Gene Regulation and Systems Biology

Libertas Academica
FREEDOM TO RESEARCH

OPEN ACCESS
Full open access to this and thousands of other papers at http://www.la-press.com.

ORIGINAL RESEARCH

Identification of *cis*-Regulatory Elements in the *dmyc* Gene of *Drosophila Melanogaster*

Jasmine Kharazmi[1,2], Cameron Moshfegh[1] and Thomas Brody[3]

[1]Biotechnopark Zurich, Molecular Biology Laboratory, [2]Institute of Molecular Life Sciences, University of Zurich-Irchel, Zurich, Switzerland. [3]Neural Cell-Fate Determinants Section, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD. Corresponding author email: jasmine.kharazmi@uzh.ch

Abstract: Myc is a crucial regulator of growth and proliferation during animal development. Many signals and transcription factors lead to changes in the expression levels of *Drosophila myc*, yet no clear model exists to explain the complexity of its regulation at the level of transcription. In this study we used *Drosophila* genetic tools to track the *dmyc cis*-regulatory elements. Bioinformatics analyses identified conserved sequence blocks in the noncoding regions of the *dmyc* gene. Investigation of *lacZ* reporter activity driven by upstream, downstream, and intronic sequences of the *dmyc* gene in embryonic, larval imaginal discs, larval brain, and adult ovaries, revealed that it is likely to be transcribed from multiple transcription initiation units including a far upstream regulatory region, a TATA box containing proximal complex and a TATA-less downstream promoter element in conjunction with an initiator within the intron 2 region. Our data provide evidence for a modular organization of *dmyc* regulatory sequences; these modules will most likely be required to generate the tissue-specific patterns of *dmyc* transcripts. The far upstream region is active in late embryogenesis, while activity of other *cis* elements is evident during embryogenesis, in specific larval imaginal tissues and during oogenesis. These data provide a framework for further investigation of the transcriptional regulatory mechanisms of *dmyc*.

Keywords: *dmyc*, *cis*-regulatory module, enhancer, promoter, downstream promoter element, *Drosophila*

Gene Regulation and Systems Biology 2012:6 15–42

doi: 10.4137/GRSB.S8044

This article is available from http://www.la-press.com.

© the author(s), publisher and licensee Libertas Academica Ltd.

This is an open access article. Unrestricted non-commercial use is permitted provided the original work is properly cited.

Introduction

In the early stages of organ development, the expression patterns of genes must be tightly spatiotemporally controlled. This function requires a set of complex interactions between the *cis* regulatory modules of each gene and the gene's regulatory proteins, which bind to these elements to modulate transcription. Indeed, the interactions between *cis* elements and their binding factors is well established as a key mechanism controlling expression of the developmental genes required for establishing the anteroposterior and dorsoventral axes in *Drosophila*.[1,2]

Myc is an important developmental gene requiring tight transcriptional regulation.[3–6] As an evolutionarily conserved bHLHZ (basic-helix-loop-helix leucine zipper) transcription factor, Myc is a master regulator of cell growth and proliferation.[7] Upon dimerization with Max, another bHLH protein, Myc binds to the E-box sequences of target genes to activate cellular growth and cell cycle progression.[8–10] Conversely, heterodimers of Myc and other Myc-associated zinc finger proteins, such as Miz1, can act negatively to regulate transcription of genes responsible for cell cycle arrest[11–13] (Fig. 1).

Myc proteins can link growth with cell cycle progression via activation of the S phase cyclins, which are required for DNA replication.[6,8,14] Regulation of cell growth and division is critical for animal

