-nt to 5-nt Variant p53RE DNA Elements in the Mouse Genome

p53RE = TGCCCGGGCA

1 2 3 4 5 6 7 8 9 10

	DNA Element	Combinatorial Counts				DNA Element Frequency in Genome	
A	B	C	D	E	F	Mouse Genome (mm10)	
k	**P53RE** (n=10)	Combinations $C = \binom{n}{k}$	4-nts $D = C \times 4^k$	Unique $E = C \times 3^k$	Total Unique $F = \Sigma E$	Unique	Total Unique
0	0-nt Variant p53RE	1	1	1	1	1,078	
1	1-nt Variant p53RE	10	40	30	31	33,992	35,070
2	2-nt Variant p53RE	45	720	405	436	641,422	676,492
3	3-nt Variant p53RE	120	7,680	3,240	3,676	5,380,457	6,056,949
4	4-nt Variant p53RE	210	53,760	17,010	20,686	29,257,379	35,314,328
5	5-nt Variant p53RE	252	258,048	61,236	81,922	113,933,936	149,248,264

A	B	C	D	E	F
k		Combinations $C = \binom{n}{k}$	4-nts $D = C \times 4^k$	Unique $E = C \times 3^k$	Total Unique $F = \Sigma E$
0	0-nt Variant	10!/(0! x 10!)	1 x 1	1 x 1	1 + 0
1	1-nt Variant	10!/(1! x 9!)	10 x 4	10 x 3	30 + 1
2	2-nt Variant	10!/(2! x 8!)	45 x 16	45 x 9	405 + 31
3	3-nt Variant	10!/(3! x 7!)	120 x 64	120 x 27	3,240 + 436
4	4-nt Variant	10!/(4! x 6!)	210 x 256	210 x 81	17,010 +3,676
5	5-nt Variant	10!/(5! x 5!)	252 x 1,024	252 x 243	61,236 + 20,686

Figure S81. (S/N) analysis of 0-nt to 5-nt Variant p53REs in p53 ChIPSeq Peaks (252 Half-Site Groups)

Figure S82-S91 Descriptions: sNR DNA-Binding at 13-nt and 15-nt DNA Elements in the Genome (AR and AR-SPARKI)

****All detailed data and statistics associated with every figure are compiled in Data S1****

It's perplexing that the inversion symmetry that underlies TF DNA-binding specificity and functionality in the genome, observed in hundreds of ChIPSeq and ChIPExo experiments, has not been detected using current DNA motif identification algorithms. Our analysis reveals that the resolution capability of PWMs-based DNA motif identification is limited. An overreliance on these PWMs-based algorithms has promoted the misconception that the majority of TF DNA-binding events in the genome are driven by mechanisms other than the DNA sequence. This overreliance has also led to the misconception that sNRs can indiscriminately bind to the *cis*-regulatory DNA element of other nuclear receptors. Furthermore, different *cis*-regulatory DNA elements have been delineated for the different members of the KR family.

Analysis of KR DNA-binding at 81,922 0-nt to 5-nt variant 13-nt HRE DNA elements and at 912,718 0-nt to 6-nt variant 15-nt HRE DNA elements in the genome (in 194 KR ChIPSeq and ChIPExo experiments, representing a wide variety of mouse tissues and human cell lines, and across multiple peak selection criteria (L4-L20)) quantitively demonstrates that all members of the KR family are following the same DNA-binding rules, and thus bind the same DNA elements. This is further illustrated by the equivalent structure of the DNA-binding curves observed in AR and an AR DBD mouse model (SPARKI), where the second zinc finger in the AR DBD is replaced by that of the GR, at 81,922 0-nt to 5-nt variant 13-nt HRE DNA elements and at 912,718 0-nt to 6-nt variant 15-nt HRE DNA elements in the genome. Two biological replicates (#1 and #2) from intact AR wild-type and AR-SPARKI mice were sequenced and used in peak calling (4). To increase the depth of the analysis, replicate ChIPSeq samples were merged (#M) and the concatenated samples were used in peak calling (4).

Figure S82. (S/N) analysis of 0-nt to 5-nt Variant 13-nt HREs in AR and AR-SPARKI ChIPSeq Peaks #1 (displayed by 252 half-site groups)
(S/N) analysis of 0-nt to 5-nt variant HREs (displayed by the 252 half-site groups) in AR (AR-wt1) [49,859 peaks, 136-nt peak length] ChIPSeq peaks and AR-SPARKI (AR-SPARKI1) [41,315 peaks, 141-nt peak length] ChIPSeq peaks. X-axis order =Reverse-Complement Vacancy Position ID 2-9 > 5-6 > 1-10 > 4-7 > 3-8. See **Table S6** for x-axis details.

Figure S83. (S/N) analysis of 0-nt to 5-nt Variant 13-nt HREs in AR and AR-SPARKI ChIPSeq Peaks #2 (displayed by 252 half-site groups)
(S/N) analysis of 0-nt to 5-nt variant HREs (displayed by the 252 half-site groups) in AR (AR-wt2) [47,112 peaks, 142-nt peak length] ChIPSeq peaks and AR-SPARKI (AR-SPARKI2) [21,735 peaks, 130-nt peak length] ChIPSeq peaks. X-axis order =Reverse-Complement Vacancy Position ID 2-9 > 5-6 > 1-10 > 4-7 > 3-8. See **Table S6** for x-axis details.

Figure S84. (S/N) analysis of 0-nt to 5-nt Variant 13-nt HREs in AR and AR-SPARKI ChIPSeq Peaks #M (displayed by 252 half-site groups)
(S/N) analysis of 0-nt to 5-nt variant HREs (displayed by the 252 half-site groups) in AR (Pro-AR-wt1) [40,737 peaks, 147-nt peak length] ChIPSeq peaks and AR-SPARKI (Pro-AR-SPARKI1) [35,846 peaks, 147-nt peak length] ChIPSeq peaks. X-axis order =Reverse-Complement Vacancy Position ID 2-9 > 5-6 > 1-10 > 4-7 > 3-8. See **Table S6** for x-axis details.

*****

Figure S85. (S/N) analysis of 0-nt to 6-nt Variant 15-nt HREs in AR and AR-SPARKI ChIPSeq Peaks #1 (displayed by 252 half-site groups)
(S/N) analysis of 0-nt to 6-nt variant HREs (displayed by the 924 half-site groups) in AR (AR-wt1) [49,859 peaks, 136-nt peak length] ChIPSeq peaks and AR-SPARKI (AR-SPARKI1) [41,315 peaks, 141-nt peak length] ChIPSeq peaks. X-axis order =Reverse-Complement Vacancy Position ID 3-10 > 1-12 > 6-7 > 2-11 > 5-8 > 4-9. See **Table S26** for x-axis details.

Figure S86. (S/N) analysis of 0-nt to 6-nt Variant 15-nt HREs in AR and AR-SPARKI ChIPSeq Peaks #2 (displayed by 252 half-site groups)
(S/N) analysis of 0-nt to 6-nt variant HREs (displayed by the 924 half-site groups) in AR (AR-wt2) [47,112 peaks, 142-nt peak length] ChIPSeq peaks and AR-SPARKI (AR-SPARKI2) [21,735 peaks, 130-nt peak length] ChIPSeq peaks. X-axis order =Reverse-Complement Vacancy Position ID 3-10 > 1-12 > 6-7 > 2-11 > 5-8 > 4-9. See **Table S26** for x-axis details.

Figure S87. (S/N) analysis of 0-nt to 6-nt Variant 15-nt HREs in AR and AR-SPARKI ChIPSeq Peaks #M (displayed by 252 half-site groups)
(S/N) analysis of 0-nt to 6-nt variant HREs (displayed by the 924 half-site groups) in AR (Pro-AR-wt1) [40,737 peaks, 147-nt peak length] ChIPSeq peaks and AR-SPARKI (Pro-AR-SPARKI1) [35,846 peaks, 147-nt peak length] ChIPSeq peaks. X-axis order =Reverse-Complement Vacancy Position ID 3-10 > 1-12 > 6-7 > 2-11 > 5-8 > 4-9. See **Table S26** for x-axis details.

*****

Figure S88. (S/N) analysis of 1-nt to 5-nt Variant 13-nt HREs in AR and AR-SPARKI ChIPSeq Peaks #1 #2 (displayed by variant position)
(S/N) analysis of 1-nt to 5-nt variant HREs (displayed by variant position) in AR (AR-wt1 and -2) [49,859 peaks, 136-nt peak length] [47,112 peaks, 142-nt peak length] ChIPSeq peaks and AR-SPARKI (AR-SPARKI1 and -2) [41,315 peaks, 141-nt peak length] [21,735 peaks, 130-nt peak length] ChIPSeq peaks.

Figure S89. (S/N) analysis of 1-nt to 5-nt Variant 13-nt HREs in AR and AR-SPARKI ChIPSeq Peaks #M (displayed by variant position)
(S/N) analysis of 1-nt to 5-nt variant HREs (displayed by variant position) in AR (Pro-AR-wt1) [40,737 peaks, 147-nt peak length] ChIPSeq peaks and AR-SPARKI (Pro-AR-SPARKI1) [35,846 peaks, 147-nt peak length] ChIPSeq peaks.

*****

Figure S90. (S/N) analysis of 1-nt to 6-nt Variant 15-nt HREs in AR and AR-SPARKI ChIPSeq Peaks #1 #2 (displayed by variant position)
(S/N) analysis of 1-nt to 6-nt variant HREs (displayed by variant position) in AR (AR-wt1 and -2) [49,859 peaks, 136-nt peak length] [47,112 peaks, 142-nt peak length] ChIPSeq peaks and AR-SPARKI (AR-SPARKI1 and -2) [41,315 peaks, 141-nt peak length] [21,735 peaks, 130-nt peak length] ChIPSeq peaks.

Figure S91. (S/N) analysis of 1-nt to 6-nt Variant 15-nt HREs in AR and AR-SPARKI ChIPSeq Peaks #M (displayed by variant position)
(S/N) analysis of 1-nt to 6-nt variant HREs (displayed by variant position) in AR (Pro-AR-wt1) [40,737 peaks, 147-nt peak length] ChIPSeq peaks and AR-SPARKI (Pro-AR-SPARKI1) [35,846 peaks, 147-nt peak length] ChIPSeq peaks.

Figure S82. (S/N) analysis of 0-nt to 5-nt Variant HREs in AR and AR-SPARKI ChIPSeq Peaks #1 (252 Half-Site Groups)

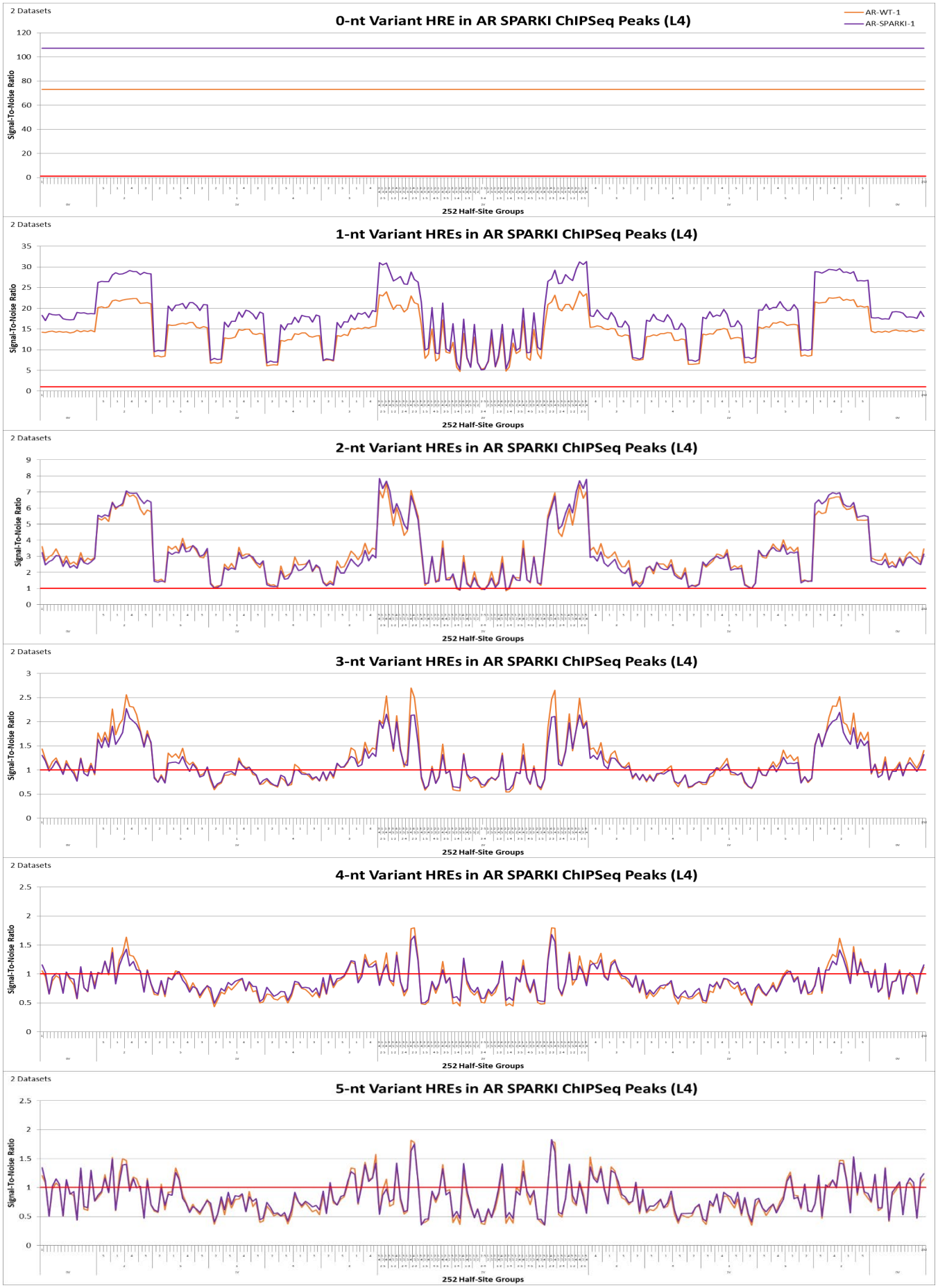


Figure S83. (S/N) analysis of 0-nt to 5-nt Variant HREs in AR and AR-SPARKI ChIPSeq Peaks #2 (252 Half-Site Groups)

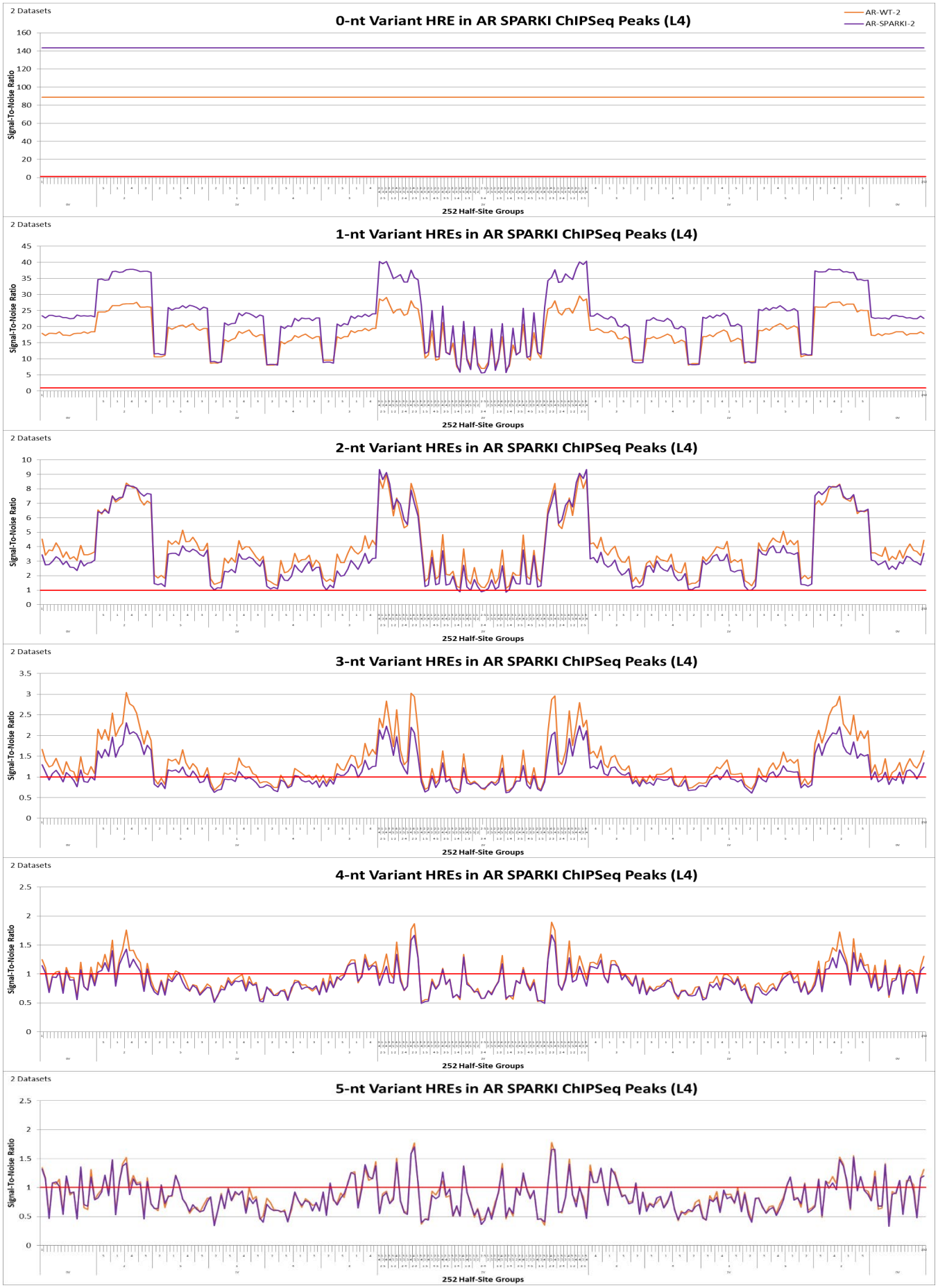


Figure S84. (S/N) analysis of 0-nt to 5-nt Variant HREs in AR and AR-SPARKI ChIPSeq Peaks #M (252 Half-Site Groups)

Figure S85. (S/N) analysis of 0-nt to 6-nt Variant HREs in AR and AR-SPARKI ChIPSeq Peaks #1 (924 Half-Site Groups)

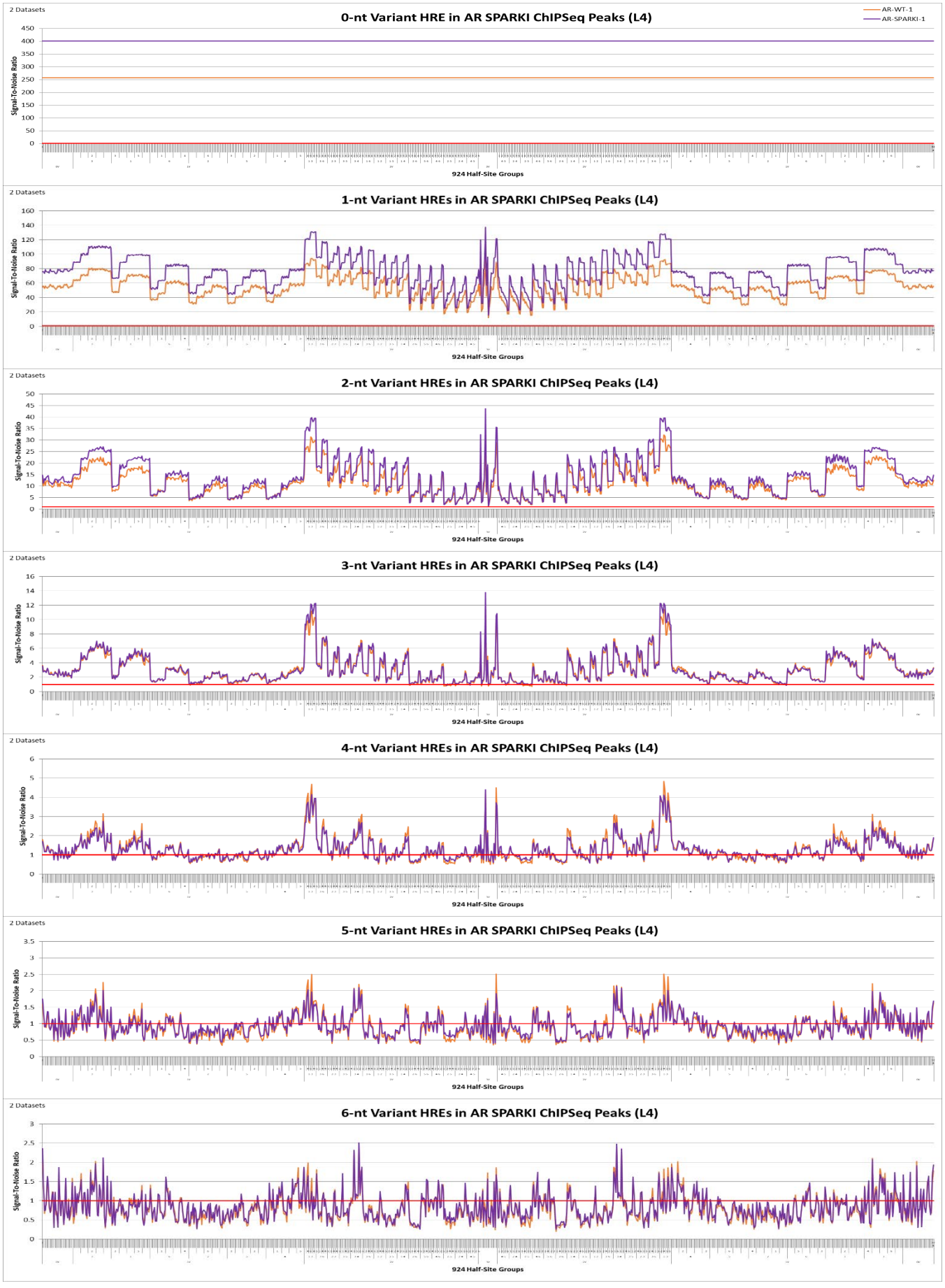


Figure S86. (S/N) analysis of 0-nt to 6-nt Variant HREs in AR and AR-SPARKI ChIPSeq Peaks #2 (924 Half-Site Groups)

