Supplemental Experimental Procedures.

ChIP.

RA treated EBs were directly dounced in 1XPBS with 0,4% formaldehyde for 10 min at room temperature and the reaction was stopped by adding glycine to final concentration 0.2M for 10 minutes at room temperature. Fixed cells were rinsed twice with PBS and resuspended in lysis buffer (10 mM EDTA pH 8, 50 mM Tris-HCl pH 8, SDS 1 %). Lysate was sonicated 30 min (30 sec on / 30 sec off) in a Diagenode water bath-sonicator and centrifuged at 14000 rpm for 10 min. The cleared supernatant was used immediately in ChIP experiments or stored at -80°C.

50-150 ug of sonicated chromatin was diluted 10 times in ChIP Dilution Buffer (SDS 0.01%, Triton X-100 1.1%, 1.2 mM EDTA pH 8, 16.7 mM Tris-HCl pH 8 and 167 mM NaCl) and pre-cleared for 1 hour, rotating at 4°C, with 50 ul blocked beads (Protein G Sepharose (PGS) 50% slurry incubated for 4 hours minimum with BSA 0.5 mg/mL and yeast tRNA 0.2 mg/mL) before the overnight incubation with 2-5ug of specific antibody. The bound material was recovered after a 2 hours incubation, rotating at 4°C, with 30 ul blocked PGS. The beads were washed, for 5 minutes, once in Low Salt Buffer (SDS 0.1%, Triton X-100 1%, 2 mM EDTA pH 8, 20 mM Tris-HCl pH 8 and 150 mM NaCl), twice in High Salt Buffer (SDS 0.1%, Triton X-100 1%, 2 mM EDTA pH 8, 20 mM Tris-HCl pH 8 and 500 mM NaCl), twice in LiCl Buffer (0.25 M LiCl, NP-40 1%, Na Deoxycholate 1%, 1 mM EDTA pH 8 and 10 mM Tris-HCl pH 8) and twice in TE. ChIPed material was eluted by two 15 minute incubations at room temperature with 250 ul Elution Buffer (SDS 1%, 0.1 M NaHCO3). Chromatin was reverse-crosslinked by adding 20 ul of NaCl 5M and incubated at 65°C for 4 hours minimum and DNA was submitted to RNase and proteinase K digestion and extracted by phenol-chloroform.

RNA-seq.

The mRNA-seq libraries were prepared following the Illumina protocol with some modifications. Briefly, mRNA was purified from total RNA using oligo-dT magnetic beads and fragmented using divalent cations at 95°C for 5 minutes. The cleaved mRNA fragments were reverse transcribed to cDNA using random primers. This was followed by second strand cDNA synthesis using Polymerase I and RNase H. The double strand cDNA fragments were blunted, phosphorylated and ligated to single-end adapter dimers follwed by PCR amplification (30 sec at 98°C; [10 sec at 98°C, 30 sec at 65°C, 30 sec at 72°C] x 13 cycles; 5 min at 72°C). After PCR amplification, surplus PCR primers and dimer adapters were removed by purification using AMPure beads (Agencourt Biosciences Corporation). Size selection was performed by electrophoresis on a 2% agarose gel and DNA fragments in the range of ~250-350bp were excised and purified using QIAquick Gel Extraction Kit (Qiagen). DNA libraries were checked for quality and quantified using 2100 Bioanalyzer (Agilent). The libraries were loaded in the flowcell at 6pM concentration and clusters generated and sequenced on the Illumina Genome Analyzer IIx as single-end 54 base reads. Image analysis and base calling were performed using the Illumina Pipeline version 1.6.

Legends to Supplemental Figures and Tables.

Supplemental Figure 1. Schematic representation of the distribution of RAR-occupied site locations with respect to gene annotation. The % of sites at each defined class of locations are indicated.

Supplemental Figure 2. RAR genomic occupancy in EBs and F9 cells. **A.** Comparative clustering of the EB and F9 cell 2 hour RAR ChIP-seq data. Tag densities from each dataset were collected in a window of 10 kb around the 13385 EB RAR-occupied sites. The collected

values were subjected to k-means clustering coupled to ranked based-normalization. Examples of read density plots for several classes showing comparable or preferential occupation in either data set are shown to the right of the figure. **B.** Venn diagramme comparison of RAR-occupied genes in embryoid bodies and F9 cells. The total number of annotated Ensembl/Refseq target transcripts/genes in each cell type is indicated along with those that are shared and specific to each data set.

