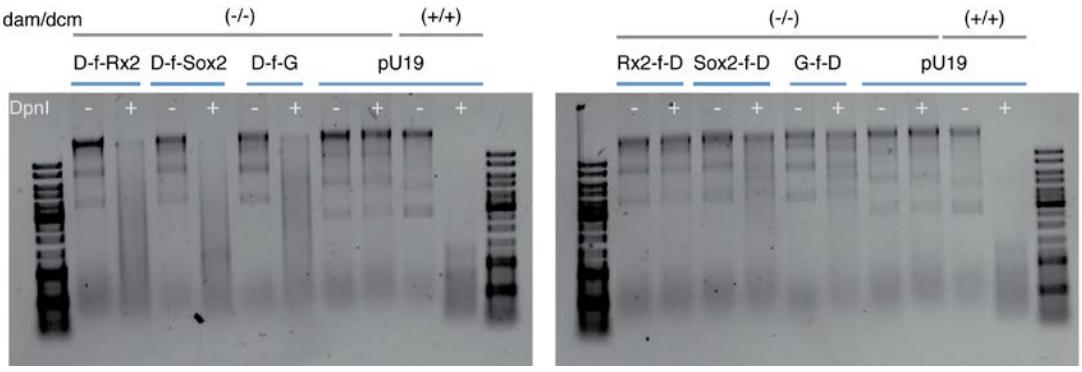
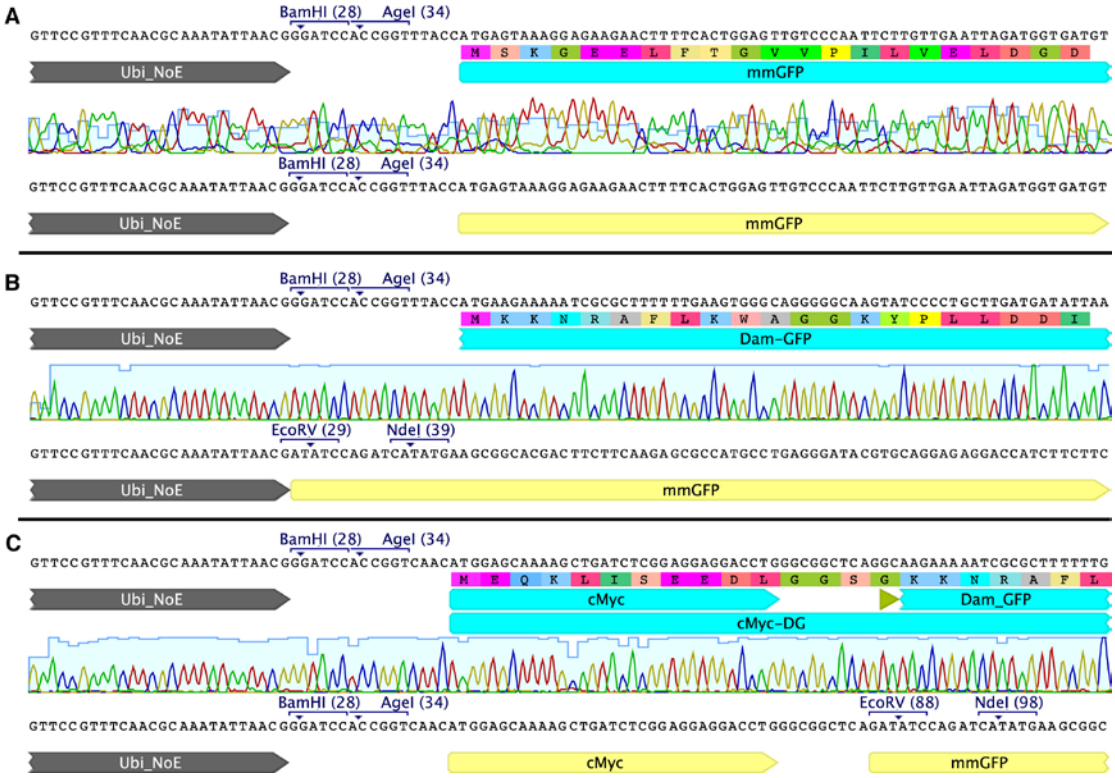
Supplementary information