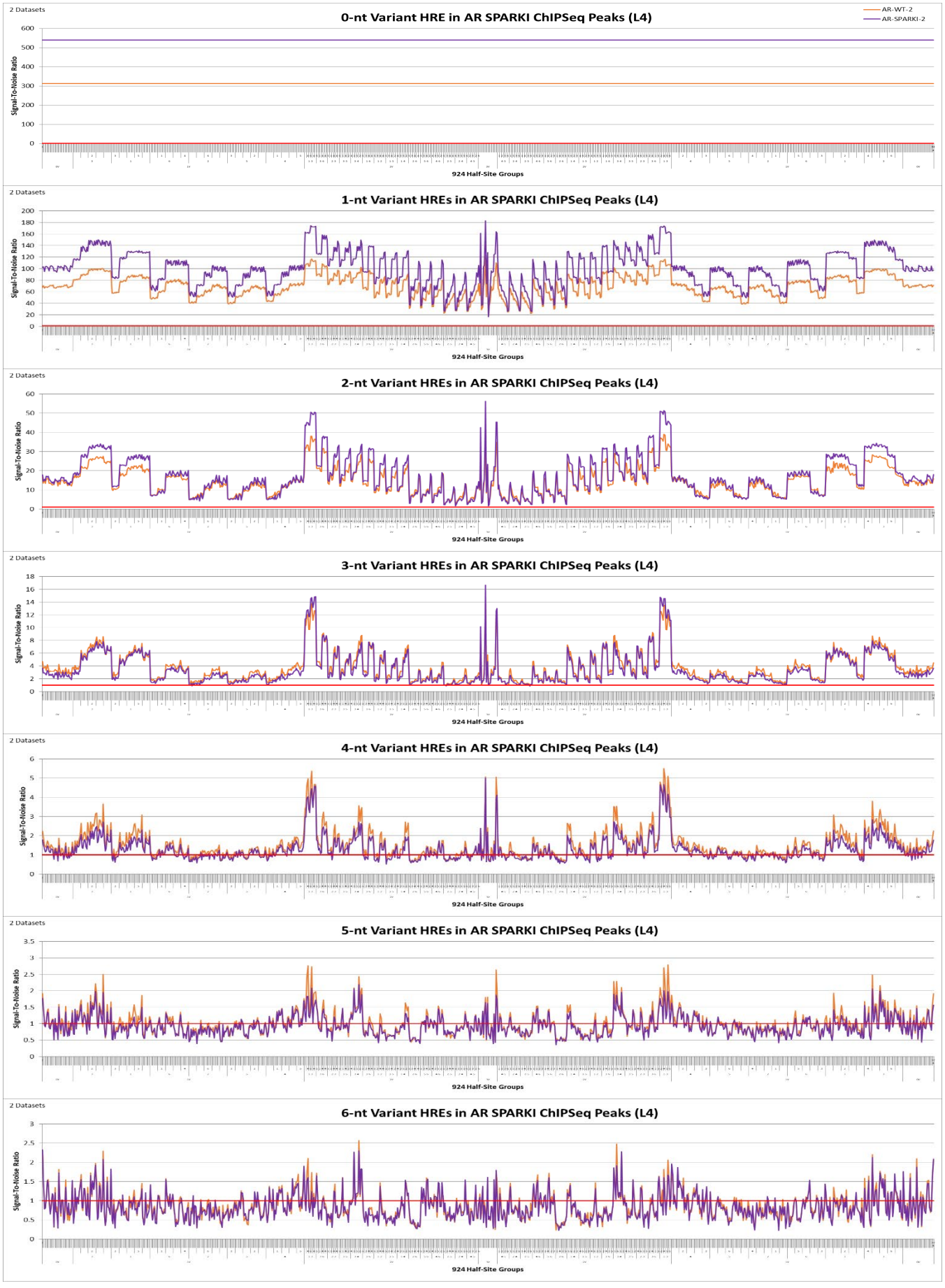


Figure S87. (S/N) analysis of 0-nt to 6-nt Variant HREs in AR and AR-SPARKI ChIPSeq Peaks #M (924 Half-Site Groups)

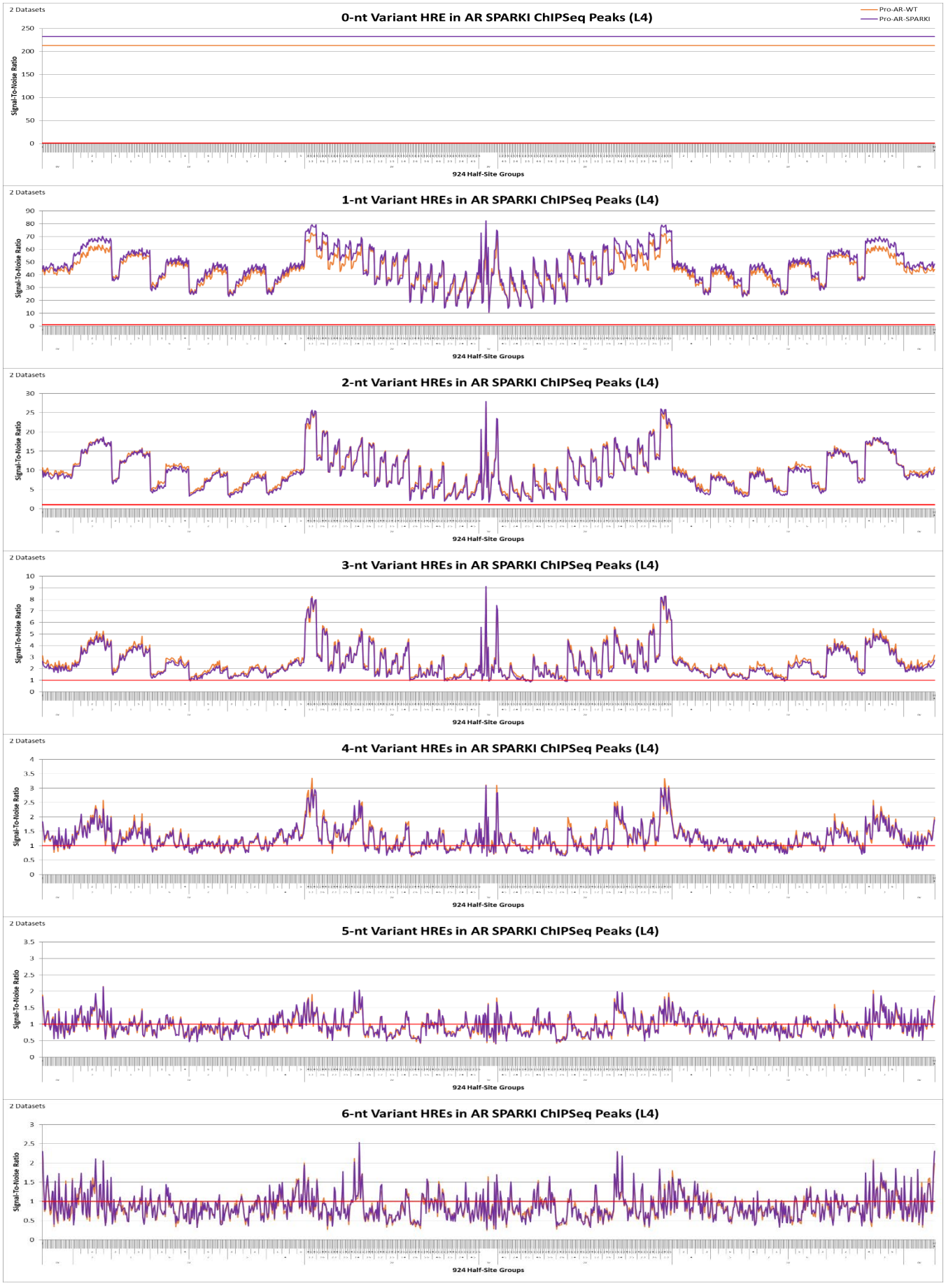
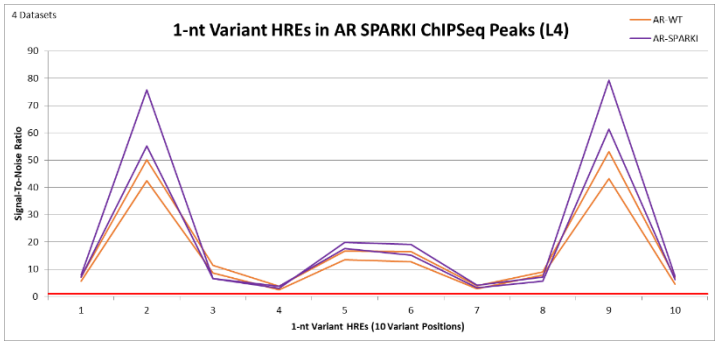
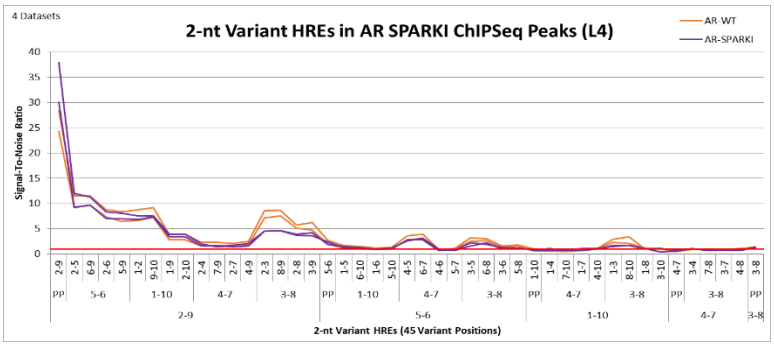
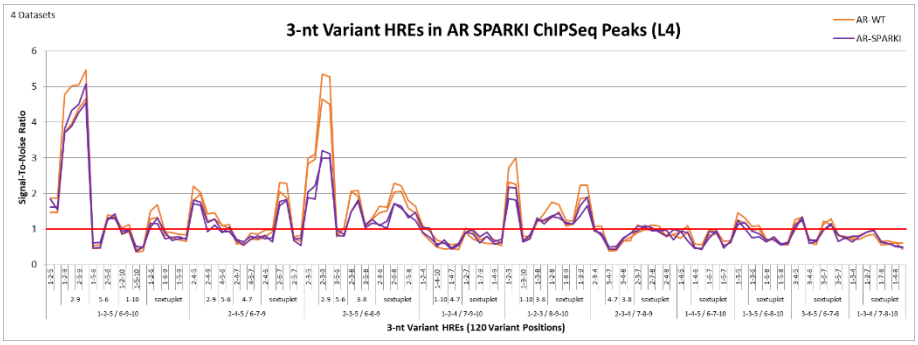
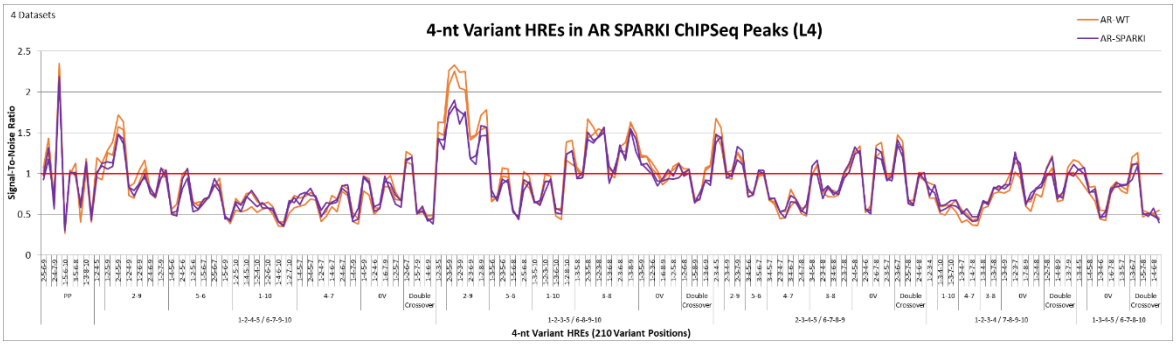


Figure S88. (S/N) analysis of 1-nt to 5-nt Variant HREs in AR and AR-SPARKI ChIPSeq Peaks #1 & #2 (Variant Position)









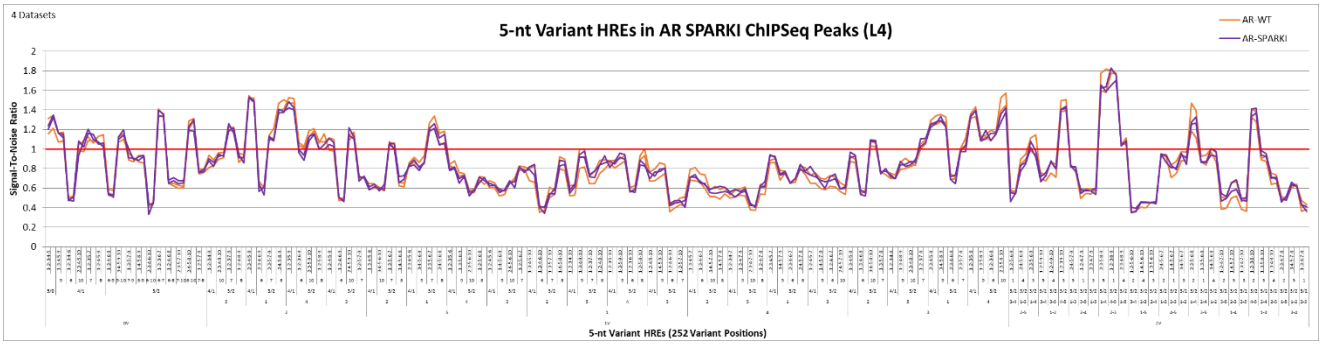
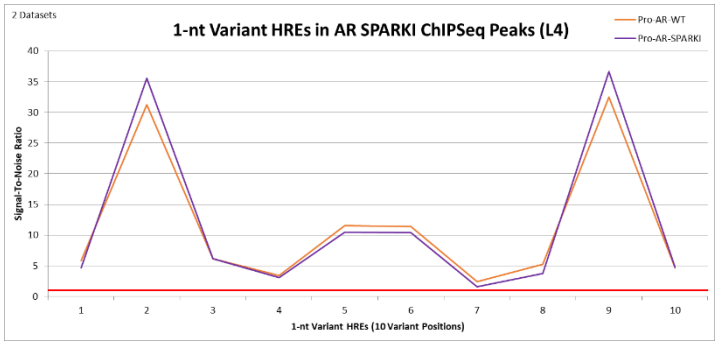
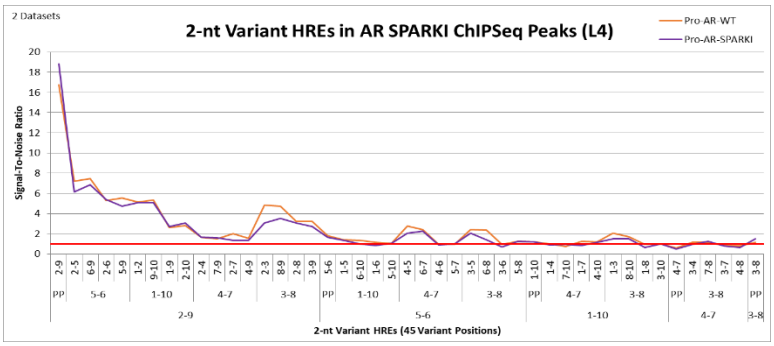
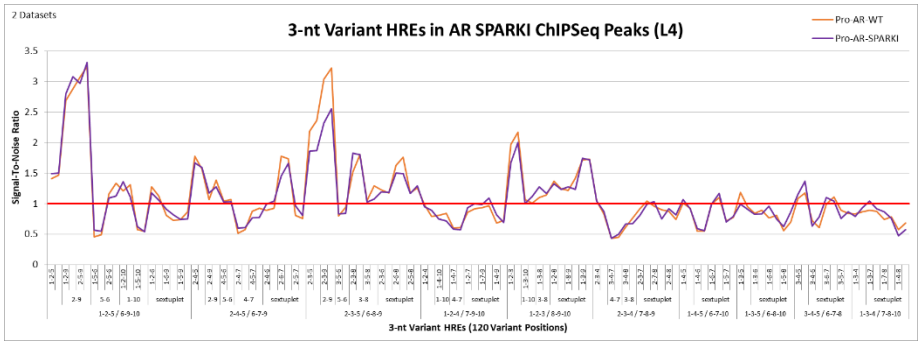
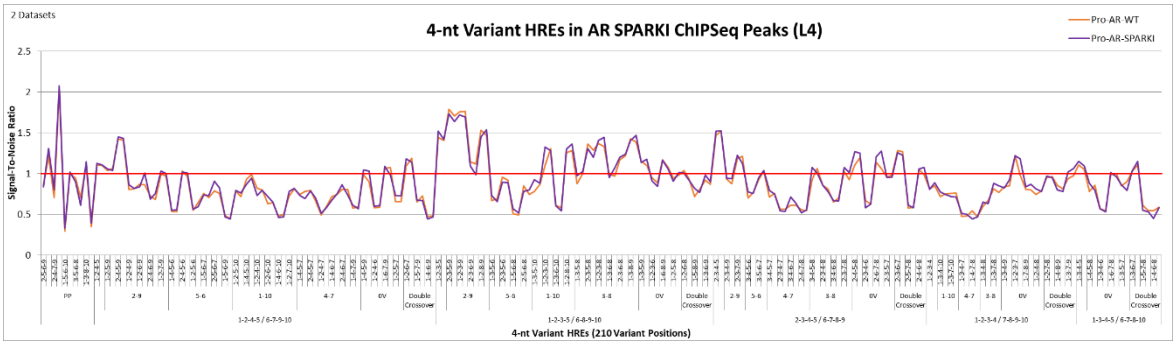


Figure S89. (S/N) analysis of 1-nt to 5-nt Variant HREs in AR and AR-SPARKI ChIPSeq Peaks #M (Variant Position)









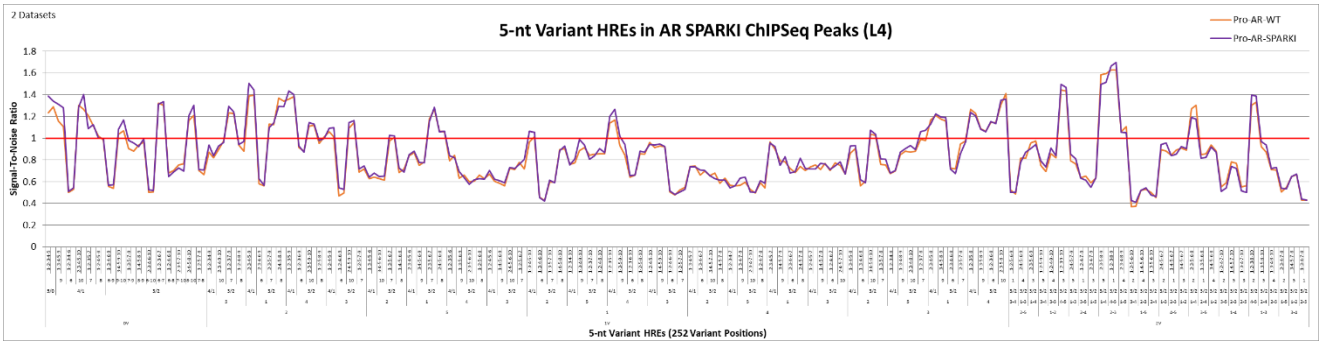
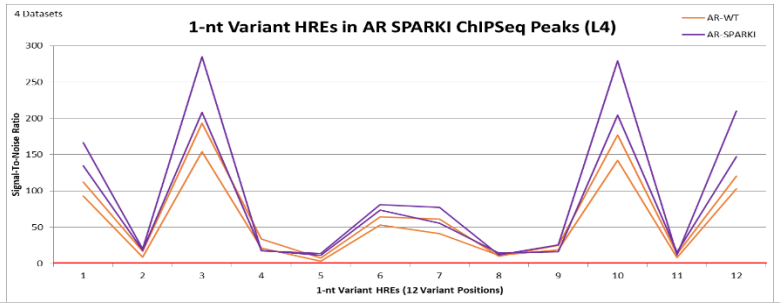
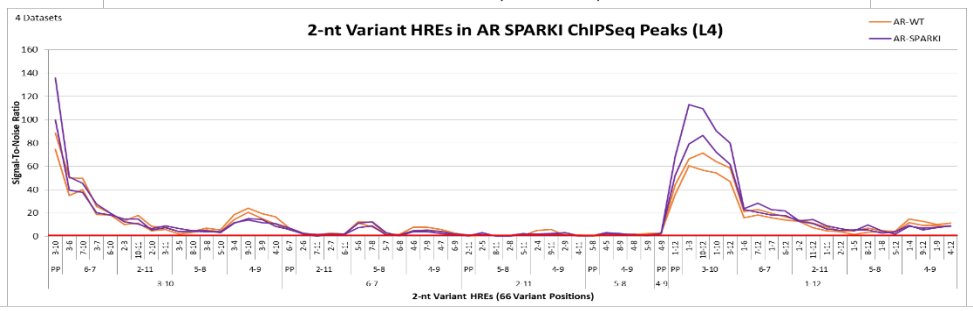
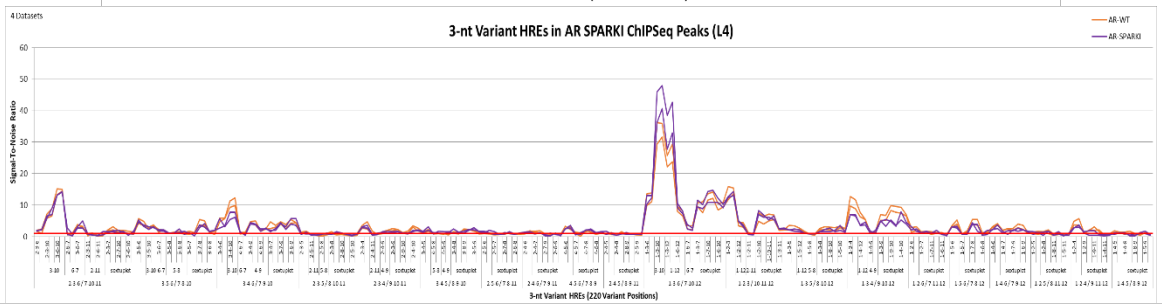
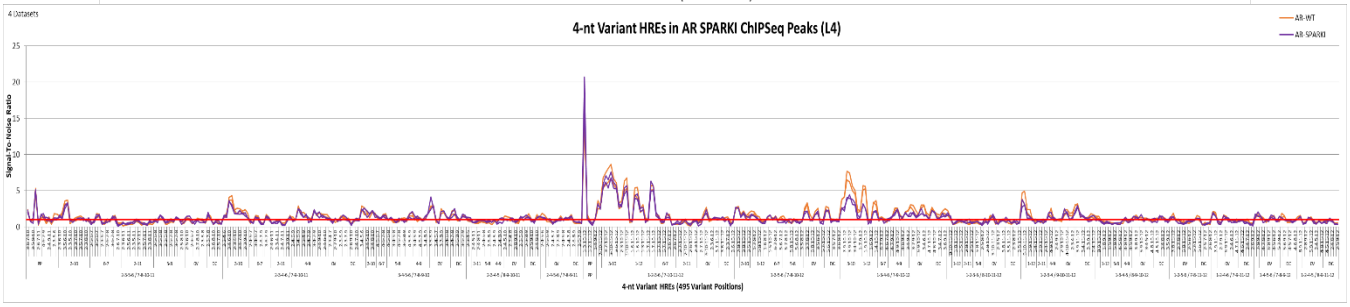


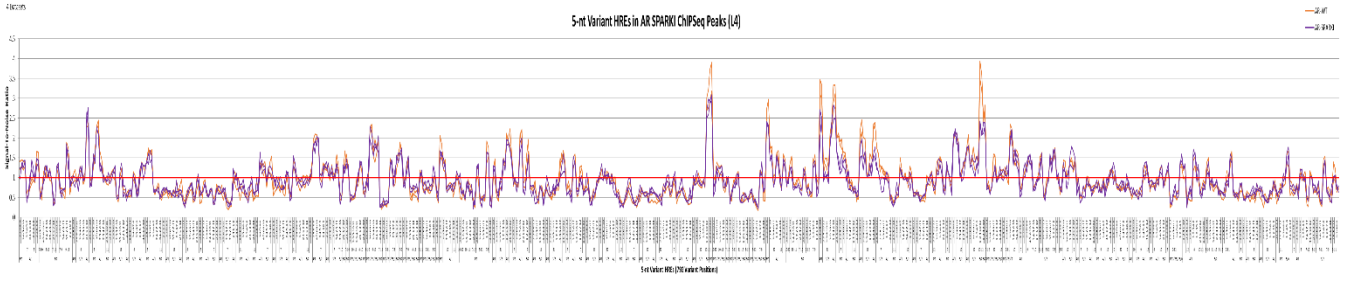
Figure S90. (S/N) analysis of 1-nt to 6-nt Variant HREs in AR and AR-SPARKI ChIPSeq Peaks #1 & #2 (Variant Position)