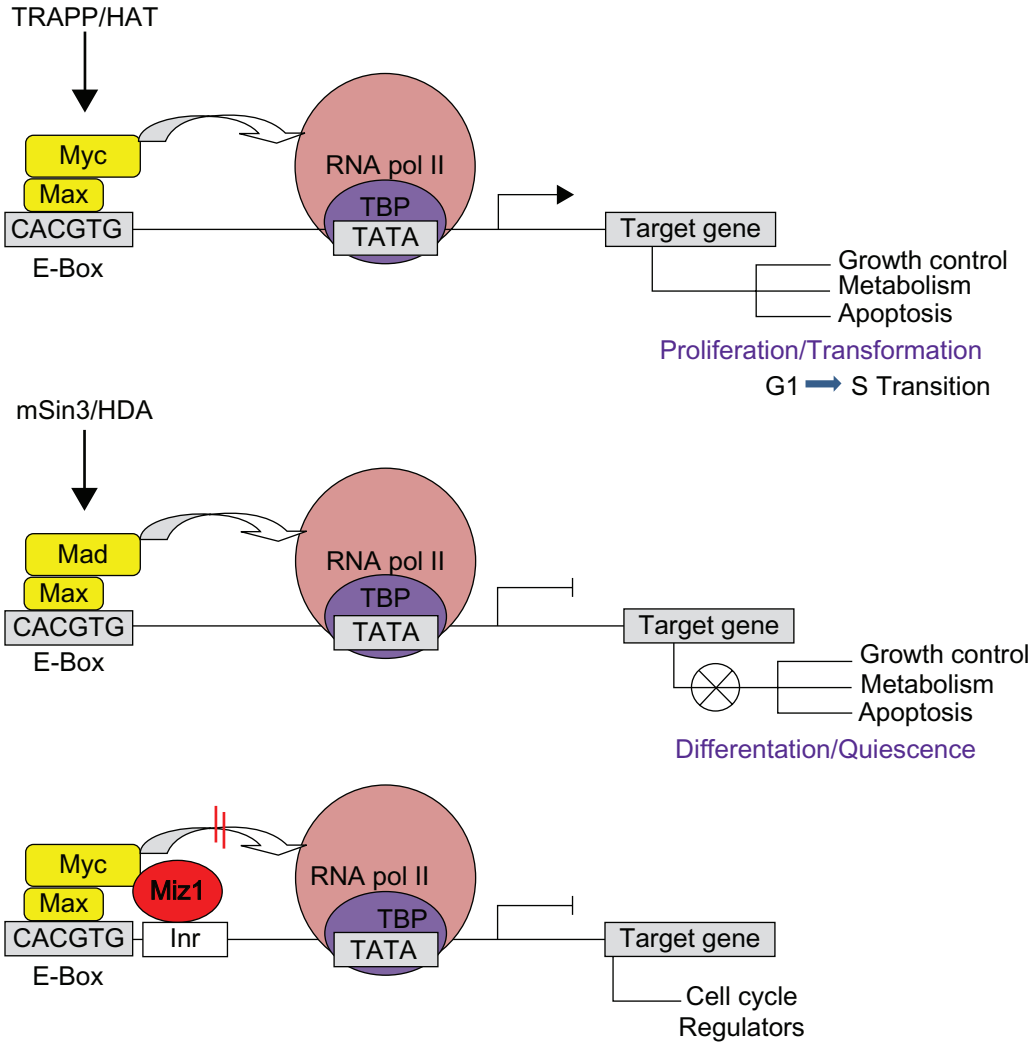


Figure 1. Simplified schematic interaction of Myc/Myc-associated bHLH proteins at the promoters of target genes. Upon binding to E-box sequences in the promoter region of target genes, heterodimers of Myc/Max can recruit chromatin remodeling complex TRAPP/Histone acetyl-transferase (HAT) and interact with the bound basal transcription machinery at the TATA region of target genes to activate transcription. Conversely, heterodimers of Mad/Max transcription factors recruit mSin3/Histone deacetylases to counteract Myc activity and repress Myc target genes by regulating differentiation and cell cycle arrest. The binding of Myc/Max dimers can interfere with the function of transcription activator Miz-1 to inhibit the recruitment of cofactor proteins like p300 to the promoters of genes responsible for cell cycle regulation.

development because too little growth leads to small organs and small body size, whilst excessive growth can lead to tissue overgrowth and initiation of cancer. For Myc, its misexpression during development perturbs normal growth; too much Myc can facilitate malignant transformation and too little Myc leads to retarded growth.[12,15–29]

Functional conservation between *Drosophila* dMyc and human c-Myc has been demonstrated in a variety of biological activities, such as the ability of dMyc to drive the cell cycle in *c-myc* null fibroblasts and to transform primary mammalian cells.[30,31] Conversely, c-Myc can rescue lethal mutations of *dmyc*.[32]

The evolutionarily conserved structure and function between dMyc and c-Myc[7] has prompted the use of *Drosophila* as a model to gain insight into many aspects of mammalian Myc biology. In particular, *Drosophila* genetic models have demonstrated that dMyc controls cell growth and cell division by regulating its direct targets involved in protein biogenesis and metabolism.[8,33–38] In response to developmental signals and mitogenic stimuli, patterned *dmyc* expression is required for a developing organ to reach its appropriate size and shape. For example, Wingless, Dpp, and Notch signaling pathways are key regulators of *dmyc* expression that are required to keep the balance between cell growth/division and differentiation.[35,37] A variety of tumor suppressor factors, including Half pint and Lethal giant larvae, negatively regulate *dmyc* transcription to achieve cell cycle inhibition and promote differentiation.[39,40] In addition, *dmyc* target genes can influence the transcription of *dmyc* in a regulatory feedback manner. For instance, the Hippo pathway transcriptional coactivator, Yorkie (yki), upregulates *dmyc*, and high levels of dMyc in turn repress yki.[36]

