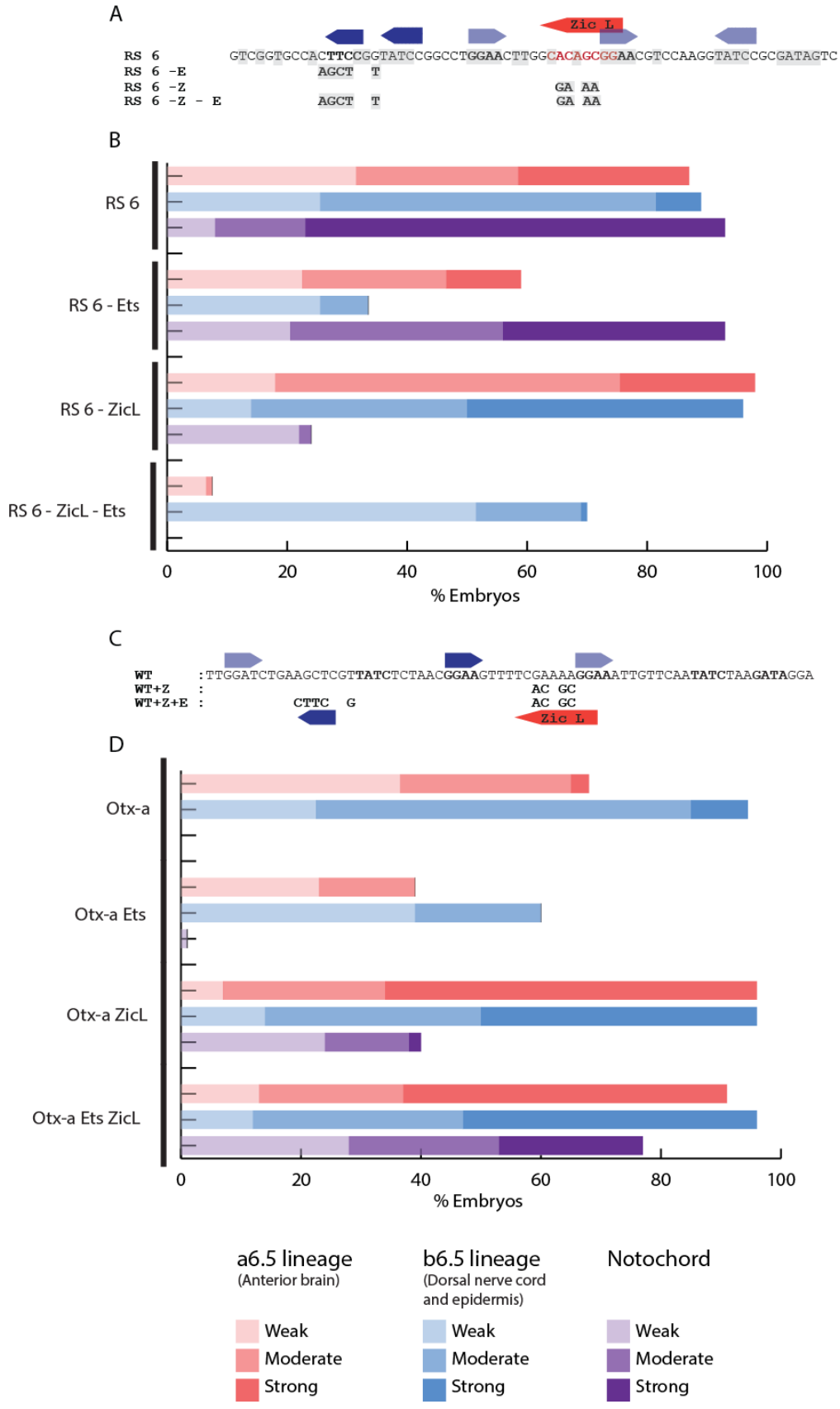
Supporting Information for

Enhancer syntax compensates for poor binding sites to encode tissue specificity of developmental enhancers

This file includes:


A
Zic L
RS 6 GTCGGTGCCACTTCCGGTATCCGGCCTGGAACTTGGCACAGCGGAACGTCCAAGGTATCCGCGATAGTC
RS 6 -E AGCT T
RS 6 -Z GA AA
RS 6 -Z - E AGCT T GA AA
B
RS 6
RS 6 - Ets
RS 6 - ZicL
RS 6 - ZicL - Ets
0 20 40 60 80 100
% Embryos
C
WT :TTGGATCTGAAGCTCGTTATCTCTAACGGAAGTTTTCGAAAAGGAAATTGTTCAATATCTAAGATAGGA
WT+Z : AC GC
WT+Z+E : CTTC G AC GC
Zic L
D
Otx-a
Otx-a Ets
Otx-a ZicL
Otx-a Ets ZicL
0 20 40 60 80 100
% Embryos
a6.5 lineage
(Anterior brain)
Weak
Moderate
Strong
b6.5 lineage
(Dorsal nerve cord
and epidermis)
Weak
Moderate
Strong
Notochord
Weak
Moderate
Strong