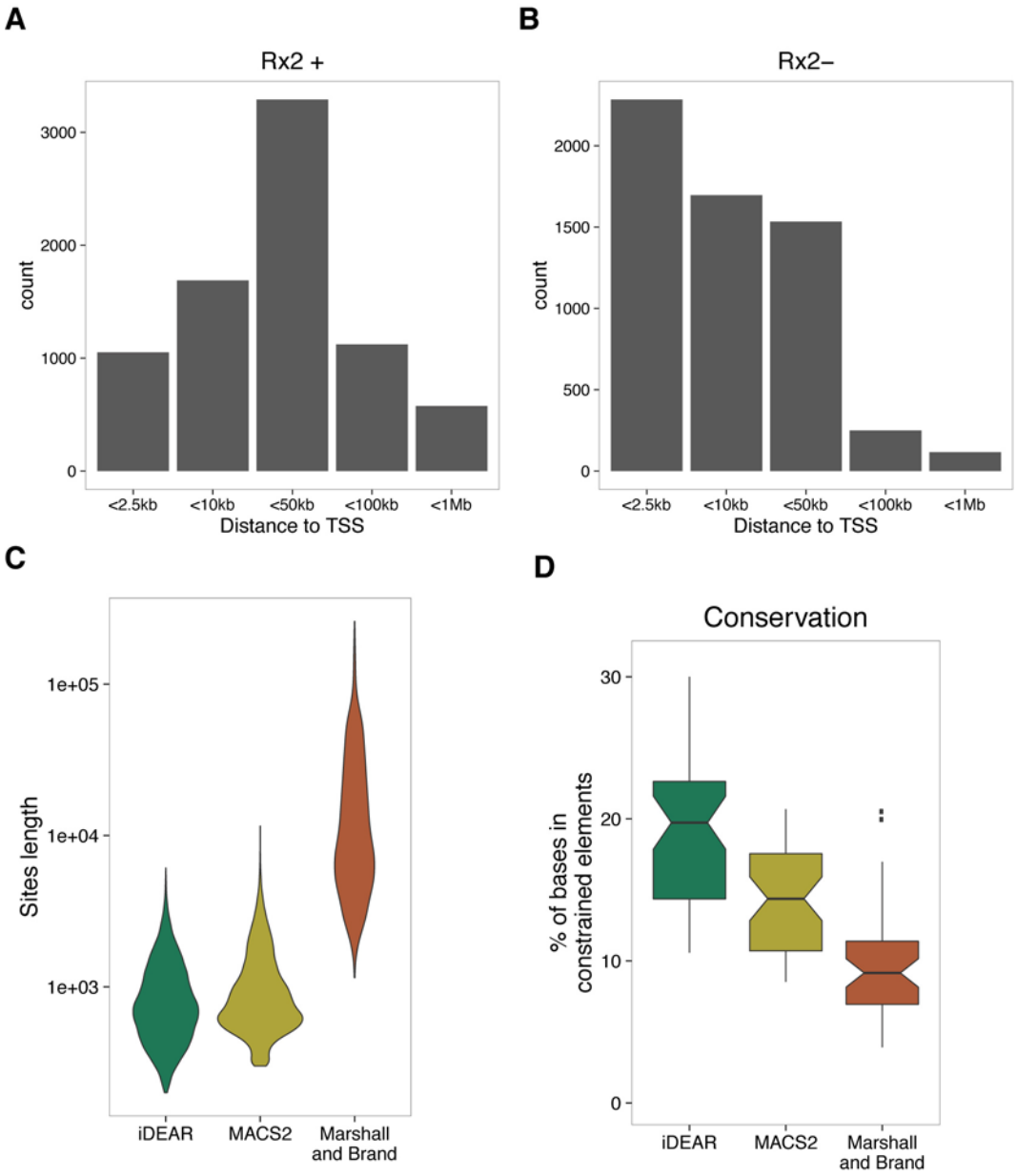


Supplementary Figure 1. Dam is mostly functional when Rx2, Sox2 and GFP are located in the carboxyterminal part of the fusion protein. DpnI protection assay of gDNA isolated from dam/dcm deficient bacteria transformed with constructs coding for the Dam protein fused, via flexilinker, to Rx2, Sox2 or GFP in amino (DRx2, DSox2 and DG) or carboxyterminal (Rx2D, Sox2D and GD) orientation. As control, pUC19 plasmid was used to transform dam/dcm deficient (-/-) or wild type (+/+) cells.