However, despite a growing list of pathways leading to altered *dmyc* transcription, we are far from unraveling the many complex interactions required at the *dmyc* promoter for patterning of *myc* transcription throughout development. Here we dissected the *dmyc* promoter and other potential regulatory regions and have drawn connections between certain domains and spatial and temporal patterning of *dmyc* expression throughout development of *Drosophila*. We find that its high expression in early embryos, larval discs and brain, and ovary is achieved through initiation of transcription at multiple sites. Additionally, *dmyc* utilizes multiple polyadenylation signals to terminate transcription and promote 3′-end formation. These findings raise the possibility that modularly structured regulatory elements of the *dmyc* gene play a key role in controlling both its high expression in growing and dividing cells as well as its downregulation during differentiation.

Our computational analyses of the *dmyc* locus reveal multiple conserved sequence blocks within the noncoding regions that show reporter activity in the tested tissues. This work will provide a basis for understanding the regulatory clusters likely to be important downstream of the developmental signals previously implicated in *myc* regulation.

Materials and Methods

Cloning and sequencing of the *dmyc* gene

The RP98-2A13 BAC clone containing the *dmyc* locus (obtained from the Children's Hospital Oakland Research Institute, Oakland, CA) was triple digested with NotI/Asp718/DraIII. The resulting genomic fragment RP27, a 27-kb insert harboring the *dmyc* locus of 12.83-kb, was cloned into the 5′ NotI and 3′ ASP718 sites in the fly transforming vector pCaSpeR4 (Drosophila Genomic Resources Center, DGRC) to obtain pC-RP27. The genomic fragment in the BAC clone RP98–2A13 harboring the *dmyc* gene was sequenced at the beginning and at the end of the *dmyc* gene, each time towards the end of the *dmyc* gene with primers BAC-F and BAC-R. The 27-kb genomic fragment in the pC-RP27, including the *dmyc* gene, was sequenced at its distal site with the primer pcaF and at its proximal end with the primer pcaR, each time towards the *dmyc* gene. Sequences for all the polymerase chain reaction and sequencing primers are listed in Supplemental Table 1.

Generation of *LacZ* reporter strains

For random P-element transformation, the ready-to-use transforming vector pCaSpeR-NLSlacZ (Drosophila Genomic Resources, originally from Tummel's laboratory) was used. The pCaSpeR4-NLSlacZ vector, free from hsp70 promoter or any other regulatory elements, contains the eye marker *white* gene, the reporter NLSlacZ, SV40 poly (A) tail, and the ampicillin resistance gene. Inserts for the reporter