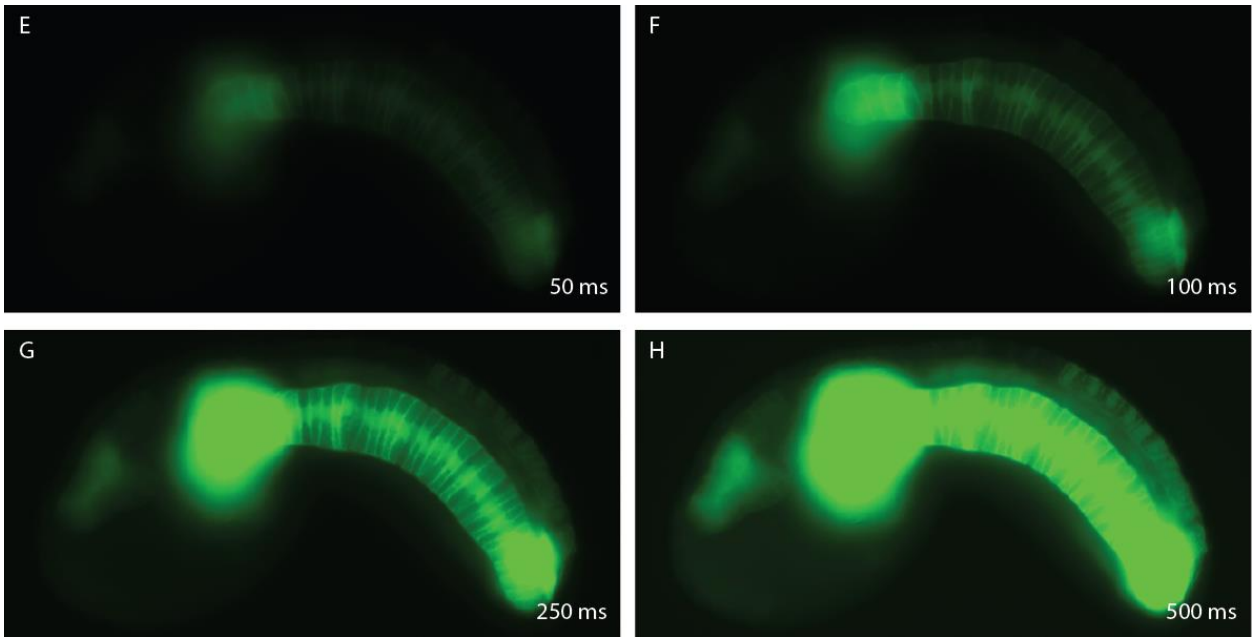


Figure S1. ZicL and ETS mediate notochord expression. Sequences and counting of RS 6 manipulations and WT Otx-a manipulations.

(A) Sequences of RS 6 and manipulations. Dark blue arrows refer to ETS binding sites with a relative binding affinity ≥0.60. Light blue arrows refer to ETS bind sites with a binding affinity < 0.60. Grey boxes show sequence conservation with WT Otx-a. **(B)** Scoring for embryos electroporated with RS 6, RS 6 – ETS, RS 6 – ZicL, and RS 6 – ZicL – ETS, as pictured in Figure 1 (n=100 embryos for each construct). a6.5 and b6.5 expression is also seen as RS6 is a synthetic variant of the Otx-a enhancer. It therefore contains GATA and ETS binding sites. **(C)** Sequences of Otx-a, Otx-a ETS, Otx-a ZicL, and Otx-a ZicL ETS. **(D)** Scoring for embryos electroporated with Otx-a, Otx-a ETS, Otx-a ZicL, and Otx-a ZicL ETS, as pictured in Figure 1 (n=100 embryos for each construct). Weak, moderate and strong expression definitions are based on exposure time required for 10% saturated pixels. Strong <500ms, moderate <800ms and weak >800ms. **(E, F, G, H**) Fig. 1A is overexposed as the same exposure time is shown for all constructs in each figure to allow for comparison. Here images of RS6 at different exposure times 50ms, 100ms, 250ms and 500ms respectively are shown.