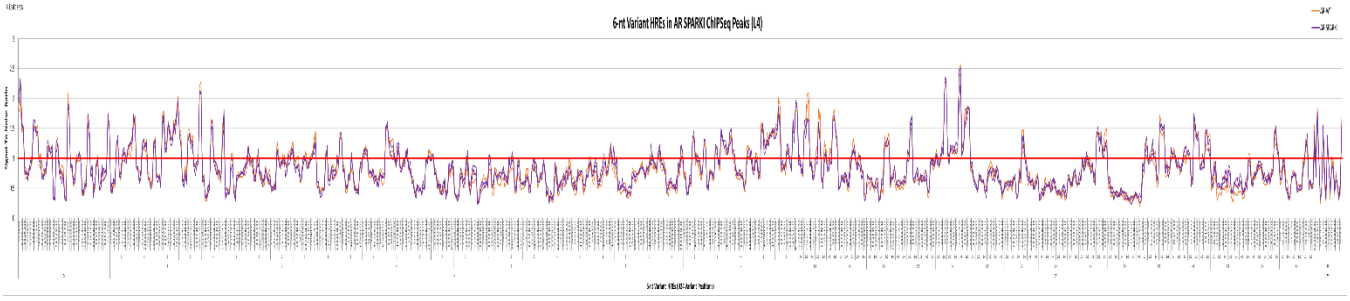
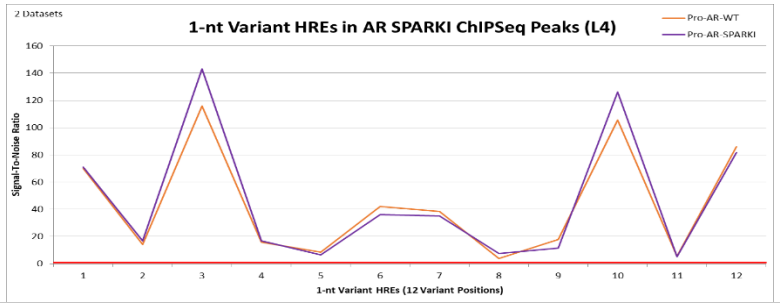
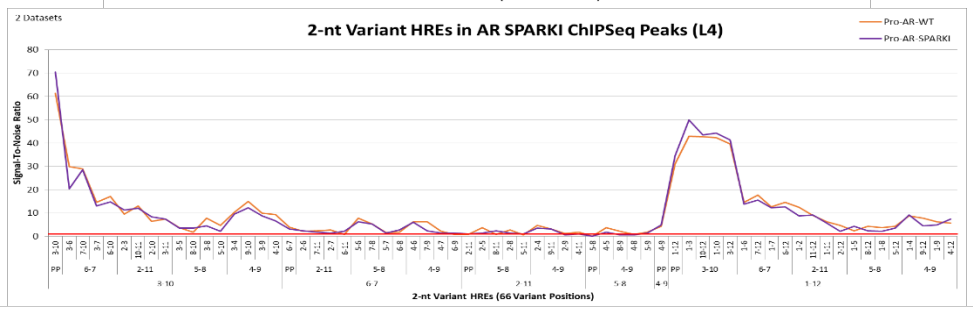
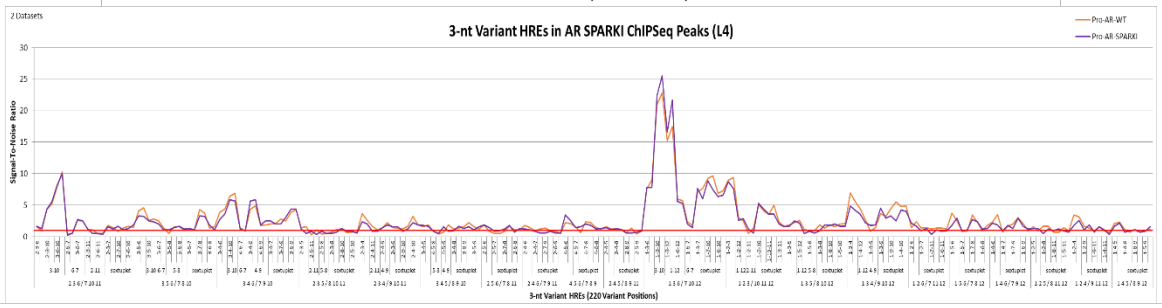
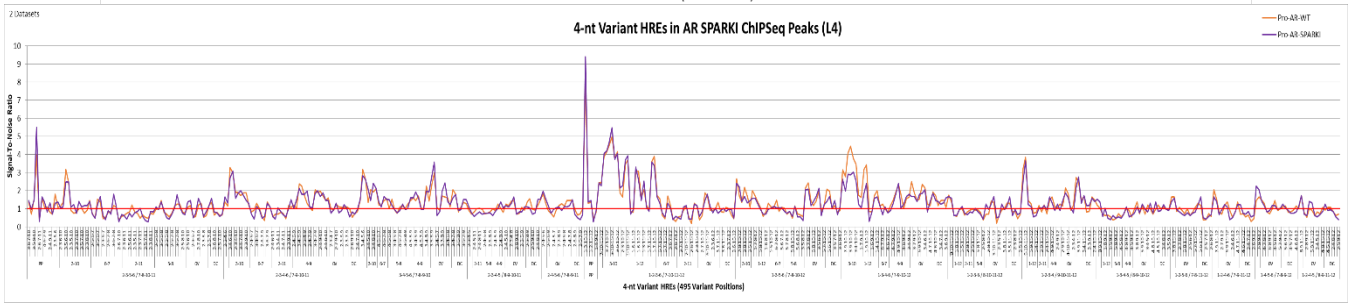


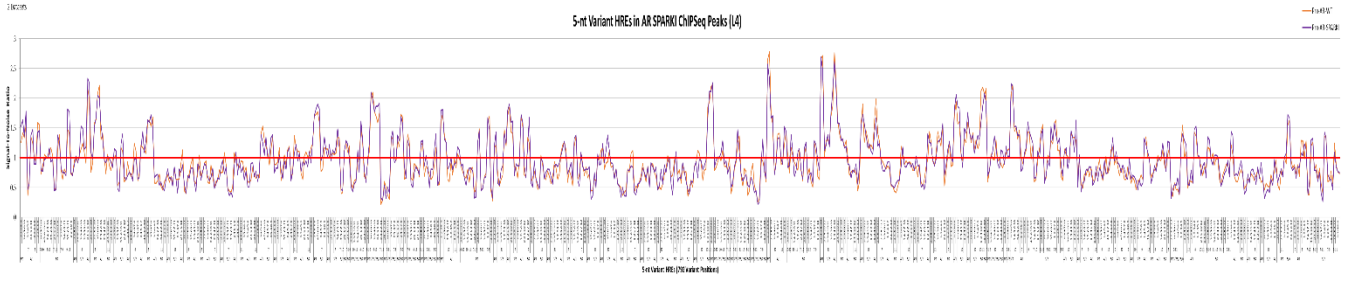
Figure S91. (S/N) analysis of 1-nt to 6-nt Variant HREs in AR and AR-SPARKI ChIPSeq Peaks #M (Variant Position)

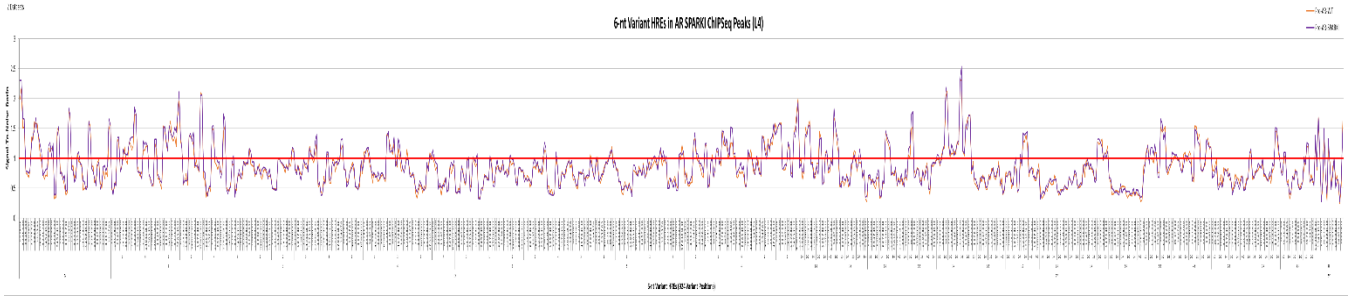












Data S1. All Detailed Data and Statistics Associated with Every Figure

- Figure 1 – Figure 5
- Table 3
- Table 5
- Figure S1 – Figure S91
- Upper and lower one-tailed Poisson significance thresholds at $p<0.001$, thus the probability of the sNR DNA-binding signal occurring outside these boundaries by chance is less than one in a thousand

Table S1-S75 Descriptions

Table S1-S20. 13-nt ERE and HRE DNA Element Analysis

The analyses in this section were completed by overlapping the location coordinates of each 0-nt to 5-nt variant 13-nt ERE or HRE DNA element in the genome and the location coordinates of the ChIPSeq or ChIPExo peaks in an experiment. The 0-nt variant consensus palindromic 13-nt ERE DNA element (5′-GGTCAnnnTGACC-3′) and 0-nt variant consensus palindromic 13-nt HRE DNA element (5′-GAACAnnnTGTTC-3′) have ten (10) primary positions. The 0-nt to 5-nt variant ERE or HRE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements (10 variant positions), 405 2-nt variant DNA elements (45 variant positions), 3,240 3-nt variant DNA elements (120 variant positions), 17,010 4-nt variant DNA elements (210 variant positions), and 61,236 5-nt variant DNA elements (252 variant positions), for a total of 81,922 DNA elements.

Table S1. Categorize 0-nt to 5-nt Variant 13-nt EREs into 252 Half-Site Groups (Variant Position)

- Categorize the 81,922 0-nt to 5-nt variant 13-nt ERE DNA elements into 252 half-site groups (5 positions are fixed, allowing for *up to* 5 positions to be varied)
- Each of the 252 ERE half-site groups contain: 1 0-nt variant consensus palindromic ERE DNA element, 15 1-nt variant ERE DNA elements, 90 2-nt variant ERE DNA elements, 270 3-nt variant ERE DNA elements, 405 4-nt variant ERE DNA elements, 243 5-nt variant ERE DNA elements, for a total of 1,024 ERE DNA elements
- 252 half-site groups x 1,024 DNA elements per half-site group =258,048 DNA elements (81,922 unique)

Table S2. Categorize 0-nt to 5-nt Variant 13-nt HREs into 252 Half-Site Groups (Variant Position)

- Categorize the 81,922 0-nt to 5-nt variant 13-nt HRE DNA elements into 252 half-site groups (5 positions are fixed, allowing for *up to* 5 positions to be varied)
- Each of the 252 HRE half-site groups contain: 1 0-nt variant consensus palindromic HRE DNA element, 15 1-nt variant HRE DNA elements, 90 2-nt variant HRE DNA elements, 270 3-nt variant HRE DNA elements, 405 4-nt variant HRE DNA elements, 243 5-nt variant HRE DNA elements, for a total of 1,024 HRE DNA elements
- 252 half-site groups x 1,024 DNA elements per half-site group =258,048 DNA elements (81,922 unique)

Table S3. Categorize 0-nt to 5-nt Variant 13-nt EREs into 252 Half-Site Groups (Sequence)

- Categorize the 81,922 0-nt to 5-nt variant 13-nt ERE DNA elements into 252 half-site groups (5 positions are fixed, allowing for *up to* 5 positions to be varied)
- Each of the 252 ERE half-site groups contain: 1 0-nt variant consensus palindromic ERE DNA element, 15 1-nt variant ERE DNA elements, 90 2-nt variant ERE DNA elements, 270 3-nt variant ERE DNA elements, 405 4-nt variant ERE DNA elements, 243 5-nt variant ERE DNA elements, for a total of 1,024 ERE DNA elements
- 252 half-site groups x 1,024 DNA elements per half-site group =258,048 DNA elements (81,922 unique)

Table S4. Categorize 0-nt to 5-nt Variant 13-nt HREs into 252 Half-Site Groups (Sequence)

- Categorize the 81,922 0-nt to 5-nt variant 13-nt HRE DNA elements into 252 half-site groups (5 positions are fixed, allowing for *up to* 5 positions to be varied)
- Each of the 252 HRE half-site groups contain: 1 0-nt variant consensus palindromic HRE DNA element, 15 1-nt variant HRE DNA elements, 90 2-nt variant HRE DNA elements, 270 3-nt variant HRE DNA elements, 405 4-nt variant HRE DNA elements, 243 5-nt variant HRE DNA elements, for a total of 1,024 HRE DNA elements
- 252 half-site groups x 1,024 DNA elements per half-site group =258,048 DNA elements (81,922 unique)

Table S5. X-Axis Order of 252 ERE Half-Site Groups

- (S/N) analysis of 0-nt to 5-nt variant EREs in ER ChIPSeq or ChIPExo peaks (displayed by 252 ERE half-site groups)
- X-axis order
    - 126 half-site groups [left-to-right: 0 vacancies, 1 vacancy, 2 vacancies] followed by their 126 reverse-complements
    - Of these 252 half-site groups, 32 have zero vacancies, 160 have one vacancy, 60 have two vacancies
    - The x-axis is labeled by the reverse-complement vacancy position ID (primary label) and reverse-complement double occupant position ID (secondary label)
    - The order of the reverse-complement vacancy position IDs for ERE is 3-8 > 1-10 > 5-6 > 4-7 > 2-9

Table S6. X-Axis Order of 252 HRE Half-Site Groups

- (S/N) analysis of 0-nt to 5-nt variant HREs in KR ChIPSeq or ChIPExo peaks (displayed by 252 HRE half-site groups)
- X-axis order
    - 126 half-site groups [left-to-right: 0 vacancies, 1 vacancy, 2 vacancies] followed by their 126 reverse-complements
    - Of these 252 half-site groups, 32 have zero vacancies, 160 have one vacancy, 60 have two vacancies

- The x-axis is labeled by the reverse-complement vacancy position ID (primary label) and reverse-complement double occupant position ID (secondary label)
- The order of the reverse-complement vacancy position IDs for HRE is 2/9 > 5/6 > 1/10 > 4/7 > 3/8

Table S7. (S/N) analysis of 0-nt to 5-nt Variant EREs in ER ChIPSeq Peaks (252 Half-Site Groups) [157]

- (S/N) analysis of 0-nt to 5-nt variant EREs in ER ChIPSeq or ChIPExo peaks (displayed by 252 ERE half-site groups)
- 157 ER experiments
- Peak selection criteria =L4-L20
- Categorize the counts of ER DNA-binding at 0-nt to 5-nt variant EREs into 252 ERE half-site groups (bottom table)
- Convert these counts to (S/N) values (top table)

Table S8. (S/N) analysis of 0-nt to 5-nt Variant HREs in KR ChIPSeq Peaks (252 Half-Site Groups) [194]

- (S/N) analysis of 0-nt to 5-nt variant HREs in KR ChIPSeq or ChIPExo peaks (displayed by 252 HRE half-site groups)
- 194 KR experiments
- Peak selection criteria =L4-L20
- Categorize the counts of KR DNA-binding at 0-nt to 5-nt variant HREs into 252 HRE half-site groups (bottom table)
- Convert these counts to (S/N) values (top table)

*****

Table S9. Quantify the 5 Discrete States of ER DNA-Binding at 1-nt Variant EREs in the Genome (3,2,0,0,0) [157]

- (S/N) analysis of 1-nt variant EREs in ER ChIPSeq or ChIPExo peaks (displayed by 252 ERE half-site groups)
- 157 ER experiments
- Peak selection criteria =L4-L20
- Categorize the counts of ER DNA-binding at 1-nt variant EREs into 252 ERE half-site groups (bottom table)
- Convert these counts to (S/N) values (top table)
- Calculate the mean value, standard deviation, and difference from the mean of the 6 groups that reach the (+2) plateau, 60 groups reach the (+1) plateau, 120 groups reach the (0) plateau, 60 groups reach the (-1) plateau, 6 groups reach the (-2) plateau, for a total of 252 half-site groups (right-hand side)

Table S10. Quantify the 3 Discrete States of KR DNA-Binding at 1-nt Variant HREs in the Genome (4,1,0,0,0) [194]

- (S/N) analysis of 1-nt variant HREs in KR ChIPSeq or ChIPExo peaks (displayed by 252 HRE half-site groups)
- 194 KR experiments
- Peak selection criteria =L4-L20
- Categorize the counts of KR DNA-binding at 1-nt variant HREs into 252 HRE half-site groups (bottom table)
- Convert these counts to (S/N) values (top table)
- Calculate the mean value, standard deviation, and difference from the mean of the 56 groups that reach the (+1) plateau, 140 groups reach the (0) plateau, 56 groups reach the (-1) plateau, for a total of 252 half-site groups (right-hand side)

*****

Table S11. (S/N) analysis of 1-nt Variant EREs in ER ChIPSeq Peaks (Variant Position) [157]

- (S/N) analysis of 1-nt variant EREs in ER ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 10 variant positions of the 30 1-nt variant EREs
- 157 ER experiments
- Peak selection criteria =L4-L20
- Categorize the counts of ER DNA-binding at 1-nt variant EREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

Table S12. (S/N) analysis of 2-nt Variant EREs in ER ChIPSeq Peaks (Variant Position) [157]

- (S/N) analysis of 2-nt variant EREs in ER ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 45 variant positions of the 405 2-nt variant EREs
  - 5 of the 45 variant positions are the five (5) palindromic position pairs: 1-10, 2-9, 3-8, 4-7, 5-6
  - 20 of the 45 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer)
  - 20 of the 45 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer)
  - Each variant position is immediately followed by its reverse-complement variant position
- 157 ER experiments
- Peak selection criteria =L4-L20
- Categorize the counts of ER DNA-binding at 2-nt variant EREs by variant position (bottom table)

- Convert these counts to (S/N) values (top table)

Table S13. (S/N) analysis of 3-nt Variant EREs in ER ChIPSeq Peaks (Variant Position) [157]

- (S/N) analysis of 3-nt variant EREs in ER ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 120 variant positions of the 3,240 3-nt variant EREs
    - 20 of the 120 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer)
    - 100 of the 120 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer)
    - Each variant position is immediately followed by its reverse-complement variant position
- 157 ER experiments
- Peak selection criteria =L4-L20
- Categorize the counts of ER DNA-binding at 3-nt variant EREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

Table S14. (S/N) analysis of 4-nt Variant EREs in ER ChIPSeq Peaks (Variant Position) [157]

- (S/N) analysis of 4-nt variant EREs in ER ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 210 variant positions of the 17,010 4-nt variant EREs
    - 10 of the 210 variant positions are the five (5) palindromic position pairs: 1-3-8-10, 3-5-6-8, 3-4-7-8, 2-3-8-9, 1-5-6-10, 1-4-7-10, 1-2-9-10, 4-5-6-7, 2-5-6-9, 2-4-7-9
    - 10 of the 210 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer)
    - 190 of the 45 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer)
    - Each variant position is immediately followed by its reverse-complement variant position
- 157 ER experiments
- Peak selection criteria =L4-L20
- Categorize the counts of ER DNA-binding at 4-nt variant EREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

Table S15. (S/N) analysis of 5-nt Variant EREs in ER ChIPSeq Peaks (Variant Position) [157]

- (S/N) analysis of 5-nt variant EREs in ER ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 252 variant positions of the 61,236 5-nt variant EREs
    - 2 of the 252 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer)
    - 250 of the 252 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer)
    - Each variant position is immediately followed by its reverse-complement variant position
- 157 ER experiments
- Peak selection criteria =L4-L20
- Categorize the counts of ER DNA-binding at 5-nt variant EREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

*****

Table S16. (S/N) analysis of 1-nt Variant HREs in KR ChIPSeq Peaks (Variant Position) [194]

- (S/N) analysis of 1-nt variant HREs in KR ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 10 variant positions of the 30 1-nt variant HREs
- 194 KR experiments
- Peak selection criteria =L4-L20
- Categorize the counts of KR DNA-binding at 1-nt variant HREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

Table S17. (S/N) analysis of 2-nt Variant HREs in KR ChIPSeq Peaks (Variant Position) [194]

- (S/N) analysis of 2-nt variant HREs in KR ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 45 variant positions of the 405 2-nt variant HREs
    - 5 of the 45 variant positions are the five (5) palindromic position pairs: 1-10, 2-9, 3-8, 4-7, 5-6
    - 20 of the 45 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer)
    - 20 of the 45 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer)
    - Each variant position is immediately followed by its reverse-complement variant position
- 194 KR experiments
- Peak selection criteria =L4-L20
- Categorize the counts of KR DNA-binding at 2-nt variant HREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

Table S18. (S/N) analysis of 3-nt Variant HREs in KR ChIPSeq Peaks (Variant Position) [194]

- (S/N) analysis of 3-nt variant HREs in KR ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 120 variant positions of the 3,240 3-nt variant HREs
    - 20 of the 120 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer)
    - 100 of the 120 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer)
    - Each variant position is immediately followed by its reverse-complement variant position
- 194 KR experiments
- Peak selection criteria =L4-L20
- Categorize the counts of KR DNA-binding at 3-nt variant HREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

Table S19. (S/N) analysis of 4-nt Variant HREs in KR ChIPSeq Peaks (Variant Position) [194]

- (S/N) analysis of 4-nt variant HREs in KR ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 210 variant positions of the 17,010 4-nt variant HREs
    - 10 of the 210 variant positions are the five (5) palindromic position pairs: 2-5-6-9, 1-2-9-10, 2-4-7-9, 2-3-8-9, 1-5-6-10, 4-5-6-7, 3-5-6-8, 1-4-7-10, 1-3-8-10, 3-4-7-8
    - 10 of the 210 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer)
    - 190 of the 45 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer)
    - Each variant position is immediately followed by its reverse-complement variant position
- 194 KR experiments
- Peak selection criteria =L4-L20
- Categorize the counts of KR DNA-binding at 4-nt variant HREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

Table S20. (S/N) analysis of 5-nt Variant HREs in KR ChIPSeq Peaks (Variant Position) [194]

- (S/N) analysis of 5-nt variant HREs in KR ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 252 variant positions of the 61,236 5-nt variant HREs
    - 2 of the 252 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer)
    - 250 of the 252 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer)
    - Each variant position is immediately followed by its reverse-complement variant position
- 194 KR experiments
- Peak selection criteria =L4-L20
- Categorize the counts of KR DNA-binding at 5-nt variant HREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

Table S21-S40. 15-nt ERE and HRE DNA Element Analysis

The analyses in this section were completed by overlapping the location coordinates of each 0-nt to 6-nt variant 15-nt ERE or HRE DNA element in the genome and the location coordinates of the ChIPSeq or ChIPExo peaks in an experiment. The 0-nt variant consensus palindromic 15-nt ERE DNA element (5′-AGGTCAnnnTGACCT-3′) and 0-nt variant consensus palindromic 15-nt HRE DNA element (5′-AGAACAnnnTGTTCT-3′) have twelve (12) primary positions. The 0-nt to 6-nt variant 15-nt ERE or HRE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 36 1-nt variant DNA elements (12 variant positions), 594 2-nt variant DNA elements (66 variant positions), 5,940 3-nt variant DNA elements (220 variant positions), 40,095 4-nt variant DNA elements (495 variant positions), 192,456 5-nt variant DNA elements (792 variant positions), and 673,596 6-nt variant DNA elements (924 variant positions), for a total of 912,718 DNA elements.