constructs, J2.1–J7, J1–J1.5, J8, J8.2, J9 (Figs. 5–8), J8.1, and J10 (Supplemental Figs. 3 and 4), are derived from the genomic sequences in pC-RP27 and were subcloned either into pCaSpeR-NLSlacZ and/or into the site-specific pattB-temp del LoxP reporter plasmid (Rainbow Transgenic Flies, original pUAST-attB vector was a gift from Basler's laboratory). The original vector pUAST-attB was engineered to remove the 5 × UAS-hsp70 and LoxP site sequences, followed by self-ligation to obtain the ready-to-use pattB-temp del LoxP transforming vector. pC-RP27 was digested with BamHI and run on 0.5% agarose gel to isolate the 7.9-kb BamHI-BamHI *dmyc* 5′ and 8.9-kb BamHI-BamHI *dmyc* large intron fragments. Each fragment was subcloned into pBS-SKII (+) (which was a generous gift from Oleg Georgiev, Walter Schaffner's laboratory) to obtain SKII-dmyc5 and SKII-dmIn2 intermediate plasmids. For subcloning the *dmyc* 5′ promoter/enhancer into pCaspeR4-NLSlacZ, the 7.9-kb BamHI-BamHI fragment in SKII-dmyc5 was first mutated by polymerase chain reaction to remove the approximately 800 bp open reading frame sequence and to introduce an Acc65I restriction site at the 3′ end of the fragment. Polymerase chain reaction amplification of the *dmyc* promoter/enhancer, the BAC clone RP98–2A13 served as the template, the amplified fragment was 2601 bp in size, and the amplifying primers had the names dm5E2F and dm5E2R (primers 17 and 18 in Supplemental Table 1). Polymerase chain reaction conditions in the thermal cycler were one cycle of initial denaturation at 98 °C for 30 seconds; 30 cycles of denaturation at 98 °C for 10 seconds, annealing at 60 °C for 30 seconds, and extension at 72 °C for 60 seconds; and one final extension cycle at 72 °C for 5 minutes, held at 4 °C. Plasmids SKII-dmyc5 and the polymerase chain reaction fragment were each separately digested with AatII and Acc65I (AatII is located 1550 bp downstream of the transcription start and Acc65I is in the SKII-dmyc5, 60 bp downstream of the *dmyc* fragment in the multiple cloning site). The vector SKII-dmyc5 was digested to excise an approximately 1150 bp AatII/Acc65I fragment. The polymerase chain reaction fragment was digested to isolate the approximately 296 bp AatII/Acc65I fragment. After isolation on 1.5% agarose gel, the vector and insert were ligated overnight to create plasmid SKII-dmyc5, free from open reading frame (10,035 bp in size). The 5′ NotI-Acc65I 3′ *dmyc* 5′ fragment in the plasmid SKII-dmyc5, free from open reading frame, was excised and subcloned in 5′ NotI-Acc65I 3′ linearized/dephosphorylated pCaspeR4-NLSlacZ in front of reporter lacZ to obtain the transgene J2.1. Constructs J3–J7 contain successive restriction deletions of the J2.1 insert, using either J2.1 or pC-RP27 as templates and digesting with NotI and another suitable enzyme. For the creation of J1, the 5′ BamHI-XbaI 3′ SV40 trailer in transgene J2.1 was replaced with an approximately 10.5-kb fragment from the 3′ end of *dmyc* in pC-RP27. The genomic fragment for JD was obtained by polymerase chain reaction amplification of the RP98–34B12 BAC clone (obtained from the Children's Hospital Oakland Research Institute) in two steps to create products D1 and D2 (using a high fidelity polymerase chain reaction kit, Finnzymes Inc, Lafayette, CO). The two polymerase chain reaction fragments were combined by blunt ligation (CIAP, Roche Diagnostics, Indianapolis, IN) into BlueScript, then removed and exchanged for the J2.1 insert by 5′ NotI and 3′ Acc65I digestion. For the creation of J8, SKII-dmIn2 served as the template. As for the creation of J2.1, the open reading frame sequences flanking the intron 2 sequence were first removed by the polymerase chain reaction, for which plasmid pC-RP27 served as a template. The resulting 5′ NotI and 3′ ASP718 full length intron 2 sequence, free from open reading frames, was exchanged for the J2.1 insert by 5′ NotI and 3′ ASP718 restriction digestion. The insert of J5 transgene was either fused proximal to distal relative to the 5′ end of the J8 fragment to obtain the J9 reporter construct, or it was combined distal to proximal with the 5′ end of the J8 construct to generate the J10 transgene. Reporters J1–J1.5 terminate transcription by full length or truncated forms of *dmyc* 3′ sequences; all the other constructs contain the SV40 poly (A) signal.

All BAC clones, intermediate plasmid constructs, engineered transforming vectors, and transgenes were sequenced by Microsynth AG, Balgach, Switzerland. Except for the standard primers (provided by Microsynth), all other oligonucleotides and sequencing primers were designed using the software tool DNASTAR Lasergene 9.1 module PrimerSelect and synthesized at Microsynth AG. Sequences for all the polymerase chain reaction and sequencing primers are listed in Supplemental Table 1.

Each reporter transgene was transfected into electrocompetent Max DH10B *Escherichia coli* cells (Invitrogen). Plasmid DNA for injection was isolated using a Qiagen large construct kit. Each plasmid was sequenced at the distal and proximal sites of the *dmyc* promoter sequences before injection. For the reporter studies based on random P-element insertion, embryos from genotype *y[1] w[*1118*]* were used and for the studies with the phage ΦC31 integrase transgenesis system we used embryos from fly lines carrying different attP attachment sites (both strains were from Bloomington Stock Center, Bloomington, IN). The attP fly stocks used in this study are listed in Supplemental Table 2.

X-Gal Staining Assays

For each construct, 4–15 independent transgenic lines (except the largest construct [J1] for which only two independent lines were obtained) were dissected and X-Gal staining was performed at standardized reaction conditions.[41] The incubation temperature for different tissues was as follows: discs 29 °C, embryos 37 °C, and ovaries at room temperature. For each construct, a representative disc, brain, embryo, and ovary was chosen for presentation. *dpp-lacZ* fly stocks (gifted by Dragan Gligorov, Karch's laboratory) were used as a positive control, while *y[1] w[*1118*]* and attP-fly stocks were used as negative controls.

