Supplementary Table 1. Used core databases and genome builds. The table indicates, for each genome used in the analyses, the name of the respective ensembl core database and the relative genome build.

Species	**Ensembl core database**	**Genome build**
Homo sapiens	homo_sapiens_core_49_36k	NCBI36
Mus musculus	mus_musculus_core_49_37b	NCBIM37
Cani familiaris	canis_familiaris_core_49_2g	BROADD2
Takifugu rubripes	takifugu_rubripes_core_49_4i	FUGU4
Danio rerio	danio_rerio_core_49_7c	ZFISH7
Gasterosteus aculeatus	gasterosteus_aculeatus_core_49_1f	BROADS1
Oryzias latipes	oryzias_latipes_core_49_1e	MEDAKA1
Ciona intestinalis	ciona_intestinalis_core_49_2i	JGI2
Ciona savignyi	ciona_savignyi_core_49_2f	CSAV2.0

Supplementary Table 2. Amplifications of the fragments containing E1, E2, E3.

Primers used for Ciona enhancer validations
E1_f: CCGCCAATTTTTCCGG
E1_r: CACATTACCAGCATAA
E2_f: CCGTGTCTTCGTTCAC
E2_r: CCAAACGCAAAGTCC
E3_f: AATGCGTTGGTTACCC
E3_r: CGTCATCTTGTTACGT

Length of amplified fragments
E1: 600 bp
E2: 940 bp
E3: 300 bp

Primers used for zebrafish enhancer validations
E1_f: TTGTTAAGCGCAGTTGACAGAAG
E1_r: TACTTGAGTTTCCCCCTTTGTG
E2_f: TTAGAGATAGCCACAGGGGTTGAG
E2_r: GTTGGGGATTCACACACATAAGC
E3_f: AGTTTCACTCCCACCGTCAATC
E3_r: GCCGTGGTACTGAAAGTTTCTGAG

Length of amplified elements
E1: 490 bp
E2: 480 bp
E3: 1110

Supplementary Table 3. Primers used in RT-PCR experiments.

Primers used for Ciona RT-PCRs
E1_f: GAATGGCCCAGCTTATCTCA
E1_r: GAGACTTCGGGCTCATTCTG
E2_f: GCCGCGCAGAAATTTAATAC
E2_r: CAGCATAGGCCCTGATTCAT
E3_f: GTGCCTTGGGCATAATTGTT
E3_r: TGCACTTAATCAAGTTTTGCAT
Ci-pans_f: ATGAATGAAGGCACCTACGC
Ci-pans_r: TACGAGCGGAAGCAACTCTT
Ci-atbf_f: TCATCATTGCCACGCAGTAT
Ci-atbf_r: CATTCGATTGTCGATGCTTG

Primers used for mouse RT-PCRs
E1_f: CAGAGGACAGCGGTGTCAATA
E1_r: TGCTGCCTGAAATGAAATTACC
E2_f: GCTGGTGCCACTACTGAAAGA
E2_r: ATAGACCATCGCGCAGGATA
E3_f: GATGCATCTTGAGTGGGTCA
E3_r: TGGCGGAAAGCCTTTATGAT
Mm-pans_f: GCAAGCAGCCTTTTCTCTTC
Mm-pans_r: CTGAGGTGATGCAGCATTCTA
Mm-Otx2_f: GGACTGCAGGGCAGAGAC
Mm-Otx2_r: GAGTGACGGAACTAGGAGCAA
Mm-bActin_f: AGATGACCCAGATCATGTTTGA
Mm-bActin_r: TGTGGTACGACCAGAGGCATA

Primers used for zebrafish RT-PCRs
E1_f: GAGCGCGAATGAAATTACCA
E1_r: GGCGTGTCAATATCCCACTC
E2_f: CACAGCCTGGATGCCTTAAT
E2_r: GGCATAGAGCGCAATACACA
E3_f: CAGTTTCCGATGCATTTTGA
E3_r: TATGGCCCCCTGCACTATTA
Dr-bActin_f: TCACAATACCAGTAGTACGACCAGA
Dr-bActin_r: TACAATGAGCTCCGTGTTGC

Supplementary Table 4. Protocols used in RT-PCR experiments in each analyzed species.