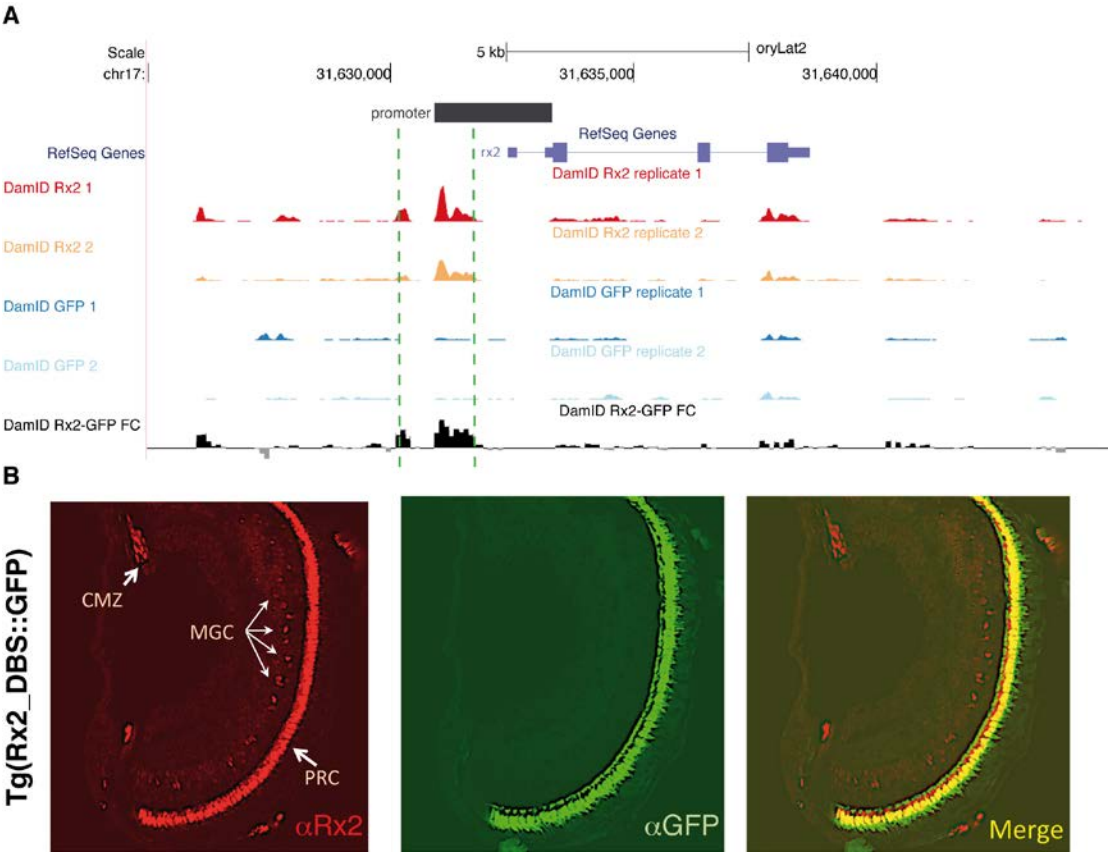


Supplementary Figure 2. Aberrant splicing events occurring in the transcripts from constructs carrying *E. coli* Dam fused to mmGFP in medaka hatchling larvae. The expected products are shown above the respective chromatograms. (A) The ubiquitin non-coding exon (NoE) uses the correct acceptor splice site upstream of the BamHI restriction site just upstream the start codon of mmGFP. (B) In Dam-GFP fusion, NoE splices to an internal cryptic acceptor site in the mmGFP sequence, skipping completely the *Dam* coding sequence. (C) In the cMyc-Dam-GFP construct, NoE splices to the correct acceptor splice site as for mmGFP. However a second splicing event occurs after the cMyc sequence between a cryptic donor site, just upstream of the *Dam* gene, and the same internal cryptic acceptor site in the mmGFP sequence (not shown here) completely skipping the Dam coding sequence.

Supplementary Figure 3. Alignment of the *Dam* gene used in this study (DamL122A) and the optimized version.



Supplementary Figure 4. Characterization of the Rx2 sites. A-B) Most Rx2 occupied sites tend to be within 10kb and 50kb of annotated transcription start sites (TSS) but Rx2 negative sites tend to be in the proximal upstream region of genes. C) Length distribution of the Rx2 bound sites identified by iDEAR, MACS2 and the Marshall and Brand pipeline. iDEAR provides regions around 1kb long, similar to MACS2, but Marshal and Brand gives regions always above that length and even some over 100kb. D) Degree of overlap of the sites identified by each tool with evolutionary constrained elements in fish. iDEAR shows a higher sensitivity with respect to conservation.