Table S21. Categorize 0-nt to 6-nt Variant 15-nt EREs into 924 Half-Site Groups (Variant Position)

- Categorize the 912,718 0-nt to 6-nt variant 15-nt ERE DNA elements into 924 half-site groups (6 positions are fixed, allowing for *up to* 6 positions to be varied)
- Each of the 924 ERE half-site groups contain: 1 0-nt variant consensus palindromic ERE DNA element, 18 1-nt variant ERE DNA elements, 135 2-nt variant ERE DNA elements, 540 3-nt variant ERE DNA elements, 1215 4-nt variant ERE DNA elements, 1458 5-nt variant ERE DNA elements, 729 6-nt variant ERE DNA elements, for a total of 4,096 ERE DNA elements
- 924 half-site groups x 4,096 DNA elements per half-site group =3,784,704 DNA elements (912,718 unique)

Table S22. Categorize 0-nt to 6-nt Variant 15-nt HREs into 924 Half-Site Groups (Variant Position)

- Categorize the 912,718 0-nt to 6-nt variant 15-nt HRE DNA elements into 924 half-site groups (6 positions are fixed, allowing for *up to* 6 positions to be varied)
- Each of the 924 HRE half-site groups contain: 1 0-nt variant consensus palindromic HRE DNA element, 18 1-nt variant HRE DNA elements, 135 2-nt variant HRE DNA elements, 540 3-nt variant HRE DNA elements, 1215 4-nt variant HRE DNA elements, 1458 5-nt variant HRE DNA elements, 729 6-nt variant HRE DNA elements, for a total of 4,096 HRE DNA elements
- 924 half-site groups x 4,096 DNA elements per half-site group =3,784,704 DNA elements (912,718 unique)

Table S23. Categorize 0-nt to 6-nt Variant 15-nt EREs into 924 Half-Site Groups (Sequence)

- Categorize the 912,718 0-nt to 6-nt variant 15-nt ERE DNA elements into 924 half-site groups (6 positions are fixed, allowing for *up to* 6 positions to be varied)
- Each of the 924 ERE half-site groups contain: 1 0-nt variant consensus palindromic ERE DNA element, 18 1-nt variant ERE DNA elements, 135 2-nt variant ERE DNA elements, 540 3-nt variant ERE DNA elements, 1215 4-nt variant ERE DNA elements, 1458 5-nt variant ERE DNA elements, 729 6-nt variant ERE DNA elements, for a total of 4,096 ERE DNA elements
- 924 half-site groups x 4,096 DNA elements per half-site group =3,784,704 DNA elements (912,718 unique)

Table S24. Categorize 0-nt to 6-nt Variant 15-nt HREs into 924 Half-Site Groups (Sequence)

- Categorize the 912,718 0-nt to 6-nt variant 15-nt HRE DNA elements into 924 half-site groups (6 positions are fixed, allowing for *up to* 6 positions to be varied)
- Each of the 924 HRE half-site groups contain: 1 0-nt variant consensus palindromic HRE DNA element, 18 1-nt variant HRE DNA elements, 135 2-nt variant HRE DNA elements, 540 3-nt variant HRE DNA elements, 1215 4-nt variant HRE DNA elements, 1458 5-nt variant HRE DNA elements, 729 6-nt variant HRE DNA elements, for a total of 4,096 HRE DNA elements
- 924 half-site groups x 4,096 DNA elements per half-site group =3,784,704 DNA elements (912,718 unique)

Table S25. X-Axis Order of 924 ERE Half-Site Groups

- (S/N) analysis of 0-nt to 6-nt variant EREs in ER ChIPSeq or ChIPExo peaks (displayed by 924 ERE half-site groups)
- X-axis order
    - 452 half-site groups [left-to-right: 0 vacancies, 1 vacancy, 2 vacancies], 20 half-site groups [3 vacancies], followed by their 452 reverse-complements
    - Of these 924 half-site groups, 64 have zero vacancies, 480 have one vacancy, 360 have two vacancies, 20 have three vacancies
    - The x-axis is labeled by the reverse-complement vacancy position ID (primary label) and reverse-complement double occupant position ID (secondary label)
    - The order of the reverse-complement vacancy position IDs for ERE is 4-9 > 1-12 > 2-11 > 6-7 > 5-8 > 3-10

Table S26. X-Axis Order of 924 HRE Half-Site Groups

- (S/N) analysis of 0-nt to 6-nt variant HREs in KR ChIPSeq or ChIPExo peaks (displayed by 924 HRE half-site groups)

- X-axis order
    - 452 half-site groups [left-to-right: 0 vacancies, 1 vacancy, 2 vacancies], 20 half-site groups [3 vacancies], followed by their 452 reverse-complements
    - Of these 924 half-site groups, 64 have zero vacancies, 480 have one vacancy, 360 have two vacancies, 20 have three vacancies
    - The x-axis is labeled by the reverse-complement vacancy position ID (primary label) and reverse-complement double occupant position ID (secondary label)
    - The order of the reverse-complement vacancy position IDs for ERE is 3-10 > 1-12 > 6-7 > 2-11 > 5-8 > 4-9

Table S27. (S/N) analysis of 0-nt to 6-nt Variant EREs in ER ChIPSeq Peaks (924 Half-Site Groups) [157]

- (S/N) analysis of 0-nt to 6-nt variant EREs in ER ChIPSeq or ChIPExo peaks (displayed by 924 ERE half-site groups)
- 157 ER experiments
- Peak selection criteria =L4-L20
- Categorize the counts of ER DNA-binding at 0-nt to 6-nt variant EREs into 924 ERE half-site groups (bottom table)
- Convert these counts to (S/N) values (top table)

Table S28. (S/N) analysis of 0-nt to 6-nt Variant HREs in KR ChIPSeq Peaks (924 Half-Site Groups) [194]

- (S/N) analysis of 0-nt to 6-nt variant HREs in KR ChIPSeq or ChIPExo peaks (displayed by 924 HRE half-site groups)
- 194 KR experiments
- Peak selection criteria =L4-L20
- Categorize the counts of KR DNA-binding at 0-nt to 6-nt variant HREs into 924 HRE half-site groups (bottom table)
- Convert these counts to (S/N) values (top table)

*****

Table S29. (S/N) analysis of 1-nt Variant EREs in ER ChIPSeq Peaks (Variant Position) [157]

- (S/N) analysis of 1-nt variant EREs in ER ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 12 variant positions of the 36 1-nt variant EREs
- 157 ER experiments
- Peak selection criteria =L4-L20
- Categorize the counts of ER DNA-binding at 1-nt variant EREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

Table S30. (S/N) analysis of 2-nt Variant EREs in ER ChIPSeq Peaks (Variant Position) [157]

- (S/N) analysis of 2-nt variant EREs in ER ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 66 variant positions of the 594 2-nt variant EREs
    - 6 of the 66 variant positions are the six (6) palindromic position pairs: 1-12, 2-11, 3-10, 4-9, 5-8, 6-7
    - 30 of the 66 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer)
    - 30 of the 66 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer)
    - Each variant position is immediately followed by its reverse-complement variant position
- 157 ER experiments
- Peak selection criteria =L4-L20
- Categorize the counts of ER DNA-binding at 2-nt variant EREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

Table S31. (S/N) analysis of 3-nt Variant EREs in ER ChIPSeq Peaks (Variant Position) [157]

- (S/N) analysis of 3-nt variant EREs in ER ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 220 variant positions of the 5,940 3-nt variant EREs
    - 40 of the 220 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer)
    - 180 of the 220 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer)
    - Each variant position is immediately followed by its reverse-complement variant position
- 157 ER experiments
- Peak selection criteria =L4-L20
- Categorize the counts of ER DNA-binding at 3-nt variant EREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

Table S32. (S/N) analysis of 4-nt Variant EREs in ER ChIPSeq Peaks (Variant Position) [157]

- (S/N) analysis of 4-nt variant EREs in ER ChIPSeq or ChIPExo peaks (displayed by variant position)

- Displayed by the 495 variant positions of the 40,095 4-nt variant EREs
    - 15 of the 495 variant positions are the five (5) palindromic position pairs: 2-4-9-11, 4-6-7-9, 4-5-8-9, 3-4-9-10, 2-6-7-11, 2-5-8-11, 2-3-10-11, 5-6-7-8, 3-6-7-10, 3-5-8-10, 1-4-9-12, 1-2-11-12, 1-6-7-12, 1-5-8-12, 1-3-10-12
    - 30 of the 495 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer)
    - 450 of the 495 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer)
    - Each variant position is immediately followed by its reverse-complement variant position
- 157 ER experiments
- Peak selection criteria =L4-L20
- Categorize the counts of ER DNA-binding at 4-nt variant EREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

Table S33. (S/N) analysis of 5-nt Variant EREs in ER ChIPSeq Peaks (Variant Position) [157]

- (S/N) analysis of 5-nt variant EREs in ER ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 792 variant positions of the 192,456 5-nt variant EREs
    - 12 of the 792 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer)
    - 780 of the 792 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer)
    - Each variant position is immediately followed by its reverse-complement variant position
- 157 ER experiments
- Peak selection criteria =L4-L20
- Categorize the counts of ER DNA-binding at 5-nt variant EREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

Table S34. (S/N) analysis of 6-nt Variant EREs in ER ChIPSeq Peaks (Variant Position) [157]

- (S/N) analysis of 6-nt variant EREs in ER ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 924 variant positions of the 673,596 6-nt variant EREs
    - 20 of the 924 variant positions are the five (5) palindromic position pairs: 1-4-6-7-9-12, 2-3-4-9-10-11, 1-3-4-9-10-12, 1-2-4-9-11-12, 2-4-6-7-9-11, 1-4-5-8-9-12, 2-4-5-8-9-11, 3-4-6-7-9-10, 1-2-6-7-11-12, 4-5-6-7-8-9, 1-2-5-8-11-12, 1-3-6-7-10-12, 1-2-3-10-11-12, 3-4-5-8-9-10, 1-5-6-7-8-12, 2-3-5-8-10-11, 2-3-6-7-10-11, 1-3-5-8-10-12, 3-5-6-7-8-10, 2-5-6-7-8-11
    - 2 of the 924 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer)
    - 922 of the 924 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer)
    - Each variant position is immediately followed by its reverse-complement variant position
- 157 ER experiments
- Peak selection criteria =L4-L20
- Categorize the counts of ER DNA-binding at 6-nt variant EREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

*****

Table S35. (S/N) analysis of 1-nt Variant HREs in KR ChIPSeq Peaks (Variant Position) [194]

- (S/N) analysis of 1-nt variant HREs in KR ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 12 variant positions of the 36 1-nt variant HREs
- 194 KR experiments
- Peak selection criteria =L4-L20
- Categorize the counts of KR DNA-binding at 1-nt variant HREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

Table S36. (S/N) analysis of 2-nt Variant HREs in KR ChIPSeq Peaks (Variant Position) [194]

- (S/N) analysis of 2-nt variant HREs in KR ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 66 variant positions of the 594 2-nt variant HREs
    - 6 of the 66 variant positions are the six (6) palindromic position pairs: 1-12, 2-11, 3-10, 4-9, 5-8, 6-7
    - 30 of the 66 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer)
    - 30 of the 66 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer)
    - Each variant position is immediately followed by its reverse-complement variant position
- 194 KR experiments
- Peak selection criteria =L4-L20
- Categorize the counts of KR DNA-binding at 2-nt variant HREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

Table S37. (S/N) analysis of 3-nt Variant HREs in KR ChIPSeq Peaks (Variant Position) [194]

- (S/N) analysis of 3-nt variant HREs in KR ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 220 variant positions of the 5,940 3-nt variant HREs
    - 40 of the 220 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer)
    - 180 of the 220 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer)
    - Each variant position is immediately followed by its reverse-complement variant position
- 194 KR experiments
- Peak selection criteria =L4-L20
- Categorize the counts of KR DNA-binding at 3-nt variant HREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

Table S38. (S/N) analysis of 4-nt Variant HREs in KR ChIPSeq Peaks (Variant Position) [194]

- (S/N) analysis of 4-nt variant HREs in KR ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 495 variant positions of the 40,095 4-nt variant HREs
    - 15 of the 495 variant positions are the five (5) palindromic position pairs: 3-6-7-10, 2-3-10-11, 3-5-8-10, 3-4-9-10, 2-6-7-11, 5-6-7-8, 4-6-7-9, 2-5-8-11, 2-4-9-11, 4-5-8-9, 1-3-10-12, 1-6-7-12, 1-2-11-12, 1-5-8-12, 1-4-9-12
    - 30 of the 495 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer)
    - 450 of the 495 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer)
    - Each variant position is immediately followed by its reverse-complement variant position
- 194 KR experiments
- Peak selection criteria =L4-L20
- Categorize the counts of KR DNA-binding at 4-nt variant HREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

Table S39. (S/N) analysis of 5-nt Variant HREs in KR ChIPSeq Peaks (Variant Position) [194]

- (S/N) analysis of 5-nt variant HREs in KR ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 792 variant positions of the 192,456 5-nt variant HREs
    - 12 of the 792 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer)
    - 780 of the 792 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer)
    - Each variant position is immediately followed by its reverse-complement variant position
- 194 KR experiments
- Peak selection criteria =L4-L20
- Categorize the counts of KR DNA-binding at 5-nt variant HREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

Table S40. (S/N) analysis of 6-nt Variant HREs in KR ChIPSeq Peaks (Variant Position) [194]

- (S/N) analysis of 6-nt variant HREs in KR ChIPSeq or ChIPExo peaks (displayed by variant position)
- Displayed by the 924 variant positions of the 673,596 6-nt variant HREs
    - 20 of the 924 variant positions are the five (5) palindromic position pairs: 1-2-3-10-11-12, 3-4-5-8-9-10, 1-3-4-9-10-12, 1-2-4-9-11-12, 2-3-6-7-10-11, 1-3-5-8-10-12, 3-5-6-7-8-10, 2-5-6-7-8-11, 1-2-6-7-11-12, 4-5-6-7-8-9, 1-5-6-7-8-12, 2-3-5-8-10-11, 1-4-6-7-9-12, 2-3-4-9-10-11, 1-2-5-8-11-12, 1-3-6-7-10-12, 2-4-6-7-9-11, 1-4-5-8-9-12, 2-4-5-8-9-11, 3-4-6-7-9-10
    - 2 of the 924 variant positions are same-side variants (i.e., the variants do not crossover the 3-nt spacer)
    - 922 of the 924 variant positions are crossover variants (i.e., the variants do crossover the 3-nt spacer)
    - Each variant position is immediately followed by its reverse-complement variant position
- 194 KR experiments
- Peak selection criteria =L4-L20
- Categorize the counts of KR DNA-binding at 6-nt variant HREs by variant position (bottom table)
- Convert these counts to (S/N) values (top table)

Table S41-S46. Transform the 13-nt DNA Element Analysis to the 15-nt DNA Element Analysis

Analysis of the 13-nt ERE and HRE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements (10 variant positions), 405 2-nt variant DNA elements (45 variant positions), 3,240 3-nt variant DNA elements (120 variant positions), 17,010 4-nt variant DNA elements (210 variant positions), and 61,236 5-nt variant DNA elements (252 variant positions), for a total of 81,922 DNA elements. Analysis of the 15-nt ERE and HRE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 36 1-nt variant DNA elements (12 variant positions), 594 2-nt variant DNA elements (66 variant positions), 5,940 3-nt variant DNA elements (220 variant positions), 40,095 4-nt variant DNA elements (495 variant positions), 192,456 5-nt variant DNA elements (792 variant positions), and 673,596 6-nt variant DNA elements (924 variant positions), for a total of 912,718 DNA elements.

Table S41. Transform the 13-nt EREs to the 15-nt EREs (Variant Position)

- Each of the 81,922 0-nt to 5-nt variant 13-nt EREs splits into 16 categories of the 15-nt EREs
- The 15-nt EREs are displayed by their variant position
- 16 categories:
    - 0-nt variant difference: an adenine (A) is in position 1 and a thymine (T) is in position 12 of the 15-nt DNA element [=1 DNA element]
    - 1-nt variant difference: one of the three (3) alternative nucleotide possibilities are in position 1 or position 12 [=6 DNA elements]
    - 2-nt variant difference: one of the three (3) alternative nucleotide possibilities are in position 1 and position 12 [=9 DNA elements]
- 81,922 0-nt to 5-nt variant 13-nt EREs x 16 categories =1,310,752 15-nt EREs
- The 81,922 0-nt to 5-nt variant 13-nt EREs include 551,124 EREs that are not part of the 912,718 0-nt to 6-nt variant 15-nt EREs (these represent a 7-nt variant of the 15-nt ERE)
- Thus, of the 81,922 0-nt to 5-nt variant 13-nt EREs and the 912,718 0-nt to 6-nt variant 15-nt EREs, 759,628 EREs correspond to each other

Table S42. Transform the 13-nt HREs to the 15-nt HREs (Variant Position)

- Each of the 81,922 0-nt to 5-nt variant 13-nt HREs splits into 16 categories of the 15-nt HREs
- The 15-nt HREs are displayed by their variant position
- 16 categories:
    - 0-nt variant difference: an adenine (A) is in position 1 and a thymine (T) is in position 12 of the 15-nt DNA element [=1 DNA element]
    - 1-nt variant difference: one of the three (3) alternative nucleotide possibilities are in position 1 or position 12 [=6 DNA elements]
    - 2-nt variant difference: one of the three (3) alternative nucleotide possibilities are in position 1 and position 12 [=9 DNA elements]
- 81,922 0-nt to 5-nt variant 13-nt HREs x 16 categories =1,310,752 15-nt HREs
- The 81,922 0-nt to 5-nt variant 13-nt HREs include 551,124 HREs that are not part of the 912,718 0-nt to 6-nt variant 15-nt HREs (these represent a 7-nt variant of the 15-nt HRE)
- Thus, of the 81,922 0-nt to 5-nt variant 13-nt HREs and the 912,718 0-nt to 6-nt variant 15-nt HREs, 759,628 HREs correspond to each other

Table S43. Transform the 13-nt EREs to the 15-nt EREs (Sequence)

- Each of the 81,922 0-nt to 5-nt variant 13-nt EREs splits into 16 categories of the 15-nt EREs
- The 15-nt EREs are displayed by their DNA sequence
- 16 categories:
    - 0-nt variant difference: an adenine (A) is in position 1 and a thymine (T) is in position 12 of the 15-nt DNA element [=1 DNA element]
    - 1-nt variant difference: one of the three (3) alternative nucleotide possibilities are in position 1 or position 12 [=6 DNA elements]
    - 2-nt variant difference: one of the three (3) alternative nucleotide possibilities are in position 1 and position 12 [=9 DNA elements]
- 81,922 0-nt to 5-nt variant 13-nt EREs x 16 categories =1,310,752 15-nt EREs

- The 81,922 0-nt to 5-nt variant 13-nt EREs include 551,124 EREs that are not part of the 912,718 0-nt to 6-nt variant 15-nt EREs (these represent a 7-nt variant of the 15-nt ERE)
- Thus, of the 81,922 0-nt to 5-nt variant 13-nt EREs and the 912,718 0-nt to 6-nt variant 15-nt EREs, 759,628 EREs correspond to each other

Table S44. Transform the 13-nt HREs to the 15-nt HREs (Sequence)

- Each of the 81,922 0-nt to 5-nt variant 13-nt HREs splits into 16 categories of the 15-nt HREs
- The 15-nt HREs are displayed by their DNA sequence
- 16 categories:
    - 0-nt variant difference: an adenine (A) is in position 1 and a thymine (T) is in position 12 of the 15-nt DNA element [=1 DNA element]
    - 1-nt variant difference: one of the three (3) alternative nucleotide possibilities are in position 1 or position 12 [=6 DNA elements]
    - 2-nt variant difference: one of the three (3) alternative nucleotide possibilities are in position 1 and position 12 [=9 DNA elements]
- 81,922 0-nt to 5-nt variant 13-nt HREs x 16 categories =1,310,752 15-nt HREs
- The 81,922 0-nt to 5-nt variant 13-nt HREs include 551,124 HREs that are not part of the 912,718 0-nt to 6-nt variant 15-nt HREs (these represent a 7-nt variant of the 15-nt HRE)
- Thus, of the 81,922 0-nt to 5-nt variant 13-nt HREs and the 912,718 0-nt to 6-nt variant 15-nt HREs, 759,628 HREs correspond to each other

Table S45. Transform the 15-nt EREs to the 13-nt EREs

- The 912,718 0-nt to 6-nt variant 15-nt EREs and their corresponding 13-nt EREs
- The 912,718 0-nt to 6-nt variant 15-nt EREs include 153,090 EREs that are not part of the 81,922 0-nt to 5-nt variant 13-nt EREs (these represent a 6-nt variant of the 13-nt ERE)
- Thus, of the 81,922 0-nt to 5-nt variant 13-nt EREs and the 912,718 0-nt to 6-nt variant 15-nt EREs, 759,628 EREs correspond to each other
- Example Experiment (WT-E2-1hr.L4) [76,163 peaks, 146-nt peak length]
    - The 0-nt variant consensus palindromic 13-nt ERE (5′-GGTCAnnnTGACC-3′) occurred 1,202 times in an ER experiment
    - Each of those 1,202 EREs are part of a 15-nt ERE
    - Analysis of ER DNA-binding at all possible 15-nt EREs splits those 1,202 13-nt EREs into 16 categories (i.e., 4-nt possibilities in position 1 and 4-nt possibilities in position 12): 109 (1T), 155 (1G), 19 (1C), 13 (12G), 151 (12C), 110 (12A), 11 (1-12 TG), 67 (1-12 TC), 17 (1-12 GG), 305 (1-12 AT), 29 (1-12 TA), 107 (1-12 GC), 5 (1-12 CG), 14 (1-12 CC), 88 (1-12 GA), 2 (1-12 CA), totaling 1,202 **(Figure S33)**

Table S46. Transform the 15-nt HREs to the 13-nt HREs

- The 912,718 0-nt to 6-nt variant 15-nt HREs and their corresponding 13-nt HREs
- The 912,718 0-nt to 6-nt variant 15-nt HREs include 153,090 HREs that are not part of the 81,922 0-nt to 5-nt variant 13-nt HREs (these represent a 6-nt variant of the 13-nt HRE)
- Thus, of the 81,922 0-nt to 5-nt variant 13-nt HREs and the 912,718 0-nt to 6-nt variant 15-nt HREs, 759,628 HREs correspond to each other

Table S47-S50. Inversion Symmetry of the Single-Stranded Mouse and Human Genome

Evaluation of the number of times every DNA element (1-nt to 20-nt) occurs in the single-stranded mouse (mm10) and human (hg19) genome. This includes all DNA elements from the four (4) 1-nt DNA elements to the 1 trillion (1,099,511,627,776) 20-nt DNA elements. Comparing the population count of every DNA element to the population count of its reverse-complement DNA element revealed a near perfect inversion symmetry for all 1-nt to 20-nt DNA elements in the single-stranded mouse and human genome. This property is also maintained at the level of each individual chromosome, except for chromosome M (i.e., the mitochondrial DNA). In addition to demonstrating the inherent inversion symmetry structure for all DNA elements in the single-stranded genome, we further demonstrate the absence of analog symmetries between reverse pairs and complement pairs of DNA elements in the single-stranded genome.

Table S47. Population Count of 1-nt to 20-nt DNA Elements in the Single-Stranded Mouse Genome [1.4 T]

- Correlation coefficient, slope, y-intercept, and general error fraction (GEF) (ratio of average error to average length) between the population count of every DNA element (1-nt to 20-nt) and its reverse DNA element, its complement DNA element, and its reverse-complement DNA element (=6 comparisons) in the single-stranded mouse (mm10) genome

Table S48. Population Count of 1-nt to 20-nt DNA Elements in the Single-Stranded Human Genome [1.4 T]

- Correlation coefficient, slope, y-intercept, and general error fraction (GEF) (ratio of average error to average length) between the population count of every DNA element (1-nt to 20-nt) and its reverse DNA element, its complement DNA element, and its reverse-complement DNA element (=6 comparisons) in the single-stranded human (hg19) genome

Table S49. Population Count of 1-nt to 20-nt DNA Elements in the Single-Stranded Mouse Genome (Chromosome)

- Correlation coefficient, slope, y-intercept, and general error fraction (GEF) (ratio of average error to average length) between the population count of every DNA element (1-nt to 20-nt) and its reverse DNA element, its complement DNA element, and its reverse-complement DNA element (=6 comparisons) in the single-stranded mouse (mm10) genome

Table S50. Population Count of 1-nt to 20-nt DNA Elements in the Single-Stranded Human Genome (Chromosome)

- Correlation coefficient, slope, y-intercept, and general error fraction (GEF) (ratio of average error to average length) between the population count of every DNA element (1-nt to 20-nt) and its reverse DNA element, its complement DNA element, and its reverse-complement DNA element (=6 comparisons) in the single-stranded human (hg19) genome

Table S51-S54. p53 DNA Element Analysis

The analyses in this section were completed by overlapping the location coordinates of each 0-nt to 5-nt variant 10-nt p53RE DNA element in the genome and the location coordinates of the ChIPSeq or ChIPExo peaks in an experiment. The 0-nt variant consensus palindromic 10-nt p53RE DNA element (5′-TGCCCGGGCA-3′) has ten (10) primary positions. The 0-nt to 5-nt variant p53RE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements (10 variant positions), 405 2-nt variant DNA elements (45 variant positions), 3,240 3-nt variant DNA elements (120 variant positions), 17,010 4-nt variant DNA elements (210 variant positions), and 61,236 5-nt variant DNA elements (252 variant positions), for a total of 81,922 DNA elements.