Supplemental Figure 3. Binding of RAR-RXR to IR0 elements. **A**. UCSC view of sequence tag density in .wig file format of the RAR-occupied sites at the *Mbd6* and *4930452B06Rik* genes comprising an IR0 in EBs. **B**. EMSA competition analysis of the indicated IR0 elements. The sequences of the IR0 motifs within the competing oligonucleotides are shown with the inverted half sites indicated by arrows. Mutated nucleotides are indicated in red. All competitors were used a 100-fold excess.

Supplemental Figure 4. A. Examples of simple and composite DR8 elements that are occupied by RAR at the indicated gene loci in embryoid bodies. The frequency of each type of DR8 in the top 1000 sites in EBs and F9 cells is indicated. **B**. UCSC view of sequence tag density in .wig file format of the RAR-occupied sites at the *Mafa* and *Cd97* genes comprising a composite DR8 in embryoid bodies. **C**. EMSA competition analysis of the various indicated IR0 elements to compete for RAR/RXR complex formation. The sequences of the DR8 motifs within the competing oligonucleotides are shown with the half sites indicated by arrows. Mutated nucleotides are indicated in red. All competitors were used a 100-fold excess.

Supplemental Figure 5. EMSA competition analysis of simple and composite DR8. **A**. Competition with a simple DR8 from the *Dedd* gene locus. Competition was performed with increasing quantities (10, 25, 50, 100-fold excess) of the oligonucleotides shown above each lane. Lane 1 is the oligonucleotide probe with no recombinant RAR-RXR and lane 2 with RAR-RXR, but no competitor. The wild type and mutated oligonucleotides are shown below

the EMSA panel, and mutations are indicated in red. **B.** Competition with a DR8 comprising a variant 5'RGATCA-3' half site from the Neuropilin 1 (*Nrp1*) gene. The sequences of the *Nrp1* DR8 within the competing oligonucleotides are shown with the half sites indicated by arrows. Mutated nucleotides are indicated in red. **C.** Binding of RAR-RXR to different half sites spacings. The labelled oligonucleotides used are indicated above each lane along with the presence or absence of recombinant RAR-RXR.

Supplemental Figure 6. List of potential pseudo DR0 elements. The conserved G residues at position 2 of the half sites are boxed. Those that do not conserve both of these residues are unlikely to constitute RAR/RXR binding elements as exemplified by the *Tspan9* and *Wsb2* elements in Figure S7.

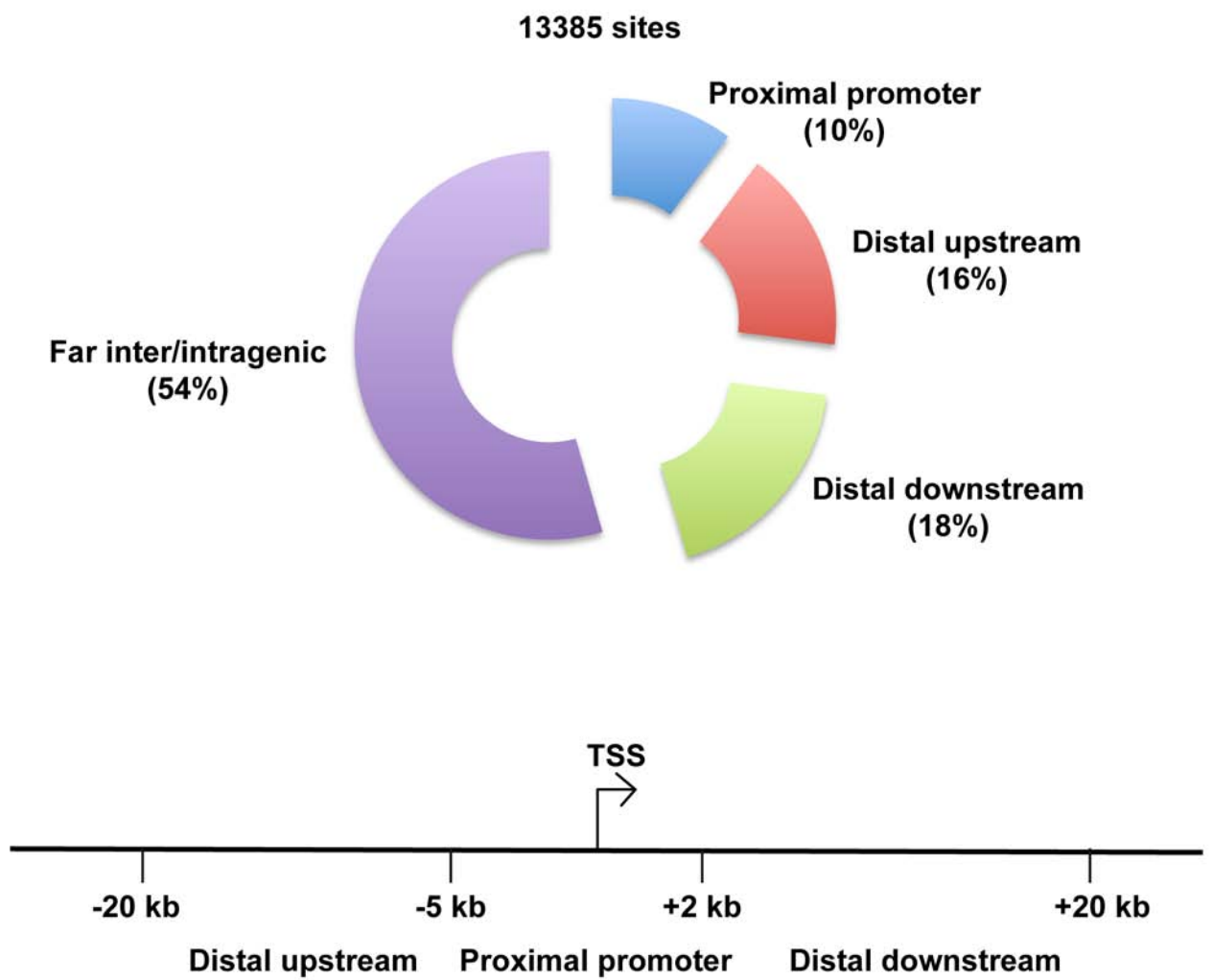
Supplemental Figure 7. DR0/8 are embedded in a longer sequence element. **A**. Results of the MEME and SPAMO analysis are summarised. The consensus sequences for each of the identified motifs are shown along with their frequency, location and distance relative to the 200 top DR0 and 186 top DR8 elements in the EB data set. **B**. Results of TOMTOM analysis showing the potential similarity of one of the motifs to the consensus PITX2 binding sites in the Transfac database. **C**. Schematic representation of the locations of the identified motifs relative to the DR0/8 element.