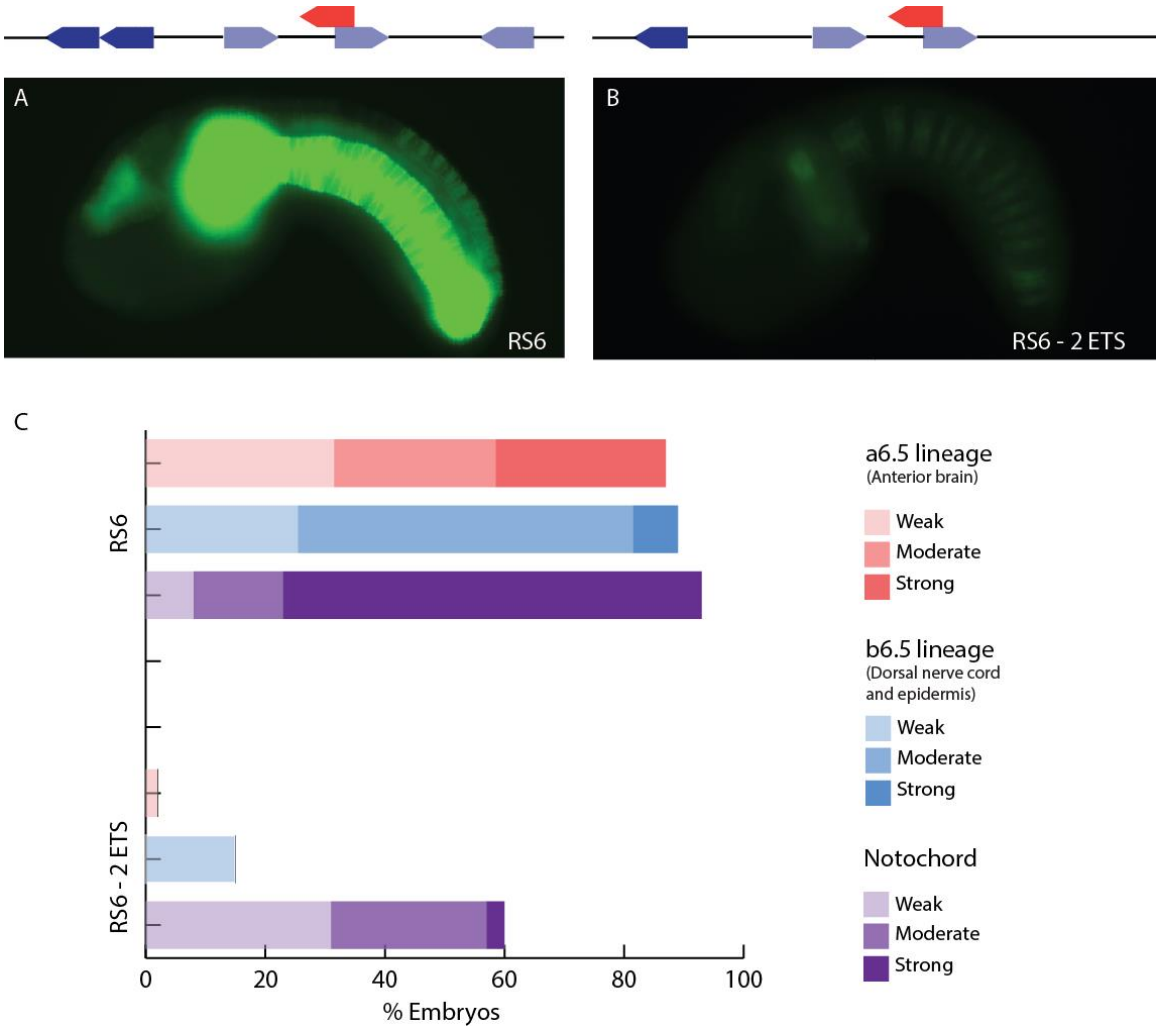


Figure S2. Extra ETS in RS6 explain elevated levels of notochord expression in RS6. Images and counting of RS 6 and RS 6 – 2 ETS.

(A) Image of RS 6 with schematic of sequence above showing five ETS binding sites and one ZicL binding site. Dark blue arrows refer to ETS binding sites with a relative binding affinity ≥0.60, and light blue arrows refer to ETS bind sites with a relative binding affinity <0.60. Expression is seen strongly in the notochord and mesenchyme and moderately in the dorsal epidermis and anterior brain. **(B)** Image of RS 6 – 2 ETS with schematic of sequence above showing three ETS binding sites and one ZicL binding site. Expression is seen only weakly in the notochord. **(C)** Scoring for embryos electroporated with RS 6, RS 6 – 2 ETS (n=100 embryos for each construct). Weak, moderate and strong expression definitions are based on exposure time required for 10% saturated pixels. Strong <500ms, moderate <800ms and weak >800ms.

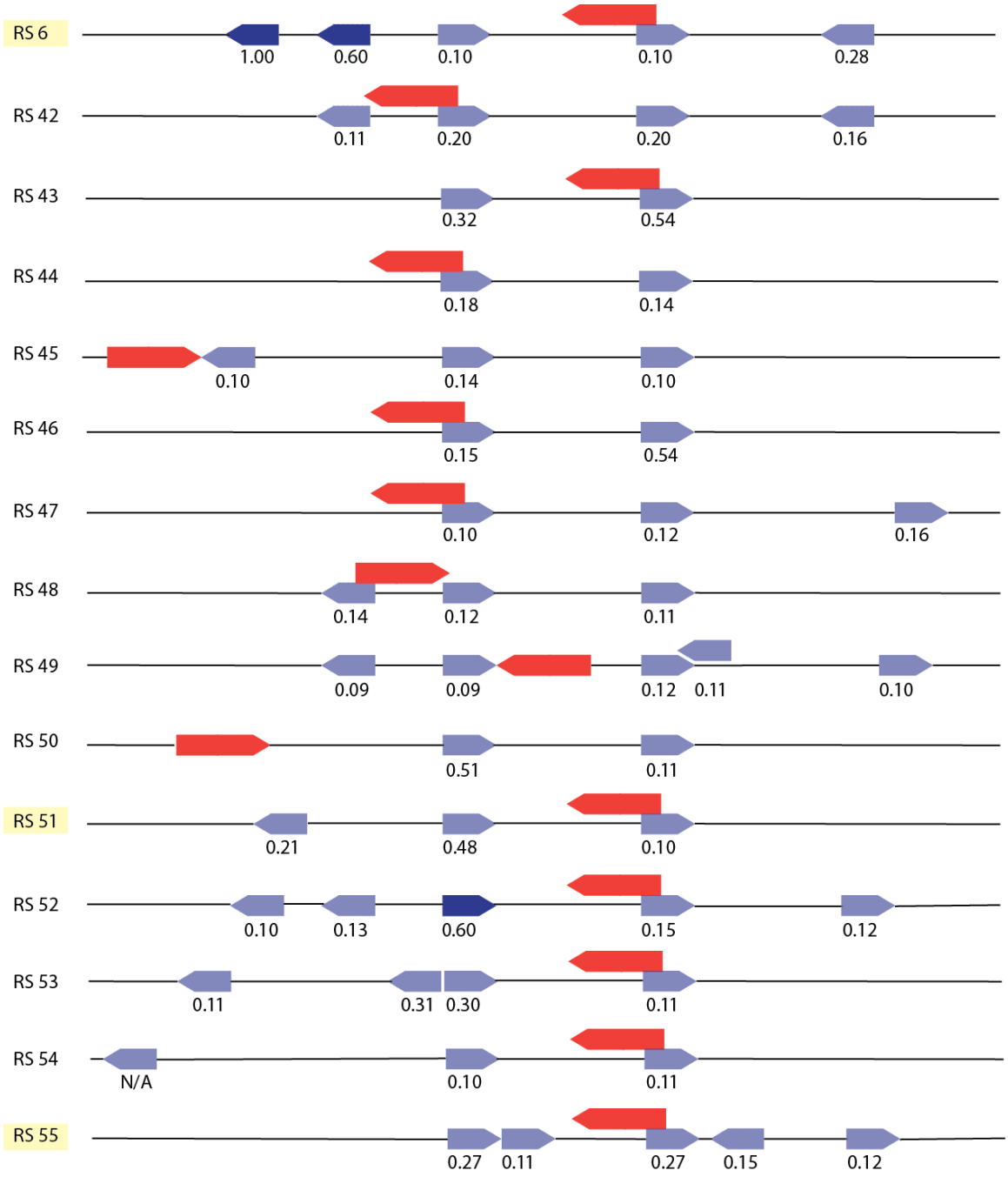


Figure S3. Schematics of Randomized synthetic (RS) enhancer variants with ETS and ZicL sites. Sequence names highlighted in yellow boxes are variants that activated notochord expression in the notochord. Dark blue arrows refer to ETS binding sites with a relative binding affinity ≥0.60, and light blue arrows refer to ETS bind sites with a relative binding affinity <0.60. Scores shown are for relative affinity. Relative affinities were calculated using median signal intensities of the universal protein binding microarray (PBM) data for mouse ETS-1 (17) proteins from UniProbe database (http://thebrain.bwh.harvard. edu/uniprobe/index.php)