Table S51. Categorize 0-nt to 5-nt Variant 10-nt p53REs into 252 Half-Site Groups (Variant Position)

- Categorize the 81,922 0-nt to 5-nt variant 10-nt p53RE DNA elements into 252 half-site groups (5 positions are fixed, allowing for *up to* 5 positions to be varied)
- Each of the 252 p53RE half-site groups contain: 1 0-nt variant consensus palindromic p53RE DNA element, 15 1-nt variant p53RE DNA elements, 90 2-nt variant p53RE DNA elements, 270 3-nt variant p53RE DNA elements, 405 4-nt variant p53RE DNA elements, 243 5-nt variant p53RE DNA elements, for a total of 1,024 p53RE DNA elements
- 252 half-site groups x 1,024 DNA elements per half-site group =258,048 DNA elements (81,922 unique)

Table S52. Categorize 0-nt to 5-nt Variant 10-nt p53REs into 252 Half-Site Groups (Sequence)

- Categorize the 81,922 0-nt to 5-nt variant 10-nt p53RE DNA elements into 252 half-site groups (5 positions are fixed, allowing for *up to* 5 positions to be varied)
- Each of the 252 p53RE half-site groups contain: 1 0-nt variant consensus palindromic p53RE DNA element, 15 1-nt variant p53RE DNA elements, 90 2-nt variant p53RE DNA elements, 270 3-nt variant p53RE DNA elements, 405 4-nt variant p53RE DNA elements, 243 5-nt variant p53RE DNA elements, for a total of 1,024 p53RE DNA elements
- 252 half-site groups x 1,024 DNA elements per half-site group =258,048 DNA elements (81,922 unique)

Table S53. X-Axis Order of 252 p53RE Half-Site Groups

- (S/N) analysis of 0-nt to 5-nt variant p53REs in p53 ChIPSeq or ChIPExo peaks (displayed by 252 p53RE half-site groups)
- X-axis order
    - 126 half-site groups [left-to-right: 0 vacancies, 1 vacancy, 2 vacancies] followed by their 126 reverse-complements
    - Of these 252 half-site groups, 32 have zero vacancies, 160 have one vacancy, 60 have two vacancies
    - The x-axis is labeled by the reverse-complement vacancy position ID (primary label) and reverse-complement double occupant position ID (secondary label)
    - The order of the reverse-complement vacancy position IDs for p53RE is 5-6 > 1-10 > 3-8 > 4-7 > 2-9

Table S54. (S/N) analysis of 0-nt to 5-nt Variant p53REs in p53 ChIPSeq Peaks (252 Half-Site Groups) [22]

- (S/N) analysis of 0-nt to 5-nt variant p53REs in p53 ChIPSeq or ChIPExo peaks (displayed by 252 p53RE half-site groups)
- 22 p53 experiments
- Peak selection criteria =L4-L20
- Categorize the counts of p53 DNA-binding at 0-nt to 5-nt variant p53REs into 252 p53RE half-site groups (bottom table)
- Convert these counts to (S/N) values (top table)

Table S55-S63. Number of DNA Elements (ERE, HRE, p53RE) in the Mouse and Human Genome

The number of 0-nt to 5-nt variants of the 13-nt ERE and HRE consensus palindromic DNA element on the positive/sense strand in the mouse (mm10) and human (hg19) genome. Here the 13-nt ERE is the estrogen response element (5′-GGTCAnnnTGACC-3′) and the 13-nt HRE is the hormone response element (5′-GAACAnnnTGTTC-3′). The 0-nt to 5-nt variant ERE or HRE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements (10 variant positions), 405 2-nt variant DNA elements (45 variant positions), 3,240 3-nt variant DNA elements (120 variant positions), 17,010 4-nt variant DNA elements (210 variant positions), and 61,236 5-nt variant DNA elements (252 variant positions), for a total of 81,922 DNA elements.

The number of 0-nt to 6-nt variants of the 15-nt ERE and HRE consensus palindromic DNA element on the positive/sense strand in the mouse (mm10) and human (hg19) genome. Here the 15-nt ERE is the estrogen response element (5′-AGGTCAnnnTGACCT-3′) and the 15-nt HRE is the hormone response element (5′-AGAACAnnnTGTTCT-3′). The 0-nt to 6-nt variant 15-nt ERE or HRE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 36 1-nt variant DNA elements (12 variant positions), 594 2-nt variant DNA elements (66 variant positions), 5,940 3-nt variant DNA elements (220 variant positions), 40,095 4-nt variant DNA elements (495 variant positions), 192,456 5-nt variant DNA elements (792 variant positions), and 673,596 6-nt variant DNA elements (924 variant positions), for a total of 912,718 DNA elements.

The number of 0-nt to 5-nt variants of the 10-nt p53RE consensus palindromic DNA element on the positive/sense strand in the mouse (mm10) genome. Here the 10-nt p53RE is the p53 response element (5′-TGCCCGGGCA-3′). The 0-nt to 5-nt variant p53RE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements (10 variant positions), 405 2-nt variant DNA elements (45 variant positions), 3,240 3-nt variant DNA elements (120 variant positions), 17,010 4-nt variant DNA elements (210 variant positions), and 61,236 5-nt variant DNA elements (252 variant positions), for a total of 81,922 DNA elements.

Table S55. Number of 0-nt to 5-nt Variant 13-nt ERE DNA Elements in the Mouse Genome (81,922)

- Population count of the 81,922 0-nt to 5-nt variant 13-nt EREs in the mouse (mm10) genome [=177,963,712]
- The 0-nt to 5-nt variant 13-nt EREs include:
    - 1 0-nt variant consensus palindromic ERE
    - 30 1-nt variant EREs (10 variant positions)
    - 405 2-nt variant EREs (45 variant positions)
    - 3,240 3-nt variant EREs (120 variant positions)
    - 17,010 4-nt variant EREs (210 variant positions)
    - 61,236 5-nt variant EREs (252 variant positions)

Table S56. Number of 0-nt to 5-nt Variant 13-nt ERE DNA Elements in the Human Genome (81,922)

- Population count of the 81,922 0-nt to 5-nt variant 13-nt EREs in the human (hg19) genome [=187,996,184]
- The 0-nt to 5-nt variant 13-nt EREs include:
    - 1 0-nt variant consensus palindromic ERE
    - 30 1-nt variant EREs (10 variant positions)
    - 405 2-nt variant EREs (45 variant positions)
    - 3,240 3-nt variant EREs (120 variant positions)
    - 17,010 4-nt variant EREs (210 variant positions)
    - 61,236 5-nt variant EREs (252 variant positions)

Table S57. Number of 0-nt to 5-nt Variant 13-nt HRE DNA Elements in the Mouse Genome (81,922)

- Population count of the 81,922 0-nt to 5-nt variant 13-nt HREs in the mouse (mm10) genome [=230,731,128]
- The 0-nt to 5-nt variant 13-nt HREs include:
    - 1 0-nt variant consensus palindromic HRE
    - 30 1-nt variant HREs (10 variant positions)
    - 405 2-nt variant HREs (45 variant positions)
    - 3,240 3-nt variant HREs (120 variant positions)
    - 17,010 4-nt variant HREs (210 variant positions)
    - 61,236 5-nt variant HREs (252 variant positions)

Table S58. Number of 0-nt to 5-nt Variant 13-nt HRE DNA Elements in the Human Genome (81,922)

- Population count of the 81,922 0-nt to 5-nt variant 13-nt HREs in the human (hg19) genome [=253,825,466]
- The 0-nt to 5-nt variant 13-nt HREs include:
    - 1 0-nt variant consensus palindromic HRE
    - 30 1-nt variant HREs (10 variant positions)
    - 405 2-nt variant HREs (45 variant positions)
    - 3,240 3-nt variant HREs (120 variant positions)
    - 17,010 4-nt variant HREs (210 variant positions)

- 61,236 5-nt variant HREs (252 variant positions)

*****

Table S59. Number of 0-nt to 6-nt Variant 15-nt ERE DNA Elements in the Mouse Genome (912,718)

- Population count of the 912,718 0-nt to 6-nt variant 15-nt EREs in the mouse (mm10) genome [=143,500,605]
- The 0-nt to 6-nt variant 15-nt EREs include:
  - 1 0-nt variant consensus palindromic ERE
  - 36 1-nt variant EREs (12 variant positions)
  - 594 2-nt variant EREs (66 variant positions)
  - 5,940 3-nt variant EREs (220 variant positions)
  - 40,095 4-nt variant EREs (495 variant positions)
  - 192,456 5-nt variant EREs (792 variant positions)
  - 673,596 6-nt variant EREs (924 variant positions)

Table S60. Number of 0-nt to 6-nt Variant 15-nt ERE DNA Elements in the Human Genome (912,718)

- Population count of the 912,718 0-nt to 6-nt variant 15-nt EREs in the human (hg19) genome [=154,794,932]
- The 0-nt to 6-nt variant 15-nt EREs include:
  - 1 0-nt variant consensus palindromic ERE
  - 36 1-nt variant EREs (12 variant positions)
  - 594 2-nt variant EREs (66 variant positions)
  - 5,940 3-nt variant EREs (220 variant positions)
  - 40,095 4-nt variant EREs (495 variant positions)
  - 192,456 5-nt variant EREs (792 variant positions)
  - 673,596 6-nt variant EREs (924 variant positions)

Table S61. Number of 0-nt to 6-nt Variant 15-nt HRE DNA Elements in the Mouse Genome (912,718)

- Population count of the 912,718 0-nt to 6-nt variant 15-nt HREs in the mouse (mm10) genome [=191,056,970]
- The 0-nt to 6-nt variant 15-nt HREs include:
  - 1 0-nt variant consensus palindromic HRE
  - 36 1-nt variant HREs (12 variant positions)
  - 594 2-nt variant HREs (66 variant positions)
  - 5,940 3-nt variant HREs (220 variant positions)
  - 40,095 4-nt variant HREs (495 variant positions)
  - 192,456 5-nt variant HREs (792 variant positions)
  - 673,596 6-nt variant HREs (924 variant positions)

Table S62. Number of 0-nt to 6-nt Variant 15-nt HRE DNA Elements in the Human Genome (912,718)

- Population count of the 912,718 0-nt to 6-nt variant 15-nt HREs in the human (hg19) genome [=212,963,770]
- The 0-nt to 6-nt variant 15-nt HREs include:
  - 1 0-nt variant consensus palindromic HRE
  - 36 1-nt variant HREs (12 variant positions)
  - 594 2-nt variant HREs (66 variant positions)
  - 5,940 3-nt variant HREs (220 variant positions)
  - 40,095 4-nt variant HREs (495 variant positions)
  - 192,456 5-nt variant HREs (792 variant positions)
  - 673,596 6-nt variant HREs (924 variant positions)

*****

Table S63. Number of 0-nt to 5-nt Variant 10-nt p53RE DNA Elements in the Mouse Genome (81,922)

- Population count of the 81,922 0-nt to 5-nt variant p53REs in the mouse (mm10) genome [=149,248,264]
- The 0-nt to 5-nt variant 10-nt p53REs include:
  - 1 0-nt variant consensus palindromic p53RE
  - 30 1-nt variant p53REs (10 variant positions)
  - 405 2-nt variant p53REs (45 variant positions)
  - 3,240 3-nt variant p53REs (120 variant positions)
  - 17,010 4-nt variant p53REs (210 variant positions)
  - 61,236 5-nt variant p53REs (252 variant positions)

Table S64-S71. Data Analysis Example (WT-E2-1hr.L4)

Step-by-step instructions of the data analysis from peak selection to (S/N) analysis of ER DNA-binding at 81,922 0-nt to 5-nt variant 13-nt EREs and at 912,718 0-nt to 6-nt variant 15-nt EREs in the genome (displayed by 252 half-site groups and 924 half-site groups)

Table S64. Count 0-nt to 5-nt Variant 13-nt EREs in ER ChIPSeq Peaks (ChIPSeq Peak)

- The experiment WT-E2-1hr.L4 contains 76,163 ChIPSeq peaks (146-nt peak length)
  - 1,201 peaks contain a 0-nt variant consensus palindromic ERE
  - 7,349 peaks contain a 1-nt variant ERE
  - 16,368 peaks contain a 2-nt variant ERE
  - 34,605 peaks contain a 3-nt variant ERE
  - 68,484 peaks contain a 4-nt variant ERE
  - 76,147 peaks contain a 5-nt variant ERE
- These 76,163 peaks contain 827,919 0-nt to 5-nt variant 13-nt EREs
- Listed by ChIPSeq peak

Table S65. Count 0-nt to 5-nt Variant 13-nt EREs in ER ChIPSeq Peaks (Variant Position)

- The experiment WT-E2-1hr.L4 contains 76,163 ChIPSeq peaks (146-nt peak length)
- These 76,163 peaks contain 827,919 0-nt to 5-nt variant 13-nt EREs
- Listed by variant position

Table S66. Categorize 0-nt to 5-nt Variant 13-nt EREs into 252 Half-Site Groups

- Categorize the 827,919 0-nt to 5-nt variant 13-nt EREs into the 252 half-site groups
- For example, the half-site group 1-2-3-4-5 contains all 1-nt variant EREs that vary positions 6 through 10 (i.e., 6A, 6C, 6G, 7A, 7C, 7T, 8C, 8G, 8T, 9A, 9G, 9T, 10A, 10G, 10T), totaling 3,724
- To convert this count to a (S/N) value, multiply the 76,163 peaks by the 146-nt peak length (=11,119,798 nucleotides), and divide 11,119,798 nucleotides by 4^{10} (i.e., the probability that any 10-nt DNA element that has a maximum possibility of 4 nucleotides in each position will occur in the genome is once every 1,048,576 nucleotides (4^{10}) at random occurrence), which yields an expected noise signal of 10.60
- There are 15 (1-nt variant DNA elements) in each 252 half-site group, thus 10.60 x 15 =159.07 expected noise signal
- The DNA element occurrence (3,724), divided by the expected noise (159.07), gives a (S/N) value of 23.41, which is displayed on the 252 half-site plots

Table S67. ER DNA-Binding at Non-NRFE Sites is Independent of NRFE-Containing ChIPSeq Peaks (13-nt ERE)

- The DNA-binding structure observed at non-NRFE DNA elements in the genome is independent of the binding peaks that contain an NRFE

- (S/N) value for all 827,919 0-nt to 5-nt variant EREs:
  - ER DNA-binding at the 0-nt variant consensus palindromic ERE =113.35
  - ER DNA-binding at 1-nt variant EREs =23.47
  - ER DNA-binding at 2-nt variant EREs =4.20
  - ER DNA-binding at 3-nt variant EREs =1.36
  - ER DNA-binding at 4-nt variant EREs =0.96
  - ER DNA-binding at 5-nt variant EREs =0.90

- (S/N) value for all 827,919 0-nt to 5-nt variant EREs (excluding peaks that have already been assigned to an ERE DNA element with less variants):
  - 1,201 peaks contain a 0-nt variant consensus palindromic ERE =7188.01
  - 7,332 peaks contain a 1-nt variant ERE =243.22
  - 15,490 peaks contain a 2-nt variant ERE =19.53
  - 24,374 peaks contain a 3-nt variant ERE =3.02
  - 25,194 peaks contain a 4-nt variant ERE =1.08
  - 2,599 peaks contain a 5-nt variant ERE =0.96

- (S/N) value for all other ERE DNA elements found within these groups
  - Of the 1,201 peaks that contain a 0-nt variant consensus palindromic ERE =7188.01
    - 17 of them also contain a 1-nt variant ERE =3.59
    - 88 of them also contain a 2-nt variant ERE =1.34
    - 419 of them also contain a 3-nt variant ERE =0.98
    - 1,055 of them also contain a 4-nt variant ERE =0.83
    - 1,201 of them also contain a 5-nt variant ERE =0.85

Table S68. Count 0-nt to 6-nt Variant 15-nt EREs in ER ChIPSeq Peaks (ChIPSeq Peak)

- The experiment WT-E2-1hr.L4 contains 76,163 ChIPSeq peaks (146-nt peak length)
    - 304 peaks contain a 0-nt variant consensus palindromic ERE
    - 2,707 peaks contain a 1-nt variant ERE
    - 7,912 peaks contain a 2-nt variant ERE
    - 15,311 peaks contain a 3-nt variant ERE
    - 31,661 peaks contain a 4-nt variant ERE
    - 62,903 peaks contain a 5-nt variant ERE
    - 75,891 peaks contain a 6-nt variant ERE
- These 76,163 peaks contain 605,166 0-nt to 6-nt variant 15-nt EREs
- Listed by ChIPSeq peak

Table S69. Count 0-nt to 6-nt Variant 15-nt EREs in ER ChIPSeq Peaks (Variant Position)

- The experiment WT-E2-1hr.L4 contains 76,163 ChIPSeq peaks (146-nt peak length)
- These 76,163 peaks contain 605,166 0-nt to 6-nt variant 15-nt EREs
- Listed by variant position

Table S70. Categorize 0-nt to 6-nt Variant 15-nt EREs into 924 Half-Site Groups

- Categorize the 605,166 0-nt to 6-nt variant 15-nt EREs into the 924 half-site groups
- For example, the half-site group 1-2-3-4-5-6 contains all 1-nt variant ERE DNA elements that vary positions 7 through 12 (i.e., 7A, 7C, 7G, 8A, 8C, 8T, 9C, 9G, 9T, 10A, 10G, 10T, 11A, 11G, 11T, 12A, 12C, 12G), totaling 1,320
- To convert this count to a (S/N) value, multiply the 76,163 peaks by the 146-nt peak length (=11,119,798 nucleotides), and divide 11,119,798 nucleotides by 4^{12} (i.e., the probability that any 12-nt DNA element that has a maximum possibility of 4 nucleotides in each position will occur in the genome is once every 16,777,216 nucleotides (4^{12}) at random occurrence), which yields an expected noise signal of 0.66
- There are 18 (1-nt variant DNA elements) in each 252 half-site group, thus 0.66 x 18 =11.93 expected noise signal
- The DNA element occurrence (1,320), divided by the expected noise (11.93), gives a (S/N) value of 110.64, which is displayed on the 924 half-site plots

Table S71. ER DNA-Binding at Non-NRFE Sites is Independent of NRFE-Containing ChIPSeq Peaks (15-nt ERE)

- The DNA-binding structure observed at non-NRFE DNA elements in the genome is independent of the binding peaks that contain an NRFE

- (S/N) value for all 605,166 0-nt to 6-nt variant EREs:
    - ER DNA-binding at the 0-nt variant consensus palindromic ERE =460.17
    - ER DNA-binding at 1-nt variant EREs =113.91
    - ER DNA-binding at 2-nt variant EREs =20.59
    - ER DNA-binding at 3-nt variant EREs =4.28
    - ER DNA-binding at 4-nt variant EREs =1.56
    - ER DNA-binding at 5-nt variant EREs =1.03
    - ER DNA-binding at 6-nt variant EREs =0.91

- (S/N) value for all 605,166 0-nt to 6-nt variant EREs (excluding peaks that have already been assigned to an ERE DNA element with less variants):
    - 304 peaks contain a 0-nt variant consensus palindromic ERE =115290.44
    - 2,705 peaks contain a 1-nt variant ERE =3204.99
    - 7,840 peaks contain a 2-nt variant ERE =198.09
    - 13,922 peaks contain a 3-nt variant ERE =21.35
    - 21,585 peaks contain a 4-nt variant ERE =3.78
    - 24,479 peaks contain a 5-nt variant ERE =1.25
    - 5,308 peaks contain a 6-nt variant ERE =0.96

- (S/N) value for all other ERE DNA elements found within these groups
    - Of the 304 peaks that contain a 0-nt variant consensus palindromic ERE =115290.44
        - 2 of them also contain a 1-nt variant ERE =21.00
        - 2 of them also contain a 2-nt variant ERE =1.27
        - 28 of them also contain a 3-nt variant ERE =1.78
        - 120 of them also contain a 4-nt variant ERE =1.43
        - 241 of them also contain a 5-nt variant ERE =0.97
        - 304 of them also contain a 6-nt variant ERE =0.92

Table S72-S74. Genome Assembly of the Mouse and Human Genome

Table S72. Genome Assembly (Chromosome Size) of the Mouse and Human Genome

- The mouse and human genome assemblies were obtained from UCSC Genome Browser
  - mouse genome build 38 (mm10)
  - human genome build 37 (hg19)
- The genome size of the mouse (mm10) genome is 2,730,871,774 nucleotides
- The genome size of the human (hg19) genome is 3,137,161,264 nucleotides

Table S73. Genome Assembly (Effective Size) of the Mouse and Human Genome

- The effective genome size of the mouse (mm10) genome is 2,647,537,730 nucleotides when excluding unreadable nucleotides, random regions, hap regions, and chromosome Un (i.e., clone contigs that cannot be confidently placed on a specific chromosome) from the reference
- The effective genome size of the human (hg19) genome is 2,861,343,702 nucleotides when excluding unreadable nucleotides, random regions, hap regions, and chromosome Un (i.e., clone contigs that cannot be confidently placed on a specific chromosome) from the reference

Table S74. Genome Assembly (N Islands) of the Mouse and Human Genome

- The unreadable nucleotides account for 2.9% [77,999,939] of the mouse (mm10) genome
- The unreadable nucleotides account for 7.5% [234,350,281] of the human (hg19) of the genome
- These unreadable nucleotides are contained within 501 unreadable regions in the mouse (mm10) genome
- These unreadable nucleotides are contained within 339 unreadable regions in the human (hg19) genome

Table S75. Script Codes for All Data Analyses

- Script codes for data analysis of sNR DNA-binding at 0-nt to 5-nt variant DNA elements in the genome
    - motif_positions
    - 258058.ERE.txt
    - 258058.HRE.txt
- Script codes for genome assembly of the mouse (mm10) and human (hg19) genomes
    - linreg.awk
    - get_error_terms_new

Datafile S1-S5 Descriptions: Compressed Datafiles

Datafile S1. ChIPSeq and ChIPExo Experiments

- 157 ER experiments
- 194 KR experiments
- 8 GR and GR-Dim experiments
- Peak selection criteria =L4, L8, L10, L15, L20
- doi:10.17632/8ks7pm2jw6.1 (Mendeley)

Datafile S2. Population Count of 1-nt to 20-nt DNA Elements in the Single-Stranded Mouse Genome [1.4 T]

- The population count of every 1-nt to 20-nt DNA elements in the single-stranded mouse (mm10) genome, including its reverse-complement DNA element, reverse DNA element, and complement DNA element
- Total =20 files (1-nt to 20-nt DNA elements)
- tar -xvzf *filename*.tar.gz to unzip files

Datafile S3. Population Count of 1-nt to 20-nt DNA Elements in the Single-Stranded Human Genome [1.4 T]

- The population count of every 1-nt to 20-nt DNA elements in the single-stranded human (hg19) genome, including its reverse-complement DNA element, reverse DNA element, and complement DNA element
- Total =20 files (1-nt to 20-nt DNA elements)
- tar -xvzf *filename*.tar.gz to unzip files

Datafile S4. Population Count of 1-nt to 20-nt DNA Elements in the Single-Stranded Mouse Genome (Chromosome)

- The population count of every 1-nt to 20-nt DNA elements (by chromosome) in the single-stranded mouse (mm10) genome, including its reverse-complement DNA element, reverse DNA element, and complement DNA element
- Total =440 files (1-nt to 20-nt DNA elements x 22)
    - 19 chromosomes
    - chromosome M
    - chromosome X
    - chromosome Y
- tar -xvzf *filename*.tar.gz to unzip files

Datafile S5. Population Count of 1-nt to 20-nt DNA Elements in the Single-Stranded Human Genome (Chromosome)

- The population count of every 1-nt to 20-nt DNA elements (by chromosome) in the single-stranded human (hg19) genome, including its reverse-complement DNA element, reverse DNA element, and complement DNA element
- Total =500 files (1-nt to 20-nt DNA elements x 25)
    - 22 chromosomes
    - chromosome M
    - chromosome X
    - chromosome Y
- tar -xvzf *filename*.tar.gz to unzip files

Transparent Methods

Quality Control and Peak Selection Criteria

In total, 381 ChIPSeq and ChIPExo experiments were obtained from Gene Expression Omnibus (GEO) and converted to fastq format using fastq-dump v2.4.5; original experiment names were retained.

Reads were first selected using a cross-correlation analysis (trim_and_filter_SE.pl) (5).

The sequencing reads were then mapped uniquely, allowing for no more than two mismatches, to the reference genome [mouse genome build 38 (mm10) or human genome build 37 (hg19)] using Bowtie v1.1.2 (6).

Mapped reads were deduplicated using MarkDuplicates.jar from the Picard tools package v1.96.