Organism	Mouse
RNA	Total
Tissue	Whole embryo
RNA amount	1µg
cDNA preparation kit	Quantitect Reverse Transcription (Qiagen)
Negative control	no retrotranscriptase
cDNA amount in PCR	50 ng
Primers amount in PCR	400 nM
Taq	Amplitaq Gold (Invitrogen) - 1.5 units
Final volume PCR	20µl
PCR conditions	95° 5' - (95° 30'' - 58° 30'' - 72° 30'') x 38 cycles - 72° 7'

Organism	Zebrafish
RNA	Total
Tissue	whole embryo
RNA amount	1 µg
cDNA preparation kit	RevertAid H minus M-MuLV RT (Fermentas) and random hexamer primer (Roche)
Negative control	DNAseI digested RNA used also as template for cDNA synthesis
cDNA amount in PCR	2 µg
Primers amount in PCR	10 pmol of each primer (left + right primer)
Taq	GoTaq (Promega) – 0.3µl
Final volume PCR	20 µl
PCR conditions	95° 2' - (95° 20'' - 60° 20'' - 72° 30'') x 38 cycles

Organism	Ciona
RNA	Total
Tissue	Whole embryo
RNA amount	1µg
cDNA preparation kit	QuantiTect Reverse Transcription kit (Qiagen)
Negative control	water; non transcribed RNA; primers combined in ways that should not give amplification
cDNA amount in PCR	10-20 ng
Primers amount in PCR	1 µM
Taq	Quantifast PCR kit (Qiagen)
Final volume PCR	25 µl
PCR conditions	95° 5' - (95° 10'' - 60° 30'') x 40 cycles – using Biorad IQ5 real time PCR apparatus

Supplementary Table 5. oCNE data and sequences. The table contains all the relevant genomic information regarding the discovered elements and their sequences (feature_pair_id = oCNE id; csn = coordinate system name; srn = sequence region name).

(see PDF file named supplementary_table_5.pdf)

Supplementary Table 6. Genes flanking or overlapping oCNE in the species analyzed. The annotation in mouse derives from Ensembl, in *Ciona* from Aniseed and in *Oikopleura* and amphioxus from Blast2GO.

(see PDF file named supplementary_table_6.pdf)

Supplementary Table 7. Type and loci of oCNEs duplications.

Vertebrate specific duplications

1. oCNE 1349788 is present in 4 copies upstream to Irx1, Irx3, Irx4, Irx6.
2. oCNE 1351400 is present in 2 copies upstream to the Sox21, Sox14.
3. oCNE 1352111 is present in 2 copies upstream to the Lmo1, Lmo3.

Mammals specific duplications

1. oCNE 1352298 is present in 2 copies downstream to the Neurod2, Neurod6.
2. oCNE 1351047 is present in 2 copies intergenic between Tshz3/Zfp536 and Tshz1/Zfp516.

Fish specific duplications

1. oCNE 1350761 is present in 3 copies intronic to ebf1a, ebf1b, ebf3.
2. oCNE 1351012 is present in 2 copies downstream to lhx1b, lhx1a.

Zebrafish specific duplications

1. oCNE 1349653 is present in 2 copies intronic to pbx3b gene and upstream to pbx3a.
2. oCNE 1350007 is present in 2 copies intronic to dachd, dachc.
3. oCNE 1350084 is present in 3 copies intronic to dacha, dachd, dachc.
4. oCNE 1353058 is present in 2 copies downstream to pax7a, pax7b.

Ascidian specific duplications

1. oCNE 1352154 is present in 4 copies in a 3kb region on the chromosome 2q, intergenic between rpl31/pigf. The same element is present in 5 copies in the syntenic region of Ciona savignyi.
2. oCNE 1351484 is present in 2 copies in a 7kb region on the chromosome 7q, intronic and upstream to two genes of unknown function. The same element is present in 1 copy in the syntenic region of Ciona savignyi.
3. oCNE 1351973 is present in 4 copies in a 22kb region on the chromosome 1q, intronic to the ccnb3 gene and intergenic between ccn3/bai1. The same element is present in 1 copy in the syntenic region of Ciona savignyi.

Supplementary Table 8. Jaspar analysis results. For each family of transcription factors is reported the number of the oCNEs and rCNEs (considered as a negative set) containing at least 1 match with the corresponding transcription factor, the p-value and the corrected p-value for the comparisons of the 2 proportions testing the enrichments in oCNEs.