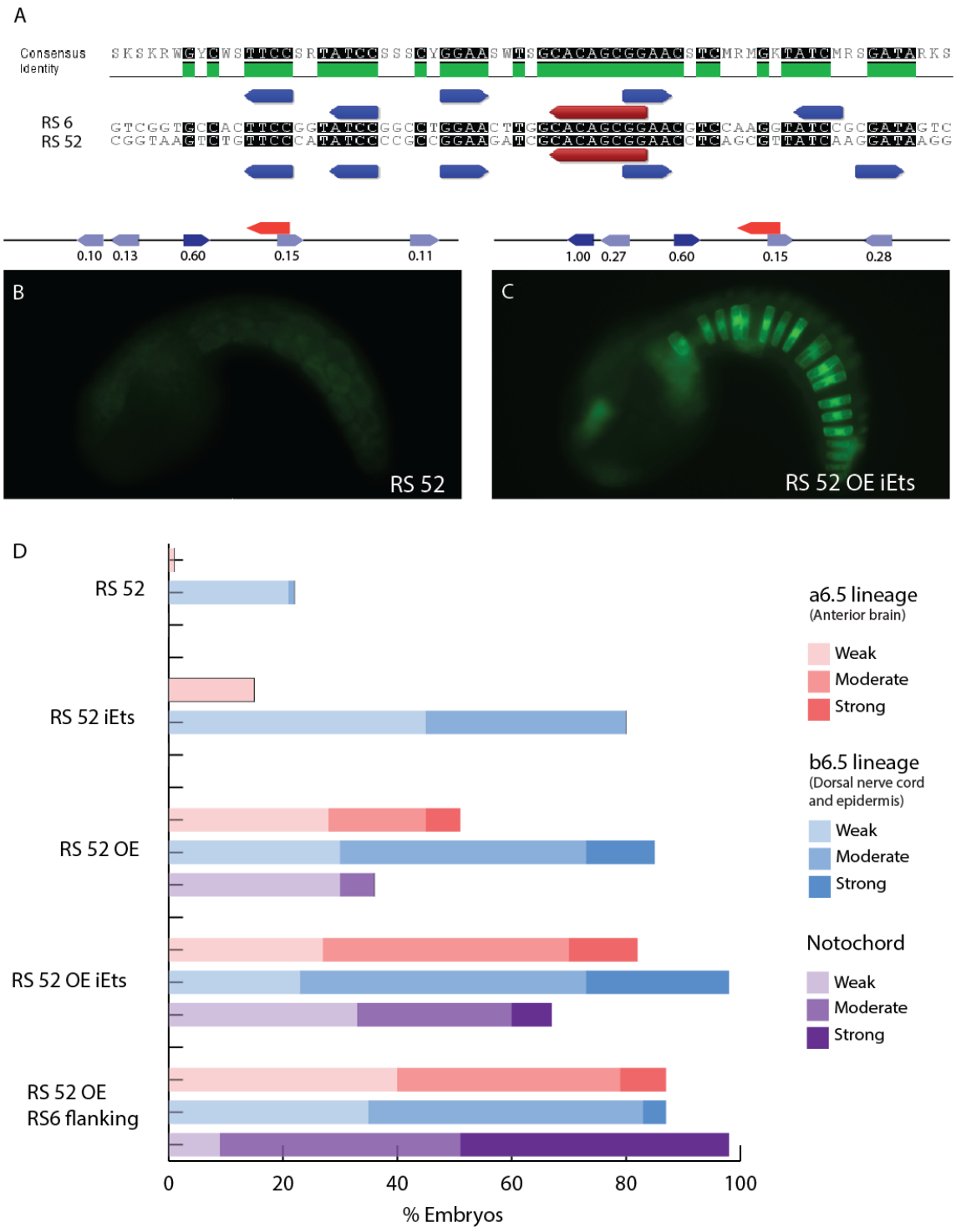


Figure S4. RS 52 manipulations show importance of the quality ETS binding sites in notochord expression.

(A) Sequence comparison of RS 6 and RS 52. There are 40 out of 69 base pairs that are conserved between the two constructs, shown in black. **(B)** Image of RS 52 with schematic of sequence above. Dark blue arrows refer to ETS binding sites with a

binding affinity above 0.60, and light blue arrows refer to ETS bind sites with a binding affinity lower than 0.60. No GFP expression is seen in the notochord. There is some auto fluorescence in the tails muscles. **(C)** Image of RS 52 with dinucleotide changes to match RS6 flanking (RS52 + OE iETS). **(D)** Scoring for embryos electroporated with RS 52, RS 52 OE, RS 52 OE iETS, and RS 52 + RS6 flanking. These constructs show the gradual increase of notochord expression as the binding affinities of ETS are improved to look more like RS 6. (n=100 embryos for each construct) Weak, moderate and strong expression definitions are based on exposure time required for 10% saturated pixels. Strong <500ms, moderate <800ms and weak >800ms.