Peak selection was performed using Hypergeometric Optimization of Motif Enrichment (HOMER) v4.7.2. Multiple peak selection criteria were used (L4, L8, L10, L15, L20), where Lx represents an x-fold greater tag density at peaks than in the surrounding 10-kb region. This performs a low-to-high stringency analysis of the data. By studying the data with respect to a multiple spectrum of peak selection criteria, we adjust for both the risk of excess background noise and the risk of filtering out any low-amplitude information.

See our previously published study for the associated script codes and data analyses for these experiments (2).

Number and Location Coordinates of DNA Elements (ERE, HRE, p53RE) in the Mouse and Human Genome

Number and Location Coordinates of 0-nt to 5-nt Variant 13-nt ERE and HRE DNA Elements in the Genome

The number and location coordinates of the 81,922 0-nt to 5-nt variants of the 13-nt ERE and HRE consensus palindromic DNA element on the positive/sense strand were identified in the mouse (mm10) and human (hg19) genome using OligoMatch (University of California, Santa Cruz). Here the 13-nt ERE is the estrogen response element (5′-GGTCAnnnTGACC-3′) and the 13-nt HRE is the hormone response element (5′-GAACAnnnTGTTC-3′). The 0-nt to 5-nt variant 13-nt ERE and HRE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements (10 variant positions), 405 2-nt variant DNA elements (45 variant positions), 3,240 3-nt variant DNA elements (120 variant positions), 17,010 4-nt variant DNA elements (210 variant positions), and 61,236 5-nt variant DNA elements (252 variant positions), for a total of 81,922 DNA elements. See our previously published study for the associated script codes (2). The population count of each of the 81,922 0-nt to 5-nt variant 13-nt ERE and HRE DNA elements in the mouse (mm10) and human (hg19) genome can be found in **Table S55-S58**.

Four nucleotide possibilities in a DNA element with 10 primary positions:

n = number of primary positions in the DNA element = **10**

L = number of nucleotide possibilities = **4**

k = number of variants

Binomial Equation:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Combinations:

$$L^n = [(L-1)+1]^n = \sum_{k=0}^{n} \binom{n}{k} (L-1)^k \, (1)^{n-k}$$

$$= \sum_{k=0}^{n} \binom{n}{k} (L-1)^k = 2^{20} = 4^{10}$$

	DNA Element	Combinatorial Counts				DNA Element Frequency in Genome			
A	B	C	D	E	F	Mouse Genome (mm10)		Human Genome (hg19)	
k	**13-nt ERE** GGTCAnnnTGACC	Combinations $C = \binom{n}{k}$	4-nts $D = C \times 4^k$	Unique $E = C \times 3^k$	Total Unique $F = \Sigma E$	Unique	Total Unique	Unique	Total Unique
0	0-nt Variant ERE	1	1	1	1	2,367		2,194	
1	1-nt Variant ERE	10	40	30	31	71,428	73,795	60,313	62,507
2	2-nt Variant ERE	45	720	405	436	898,155	971,950	914,726	977,233
3	3-nt Variant ERE	120	7,680	3,240	3,676	6,750,607	7,722,557	7,516,184	8,493,417
4	4-nt Variant ERE	210	53,760	17,010	20,686	35,508,190	43,230,747	38,222,674	46,716,091
5	5-nt Variant ERE	252	258,048	61,236	81,922	134,732,965	177,963,712	141,280,093	187,996,184

A	B	C	D	E	F	Mouse Genome (mm10)		Human Genome (hg19)	
k	**13-nt HRE** GAACAnnnTGTTC	Combinations $C = \binom{n}{k}$	4-nts $D = C \times 4^k$	Unique $E = C \times 3^k$	Total Unique $F = \Sigma E$	Unique	Total Unique	Unique	Total Unique
0	0-nt Variant HRE	1	1	1	1	3,444		3,535	
1	1-nt Variant HRE	10	40	30	31	97,039	100,483	104,767	108,302
2	2-nt Variant HRE	45	720	405	436	1,337,516	1,437,999	1,339,543	1,447,845
3	3-nt Variant HRE	120	7,680	3,240	3,676	10,461,197	11,899,196	10,771,159	12,219,004
4	4-nt Variant HRE	210	53,760	17,010	20,686	49,391,434	61,290,630	54,136,564	66,355,568
5	5-nt Variant HRE	252	258,048	61,236	81,922	169,440,498	230,731,128	187,469,898	253,825,466

A	B	C	D	E	F
k		Combinations $C = \binom{n}{k}$	4-nts $D = C \times 4^k$	Unique $E = C \times 3^k$	Total Unique $F = \Sigma E$
0	0-nt Variant	10!/(0! x 10!)	1 x 1	1 x 1	1 + 0
1	1-nt Variant	10!/(1! x 9!)	10 x 4	10 x 3	30 + 1
2	2-nt Variant	10!/(2! x 8!)	45 x 16	45 x 9	405 + 31
3	3-nt Variant	10!/(3! x 7!)	120 x 64	120 x 27	3,240 + 436
4	4-nt Variant	10!/(4! x 6!)	210 x 256	210 x 81	17,010 +3,676
5	5-nt Variant	10!/(5! x 5!)	252 x 1,024	252 x 243	61,236 + 20,686

Number and Location Coordinates of 0-nt to 6-nt Variant 15-nt ERE and HRE DNA Elements in the Genome

The number and location coordinates of the 912,718 0-nt to 6-nt variants of the 15-nt ERE and HRE consensus palindromic DNA element on the positive/sense strand were identified in the mouse (mm10) and human (hg19) genome using OligoMatch (University of California, Santa Cruz). Here the 15-nt ERE is the estrogen response element (5′-AGGTCAnnnTGACCT-3′) and the 13-nt HRE is the hormone response element (5′-AGAACAnnnTGTTCT-3′). The 0-nt to 6-nt variant 15-nt ERE and HRE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 36 1-nt variant DNA elements (12 variant positions), 594 2-nt variant DNA elements (66 variant positions), 5,940 3-nt variant DNA elements (220 variant positions), 40,095 4-nt variant DNA elements (495 variant positions), 192,456 5-nt variant DNA elements (792 variant positions), and 673,596 6-nt variant DNA elements (924 variant positions), for a total of 912,718 DNA elements. See our previously published study for the associated script codes (2). The population count of each of the 912,718 0-nt to 6-nt variant 15-nt ERE and HRE DNA elements in the mouse (mm10) and human (hg19) genome can be found in **Table S59-S62**.

Four nucleotide possibilities in a DNA element with 12 primary positions:

n = number of primary positions in the DNA element = **12**

L = number of nucleotide possibilities = **4**

k = number of variants

Binomial Equation:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Combinations:

$$L^n = [(L-1)+1]^n = \sum_{k=0}^{n} \binom{n}{k} (L-1)^k \, (1)^{n-k}$$

$$= \sum_{k=0}^{n} \binom{n}{k} (L-1)^k = 2^{24} = 4^{12}$$

	DNA Element	Combinatorial Counts				DNA Element Frequency in Genome			
A	B	C	D	E	F	Mouse Genome (mm10)		Human Genome (hg19)	
k	**15-nt ERE** AGGTCAnnnTGACCT	Combinations $C = \binom{n}{k}$	4-nts $D = C \times 4^k$	Unique $E = C \times 3^k$	Total Unique $F = \Sigma E$	Unique	Total Unique	Unique	Total Unique
0	0-nt Variant ERE	1	1	1	1	404		342	
1	1-nt Variant ERE	12	48	36	37	10,964	11,368	9,223	9,565
2	2-nt Variant ERE	66	1056	594	631	152,140	163,508	118,354	127,919
3	3-nt Variant ERE	220	14,080	5,940	6,571	1,256,647	1,420,155	1,213,963	1,341,882
4	4-nt Variant ERE	495	126,720	40,095	46,666	6,916,002	8,336,157	7,349,740	8,691,622
5	5-nt Variant ERE	792	811,008	192,456	239,122	30,575,777	38,911,934	33,052,166	41,743,788
6	6-nt Variant ERE	924	3,784,704	673,596	912,718	104,588,671	143,500,605	113,051,144	154,794,932

A	B	C	D	E	F	Mouse Genome (mm10)		Human Genome (hg19)	
k	**15-nt HRE** AGAACAnnnTGTTCT	Combinations $C = \binom{n}{k}$	4-nts $D = C \times 4^k$	Unique $E = C \times 3^k$	Total Unique $F = \Sigma E$	Unique	Total Unique	Unique	Total Unique
0	0-nt Variant HRE	1	1	1	1	639		565	
1	1-nt Variant HRE	12	48	36	37	17,847	18,486	14,639	15,204
2	2-nt Variant HRE	66	1056	594	631	256,540	275,026	206,983	222,187
3	3-nt Variant HRE	220	14,080	5,940	6,571	2,188,277	2,463,303	1,978,470	2,200,657
4	4-nt Variant HRE	495	126,720	40,095	46,666	11,128,133	13,591,436	11,661,975	13,862,632
5	5-nt Variant HRE	792	811,008	192,456	239,122	43,400,607	56,992,043	48,017,520	61,880,152
6	6-nt Variant HRE	924	3,784,704	673,596	912,718	134,064,927	191,056,970	151,083,618	212,963,770

A	B	C	D	E	F
k		Combinations $C = \binom{n}{k}$	4-nts $D = C \times 4^k$	Unique $E = C \times 3^k$	Total Unique $F = \Sigma E$
0	0-nt Variant	12!/(0! x 12!)	1 x 1	1 x 1	1 + 0
1	1-nt Variant	12!/(1! x 11!)	12 x 4	12 x 3	36 + 1
2	2-nt Variant	12!/(2! x 10!)	66 x 16	66 x 9	594 + 37
3	3-nt Variant	12!/(3! x 9!)	220 x 64	220 x 27	5,940 + 631
4	4-nt Variant	12!/(4! x 8!)	495 x 256	495 x 81	40,095 + 6,571
5	5-nt Variant	12!/(5! x 7!)	792 x 1,024	792 x 243	192,456 + 46,666
6	6-nt Variant	12!/(6! x 6!)	924 x 4,096	924 x 729	673,596 + 239,122

Number and Location Coordinates of 0-nt to 5-nt Variant 10-nt p53RE DNA Elements in the Genome

The number and location coordinates of the 81,922 0-nt to 5-nt variant of the 10-nt p53RE consensus palindromic DNA element on the positive/sense strand were identified in the mouse (mm10) genome using OligoMatch (University of California, Santa Cruz). Here the 10-nt p53RE is the p53 response element (5′-TGCCCGGGCA-3′). The 0-nt to 5-nt variant 10-nt p53RE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements (10 variant positions), 405 2-nt variant DNA elements (45 variant positions), 3,240 3-nt variant DNA elements (120 variant positions), 17,010 4-nt variant DNA elements (210 variant positions), and 61,236 5-nt variant DNA elements (252 variant positions), for a total of 81,922 DNA elements. See our previously published study for the associated script codes (2). The population count of each of the 81,922 0-nt to 5-nt variant 10-nt p53RE DNA elements in the mouse (mm10) genome can be found in **Table S63**.

Four nucleotide possibilities in a DNA element with 10 primary positions:

n = number of primary positions in the DNA element = **10**

L = number of nucleotide possibilities = **4**

k = number of variants

Binomial Equation:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Combinations:

$$L^n = [(L-1)+1]^n = \sum_{k=0}^{n} \binom{n}{k} (L-1)^k \, (1)^{n-k}$$

$$= \sum_{k=0}^{n} \binom{n}{k} (L-1)^k = 2^{20} = 4^{10}$$

	DNA Element	Combinatorial Counts				DNA Element Frequency in Genome	
A	**B**	**C**	**D**	**E**	**F**	**Mouse Genome (mm10)**	
k	**P53RE TGCCCGGGCA**	**Combinations** $C = \binom{n}{k}$	**4-nts** $D = C \times 4^k$	**Unique** $E = C \times 3^k$	**Total Unique** $F = \Sigma E$	**Unique**	**Total Unique**
0	0-nt Variant p53RE	1	1	1	1	1,078	
1	1-nt Variant p53RE	10	40	30	31	33,992	35,070
2	2-nt Variant p53RE	45	720	405	436	641,422	676,492
3	3-nt Variant p53RE	120	7,680	3,240	3,676	5,380,457	6,056,949
4	4-nt Variant p53RE	210	53,760	17,010	20,686	29,257,379	35,314,328
5	5-nt Variant p53RE	252	258,048	61,236	81,922	113,933,936	149,248,264

A	**B**	**C**	**D**	**E**	**F**
k		**Combinations** $C = \binom{n}{k}$	**4-nts** $D = C \times 4^k$	**Unique** $E = C \times 3^k$	**Total Unique** $F = \Sigma E$
0	0-nt Variant	10!/(0! x 10!)	1 x 1	1 x 1	1 + 0
1	1-nt Variant	10!/(1! x 9!)	10 x 4	10 x 3	30 + 1
2	2-nt Variant	10!/(2! x 8!)	45 x 16	45 x 9	405 + 31
3	3-nt Variant	10!/(3! x 7!)	120 x 64	120 x 27	3,240 + 436
4	4-nt Variant	10!/(4! x 6!)	210 x 256	210 x 81	17,010 +3,676
5	5-nt Variant	10!/(5! x 5!)	252 x 1,024	252 x 243	61,236 + 20,686

Overlap DNA Elements (ERE, HRE, p53RE) in the Genome with ChIPSeq and ChIPExo Experiments

A customized C++ program, cppmatch_custom_full.cpp, was used to overlap location coordinates of each 0-nt to 5-nt/6-nt variant ERE, HRE, or p53RE DNA element in the genome and the location coordinates of the ChIPSeq or ChIPExo peaks in an experiment to determine the absolute number of times each 0-nt to 5-nt/6-nt variant ERE, HRE, or p53RE DNA element occurred within an experiment (the entire 10-nt, 13-nt or 15-nt DNA element was required to be within the peak boundaries).

See our previously published study for the associated script codes (2).

The complete experiments, including individual peaks and 0-nt to 5-nt variant overlaps, have also been provided (tar -xvzf *filename*.tar.gz to unzip files) **(Datafile S1)**.

Convert 0-nt to 5-nt/6-nt Variant DNA Elements to Variant Positions

Convert 0-nt to 5-nt Variant 13-nt ERE and HRE DNA Elements to Variant Positions

To evaluate the DNA-binding of sNRs at every DNA element in the 0-nt to 5-nt variant groups, the absolute number of times each of the 81,922 0-nt to 5-nt variant 13-nt ERE or HRE DNA elements occurred in an experiment was counted.

The 0-nt variant consensus palindromic 13-nt ERE DNA element (5′-GGTCAnnnTGACC-3′) and 0-nt variant consensus palindromic 13-nt HRE DNA element (5′-GAACAnnnTGTTC-3′) have ten (10) primary positions.

The 0-nt to 5-nt variant ERE or HRE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements (10 variant positions), 405 2-nt variant DNA elements (45 variant positions), 3,240 3-nt variant DNA elements (120 variant positions), 17,010 4-nt variant DNA elements (210 variant positions), and 61,236 5-nt variant DNA elements (252 variant positions), for a total of 81,922 DNA elements.

See **Table S75** for script codes to convert the counts of all overlapped 0-nt to 5-nt DNA elements to variant positions.

13-nt ERE = G<u>GT</u>CAnnnTG<u>AC</u>C
1 2 3 4 5 6 7 8 9 10

13-nt HRE = G<u>AA</u>CAnnnTG<u>TT</u>C
1 2 3 4 5 6 7 8 9 10

13-nt ERE and HRE DNA Elements						
A	**B**	**C**	**D**	**E**	**F**	**G**
	DNA Element	**Variant Position Combinations**		**Palindromic Position Pairs (PP)**	**Same-Side Variants**	**Crossover Variants**
k	**13-nt ERE or HRE (*n*=10)**	**C** = $\binom{n}{k}$	**D** = $\binom{n}{k}$			
0	0-nt Variant ERE or HRE	10!/(0! x 10!)	1			
1	1-nt Variant ERE or HRE	10!/(1! x 9!)	10			
2	2-nt Variant ERE or HRE	10!/(2! x 8!)	45	5	20	20
3	3-nt Variant ERE or HRE	10!/(3! x 7!)	120		20	100
4	4-nt Variant ERE or HRE	10!/(4! x 6!)	210	10	10	190
5	5-nt Variant ERE or HRE	10!/(5! x 5!)	252		2	250

	Palindromic Position Pairs (PP)			
	13-nt ERE and HRE DNA Elements			
	2-nt ERE	**2-nt HRE**	**4-nt ERE**	**4-nt HRE**
1	3-8	2-9	1-3-8-10	2-5-6-9
2	1-10	5-6	3-5-6-8	1-2-9-10
3	5-6	1-10	3-4-7-8	2-4-7-9
4	4-7	4-7	2-3-8-9	2-3-8-9
5	2-9	3-8	1-5-6-10	1-5-6-10
6			1-4-7-10	4-5-6-7
7			1-2-9-10	3-5-6-8
8			4-5-6-7	1-4-7-10
9			2-5-6-9	1-3-8-10
10			2-4-7-9	3-4-7-8

Convert 0-nt to 6-nt Variant 15-nt ERE and HRE DNA Elements to Variant Positions

To evaluate the DNA-binding of sNRs at every DNA element in the 0-nt to 6-nt variant groups, the absolute number of times each of the 912,718 0-nt to 6-nt variant 15-nt ERE or HRE DNA elements occurred in an experiment was counted.

The 0-nt variant consensus palindromic 15-nt ERE DNA element (5′-AGGTCAnnnTGACCT-3′) and 0-nt variant consensus palindromic 15-nt HRE DNA element (5′-AGAACAnnnTGTTCT-3′) have twelve (12) primary positions.

The 0-nt to 6-nt variant 15-nt ERE or HRE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 36 1-nt variant DNA elements (12 variant positions), 594 2-nt variant DNA elements (66 variant positions), 5,940 3-nt variant DNA elements (220 variant positions), 40,095 4-nt variant DNA elements (495 variant positions), 192,456 5-nt variant DNA elements (792 variant positions), and 673,596 6-nt variant DNA elements (924 variant positions), for a total of 912,718 DNA elements.

See **Table S75** for script codes to convert the counts of all overlapped 0-nt to 6-nt DNA elements to variant positions.

15-nt ERE = AGGTCAnnnTGACCT
1 2 3 4 5 6 7 8 9 10 11 12

15-nt HRE = AGAACAnnnTGTTCT
1 2 3 4 5 6 7 8 9 10 11 12

15-nt ERE and HRE DNA Elements						
A	**B**	**C**	**D**	**E**	**F**	**G**
	DNA Element	**Variant Position Combinations**		**Palindromic Position Pairs (PP)**	**Same-Side Variants**	**Crossover Variants**
k	**15-nt ERE or HRE (*n*=12)**	**C** = $\binom{n}{k}$	**D** = $\binom{n}{k}$			
0	0-nt Variant ERE or HRE	12!/(0! x 12!)	1			
1	1-nt Variant ERE or HRE	12!/(1! x 11!)	12			
2	2-nt Variant ERE or HRE	12!/(2! x 10!)	66	6	30	30
3	3-nt Variant ERE or HRE	12!/(3! x 9!)	220		40	180
4	4-nt Variant ERE or HRE	12!/(4! x 8!)	495	15	30	450
5	5-nt Variant ERE or HRE	12!/(5! x 7!)	792		12	780
6	6-nt Variant ERE or HRE	12!/(6! x 6!)	924	20	2	902

	Palindromic Position Pairs (PP)					
	15-nt ERE and HRE DNA Elements					
	2-nt ERE	**2-nt HRE**	**4-nt ERE**	**4-nt HRE**	**6-nt ERE**	**6-nt HRE**
1	4-9	3-10	2-4-9-11	3-6-7-10	1-4-6-7-9-12	1-2-3-10-11-12
2	2-11	6-7	4-6-7-9	2-3-10-11	2-3-4-9-10-11	3-4-5-8-9-10
3	6-7	2-11	4-5-8-9	3-5-8-10	1-3-4-9-10-12	1-3-4-9-10-12
4	5-8	5-8	3-4-9-10	3-4-9-10	1-2-4-9-11-12	1-2-4-9-11-12
5	3-10	4-9	2-6-7-11	2-6-7-11	2-4-6-7-9-11	2-3-6-7-10-11
6	1-12	1-12	2-5-8-11	5-6-7-8	1-4-5-8-9-12	1-3-5-8-10-12
7			2-3-10-11	4-6-7-9	2-4-5-8-9-11	3-5-6-7-8-10
8			5-6-7-8	2-5-8-11	3-4-6-7-9-10	2-5-6-7-8-11
9			3-6-7-10	2-4-9-11	1-2-6-7-11-12	1-2-6-7-11-12
10			3-5-8-10	4-5-8-9	4-5-6-7-8-9	4-5-6-7-8-9
11			1-4-9-12	1-3-10-12	1-2-5-8-11-12	1-5-6-7-8-12
12			1-2-11-12	1-6-7-12	1-3-6-7-10-12	2-3-5-8-10-11
13			1-6-7-12	1-2-11-12	1-2-3-10-11-12	1-4-6-7-9-12
14			1-5-8-12	1-5-8-12	3-4-5-8-9-10	2-3-4-9-10-11
15			1-3-10-12	1-4-9-12	1-5-6-7-8-12	1-2-5-8-11-12
16					2-3-5-8-10-11	1-3-6-7-10-12
17					2-3-6-7-10-11	2-4-6-7-9-11
18					1-3-5-8-10-12	1-4-5-8-9-12
19					3-5-6-7-8-10	2-4-5-8-9-11
20					2-5-6-7-8-11	3-4-6-7-9-10

Convert 0-nt to 5-nt Variant 10-nt p53RE DNA Elements to Variant Positions

To evaluate the DNA-binding of p53 at every DNA element in the 0-nt to 5-nt variant groups, the absolute number of times each of the 81,922 0-nt to 5-nt variant 10-nt p53RE DNA elements occurred in an experiment was counted.

The 0-nt variant consensus palindromic 10-nt p53RE DNA element (5′-TGCCCGGGCA-3′) has ten (10) primary positions.

The 0-nt to 5-nt variant p53RE DNA elements include the 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements (10 variant positions), 405 2-nt variant DNA elements (45 variant positions), 3,240 3-nt variant DNA elements (120 variant positions), 17,010 4-nt variant DNA elements (210 variant positions), and 61,236 5-nt variant DNA elements (252 variant positions), for a total of 81,922 DNA elements.

See **Table S75** for script codes to convert the counts of all overlapped 0-nt to 5-nt DNA elements to variant positions.

10-nt p53RE = TGCCCGGGCA

1 2 3 4 5 6 7 8 9 10

10-nt p53RE DNA Elements						
A	**B**	**C**	**D**	**E**	**F**	**G**
	DNA Element	**Variant Position Combinations**		**Palindromic Position Pairs (PP)**	**Same-Side Variants**	**Crossover Variants**
k	**P53RE (*n*=10)**	**C** = $\binom{n}{k}$	**D** = $\binom{n}{k}$			
0	0-nt Variant p53RE	10!/(0! x 10!)	1			
1	1-nt Variant p53RE	10!/(1! x 9!)	10			
2	2-nt Variant p53RE	10!/(2! x 8!)	45	5	20	20
3	3-nt Variant p53RE	10!/(3! x 7!)	120		20	100
4	4-nt Variant p53RE	10!/(4! x 6!)	210	10	10	190
5	5-nt Variant p53RE	10!/(5! x 5!)	252		2	250

	Palindromic Position Pairs (PP)	
	10-nt p53RE DNA Elements	
	2-nt p53RE	**4-nt p53RE**
1	5-6	1-5-6-10
2	1-10	3-5-6-8
3	3-8	4-5-6-7
4	4-7	2-5-6-9
5	2-9	1-3-8-10
6		1-4-7-10
7		1-2-9-10
8		3-4-7-8
9		2-3-8-9
10		2-4-7-9

Categorize 0-nt to 5-nt/6-nt Variant DNA Elements into Half-Site Groups

Categorize 0-nt to 5-nt Variant 13-nt ERE and HRE DNA Elements Into 252 Half-Site Groups

To evaluate the DNA-binding of sNRs at every DNA element in the 0-nt to 5-nt variant groups, the absolute number of times each of the 81,922 0-nt to 5-nt variant DNA elements occurred in an experiment were counted. For display purposes, we have defined a 5-nt variant ERE and HRE DNA element by its five (5) fixed positions, resulting in 252 half-site groups (i.e., categorizing the 61,236 5-nt variant DNA elements into 252 half-site groups, defined by the 5 positions that are fixed/not varied). The remaining 0-nt to 4-nt variant DNA elements (1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements, 405 2-nt variant DNA elements, 3,240 3-nt variant DNA elements, and 17,010 4-nt variant DNA elements, for a total of 20,686 DNA elements) can be categorized into these same 252 half-site groups (i.e., 5 positions are fixed, allowing for *up to* 5 positions to be varied). For example, the half-site group 1-2-3-4-5 contains all 1-nt variant DNA elements that vary positions 6 through 10. In the case of the ERE, this includes the following 1-nt variant ERE DNA elements: 6A, 6C, 6G, 7A, 7C, 7T, 8C, 8G, 8T, 9A, 9G, 9T, 10A, 10G, 10T. Thus, each of the 252 half-site groups contain 1 (0-nt variant consensus palindromic DNA element), 15 (1-nt variant DNA elements), 90 (2-nt variant DNA elements), 270 (3-nt variant DNA elements), 405 (4-nt variant DNA elements), and 243 (5-nt variant DNA elements), for a total of 1,024 DNA elements per half-site group. This allows all 81,922 0-nt to 5-nt variant DNA elements to be categorized into the 252 half-site groups, and thus providing the ability to sequentially track sNR DNA-binding at all 81,922 0-nt to 5-nt variant ERE or HRE DNA elements in the genome **(Table S1-S4)**.