Supplemental Figure 8. Model of RAR-RXR binding to IR0 and DR0. **A.** Architecture of RAR/RXR/DR5 complex in solution obtained by small angle X-ray scattering refinement. **B** 3D model of RAR/RXR/IR0 complex based on ecdysone receptor heterodimer bound to an IR1. The region involving a steric clash bewteen the two DBDs is circled. **C.** 3D model of RAR/RXR/DR0 complex.

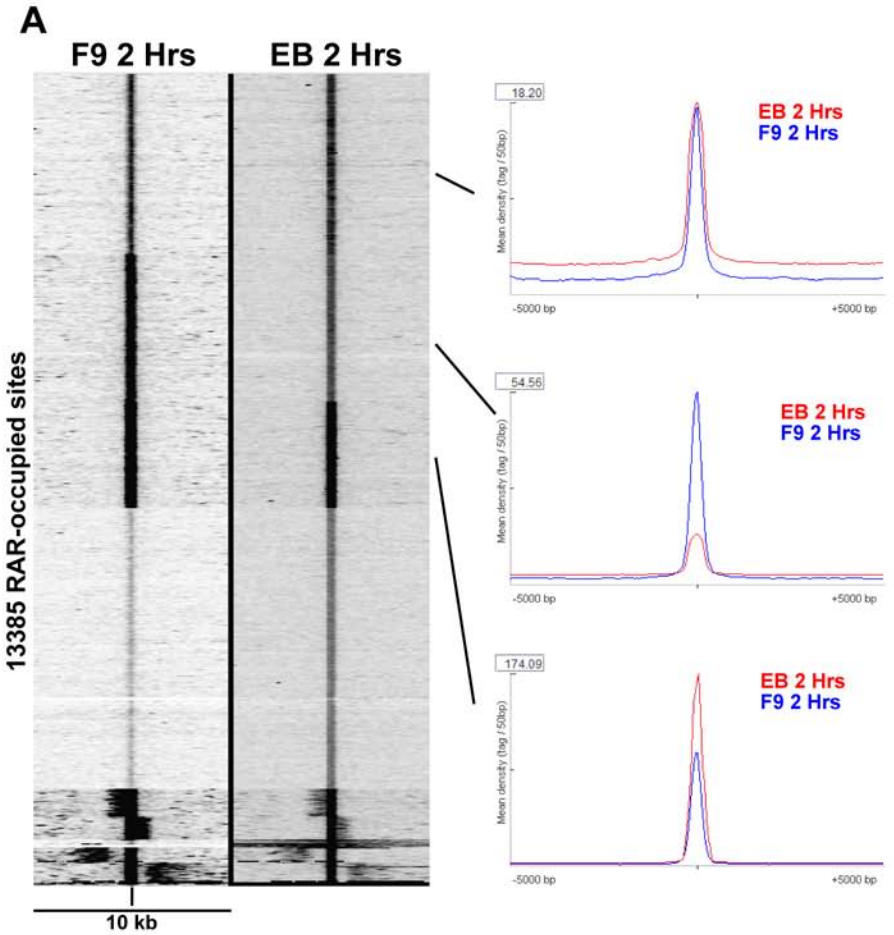
Supplemental Table 1. ChIP-seq identification of RAR-occupied loci in EB after 2 hours of RA-treatment. Excel table showing the output from the MACS-GPAT analysis. Page 1 shows the results of the peak detection and annotation. The raw sequence data was analysed first by