MOUSE				
class	**oCNEs**	**rCNEs**	**p**	**p.adj**
Homeobox class	140	91	1.67E-10	1.83E-09
bZIP cEBP-like subclass	85	58	1.85E-04	1.02E-03
HMG class	150	142	3.44E-03	1.26E-02
Forkhead class	96	83	1.33E-02	3.66E-02
TRP(MYB) class	137	136	4.90E-02	1.08E-01
MADS class	31	33	4.74E-01	8.69E-01
Nuclear Receptor class	70	118	1.00E+00	1.00E+00
bZIP CREB/G-box-like subclass	37	46	6.68E-01	1.00E+00
ETS class	17	38	9.93E-01	1.00E+00
REL class	8	15	8.47E-01	1.00E+00
bHLH(zip) class	21	44	9.94E-01	1.00E+00

FUGU				
class	**oCNEs**	**rCNEs**	**p**	**adj.p**
Homeobox class	138	81	1.56E-12	1.71E-11
bZIP cEBP-like subclass	92	47	2.26E-08	1.24E-07
Forkhead class	96	54	1.41E-07	5.18E-07
HMG class	144	128	4.58E-04	1.26E-03
TRP(MYB) class	126	122	4.04E-02	8.88E-02
REL class	11	9	3.16E-01	5.79E-01
bZIP CREB/G-box-like subclass	35	43	6.37E-01	8.76E-01
MADS class	32	39	6.11E-01	8.76E-01
ETS class	15	34	9.91E-01	9.99E-01
Nuclear Receptor class	73	114	9.99E-01	9.99E-01
bHLH(zip) class	24	44	9.80E-01	9.99E-01

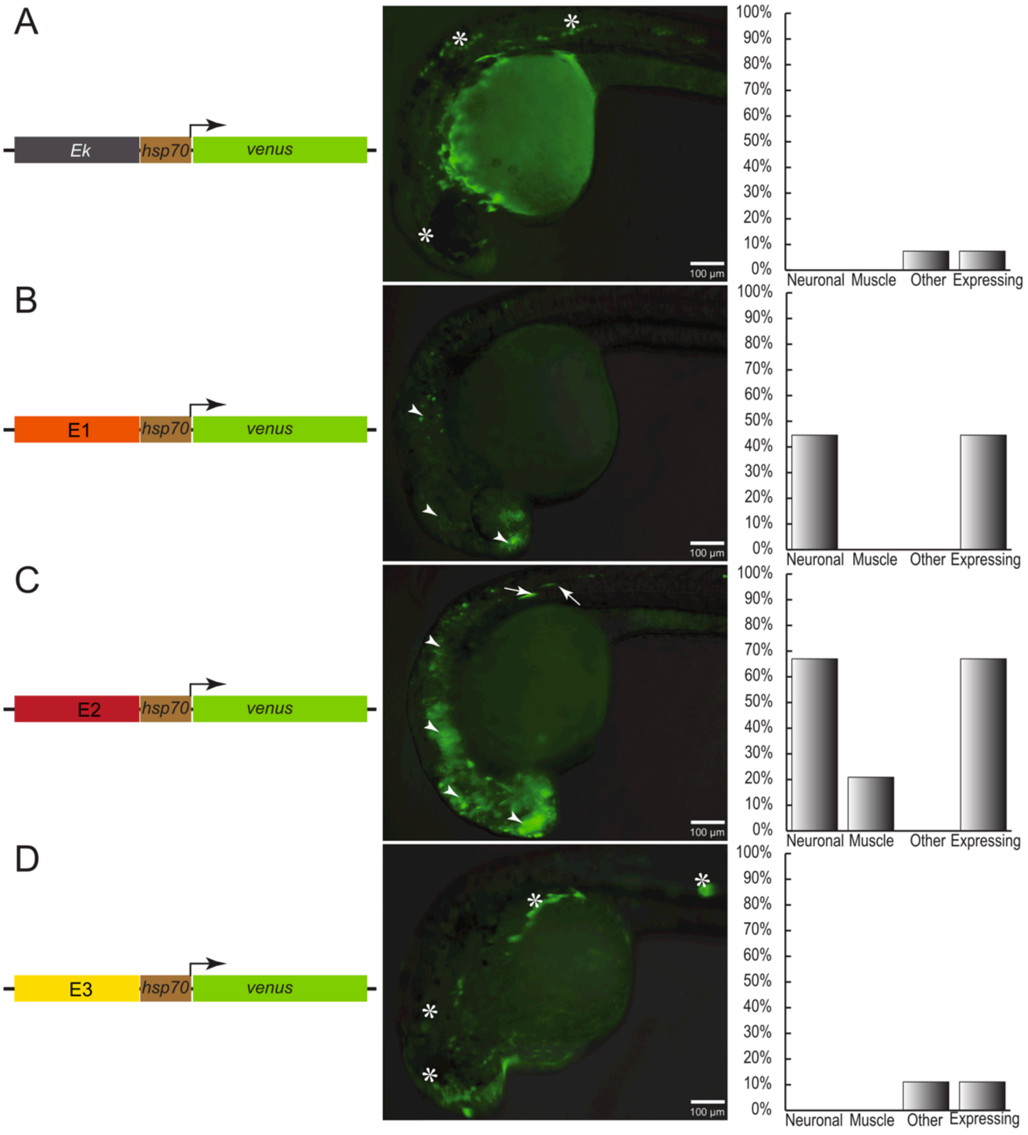
CIONA				
class	**oCNEs**	**rCNEs**	**p**	**p.adj**
Homeobox class	131	65	9.16E-15	1.01E-13
HMG class	153	129	7.78E-06	4.28E-05
Forkhead class	96	71	3.91E-04	1.43E-03
bZIP cEBP-like subclass	84	64	2.59E-03	7.13E-03
TRP(MYB) class	146	146	4.67E-02	1.03E-01
MADS class	45	37	8.05E-02	1.48E-01
ETS class	26	25	3.45E-01	5.43E-01
REL class	6	7	5.00E-01	6.88E-01
bZIP CREB/G-box-like subclass	55	67	6.96E-01	8.51E-01
bHLH(zip) class	31	43	8.23E-01	9.06E-01
Nuclear Receptor class	64	86	9.17E-01	9.17E-01