Reverse Transcriptase Polymerase Chain Reaction Analysis

Total RNA was isolated from the sample discs (Fig. 4, c–e; Fig. 5B, b; 5C, k, l; Supplement Fig. 5g) using an aMResco phenol-free total RNA purification kit (Code N788 kit) with RNase-free DNase treatment (Promega, Basel, Switzerland) following the manufacturer's protocol. Total RNA 1 µg was reverse transcribed into cDNA in a reaction volume of 30 µL using SuperScript™ III reverse transcriptase oligo(dT) 20 primers and reverse transcription reagents from Invitrogen (Carlsbad, CA). Semiquantitative polymerase chain reactions were performed on the resulting cDNA using a high fidelity Phusion DNA polymerase kit (Finnzymes, Bioconcept, Switzerland). Polymerase chain reaction primers were designed for the *lacZ* reporter gene and Drosophila actin gene using PrimerSelect from the Lasergene software suite (DNASTAR, Madison, WI). All primers were synthesized at Microsynth. Primer sequences are indicated from 5′ to 3′ in Supplemental data, Table 1. The polymerase chain reaction conditions in the thermal cycler were one cycle of initial denaturation at 98 °C for 20 seconds; 30 cycles denaturation at 98 °C for 10 seconds, annealing at 61.3 °C for 30 seconds, and extension at 72 °C for 50 seconds; one cycle final extension at 72 °C for 2 minutes, held at 4 °C. The polymerase chain reactions, 5 µL per lane, were run on 1% agarose gel for 90 minutes at a voltage of 120.

Bioinformatics Analyses of Regulatory Elements

For defining the genomic organization of the *dmyc* gene in twelve sequenced *Drosophila* species, the bioinformatics tool DNASTAR Lasergene 9.1 MegAlign module was used to edit the sequences taken from FlyBase. We used the Lasergene 9.1 MegAlign module with the settings "Multiple Alignment, ClustalW".[42] The *dmyc* 5′ end was searched with the DNASTAR Lasergene 9.1 GeneQuest module for prediction of TATA box-Inr elements, the intron 2 region for the existence of an Inr-downstream core promoter element and the 3′ end was searched for the prediction of polyadenylation sequence motifs. The phylogenetic footprinting tools, *EvoPrinter* and *cis*-Decoder[43,44] were used to detect E-boxes, bHLH binding sites, and multiple conserved sequence blocks in the *dmyc* region common to most *Drosophila* species. The neural network genetic algorithm PROMOTER 2.0 was used to predict promoter regions using CCAAT or bHLH recognition motifs.

Results

Developmental expression patterns of *dmyc*

dmyc is synthesized in a dynamic spatial and temporal pattern during development of *Drosophila*, as determined by in situ hybridization and Northern blot analysis.[7,35] We first sought to compare the pattern of *dmyc* promoter activity from our different *dmyc* deletion constructs generated from upstream, intronic, and downstream regions, with both the endogenous pattern of *dmyc* transcription previously reported and with the pattern resulting from the *dmyc-lacZ* enhancer trap line w^{67c23} *P{lacW}*dm^{G0354}/*FM7c* (Fig. 2A). The enhancer trap line w^{67c23} *P{lacW}*dm^{G0354}/*FM7c* has