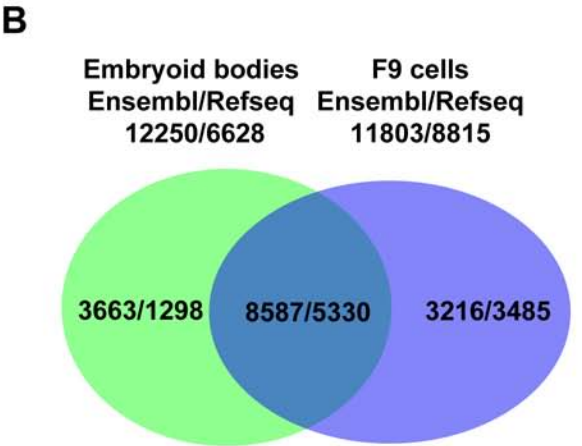
the MACs programme to detect peaks. The results of this analysis are indicated in columns A-I. The regions with significant peaks were then annotated using the GPAT programme to detect nearby annotated transcripts. The results of annotation are shown in the remaining columns. For all peaks chromosome location, length, number of reads, pvalue, and local fold enrichment are indicated. GPAT annotation was performed using a window of +/- 20kb with respect to the Ensembl gene coordinates. The identity of the gene closest to the peak and of any secondary genes located within the 20kB window are indicated along with the distance of the peak to the TSS. Page 2 shows genes with an RAR-occupied site.

Supplemental Table 2. RA-regulated genes in EBs. Pages 1 and 2 of the table show the Ensembl IDs, gene names and descriptions for transcripts that are induced or repressed by RA in EBs, respectively. The normalised expression values for the duplicate samples are indicated along with the mean value, the Log2 change in expression and adjusted pvalue.