Supplementary Table 9. The table shows for amphioxus and Oikopleura the number and the score of oCNEs associated genes classified for the top scoring GO biological processes by Blast2GO analysis.

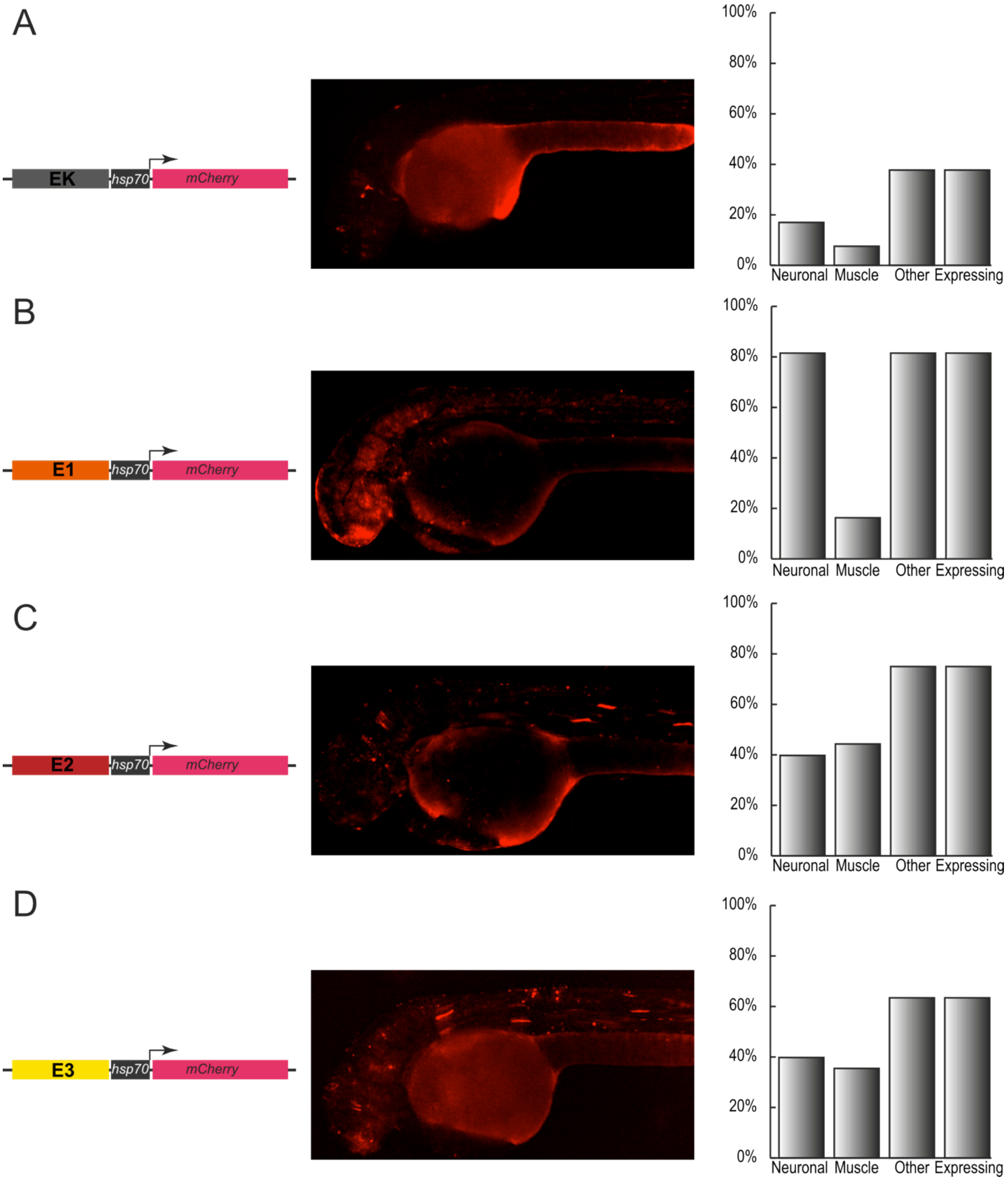
Amphioxus			
Level	**GO Term**	**#Seq**	**Score**
3	regulation of biological process	20	16.11
4	regulation of cellular process	17	15.05
3	anatomical structure development	15	11.38
3	multicellular organismal development	15	10.03
4	transport	11	9.67
5	organ development	12	8.28
3	macromolecule metabolic process	14	8.25
4	signal transduction	9	7.42
4	anatomical structure morphogenesis	8	6.84
3	cellular metabolic process	15	6.84