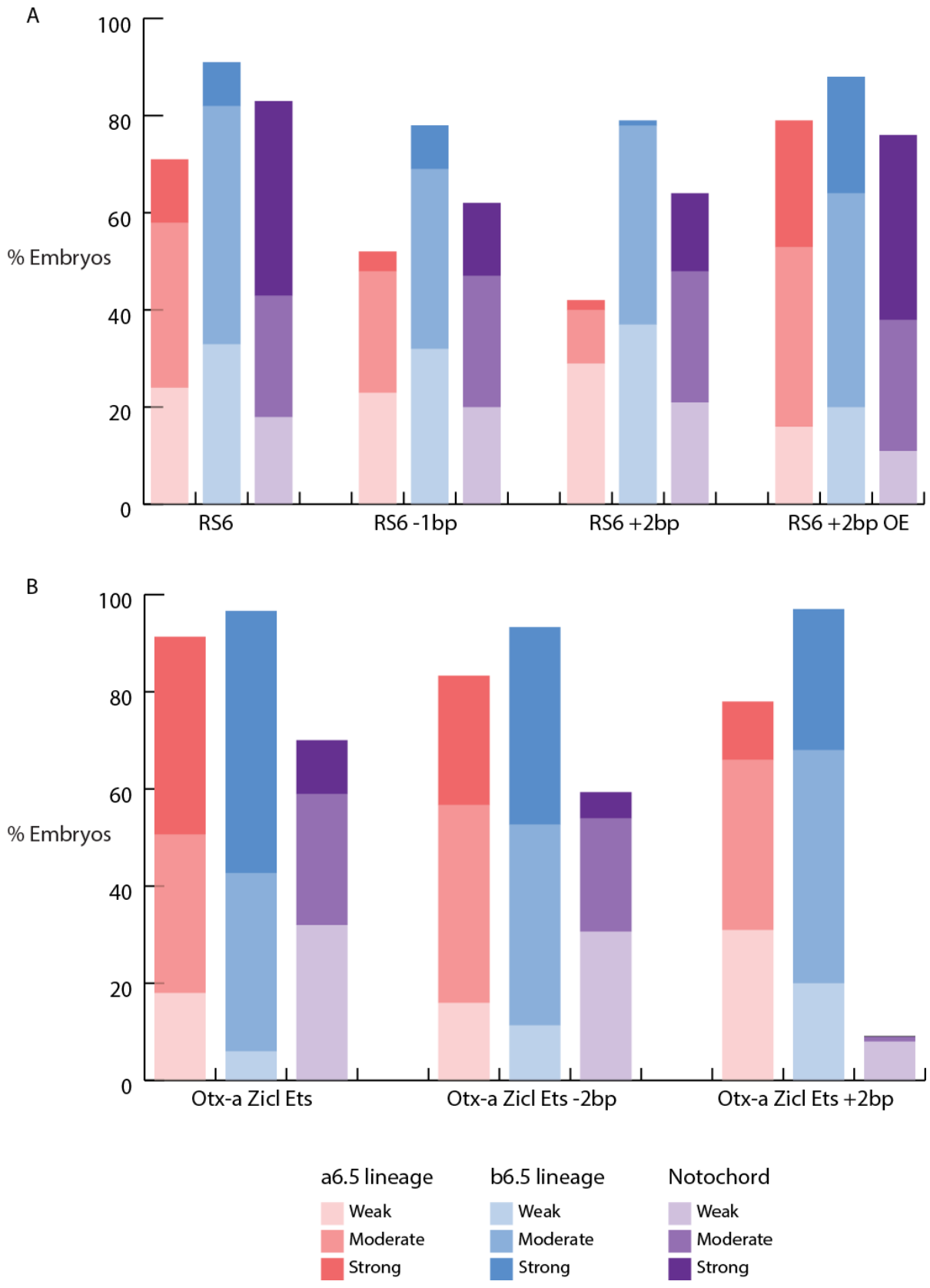

A
100
80
60
40
20
0
% Embryos
RS6
RS6 -1bp
RS6 +2bp
RS6 +2bp OE
B
100
80
60
40
20
0
% Embryos
Otx-a Zicl Ets
Otx-a Zicl Ets -2bp
Otx-a Zicl Ets +2bp
a6.5 lineage
Weak
Moderate
Strong
b6.5 lineage
Weak
Moderate
Strong
Notochord
Weak
Moderate
Strong

Figure S5. Optimal Spacing between ETS and ZicL for notochord expression is 11 base pairs.

(A) Scoring for embryos electroporated with RS 6, RS 6 -1bp, and RS 6 +2bp. RS 6 -1bp and RS 6 +2bp show less notochord expression than RS 6 (n=100 embryos for each construct). RS6+2bp OE shows the similar levels of expression as RS6 **(B)** Scoring for embryos electroporated with Otx-a ZicL ETS, Otx-a ZicL ETS –2bp, and Otx-a ZicL ETS +2bp. Otx-a ZicL ETS –2bp and Otx-a ZicL ETS +2bp show less notochord expression than Otx-a ZicL ETS. (n=100 embryos for each construct). Weak, moderate and strong expression definitions are based on exposure time required for 10% saturated pixels. Strong <500ms, moderate <800ms and weak >800ms. a6.5 and b 6.5 lineage expression is seen as both of these constructs are Otx-a variants and so contain GATA binding sites as well.