Of these 252 half-site groups, thirty-two (32) have zero vacancies, one hundred and sixty (160) have one vacancy, and sixty (60) have two vacancies.

See **Table S75** for script codes to convert the counts of all overlapped 0-nt to 5-nt DNA elements into the half-site groups.

Four nucleotide possibilities in a DNA element with 5 primary positions:

n = number of primary positions in the DNA element = **5**

L = number of nucleotide possibilities = **4**

k = number of variants

Binomial Equation:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Combinations:

$$L^n = [(L-1)+1]^n = \sum_{k=0}^{n} \binom{n}{k} (L-1)^k \, (1)^{n-k}$$

$$= \sum_{k=0}^{n} \binom{n}{k} (L-1)^k = 4^5 = 2^{10}$$

	DNA Element	Combinatorial Counts			
A	**B**	**C**	**D**	**E**	**F**
k	**5-nt Half-Site**	**Combinations** $C = \binom{n}{k}$	**4-nts** $D = C \times 4^k$	**Unique** $E = C \times 3^k$	**Total Unique** $F = \Sigma E$
0	0-nt Variant Half-Site	1	1	**1**	1
1	1-nt Variant Half-Site	5	20	**15**	16
2	2-nt Variant Half-Site	10	160	**90**	106
3	3-nt Variant Half-Site	10	640	**270**	376
4	4-nt Variant Half-Site	5	1,280	**405**	781
5	5-nt Variant Half-Site	1	1,024	**243**	1,024

A	**B**	**C**	**D**	**E**	**F**
k	**GGTCAnnnTGACC GAACAnnnTGTTC**	**Combinations** $C = \binom{n}{k}$	**4-nts** $D = C \times 4^k$	**Unique** $E = C \times 3^k$	**Total Unique** $F = \Sigma E$
0	0-nt Variant	5!/(0! x 5!)	1 x 1	1 x 1	1 + 0
1	1-nt Variant	5!/(1! x 4!)	5 x 4	5 x 3	15 + 1
2	2-nt Variant	5!/(2! x 3!)	10 x 16	10 x 9	90 + 16
3	3-nt Variant	5!/(3! x 2!)	10 x 64	10 x 27	270 + 106
4	4-nt Variant	5!/(4! x 1!)	5 x 256	5 x 81	405 + 376
5	5-nt Variant	5!/(5! x 0!)	1 x 1,024	1 x 243	243 + 781

Categorize 0-nt to 6-nt Variant 15-nt ERE and HRE DNA Elements Into 924 Half-Site Groups

To evaluate the DNA-binding of sNRs at every DNA element in the 0-nt to 6-nt variant groups, the absolute number of times each of the 912,718 0-nt to 6-nt variant DNA elements occurred in an experiment were counted. For display purposes, we have defined a 6-nt variant ERE and HRE DNA element by its six (6) fixed positions, resulting in 924 half-site groups (i.e., categorizing the 673,596 6-nt variant DNA elements into 924 half-site groups, defined by the 6 positions that are fixed/not varied). The remaining 0-nt to 5-nt variant DNA elements (1 0-nt variant consensus palindromic DNA element, 36 1-nt variant DNA elements, 594 2-nt variant DNA elements, 5,940 3-nt variant DNA elements, and 40,095 4-nt variant DNA elements, 192,456 5-nt variant DNA elements, for a total of 239,122 elements) can be categorized into these same 924 half-site groups (i.e., 6 positions are fixed, allowing for *up to* 6 positions to be varied). For example, the half-site group 1-2-3-4-5-6 contains all 1-nt variant DNA elements that vary positions 7 through 12. In the case of the ERE, this includes the following 1-nt variant ERE DNA elements: 7A, 7C, 7G, 8A, 8C, 8T, 9C, 9G, 9T, 10A, 10G, 10T, 11A, 11G, 11T, 12A, 12C, 12G. Thus, each of the 924 half-site groups contain 1 (0-nt variant consensus palindromic DNA element), 18 (1-nt variant DNA elements), 135 (2-nt variant DNA elements), 540 (3-nt variant DNA elements), 1,215 (4-nt variant DNA elements), 1,458 (5-nt variant DNA elements), and 729 (6-nt variant DNA elements), for a total of 4,096 DNA elements per half-site group. This allows all 912,718 0-nt to 6-nt variant DNA elements to be categorized into the 924 half-site groups, and thus providing the ability to sequentially track sNR DNA-binding at 0-nt to 6-nt variant ERE or HRE DNA elements in the genome **(Table S21-S24)**.

Of these 924 half-site groups, sixty-four (64) have zero vacancies, four hundred and eighty (480) have one vacancy, three hundred and sixty (360) have two vacancies, and twenty (20) have three vacancies.

See **Table S75** for script codes to convert the counts of all overlapped 0-nt to 6-nt DNA elements into the half-site groups.

Four nucleotide possibilities in a DNA element with 6 primary positions:

n = number of primary positions in the DNA element = **6**

L = number of nucleotide possibilities = **4**

k = number of variants

Binomial Equation:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Combinations:

$$L^n = [(L-1)+1]^n = \sum_{k=0}^{n} \binom{n}{k} (L-1)^k \, (1)^{n-k}$$

$$= \sum_{k=0}^{n} \binom{n}{k} (L-1)^k = 4^6 = 2^{12}$$

	DNA Element	Combinatorial Counts			
A	**B**	**C**	**D**	**E**	**F**
k	**6-nt Half-Site**	**Combinations** $C = \binom{n}{k}$	**4-nts** $D = C \times 4^k$	**Unique** $E = C \times 3^k$	**Total Unique** $F = \Sigma E$
0	0-nt Variant Half-Site	1	1	**1**	1
1	1-nt Variant Half-Site	6	24	**18**	19
2	2-nt Variant Half-Site	15	240	**135**	154
3	3-nt Variant Half-Site	20	1,280	**540**	694
4	4-nt Variant Half-Site	15	3,840	**1,215**	1,909
5	5-nt Variant Half-Site	6	6,144	**1,458**	3,367
6	6-nt Variant Half-Site	1	4,096	**729**	4,096

A	**B**	**C**	**D**	**E**	**F**
k	**AGGTCAnnnTGACCT AGAACAnnnTGTTCT**	**Combinations** $C = \binom{n}{k}$	**4-nts** $D = C \times 4^k$	**Unique** $E = C \times 3^k$	**Total Unique** $F = \Sigma E$
0	0-nt Variant	6!/(0! x 6!)	1 x 1	1 x 1	1 + 0
1	1-nt Variant	6!/(1! x 5!)	6 x 4	6 x 3	18 + 1
2	2-nt Variant	6!/(2! x 4!)	15 x 16	15 x 9	135 + 19
3	3-nt Variant	6!/(3! x 3!)	20 x 64	20 x 27	540 + 154
4	4-nt Variant	6!/(4! x 2!)	15 x 256	15 x 81	1,215 + 694
5	5-nt Variant	6!/(5! x 1!)	6 x 1,024	6 x 243	1,458 + 1,909
6	6-nt Variant	6!/(6! x 0!)	1 x 4,096	1 x 729	729 + 3,367

Categorize 0-nt to 5-nt Variant 10-nt p53RE DNA Elements Into 252 Half-Site Groups

To evaluate the DNA-binding of p53 at every DNA element in the 0-nt to 5-nt variant groups, the absolute number of times each of the 81,922 0-nt to 5-nt variant DNA elements occurred in an experiment were counted. For display purposes, we have defined a 5-nt variant p53RE DNA element by its five (5) fixed positions, resulting in 252 half-site groups (i.e., categorizing the 61,236 5-nt variant DNA elements into 252 half-site groups, defined by the 5 positions that are fixed/not varied). The remaining 0-nt to 4-nt variant DNA elements (1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements, 405 2-nt variant DNA elements, 3,240 3-nt variant DNA elements, and 17,010 4-nt variant DNA elements, for a total of 20,686 DNA elements) can be categorized into these same 252 half-site groups (i.e., 5 positions are fixed, allowing for *up to* 5 positions to be varied). For example, the half-site group 1-2-3-4-5 contains all 1-nt variant DNA elements that vary positions 6 through 10. In the case of the p53RE, this includes the following 1-nt variant p53RE DNA elements: 6A, 6C, 6T, 7A, 7C, 7T, 8A, 8C, 8T, 9A, 9G, 9T, 10C, 10G, 10T. Thus, each of the 252 half-site groups contain 1 (0-nt variant consensus palindromic DNA element), 15 (1-nt variant DNA elements), 90 (2-nt variant DNA elements), 270 (3-nt variant DNA elements), 405 (4-nt variant DNA elements), and 243 (5-nt variant DNA elements), for a total of 1,024 DNA elements per half-site group. This allows all 81,922 0-nt to 5-nt variant DNA elements to be categorized into the 252 half-site groups, and thus providing the ability to sequentially track p53 DNA-binding at all 81,922 0-nt to 5-nt variant p53RE DNA elements in the genome **(Table S51-S52)**.

Of these 252 half-site groups, thirty-two (32) have zero vacancies, one hundred and sixty (160) have one vacancy, and sixty (60) have two vacancies.

See **Table S75** for script codes to convert the counts of all overlapped 0-nt to 5-nt DNA elements into the half-site groups.

Four nucleotide possibilities in a DNA element with 5 primary positions:

n = number of primary positions in the DNA element = **5**

L = number of nucleotide possibilities = **4**

k = number of variants

Binomial Equation:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Combinations:

$$L^n = [(L-1)+1]^n = \sum_{k=0}^{n} \binom{n}{k} (L-1)^k \, (1)^{n-k}$$

$$= \sum_{k=0}^{n} \binom{n}{k} (L-1)^k = 4^5 = 2^{10}$$

	DNA Element	Combinatorial Counts			
A	**B**	**C**	**D**	**E**	**F**
k	**5-nt Half-Site**	**Combinations** $C = \binom{n}{k}$	**4-nts** $D = C \times 4^k$	**Unique** $E = C \times 3^k$	**Total Unique** $F = \Sigma E$
0	0-nt Variant Half-Site	1	1	**1**	1
1	1-nt Variant Half-Site	5	20	**15**	16
2	2-nt Variant Half-Site	10	160	**90**	106
3	3-nt Variant Half-Site	10	640	**270**	376
4	4-nt Variant Half-Site	5	1,280	**405**	781
5	5-nt Variant Half-Site	1	1,024	**243**	1,024

A	**B**	**C**	**D**	**E**	**F**
k	**TGCCCGGGCA**	**Combinations** $C = \binom{n}{k}$	**4-nts** $D = C \times 4^k$	**Unique** $E = C \times 3^k$	**Total Unique** $F = \Sigma E$
0	0-nt Variant	5!/(0! x 5!)	1 x 1	1 x 1	1 + 0
1	1-nt Variant	5!/(1! x 4!)	5 x 4	5 x 3	15 + 1
2	2-nt Variant	5!/(2! x 3!)	10 x 16	10 x 9	90 + 16
3	3-nt Variant	5!/(3! x 2!)	10 x 64	10 x 27	270 + 106
4	4-nt Variant	5!/(4! x 1!)	5 x 256	5 x 81	405 + 376
5	5-nt Variant	5!/(5! x 0!)	1 x 1,024	1 x 243	243 + 781

Signal-To-Noise (S/N) Analysis

sNR DNA-Binding at 0-nt to 5-nt Variant 13-nt ERE and HRE DNA Elements in the Genome

The signal-to-noise ratio (S/N) is the absolute number of times a 0-nt to 5-nt variant 13-nt ERE or HRE DNA element occurs in an experiment [defined by the total number of peaks in the experiment and the peak length] compared to the random frequency expectation of that DNA element occurring in the genome (i.e., the probability that any 10-nt DNA element that has a maximum possibility of 4 nucleotides in each position will occur in the genome is once every 1,048,576 nucleotides (4^{10}) at random occurrence). Thus, the expected noise is calculated by multiplying the total number of peaks in the experiment by the length of the peaks over 4^{10}.

For (S/N) analysis of sNR DNA-binding at 0-nt to 5-nt variant 13-nt ERE or HRE DNA elements in the genome (displayed by the number of variants in the DNA element), the expected noise is multiplied by the number of DNA elements in each group: 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements, 405 2-nt variant DNA elements, 3,240 3-nt variant DNA elements, 17,010 4-nt variant DNA elements, and 61,236 5-nt variant DNA elements, for a total of 81,922 0-nt to 5-nt variant DNA elements.

For (S/N) analysis of sNR DNA-binding at 0-nt to 5-nt variant 13-nt ERE or HRE DNA elements in the genome (displayed by the variant position in the DNA element) (i.e., 10 variant positions in 1-nt variant DNA elements, 45 variant positions in 2-nt variant DNA elements, 120 variant positions in 3-nt variant DNA elements, 210 variant positions in 4-nt variant DNA elements, 252 variant positions in 5-nt variant DNA elements), the expected noise is multiplied by the number of DNA elements in each group: 1 (0-nt variant consensus palindromic DNA element), 3 (1-nt variant DNA elements), 9 (2-nt variant DNA elements), 27 (3-nt variant DNA elements), 81 (4-nt variant DNA elements), and 243 (5-nt variant DNA elements), for a total of 81,922 0-nt to 5-nt variant DNA elements.

	DNA Element	Combinatorial Counts			
A	**B**	**C**	**D**	**E**	**F**
k	**13-nt ERE or HRE (*n*=10)**	**Combinations** $C = \binom{n}{k}$	**4-nts** $D = C \times 4^k$	**Unique** $E = C \times 3^k$	**Total Unique** $F = \Sigma E$
0	0-nt Variant ERE or HRE	**1**	1	**1**	1
1	1-nt Variant ERE or HRE	**10**	40	**30**	31
2	2-nt Variant ERE or HRE	**45**	720	**405**	436
3	3-nt Variant ERE or HRE	**120**	7,680	**3,240**	3,676
4	4-nt Variant ERE or HRE	**210**	53,760	**17,010**	20,686
5	5-nt Variant ERE or HRE	**252**	258,048	**61,236**	**81,922**

A	**B**	**C**	**D**	**E**	**F**
k		**Combinations** $C = \binom{n}{k}$	**4-nts** $D = C \times 4^k$	**Unique** $E = C \times 3^k$	**Total Unique** $F = \Sigma E$
0	0-nt Variant	10!/(0! x 10!)	1 x 1	1 x **1**	1 + 0
1	1-nt Variant	10!/(1! x 9!)	10 x 4	10 x **3**	30 + 1
2	2-nt Variant	10!/(2! x 8!)	45 x 16	45 x **9**	405 + 31
3	3-nt Variant	10!/(3! x 7!)	120 x 64	120 x **27**	3,240 + 436
4	4-nt Variant	10!/(4! x 6!)	210 x 256	210 x **81**	17,010 +3,676
5	5-nt Variant	10!/(5! x 5!)	252 x 1,024	252 x **243**	61,236 + 20,686

For (S/N) analysis of sNR DNA-binding at 0-nt to 5-nt variant 13-nt ERE or HRE DNA elements in the genome (displayed by the 252 half-site groups), the expected noise is multiplied by the number of DNA elements in each group: 1 (0-nt variant consensus palindromic DNA element), 15 (1-nt variant DNA elements), 90 (2-nt variant DNA elements), 270 (3-nt variant DNA elements), 405 (4-nt variant DNA elements), and 243 (5-nt variant DNA elements), for a total of 1,024 0-nt to 5-nt variant DNA elements per half-site group.

	DNA Element	Combinatorial Counts			
A	**B**	**C**	**D**	**E**	**F**
k	**13-nt Half-Site (*n*=5)**	**Combinations** $C = \binom{n}{k}$	**4-nts** $D = C \times 4^k$	**Unique** $E = C \times 3^k$	**Total Unique** $F = \Sigma E$
0	0-nt Variant Half-Site	1	1	**1**	1
1	1-nt Variant Half-Site	5	20	**15**	16
2	2-nt Variant Half-Site	10	160	**90**	106
3	3-nt Variant Half-Site	10	640	**270**	376
4	4-nt Variant Half-Site	5	1,280	**405**	781
5	5-nt Variant Half-Site	1	1,024	**243**	1,024

A	**B**	**C**	**D**	**E**	**F**

k		**Combinations** $C = \binom{n}{k}$	**4-nts** $D = C \times 4^k$	**Unique** $E = C \times 3^k$	**Total Unique** $F = \Sigma E$
0	0-nt Variant	5!/(0! x 5!)	1 x 1	1 x 1	1 + 0
1	1-nt Variant	5!/(1! x 4!)	5 x 4	5 x 3	15 + 1
2	2-nt Variant	5!/(2! x 3!)	10 x 16	10 x 9	90 + 16
3	3-nt Variant	5!/(3! x 2!)	10 x 64	10 x 27	270 + 106
4	4-nt Variant	5!/(4! x 1!)	5 x 256	5 x 81	405 + 376
5	5-nt Variant	5!/(5! x 0!)	1 x 1,024	1 x 243	243 + 781

The upper and lower one-tailed Poisson significance thresholds at $p<0.001$ were calculated using R: qpois(c(0.001,0.999),((total number of peaks in the experiment x peak length x number of DNA elements)/4^{10})).

sNR DNA-Binding at 0-nt to 6-nt Variant 15-nt ERE and HRE DNA Elements in the Genome

The signal-to-noise ratio (S/N) is the absolute number of times a 0-nt to 6-nt variant 15-nt ERE or HRE DNA element occurs in an experiment [defined by the total number of peaks in the experiment and the peak length] compared to the random frequency expectation of that DNA element occurring in the genome (i.e., the probability that any 12-nt DNA element that has a maximum possibility of 4 nucleotides in each position will occur in the genome is once every 16,777,216 nucleotides (4^{12}) at random occurrence). Thus, the expected noise is calculated by multiplying the total number of peaks in the experiment by the length of the peaks over 4^{12}.

For (S/N) analysis of sNR DNA-binding at 0-nt to 6-nt variant 15-nt ERE or HRE DNA elements in the genome (displayed by the number of variants in the DNA element), the expected noise is multiplied by the number of DNA elements in each group: 1 0-nt variant consensus palindromic DNA element, 36 1-nt variant DNA elements, 594 2-nt variant DNA elements, 5,940 3-nt variant DNA elements, 40,095 4-nt variant DNA elements, 192,456 5-nt variant DNA elements, and 673,596 6-nt variant DNA elements, for a total of 912,718 0-nt to 6-nt variant DNA elements.

For (S/N) analysis of sNR DNA-binding at 0-nt to 6-nt variant 15-nt ERE or HRE DNA elements in the genome (displayed by the variant position in the DNA element) (i.e., 12 variant positions in 1-nt variant DNA elements, 66 variant positions in 2-nt variant DNA elements, 220 variant positions in 3-nt variant DNA elements, 495 variant positions in 4-nt variant DNA elements, 792 variant positions in 5-nt variant DNA elements, 924 variant positions in 6-nt variant DNA elements), the expected noise is multiplied by the number of DNA elements in each group: 1 (0-nt variant consensus palindromic DNA element), 3 (1-nt variant DNA elements), 9 (2-nt variant DNA elements), 27 (3-nt variant DNA elements), 81 (4-nt variant DNA elements), 243 (5-nt variant DNA elements), and 729 (6-nt variant DNA elements), for a total of 912,718 0-nt to 6-nt variant DNA elements.

	DNA Element	Combinatorial Counts			
A	**B**	**C**	**D**	**E**	**F**
k	**15-nt ERE or HRE (n=12)**	**Combinations $C = \binom{n}{k}$**	**4-nts $D = C \times 4^k$**	**Unique $E = C \times 3^k$**	**Total Unique $F = \Sigma E$**
0	0-nt Variant ERE or HRE	**1**	1	**1**	1
1	1-nt Variant ERE or HRE	**12**	48	**36**	37
2	2-nt Variant ERE or HRE	**66**	1056	**594**	631
3	3-nt Variant ERE or HRE	**220**	14,080	**5,940**	6,571
4	4-nt Variant ERE or HRE	**495**	126,720	**40,095**	46,666
5	5-nt Variant ERE or HRE	**792**	811,008	**192,456**	239,122
6	6-nt Variant ERE or HRE	**924**	3,784,704	**673,596**	**912,718**

A	**B**	**C**	**D**	**E**	**F**
k		**Combinations $C = \binom{n}{k}$**	**4-nts $D = C \times 4^k$**	**Unique $E = C \times 3^k$**	**Total Unique $F = \Sigma E$**
0	0-nt Variant	12!/(0! x 12!)	1 x 1	1 x **1**	1 + 0
1	1-nt Variant	12!/(1! x 11!)	12 x 4	12 x **3**	36 + 1
2	2-nt Variant	12!/(2! x 10!)	66 x 16	66 x **9**	594 + 37
3	3-nt Variant	12!/(3! x 9!)	220 x 64	220 x **27**	5,940 + 631
4	4-nt Variant	12!/(4! x 8!)	495 x 256	495 x **81**	40,095 + 6,571
5	5-nt Variant	12!/(5! x 7!)	792 x 1,024	792 x **243**	192,456 + 46,666
6	6-nt Variant	12!/(6! x 6!)	924 x 4,096	924 x **729**	673,596 + 239,122

For (S/N) analysis of sNR DNA-binding at 0-nt to 6-nt variant 15-nt ERE or HRE DNA elements in the genome (displayed by the 924 half-site groups), the expected noise is multiplied by the number of DNA elements in each group: 1 (0-nt variant consensus palindromic DNA element), 18 (1-nt variant DNA elements), 135 (2-nt variant DNA elements), 540 (3-nt variant DNA elements), 1,215 (4-nt variant DNA elements), 1,458 (5-nt variant DNA elements), and 729 (6-nt variant DNA elements), for a total of 4,096 0-nt to 6-nt variant DNA elements per half-site group.