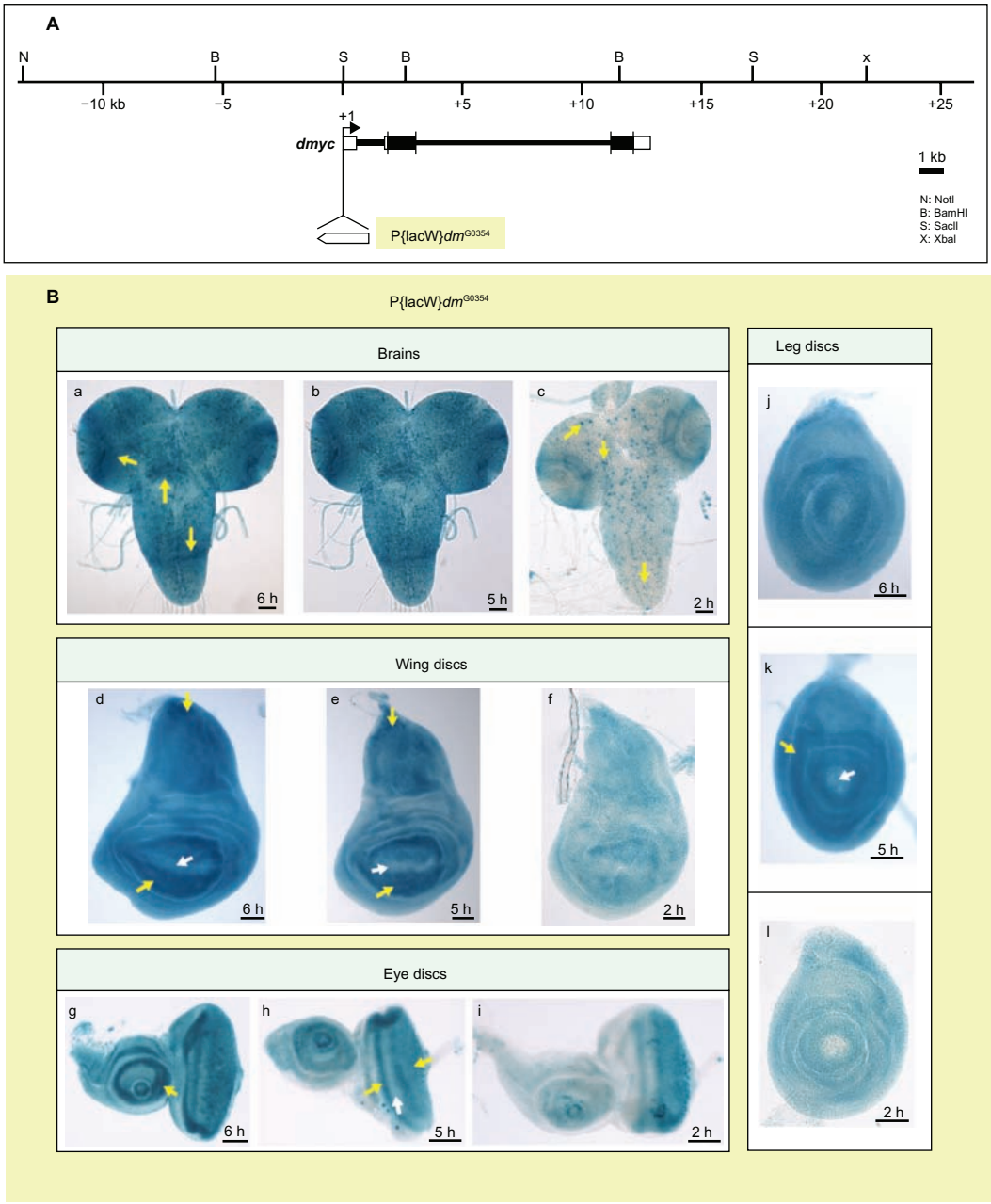


Figure 2. The *dmyc-lacZ* enhancer trap line w^{67c23} *P{lacW}*dm^{G0354}/*FM7c* reflects the endogenous expression of *dmyc* mRNA in larval brain and imaginal tissues. (**A**) Insertion site of the P-element containing reporter *lacZ* at the *dmyc* locus is shown. Breakpoints of the insertion are as follows: 3D2, X:3267141..3267197, which maps to the region 213 nucleotides upstream of *dmyc* exon 1 start site. (Bloomington Stock number and donor of the Stock: 11981, Ulrich Schaefer and Herbert Jackle). (**B**) Third instar larval brain and discs were assayed with β-gal reaction for *lacZ* expression. In all the tested imaginal discs and larval brain, *lacZ* patterning reflects the pattern reported for *dmyc* endogenous mRNA distribution (**a–c**: brain, **d–f**: wing discs, **g–i**: eye discs, **j–l**: leg discs). Details on *dmyc* patterning are explained in Discussion section.
Note: Yellow arrow indicates lacZ expression and white arrow indicates lack of lacZ activity. Staining time is indicated above the scale bar. Scale bar in (**a–l**) indicates 50 μm.

been established to give patterns of endogenous *dmyc* expression (Figs. 2B and 3) and is known to be responsive to *dmyc* regulators.[6,39,45,46] Beta-gal assays for this *dmyc-lacZ* enhancer trap line revealed ubiquitous *dmyc* promoter activity in the larval brain, with an increased level of expression in the distal and middle parts of the lobes and within dividing neuroblasts in the middle parts of the ventral ganglion (Fig. 2B, a, b, yellow arrows). Additionally, lacZ activity is restricted to a limited number of cells distributed in the two proximal halves of the hemispheres and along the ventral ganglion (Fig. 2B, c, yellow arrows).