Supplementary Figure 5. Functional validation of the Rx2 binding site on its own promoter. A) Representation of the iDamIDseq signal around the *Rx2* locus showing the normalized read coverage for each one of the replicates of Rx2 and GFP fusions and the log2 fold change of the average signal per replica (track in black at the bottom). The black bar over the RefSeq Genes track represents the previously characterized Rx2 promoter. The area of maximum Rx2 enrichment is included in the 5' end of this promoter. B) The Rx2_DBS element, sequence delimited by the green dashed lines in A, displays regulatory activity in the photoreceptor cell layer (PRC). The iDamID Rx2_DBS fragment was cloned in a GFP reporter plasmid. Immunofluorescence of retinal sections of hatchling animals from the stable line Tg(Rx2_DBS::GFP) shows complete colocalization of the GFP (green) with the expression domain of Rx2 (red) in the PRC. CMZ, ciliary marginal zone; MGC, Müller glia cells.

Table S1. List of primers. All primers except AdR_PCR are unmodified. *Bold-face letters of the AdR_PCR primer represent the bases modified with phosphorothioate bonds.

Primer (5´-3´)	
AATTTACCGGTAACATGAAGAAAAATCGCGCTTTTTTG	N_Dam fwd
AATTTGCTAGCTTTTTTCGCGGGTGAAACGACTCCTG	N_Dam rev
CTACGGTTACAACGGCgctTGTCGTTACAATCTG	L122A fwd
CAGATTGTAACGACAagcGCCGTTGTAACCGTAG	L122A rev
AATTTACCGGTCAACATGGAGCAAAAGCTGATCTCGGAGGAGGACCTGGGCGGCTCAGGCAAGAAAAATCGCGCTTTTTTGAAGTG	N_cMyc_Dam fwd
AATTTACTAGTGAACAGAAACTCATCTCTGAAGAGGATCTGGCACCCGGCAAGAAAAATCGCGCTTTTTTGAAGTG	C_cMyc_Dam fwd
CTAGCCTGAGCGGTGGAGGCGGTTCAGGCGGAGGTGGCTCTGGCGGTGGCGGATCGGGAGGCGGTGGAAGTGCAGCCGCGA	Flexilinker
CTAGTCGCGGCTGCACTTCCACCGCCTCCCGATCCGCCACCGCCAGAGCCACCTCCGCCTGAACCGCCTCCACCGCTCAGG	Flexilinker
AATTTACTAGTATGAGTAAAGGAGAAGAACTTTTC	C_mmGFP fwd
AATTTTCTAGATTATTTGTATAGTTCATCCATGC	C_mmGFP rev
AATTTTCTAGATCTAGATTAGTCCAGCCTGGGCCTCTTGGGAGGTGGTTTGTATAGTTCATCCATGCCATGTG	C_oNLS-mmGFP rev
AATTTACCGGTACCATGAGTAAAGGAGAAGAACTTTTCAC	N_mmGFP fwd
AATTTGCTAGCTTTGTATAGTTCATCCATGCCATG	N_mmGFP rev
AATTTACTAGTATGCTTTCGCATGCCGACCTGCTG	C_Dr-OtpA fwd
AATTTTCTAGATTAGGTGAAGCTCATGGACACTGTGTG	C_ Dr-OtpA rev
AATTTACCGGTACCATGCTTTCGCATGCCGACCTGCTG	N_Dr-OtpA fwd
AATTTGCTAGCGGTGAAGCTCATGGACACTGTGTG	N_ Dr-OtpA rev
AATTTACTAGTATGCATTTGTCAATGGATACCCTG	C_Ol-Rx2 fwd
AATTTTCTAGATTACATGTGCTGCCAGGTTTTATC	C_Ol-Rx2 rev
AATTTACCGGTACCATGCATTTGTCAATGGATACCCTG	N_Ol-Rx2 fwd
AATTTGCTAGCCATGTGCTGCCAGGTTTTATCCATG	N_Ol-Rx2 rev
AATTTACTAGTATGTATAACATGATGGAGACTGAAC	C_Ol-Sox2b fwd
AATTTTCTAGATTACATGTGTGTTAACGGCAGCG	C_Ol-Sox2b rev
AATTTACCGGTACCATGTATAACATGATGGAGACTGAAC	N_Ol-Sox2b fwd
AATTTGCTAGCCATGTGTGTTAACGGCAGCGTGC	N_Ol-Sox2b rev
AATTTGGCGCGCCGATCTCAGGGACTTACACTGTCATC	Rx2(Rx2_BDS) fwd
AATTTACTAGTCAGCTACAGTAGGTGTTGCAGAG	Rx2(Rx2_BDS) rev
CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAGGA	AdRt
TCCTCGGCCG	AdRb
GGTCGCGGCCGAGGATC*	AdR_PCR
GAGCCGTACTGTTCCGTTTCAAC	RT_Ubi_fwd
ATGTGTAATCCCAGCAGCCG	RT_mmGFP_rev

Table S2. Rx2-enriched sites and Rx2-depleted sites

Click here to download Table S2