Oikopleura			
Level	**GO Term**	**#Seq**	**Score**
3	anatomical structure development	5	9.32
4	regulation of cellular process	6	7.12
5	organ development	4	6.87
3	multicellular organismal development	5	6.4
3	regulation of biological process	6	5.32
4	anatomical structure morphogenesis	3	5.15
4	system development	4	5.1

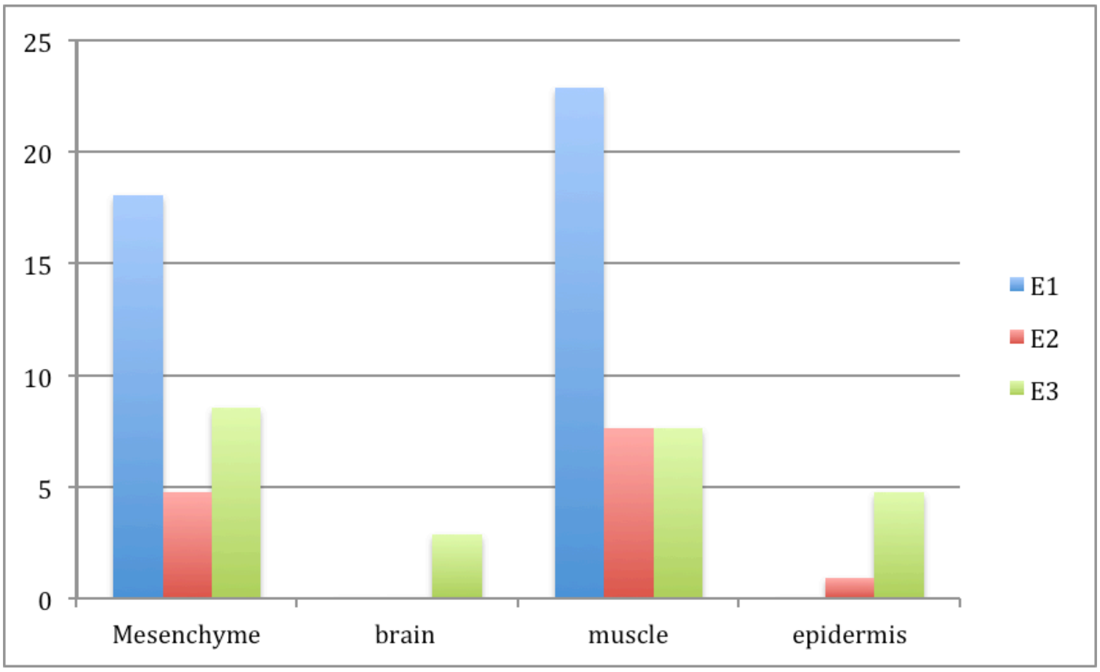
Supplementary Figure 1. Validation of zebrafish oCNEs enhancer function in zebrafish. The figure reports, for each validation, a schema representing the injected construct, a representative image of the screened embryos and a chart indicating the percentage of embryo expressing the reporter in the relative tissue. A is the construct containing the negative control, B-D are relative to the tested oCNEs.