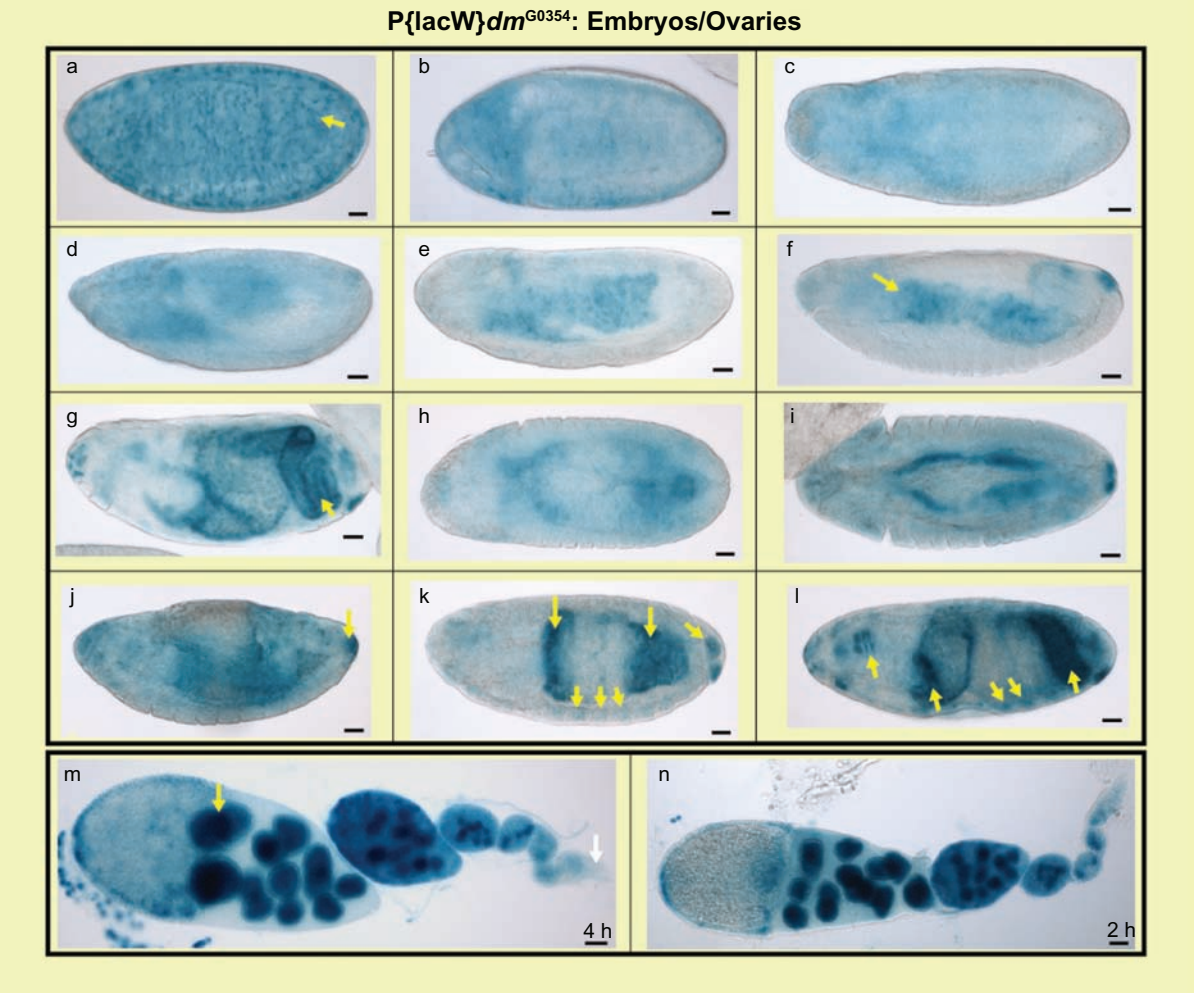


Figure 3. The *dmyc-lacZ* enhancer trap line w^{67c23} *P{lacW}*dm^{G0354}/*FM7c* produces the endogenous *dmyc* mRNA pattern during embryogenesis and in adult ovaries. In the developing embryonic tissues the *dmyc-lacZ* enhancer trap expresses *lacZ* with high similarities to *dmyc* mRNA distribution, predominantly in midgut, hindgut, pharynx, anal pad and partly in mesodermal tissues. **a–l**: embryos (embryo stages are as follows: **a, b**, stage 1–4, **c, d**, stage 6–8, **e, f**, stage 9–12, **g–l** stage 13–16). The staining in the ovaries (**m, n**) reflects *dmyc* mRNA localization, which has been reported in nurse cells, but is weakly expressed at the tip of germarium.
Note: Yellow arrow indicates lacZ expression and white arrow indicates lack of lacZ activity. Staining time for ovaries is indicated above the scale bar, embryo staining took place over-night. Scale bar in (**a–n**) indicates 50 µm.