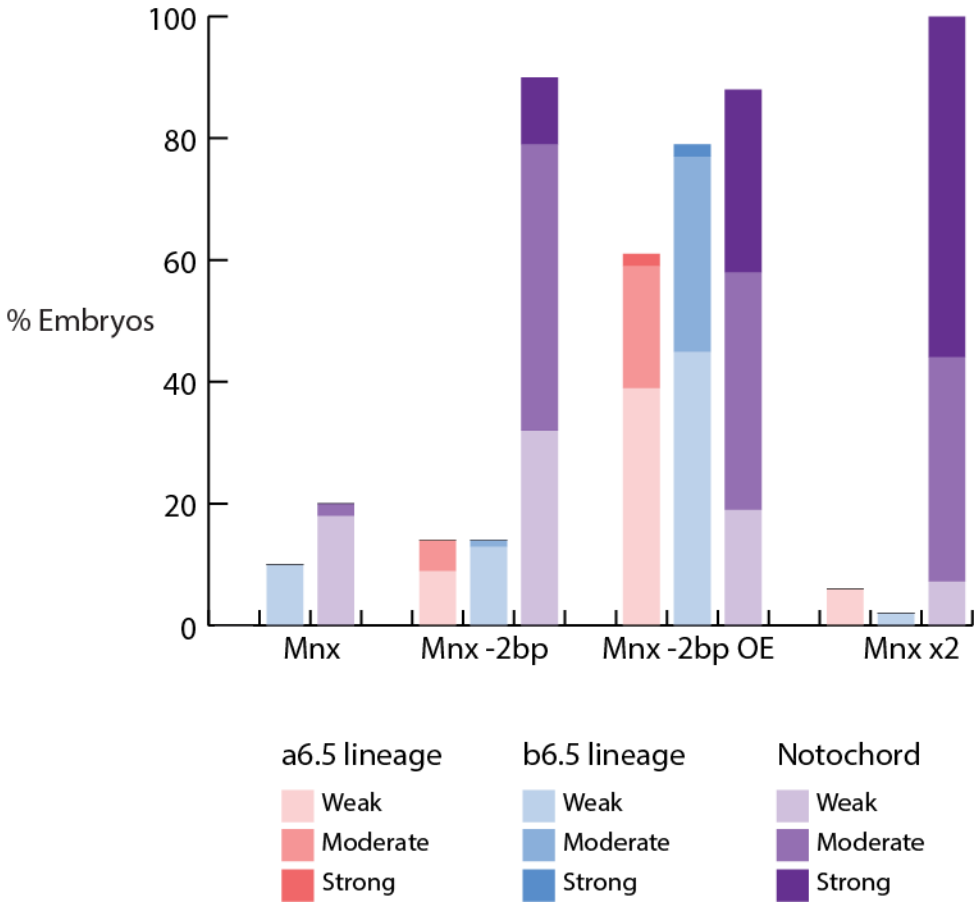


Figure S6. Manipulation of spacing and affinity modulate expression and tissue specificity.

Scoring for embryos electroporated with Mnx, Mnx –2bp, and Mnx –2bp OE. Decreasing the spacing from 13 bp to 11 bp in Mnx to Mnx –2bp results in a dramatic increase in notochord expression. Increasing the binding affinity of the distal ETS from 0.25 to 1.00 in Mnx –2bp to Mnx –2bp OE results in stronger notochord expression and ectopic expression in the a6.5 lineage and b6.5 lineage. Multimerizing the enhancer lead to higher levels of notochord expression (Mnx x2) (n=100 embryos for each construct). Weak, moderate and strong expression definitions are based on exposure time required for 10% saturated pixels. Strong <500ms, moderate <800ms and weak >800ms.

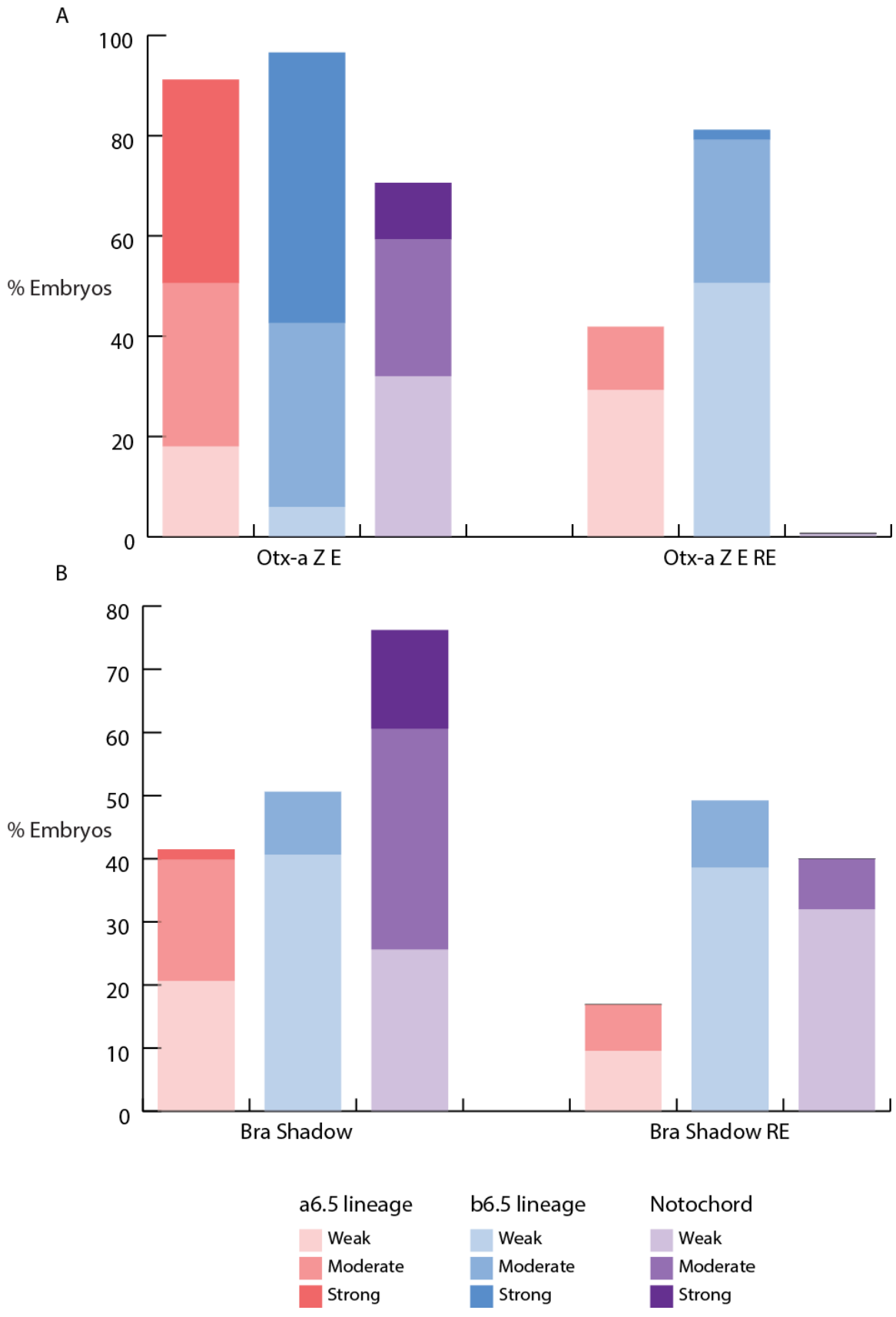

A
100
80
60
40
20
0
% Embryos
Otx-a Z E
Otx-a Z E RE
B
80
70
60
50
40
30
20
10
0
% Embryos
Bra Shadow
Bra Shadow RE
a6.5 lineage
Weak
Moderate
Strong
b6.5 lineage
Weak
Moderate
Strong
Notochord
Weak
Moderate
Strong