	DNA Element	Combinatorial Counts			
A	**B**	**C**	**D**	**E**	**F**
k	**15-nt Half-Site (n=6)**	**Combinations $C = \binom{n}{k}$**	**4-nts $D = C \times 4^k$**	**Unique $E = C \times 3^k$**	**Total Unique $F = \Sigma E$**
0	0-nt Variant Half-Site	1	1	**1**	1
1	1-nt Variant Half-Site	6	24	**18**	19
2	2-nt Variant Half-Site	15	240	**135**	154
3	3-nt Variant Half-Site	20	1,280	**540**	694
4	4-nt Variant Half-Site	15	3,840	**1,215**	1,909
5	5-nt Variant Half-Site	6	6,144	**1,458**	3,367
6	6-nt Variant Half-Site	1	4,096	**729**	4,096

A	**B**	**C**	**D**	**E**	**F**

k		**Combinations** $C = \binom{n}{k}$	**4-nts** $D = C \times 4^k$	**Unique** $E = C \times 3^k$	**Total Unique** $F = \Sigma E$
0	0-nt Variant	6!/(0! x 6!)	1 x 1	1 x 1	1 + 0
1	1-nt Variant	6!/(1! x 5!)	6 x 4	6 x 3	18 + 1
2	2-nt Variant	6!/(2! x 4!)	15 x 16	15 x 9	135 + 19
3	3-nt Variant	6!/(3! x 3!)	20 x 64	20 x 27	540 + 154
4	4-nt Variant	6!/(4! x 2!)	15 x 256	15 x 81	1,215 + 694
5	5-nt Variant	6!/(5! x 1!)	6 x 1,024	6 x 243	1,458 + 1,909
6	6-nt Variant	6!/(6! x 0!)	1 x 4,096	1 x 729	729 + 3,367

The upper and lower one-tailed Poisson significance thresholds at $p<0.001$ were calculated using R: qpois(c(0.001,0.999),((total number of peaks in the experiment x peak length x number of DNA elements)/4^{12})).

p53 DNA-Binding at 0-nt to 5-nt Variant 10-nt p53RE DNA Elements in the Genome

The signal-to-noise ratio (S/N) is the absolute number of times a 0-nt to 5-nt variant 10-nt p53RE DNA element occurs in an experiment [defined by the total number of peaks in the experiment and the peak length] compared to the random frequency expectation of that DNA element occurring in the genome (i.e., the probability that any 10-nt DNA element that has a maximum possibility of 4 nucleotides in each position will occur in the genome is once every 1,048,576 nucleotides (4^{10}) at random occurrence). Thus, the expected noise is calculated by multiplying the total number of peaks in the experiment by the length of the peaks over 4^{10}.

For (S/N) analysis of p53 DNA-binding at 0-nt to 5-nt variant 10-nt p53RE DNA elements in the genome (displayed by the number of variants in the DNA element), the expected noise is multiplied by the number of DNA elements in each group: 1 0-nt variant consensus palindromic DNA element, 30 1-nt variant DNA elements, 405 2-nt variant DNA elements, 3,240 3-nt variant DNA elements, 17,010 4-nt variant DNA elements, and 61,236 5-nt variant DNA elements, for a total of 81,922 0-nt to 5-nt variant DNA elements.

For (S/N) analysis of p53 DNA-binding at 0-nt to 5-nt variant 10-nt p53RE DNA elements in the genome (displayed by the variant position in the DNA element) (i.e., 10 variant positions in 1-nt variant DNA elements, 45 variant positions in 2-nt variant DNA elements, 120 variant positions in 3-nt variant DNA elements, 210 variant positions in 4-nt variant DNA elements, 252 variant positions in 5-nt variant DNA elements), the expected noise is multiplied by the number of DNA elements in each group: 1 (0-nt variant consensus palindromic DNA element), 3 (1-nt variant DNA elements), 9 (2-nt variant DNA elements), 27 (3-nt variant DNA elements), 81 (4-nt variant DNA elements), and 243 (5-nt variant DNA elements), for a total of 81,922 0-nt to 5-nt variant DNA elements.

	DNA Element	Combinatorial Counts			
A	**B**	**C**	**D**	**E**	**F**
k	**p53RE (n=10)**	**Combinations $C = \binom{n}{k}$**	**4-nts $D = C \times 4^k$**	**Unique $E = C \times 3^k$**	**Total Unique $F = \Sigma E$**
0	0-nt Variant p53RE	**1**	1	**1**	1
1	1-nt Variant p53RE	**10**	40	**30**	31
2	2-nt Variant p53RE	**45**	720	**405**	436
3	3-nt Variant p53RE	**120**	7,680	**3,240**	3,676
4	4-nt Variant p53RE	**210**	53,760	**17,010**	20,686
5	5-nt Variant p53RE	**252**	258,048	**61,236**	**81,922**

A	**B**	**C**	**D**	**E**	**F**
k		**Combinations $C = \binom{n}{k}$**	**4-nts $D = C \times 4^k$**	**Unique $E = C \times 3^k$**	**Total Unique $F = \Sigma E$**
0	0-nt Variant	10!/(0! x 10!)	1 x 1	1 x **1**	1 + 0
1	1-nt Variant	10!/(1! x 9!)	10 x 4	10 x **3**	30 + 1
2	2-nt Variant	10!/(2! x 8!)	45 x 16	45 x **9**	405 + 31
3	3-nt Variant	10!/(3! x 7!)	120 x 64	120 x **27**	3,240 + 436
4	4-nt Variant	10!/(4! x 6!)	210 x 256	210 x **81**	17,010 +3,676
5	5-nt Variant	10!/(5! x 5!)	252 x 1,024	252 x **243**	61,236 + 20,686

For (S/N) analysis of p53 DNA-binding at 0-nt to 5-nt variant 10-nt p53RE DNA elements in the genome (displayed by the 252 half-site groups), the expected noise is multiplied by the number of DNA elements in each group: 1 (0-nt variant consensus palindromic DNA element), 15 (1-nt variant DNA elements), 90 (2-nt variant DNA elements), 270 (3-nt variant DNA elements), 405 (4-nt variant DNA elements), and 243 (5-nt variant DNA elements), for a total of 1,024 0-nt to 5-nt variant DNA elements per half-site group.

	DNA Element	Combinatorial Counts			
A	**B**	**C**	**D**	**E**	**F**
k	**10-nt Half-Site (n=5)**	**Combinations $C = \binom{n}{k}$**	**4-nts $D = C \times 4^k$**	**Unique $E = C \times 3^k$**	**Total Unique $F = \Sigma E$**
0	0-nt Variant Half-Site	1	1	**1**	1
1	1-nt Variant Half-Site	5	20	**15**	16
2	2-nt Variant Half-Site	10	160	**90**	106
3	3-nt Variant Half-Site	10	640	**270**	376
4	4-nt Variant Half-Site	5	1,280	**405**	781
5	5-nt Variant Half-Site	1	1,024	**243**	1,024

A	**B**	**C**	**D**	**E**	**F**
k		**Combinations $C = \binom{n}{k}$**	**4-nts $D = C \times 4^k$**	**Unique $E = C \times 3^k$**	**Total Unique $F = \Sigma E$**
0	0-nt Variant	5!/(0! x 5!)	1 x 1	1 x 1	1 + 0
1	1-nt Variant	5!/(1! x 4!)	5 x 4	5 x 3	15 + 1
2	2-nt Variant	5!/(2! x 3!)	10 x 16	10 x 9	90 + 16

3	3-nt Variant	5!/(3! x 2!)	10 x 64	10 x 27	270 + 106
4	4-nt Variant	5!/(4! x 1!)	5 x 256	5 x 81	405 + 376
5	5-nt Variant	5!/(5! x 0!)	1 x 1,024	1 x 243	243 + 781

The upper and lower one-tailed Poisson significance thresholds at $p<0.001$ were calculated using R: qpois(c(0.001,0.999),((total number of peaks in the experiment x peak length x number of DNA elements)/4^{10})).

Data Analysis Example (WT-E2-1hr.L4)

ER DNA-Binding at 0-nt to 5-nt Variant 13-nt ERE DNA Elements in the Genome

Using WT-E2-1hr.L4 [76,163 peaks, 146-nt peak length] as an example experiment, we provide step-by-step instructions of the data analysis from peak selection to (S/N) analysis of ER DNA-binding at 81,922 0-nt to 5-nt variant 13-nt ERE DNA elements in the genome, displayed by the 252 half-site groups. The WT-E2-1hr.L4 experiment contains 76,163 ChIPSeq peaks **(Table S64)**. These 76,163 peaks contain 827,919 0-nt to 5-nt variant 13-nt ERE DNA elements **(Table S64, Column E)**. The 0-nt to 5-nt variant DNA elements include the 1 0-nt variant consensus palindromic ERE DNA element, 30 1-nt variant DNA elements (10 variant positions), 405 2-nt variant DNA elements (45 variant positions), 3,240 3-nt variant DNA elements (120 variant positions), 17,010 4-nt variant DNA elements (210 variant positions), and 61,236 5-nt variant DNA elements (252 variant positions), for a total of 81,922 DNA elements **(Table S65)**. First, we categorize these 827,919 0-nt to 5-nt variant ERE DNA elements by variant position **(Table S65)**. Then we categorize them into the 252 half-site groups, defined by the 5 positions that are fixed/not varied **(Table S66)**. For example, the half-site group 1-2-3-4-5 contains all 1-nt variant ERE DNA elements that vary positions 6 through 10 (i.e., 6A, 6C, 6G, 7A, 7C, 7T, 8C, 8G, 8T, 9A, 9G, 9T, 10A, 10G, 10T) **(Table S1)**, totaling 3,724 in this experiment **(Table S65-S66, Yellow)**. To convert this count to a (S/N) value, multiply the 76,163 peaks by the 146-nt peak length (=11,119,798 nucleotides), and divide 11,119,798 nucleotides by 4^{10} (i.e., the probability that any 10-nt DNA element that has a maximum possibility of 4 nucleotides in each position will occur in the genome is once every 1,048,576 nucleotides (4^{10}) at random occurrence), which yields an expected noise signal of 10.60 **(Table S66, Orange)**. There are 15 (1-nt variant DNA elements) in each 252 half-site group, thus 10.60 x 15 =159.07 expected noise signal **(Table S66, Orange)**. The DNA element occurrence (3,724), divided by the expected noise (159.07), gives a (S/N) value of 23.41, which is displayed on the 252 half-site plots **(Figure 2; Table S66, Yellow)**.

Next, to ensure that the sNR DNA-binding signal (i.e., (S/N) value) at non-NRFE DNA elements (i.e., 3-nt to 5-nt variant ERE DNA elements) is not due to non-NRFE DNA elements existing within peaks that also contain an NRFE, we first determined the (S/N) value for all 827,919 0-nt to 5-nt variant ERE DNA elements: ER DNA-binding at the 0-nt variant consensus palindromic ERE DNA element =113.35, ER DNA-binding at 1-nt variant ERE DNA elements =23.47, ER DNA-binding at 2-nt variant ERE DNA elements =4.20, ER DNA-binding at 3-nt variant ERE DNA elements =1.36, ER DNA-binding at 4-nt variant ERE DNA elements =0.96, ER DNA-binding at 5-nt variant ERE DNA elements =0.90 **(Table S67)**. Next, we calculated the (S/N) value for all 827,919 0-nt to 5-nt variant ERE DNA elements while excluding peaks that have already been assigned to an ERE DNA element with less variants. Thus, peak assignment is given to the 0-nt variant consensus palindromic ERE DNA element or the ERE DNA element with the least number of variants relative to the 0-nt variant consensus palindromic ERE DNA element (i.e., defining each peak by a single DNA element with the total number of peaks in an experiment =100%). Of 76,163 Peaks, 1,201 peaks contain a 0-nt variant consensus palindromic ERE DNA element ((S/N) value =7188.01), 7,332 peaks contain a 1-nt variant ERE DNA element ((S/N) value =243.22), 15,490 peaks contain a 2-nt variant ERE DNA element ((S/N) value =19.53), 24,374 peaks contain a 3-nt variant ERE DNA element ((S/N) value =3.02), 25,194 peaks contain a 4-nt variant ERE DNA element ((S/N) value =1.08), 2,599 peaks contain a 5-nt variant ERE DNA element ((S/N) value =0.96) **(Table S67)**. Next, we calculated the (S/N) value for all the other ERE DNA elements found within these groups **(Table S67)**. For example, of the 1,201 peaks that contain a 0-nt variant consensus palindromic ERE DNA element ((S/N) value =7188.01), 17 of them also contain a 1-nt variant ERE DNA element ((S/N) value =3.59), 88 of them also contain a 2-nt variant ERE DNA element ((S/N) value =1.34), 419 of them also contain a 3-nt variant ERE DNA element ((S/N) value =0.98), 1,055 of them also contain a 4-nt variant ERE DNA element ((S/N) value =0.83), and all 1,201 of them also contain a 5-nt variant ERE DNA element ((S/N) value =0.85) **(Table S67)**. The large (S/N) value for the ERE DNA element with the least number of variants, while the presence of other ERE DNA elements within those peaks are not enriched over what would be expected at random occurrence, was also observed in the 1-nt to 5-nt variant DNA element groups **(Table S67)**. This confirms that the DNA-binding structure observed at non-NRFE DNA elements in the genome is independent of the binding peaks that contain an NRFE DNA element.

(S/N) Value 13-nt ERE	Peaks Assigned to the ERE DNA Element with the Least Number of Variants					
	0-nt Variant ERE	1-nt Variant ERE	2-nt Variant ERE	3-nt Variant ERE	4-nt Variant ERE	5-nt Variant ERE
Other ERE DNA Elements Also Within Peaks	**1,201 Peaks**	**7,332 Peaks**	**15,490 Peaks**	**24,347 Peaks**	**25,194 Peaks**	**2,599 Peaks**
0-nt Variant ERE	**7188.01**					
1-nt Variant ERE	3.59	**243.22**				
2-nt Variant ERE	1.34	2.09	**19.53**			
3-nt Variant ERE	0.98	1.11	1.32	**3.02**		
4-nt Variant ERE	0.83	0.89	0.93	0.98	**1.08**	
5-nt Variant ERE	0.85	0.85	0.87	0.89	0.92	**0.96**

ER DNA-Binding at 0-nt to 6-nt Variant 15-nt ERE DNA Elements in the Genome

Using this same experiment, WT-E2-1hr.L4, we provide step-by-step instructions of the data analysis from peak selection to (S/N) analysis of ER DNA-binding at 912,718 0-nt to 6-nt variant 15-nt ERE DNA elements in the genome, displayed by the 924 half-site groups. The WT-E2-1hr.L4 experiment contains 76,163 ChIPSeq peaks **(Table S68)**. These 76,163 peaks contain 605,166 0-nt to 6-nt variant 15-nt ERE DNA elements **(Table S68, Column E)**. The 0-nt to 6-nt variant DNA elements include the 1 0-nt variant consensus palindromic DNA element, 36 1-nt variant DNA elements (12 variant positions), 594 2-nt variant DNA elements (66 variant positions), 5,940 3-nt variant DNA elements (220 variant positions), 40,095 4-nt variant DNA elements (495 variant positions), 192,456 5-nt variant DNA elements (792 variant positions), and 673,596 6-nt variant DNA elements (924 variant positions), for a total of 912,718 DNA elements **(Table S69)**. First, we categorize these 605,166 0-nt to 6-nt variant ERE DNA elements by variant position **(Table S69)**. Then we categorize them into the 924 half-site groups, defined by the 6 positions that are fixed/not varied **(Table S70)**. For example, the half-site group 1-2-3-4-5-6 contains all 1-nt variant ERE DNA elements that vary positions 7 through 12 (i.e., 7A, 7C, 7G, 8A, 8C, 8T, 9C, 9G, 9T, 10A, 10G, 10T, 11A, 11G, 11T, 12A, 12C, 12G) **(Table S21)**, totaling 1,320 in this experiment **(Table S69-S70, Yellow)**. To convert this count to a (S/N) value, multiply the 76,163 peaks by the 146-nt peak length (=11,119,798 nucleotides), and divide 11,119,798 nucleotides by 4^{12} (i.e., the probability that any 12-nt DNA element that has a maximum possibility of 4 nucleotides in each position will occur in the genome is once every 16,777,216 nucleotides (4^{12}) at random occurrence), which yields an expected noise signal of 0.66 **(Table S70, Orange)**. There are 18 (1-nt variant DNA elements) in each 924 half-site group, thus 0.66 x 18 =11.93 expected noise signal **(Table S70, Orange)**. The DNA element occurrence (1,320), divided by the expected noise (11.93), gives a (S/N) value of 110.64, which is displayed on the 924 half-site plots **(Figure S40; Table S70, Yellow)**.

Next, to ensure that the sNR DNA-binding signal (i.e., (S/N) value) at non-NRFE DNA elements is not due to non-NRFE DNA elements existing within peaks that also contain an NRFE, we first determined the (S/N) value for all 605,166 0-nt to 6-nt variant ERE DNA elements: ER DNA-binding at the 0-nt variant consensus palindromic ERE DNA element =460.17, ER DNA-binding at 1-nt variant ERE DNA elements =113.91, ER DNA-binding at 2-nt variant ERE DNA elements =20.59, ER DNA-binding at 3-nt variant ERE DNA elements =4.28, ER DNA-binding at 4-nt variant ERE DNA elements =1.56, ER DNA-binding at 5-nt variant ERE DNA elements =1.03, ER DNA-binding at 6-nt variant ERE DNA elements =0.91 **(Table S71)**. Next, we calculated the (S/N) value for all 605,166 0-nt to 6-nt variant ERE DNA elements, while excluding peaks that have already been assigned to an ERE DNA element with less variants. Thus, peak assignment is given to the 0-nt variant consensus palindromic ERE DNA element or the ERE DNA element with the least number of variants relative to the 0-nt variant consensus palindromic ERE DNA element (i.e., defining each peak by a single DNA element with the total number of peaks in an experiment =100%). Of 76,163 Peaks, 304 peaks contain a 0-nt variant consensus palindromic ERE DNA element ((S/N) value =115290.44), 2,705 peaks contain a 1-nt variant ERE DNA element ((S/N) value =3204.99), 7,840 peaks contain a 2-nt variant ERE DNA element ((S/N) value =198.09), 13,922 peaks contain a 3-nt variant ERE DNA element ((S/N) value =21.35), 21,585 peaks contain a 4-nt variant ERE DNA element ((S/N) value =3.78), 24,479 peaks contain a 5-nt variant ERE DNA element ((S/N) value =1.25), 5,308 peaks contain a 6-nt variant ERE DNA element ((S/N) value =0.96) **(Table S71)**. Next, we calculated the (S/N) value for all the other ERE DNA elements found within these groups **(Table S71)**. For example, of the 304 peaks that contain a 0-nt variant consensus palindromic ERE DNA element ((S/N) value =115290.44), 2 of them also contain a 1-nt variant ERE DNA element ((S/N) value =21.00), 2 of them also contain a 2-nt variant ERE DNA element ((S/N) value =1.27), 28 of them also contain a 3-nt variant ERE DNA element ((S/N) value =1.78), 120 of them also contain a 4-nt variant ERE DNA element ((S/N) value =1.43), 241 of them also contain a 5-nt variant ERE DNA element ((S/N) value =0.97), and all 304 of them also contain a 6-nt variant ERE DNA element ((S/N) value =0.92) **(Table S71)**. The large (S/N) value for the ERE DNA element with the least number of variants, while the presence of other ERE DNA elements within those peaks are not enriched over what would be expected at random occurrence, was also observed in the 1-nt to 6-nt variant DNA element groups **(Table S71)**. This confirms that the DNA-binding structure observed at non-NRFE DNA elements in the genome is independent of the binding peaks that contain an NRFE DNA element.

(S/N) Value 15-nt ERE	Peaks Assigned to the ERE DNA Element with the Least Number of Variants						
	0-nt Variant ERE	1-nt Variant ERE	2-nt Variant ERE	3-nt Variant ERE	4-nt Variant ERE	5-nt Variant ERE	6-nt Variant ERE
Other ERE DNA Elements Also Within Peaks	**304 Peaks**	**2,705 Peaks**	**7,840 Peaks**	**13,922 Peaks**	**21,585 Peaks**	**24,479 Peaks**	**5,308 Peaks**
0-nt Variant ERE	**115290.44**						
1-nt Variant ERE	21.00	**3204.99**					
2-nt Variant ERE	1.27	5.51	**198.09**				
3-nt Variant ERE	1.78	2.29	2.86	**21.35**			
4-nt Variant ERE	1.43	1.30	1.40	1.58	**3.78**		
5-nt Variant ERE	0.97	0.99	1.00	1.04	1.05	**1.25**	
6-nt Variant ERE	0.92	0.88	0.89	0.89	0.90	0.92	**0.96**

Mouse (mm10) and Human (hg19) Genome Assembly

The mouse and human genome assemblies were obtained from UCSC Genome Browser: mouse genome build 38 (mm10) and human genome build 37 (hg19). The genome size of the mouse (mm10) genome is 2,730,871,774 nucleotides and the genome size of the human (hg19) genome is 3,137,161,264 nucleotides **(Table S72)**.

Their effective sizes are 2,647,537,730 nucleotides (mm10) and 2,861,343,702 nucleotides (hg19) when excluding unreadable nucleotides, random regions, hap regions, and chromosome Un (i.e., clone contigs that cannot be confidently placed on a specific chromosome) from the reference **(Table S73)**.

Unreadable nucleotides account for 2.9% [77,999,939] (mm10) and 7.5% [234,350,281] (hg19) of the genome **(Table S73)**. These unreadable nucleotides are contained within 501 (mm10) and 339 (hg19) unreadable regions in the genome **(Table S74)**.

See **Table S75** for associated script codes.

Population Count of 1-nt to 20-nt DNA Elements in the Single-Stranded Mouse and Human Genome

Perl was used to obtain the counts of all 1-nt to 20-nt DNA elements in the single-stranded mouse and human genome **(Table S47-S50)**.

The equivalence was determined using the correlation coefficient, slope and y-intercept. Secondly, we took the measure of the variance of the points normalized by the overall length of the vector (i.e., sets of points), named the general error fraction (GEF). That is, the distance between the points normalized with respect to the overall length of the vector.

The population count of every 1-nt to 20-nt DNA elements in the single-stranded mouse and human genome, including its reverse-complement DNA element, reverse DNA element, and complement DNA element, have been provided (tar -xvzf *filename*.tar.gz to unzip files) **(Datafiles S2-S5)**.

See **Table S75** for associated script codes.

<u>Global Error Fraction</u>

Define:

$$E_K = \left|\frac{1}{2}\left(X_K - Y_K\right)\right|$$

$$A_K = \left|\frac{1}{2}\left(X_K + Y_K\right)\right|$$

$$Where\ X_K = DNA\ Element\ Frequency\ of\ X$$

$$Where\ Y_K = DNA\ Element\ Frequency\ of\ Y$$

$$\langle {E_K}^2 \rangle = \frac{1}{M}\sum_{K=1}^{M} {E_K}^2$$

$$\langle {A_K}^2 \rangle = \frac{1}{M}\sum_{K=1}^{M} {A_K}^2$$

$$\text{Global Error Fraction (GEF)} = \left[\frac{\langle {E_K}^2 \rangle}{\langle {A_K}^2 \rangle}\right]^{1/2}$$

Position Weight Matrices (PWMs) for DNA Motif Identification

Finding enriched DNA motifs in genomic regions (findMotifsGenome.pl)

The findMotifsGenome.pl program is a wrapper that helps set up the data for analysis using the HOMER motif discovery algorithm.

See **Table S75** for associated script codes.

Supplemental References

1. Aerts S. Chapter five - Computational Strategies for the Genome-Wide Identification of cis-Regulatory Elements and Transcriptional Targets. In: Plaza S, Payre F, editors. Current Topics in Developmental Biology: Academic Press; 2012. p. 121-45.

2. Coons LA, Hewitt SC, Burkholder AB, McDonnell DP, Korach KS. DNA Sequence Constraints Define Functionally Active Steroid Nuclear Receptor Binding Sites in Chromatin. Endocrinology. 2017;158(10):3212-34. Epub 2017/10/05. doi: 10.1210/en.2017-00468. PubMed PMID: 28977594; PMCID: Pmc5659708.

3. Tonelli C, Morelli MJ, Sabo A, Verrecchia A, Rotta L, Capra T, Bianchi S, Campaner S, Amati B. Genome-wide analysis of p53-regulated transcription in Myc-driven lymphomas. Oncogene. 2017;36(21):2921-9. Epub 2017/01/17. doi: 10.1038/onc.2016.443. PubMed PMID: 28092679; PMCID: PMC5454316.

4. Sahu B, Pihlajamaa P, Dubois V, Kerkhofs S, Claessens F, Janne OA. Androgen receptor uses relaxed response element stringency for selective chromatin binding and transcriptional regulation in vivo. Nucleic acids research. 2014;42(7):4230-40. Epub 2014/01/25. doi: 10.1093/nar/gkt1401. PubMed PMID: 24459135; PMCID: Pmc3985627.

5. Kharchenko PV, Tolstorukov MY, Park PJ. Design and analysis of ChIP-seq experiments for DNA-binding proteins. Nat Biotech. 2008;26(12):1351-9.

6. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nature methods. 2012;9(4):357-9. Epub 2012/03/06. doi: 10.1038/nmeth.1923. PubMed PMID: 22388286; PMCID: Pmc3322381.