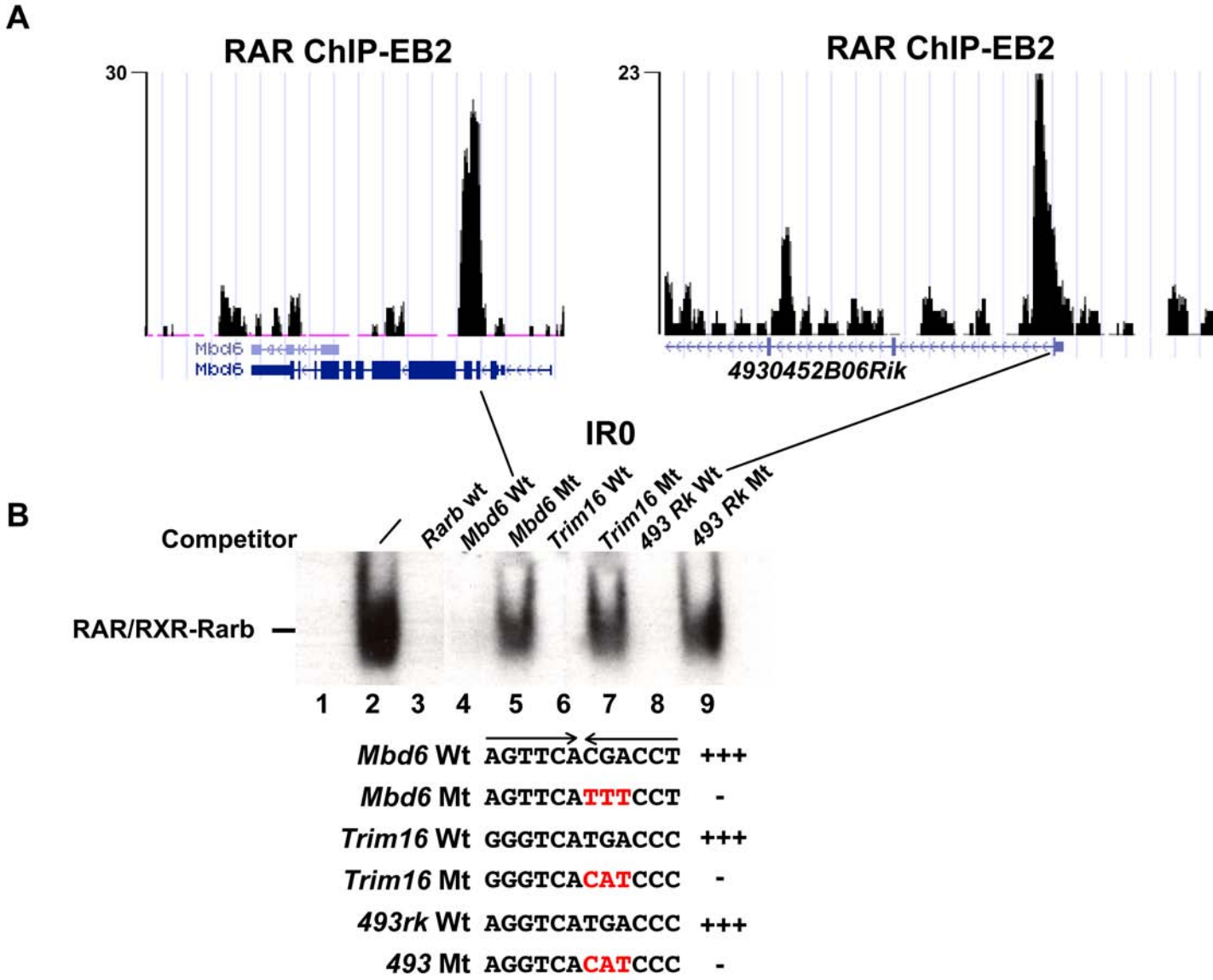


Moutier et al., Supplemental Fig. 1





Moutier et al., Supplemental Fig. 2



Moutier et al., Supplemental Fig. 3

A

Simple DR8

Rqcd1 GGGTCAgaggtgagAGGTCA

Dedd AGGTCAcgatctggAGTTCA

Frequency in top 1000 ChIP-seq
EB/F9
10/11

Composite DR8

DR8

DR2 DR0

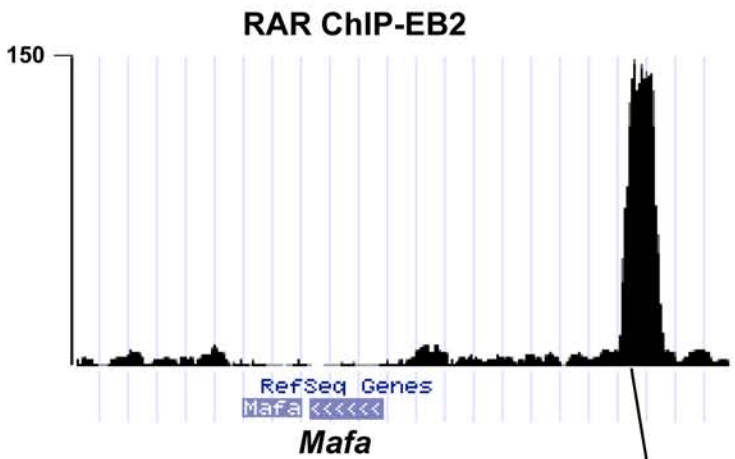
Spc25 GGGTCAcaAGTTCAAGGTCA