Figure S7. Syntax an important aspect of enhancer function in modified Otx-a enhancer and Bra Shadow enhancer

(A) Scoring for embryos electroporated with modified Otx-a enhancer with addition of ZicL and ETS sites (Otx-a Z E) and Otx-a Z E with reversed ETS (Otx-a Z E RE). Reversing the ETS results in a dramatic decrease in notochord expression. (n=100 embryos for each construct). Weak, moderate and strong expression definitions are based on exposure time required for 10% saturated pixels. Strong <500ms, moderate <800ms and weak >800ms. **(B)** Scoring for embryos electroporated Bra shadow enhancer and Bra shadow enhancer with reversed ETS (Bra Shadow RE). Reversing the ETS results in a dramatic decrease in notochord expression. (n=100 embryos for each construct). Weak, moderate and strong expression definitions are based on exposure time required for 10% saturated pixels. Strong <500ms, moderate <800ms and weak >800ms

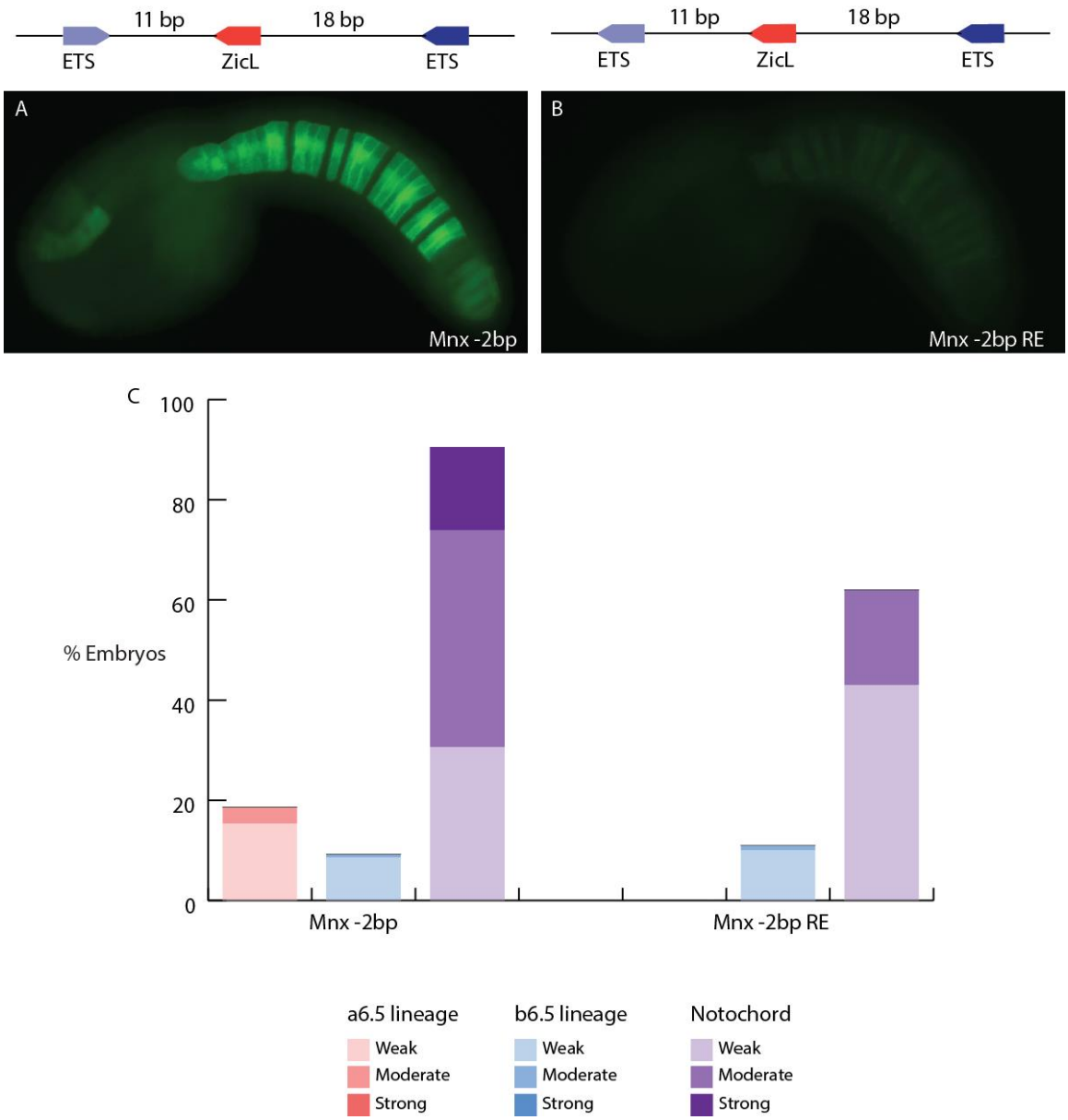


Figure S8. Syntax an important aspect of enhancer function in Mnx enhancer

(A) Embryo electroporated with Mnx enhancer with 11bp spacing between ETS and ZicL (Mnx -2bp) construct; GFP expression can be seen in the notochord. **(B)** Embryo electroporated with Mnx -2bp with ETS reversed (Mnx -2bp RE); GFP expression is diminished in the notochord. A schematic of the sequence electroporated is shown above each image. Dark blue boxes refer to ETS binding sites with a binding affinity above 0.60, light blue boxes refer to ETS binding sites with a binding affinity lower than 0.60, and red boxes refer to a ZicL binding site. All images were taken at the same exposure time. **(C)** Scoring for embryos electroporated Mnx – 2bp enhancer and Mnx -2bp with reversed ETS (Mnx -2bp RE). Reversing the ETS results in a dramatic decrease in notochord expression. (n=100 embryos for each construct). Weak, moderate and strong expression definitions are based on exposure time required for 10% saturated pixels. Strong <500ms, moderate <800ms and weak >800ms