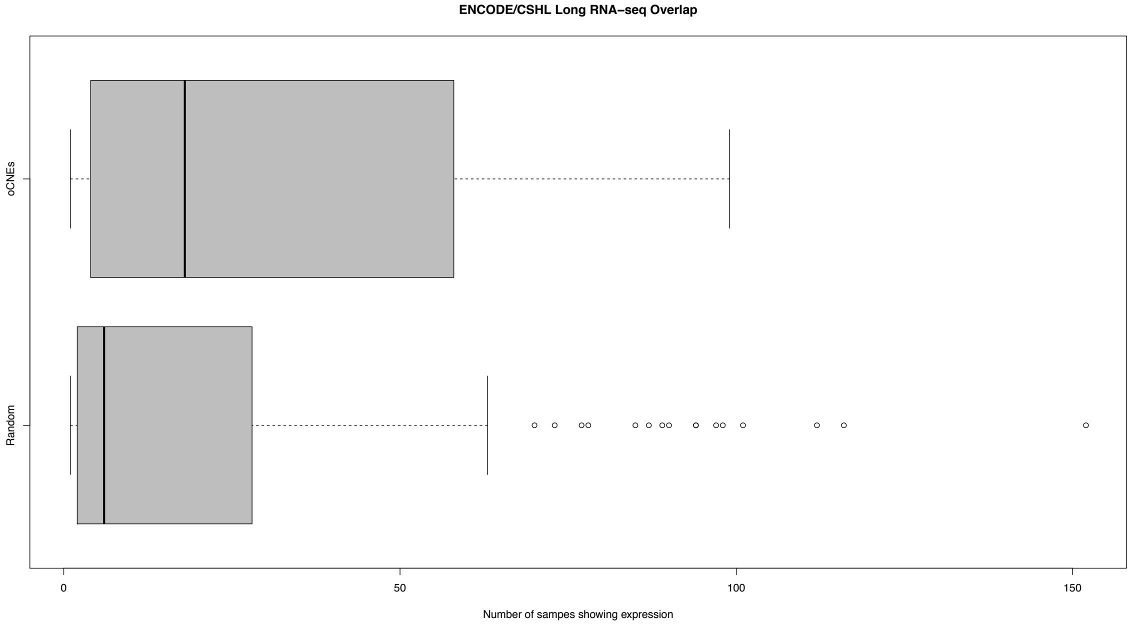
Supplementary Figure 2. Validation of ciona oCNEs enhancer function in zebrafish. The figure reports, for each transgenesis validation, a schema representing the injected construct, a representative image of the screened embryos and a chart indicating the percentage of embryo expressing the reporter in the relative tissue. A is the construct containing the negative control, B-D are relative to the tested oCNEs.



Supplementary Figure 3. Chart indicating the percentages of Ciona embryos and the tissues expressing the reporter in transgenesis experiments. For each element more than 100 embryos were tested.



Supplementary Figure 4. Boxplot showing the distributions of the number of the 183 oCNEs and 183 randomly selected genomic regions overlapping contigs from the ENCODE/CSHL Long and Short RNA-seq tracks in the UCSC hg19 huan genome browser.



Supplementary Figure 5. Window lengths and common domain enrichments. The chart shows, for 3 different genomic intervals, the difference between the percentage of oCNEs (light grey) and randomly coupled vCNEs/rCNEs (dark grey) containing at least one common protein domain in all the species analyzed. Adjusted p-values are reported on the side of each couple of bars. Window lengths are in Kb.