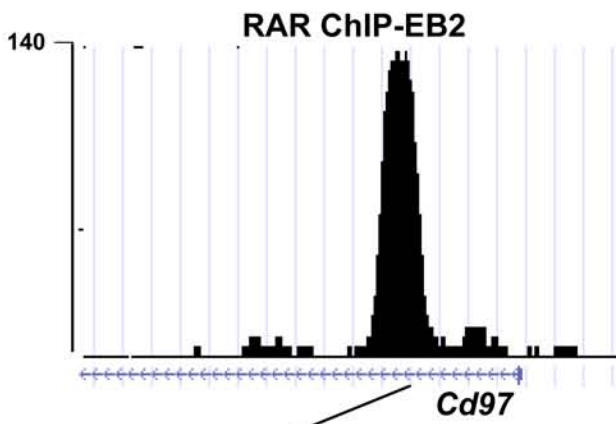
Spint4 AGGTCAgaAGTTCAAGGTCA

Frequency in top 1000 ChIP-seq
EB/F9
183/63

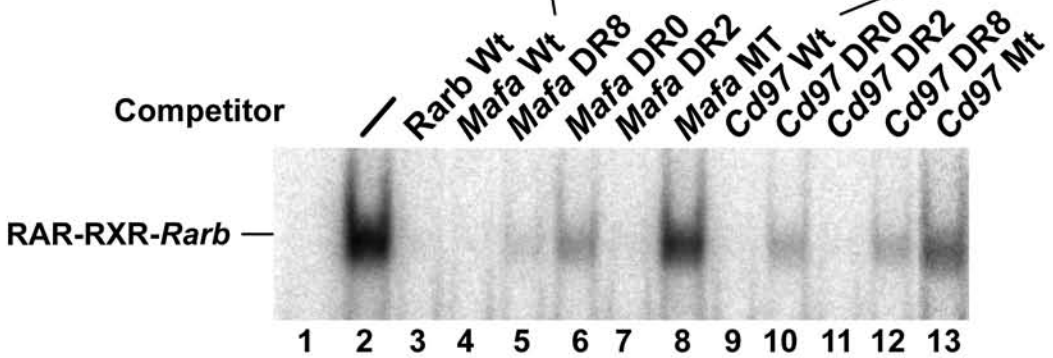
DR2 DR0

B





C



Composite DR8

DR8

DR2 DR0

Mafa Wt	AGGTCAgaAGTTCAAGGTCA	+++
Mafa DR8	AGGTCAgaAGACTAAGGTCA	+++
Mafa DR0	AGACTAgaAGTTCAAGGTCA	++
Mafa DR2	AGGTCAgaAGTTCAAGACTA	+++
Mafa Mt	AGACTAgaAGTTCAAGACTA	-