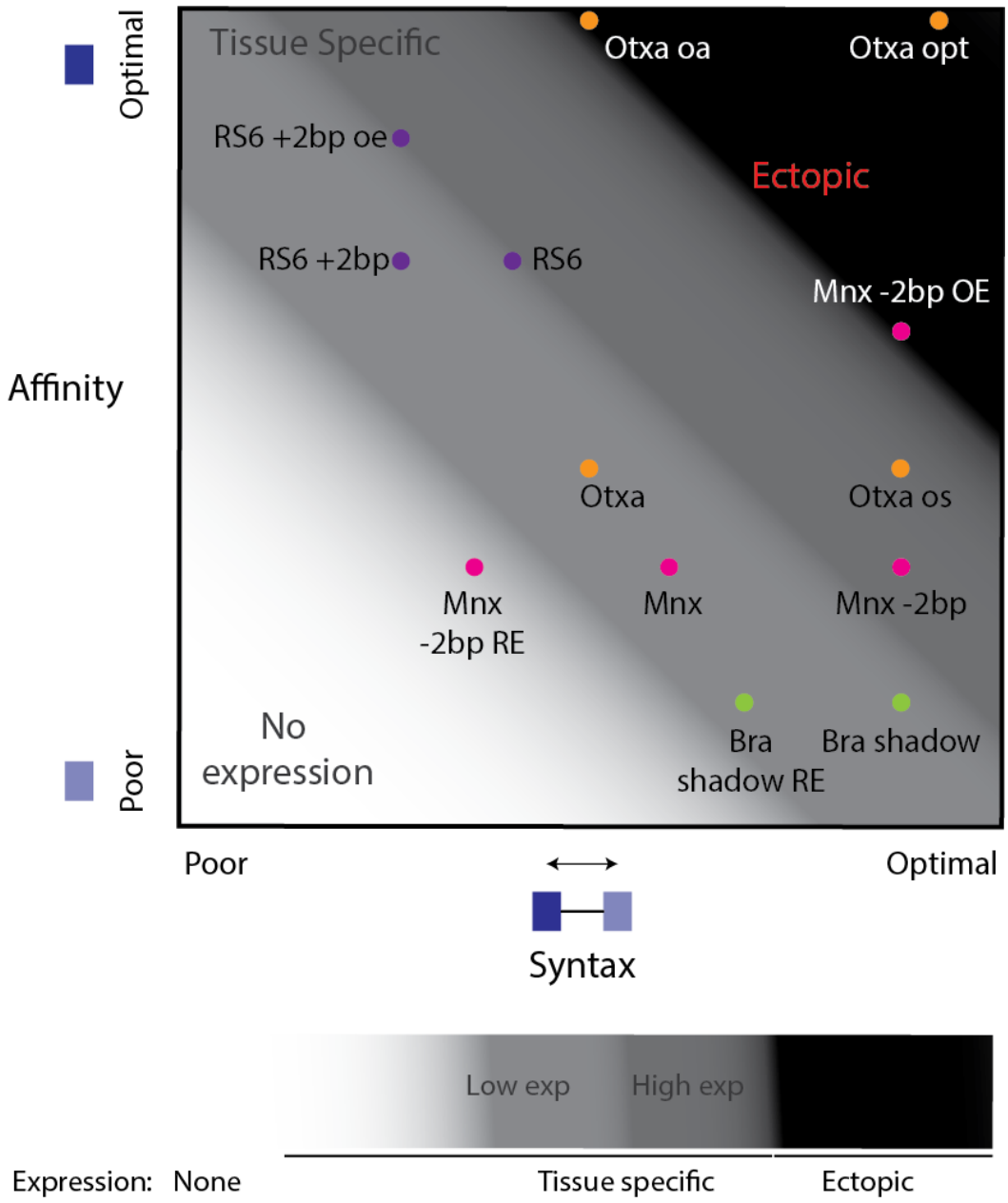


Figure S9. Balance of binding affinity and spacing encodes enhancer specificity

Schematic showing affinity of binding sites on Y-axis and syntax on X-axis. Optimal flanking and spacing leads to ectopic expression, while poor spacing and flanking results in no expression. Balancing suboptimal and optimal sequence and syntax mediates tissue specific expression. Purple dots correspond to RS and manipulations. Pink dots to Mnx and manipulations. Green dots to Bra and manipulations. Orange corresponds to Otxa, Otxa with optimal spacing (Otxa os), Otxa with optimal affinity (Otxa oa) and Otxa with optimal affinity and spacing (Otxa opt) from (2).

Dataset Legends

Dataset 1. Synthetic enhancer variants with ZicL and two Ets binding sites. We surveyed the Randomized Synthetic Otx-a library for enhancer variants containing ZicL and two ETS sites. There are 15 such variants. Table S1 lists the sequence, expression levels, and affinity of ETS binding sites within these enhancer variants.

Dataset 2. Manipulated enhancer variants used in this study. Table lists the sequence, expression levels, and affinity of ETS binding sites within these enhancer variants.

Dataset 3. List of putative notchord enhancers. We scanned the genome of *Ciona intestinalis* for a ZicL site and 2 ETS sites within a 200 bp region. We searched for sequences where the most proximal ETS points towards the ZicL and these two binding sites are between 9-15bp apart. We required one ETS to have a binding affinity of at least 0.10, and another ETS affinity of at least 0.20. Totally, 69 such regions were found in *Ciona* intestinalis. Two were tested, both worked. Table shows location of putative notochord enhancer, closest gene and putative enhancer variant sequence