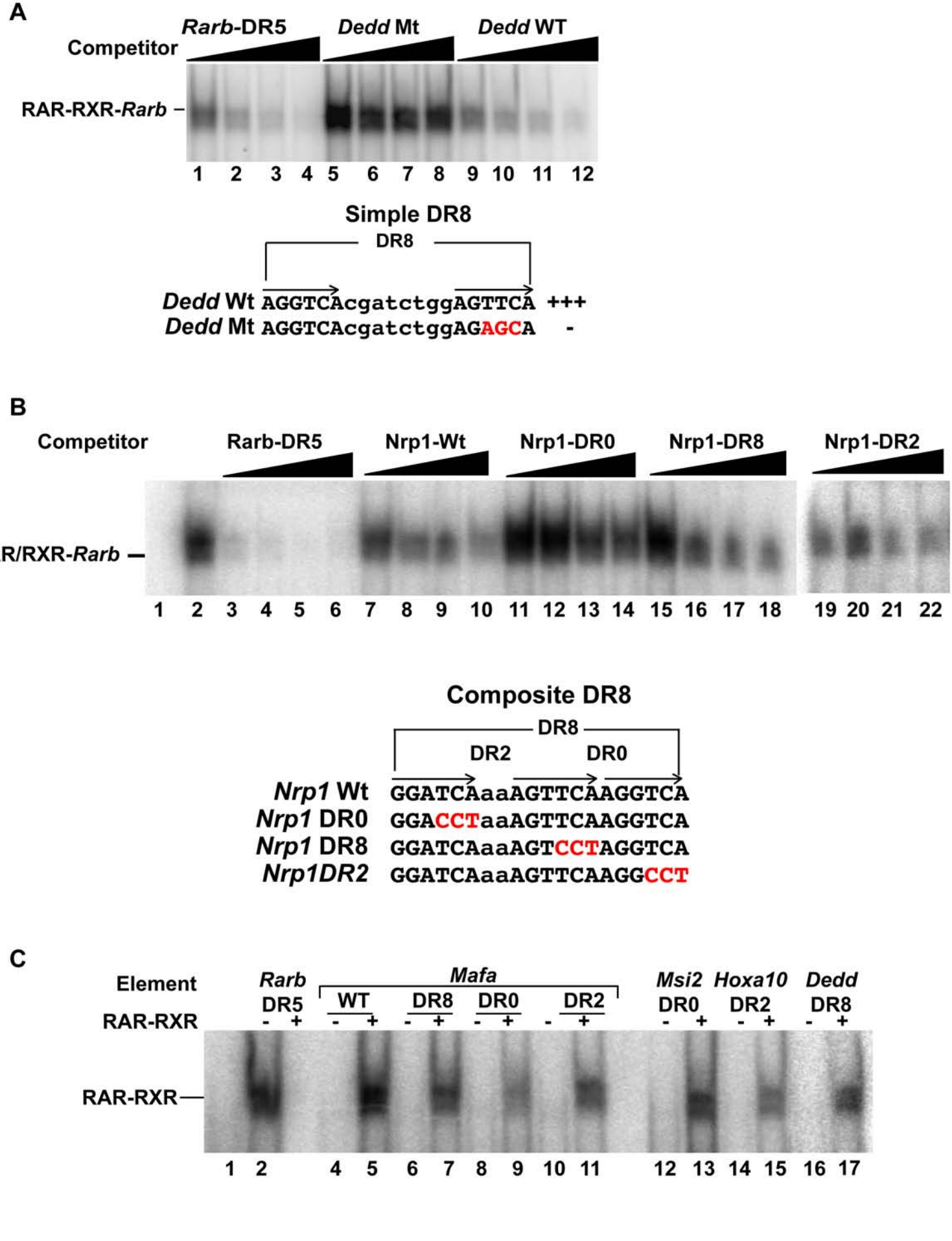
Composite DR8

DR8

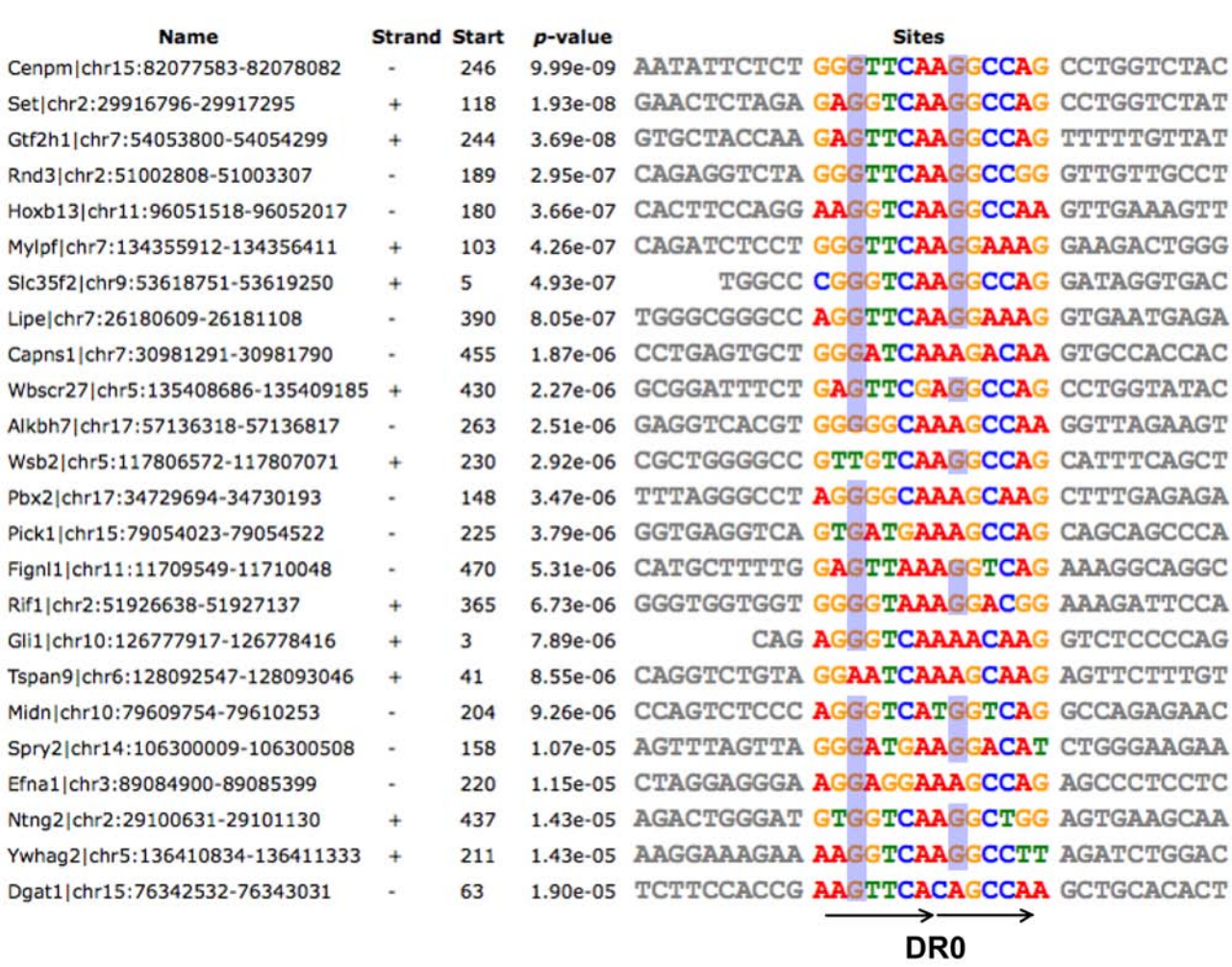
DR2 DR0

Cd97 Wt	GGGTCAaaAGTTCAAGGTCA	+++
Cd97 DR8	GGGTCAaaAGACTAAGGTCA	++
Cd97 DR0	GGACTAaaAGTTCAAGGTCA	++
Cd97 DR2	GGGTCAaaAGTTCAAGACTA	+++
Cd97 MT	GGACTAaaAGTTCAAGACTA	+

Moutier et al., Supplemental Fig. 4



Moutier et al., Supplemental Fig. 5

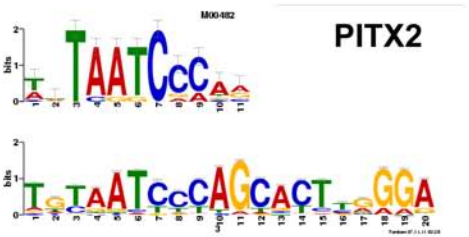


Moutier et al., Supplemental Fig. 6

A

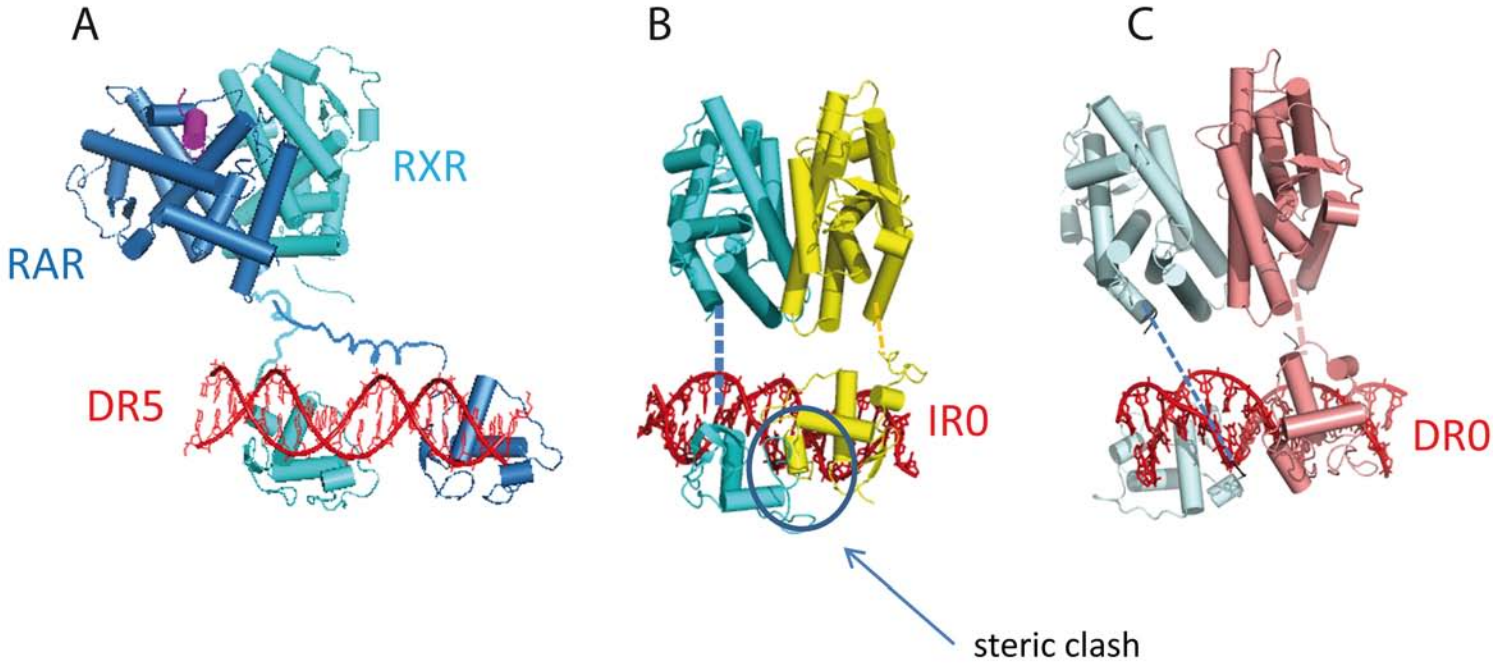
	Motif	# Motifs /200 DR0	# Motifs /186 DR8	Distance
		200	186	NA
1		126	109	5-15 bp 5’
2		120	114	19-21 bp 5’
3		98	56	20-22 bp 5’
4		48	113	0 bp 3’
5		127	132	15-17 bp 3’
6		57	NA	25-30 bp 3’

B



C

Moutier et al., Supplemental Fig. 7



Moutier et al., Supplemental Fig. 8