The enhancer trap line detected lacZ activity around the wing pouch and in the notum region (Fig. 2B, d–f), anterior and posterior to the morphogenetic furrow in the eye disc and around the center of the antennal disc (Fig. 2B, g–i) , and in the center of the leg disc (Fig. 2B, j–l). dMyc antibody staining of the tissues taken from the above enhancer trap line results in the same pattern as observed for the endogenous reporter (L Quinn, personal communication).

Multiple conserved *cis*-regulatory sequences within the 40 kb *dmyc* locus

The occurrence of conserved regions and repetitive sequence motifs in noncoding DNA has been of great value for the identification and characterization of *cis*-regulatory elements. The phylogenetic footprinting tools *EvoPrinter*[43] and *cis*-Decoder[44] can be used to identify conserved sequence blocks in developmental genes. *EvoPrinter* facilitates the multialignment and rapid identification of evolutionarily conserved sequence blocks as they exist in the species of interest. The *cis*-Decoder then characterizes repeat motifs within the conserved sequence blocks and detects conserved elements among functionally related enhancers. It is important to mention that *EvoPrinter* and *cis*-Decoder do not detect polyadenylation signals or core promoter elements, such as TATA boxes.

To identify *cis*-regulatory regions for *dmyc*, a 40-kb fragment on the X chromosome harboring the *dmyc* gene of *Drosophila melanogaster* (Fig. 4A) was used as the reference DNA to align orthologous sequences from 12 sequenced Drosophilids to test for the existence of multispecies-conserved sequence blocks in the noncoding regions. We identified several putative enhancer regions with conserved sequences, including several conserved E-boxes, which are preferred Myc binding sites[47–49] of both the CACGTG and CACTTG type (Supplemental Figs. 1 and 2). Like the *c-myc* promoter,[50,51] *dmyc* has previously been shown to undergo autoregulation.[52] The identified E-box sequences may represent the basis for the previously demonstrated ability of dMyc to undergo autoregulation.[53]

Comparison between the sequences from intron 2 identified multiple clusters of conserved sequence blocks upstream of the predicted intronic promoter (Supplemental Fig. 2). In addition to two conserved E-box sequences (CACGTG and CACTTG), we identified a repeat sequence element (ATGTTGCCA) where the core TGTTGC is repeated three times (Supplemental Fig. 2). In the large, approximately 10-kb 3′-UTR region, we identified clusters of conserved sequence blocks, but no E-boxes as was the case for the intronic region (Supplemental Fig. 2).

Far upstream region determines activity of *dmyc* in late embryogenesis

As noted above, we have identified a diverse array of potential promoter regions and enhancer elements, which represent sequences that are likely to be required to achieve the patterning of *dmyc* transcription throughout development. To determine which of these domains participate in the regulation of *dmyc* expression, we generated overlapping deletion constructs spanning the *dmyc* gene and examined the activity of each fragment throughout development. The creation of overlapping deletions was based on the sequence conservation and the existence of suitable exonuclease recognition sites in the region. We decided to first test the 8-kb far upstream fragment (JD, Fig. 4B) for its ability to self-initiate transcription, as reported for the far upstream P0 promoter in human *c-myc*.[51,54,55]

No activity was detected in the tissues known to normally express *dmyc*, including larval brain and imaginal tissues (Fig. 4C, a–e), the earlier embryonic stages (Fig. 4D, f, g), and adult ovaries (Fig. 4E, k, l). Reverse transcription polymerase chain reaction on the discs shown in Figure 4C (c–e), did not detect any lacZ transcripts beyond background level in these tissues that were negative for lacZ staining (Fig. 4F, lanes –RT-e). Activity of the JD-*lacZ* reporter was confined to presumptive mesodermal tissues in body segments and the head regions of embryos during late development (Fig. 4D, h–j, arrow heads). Previous studies have reported *dmyc* activity in putative neuromuscular tissues by in situ hybridization experiments on *dmyc* endogenous mRNA. However, the enhancer trap line used as the control in this study shows only partial expression in these tissues (Fig. 3, k, l). This result suggests that the regulatory sequences in this far upstream region are only sufficient for activation of *dmyc* expression during late embryogenesis. Due to its far upstream position, we have dubbed the remote *cis*-regulatory elements in this region "P0 (putative)", analogous to the human *c-myc* P0 promoter.

Proximal upstream region controls *dmyc* expression in larval and adult female tissues

The two main promoters required for activation of mammalian *c-myc* transcription, P1 and P2, are located in the 5′-UTR, where the majority of transcripts are initiated.[51,56–58] Thus, our first efforts were directed toward identifying regulatory elements in the *dmyc* 5′-UTR, capable of activating endogenous patterns of *dmyc* gene expression during development. Analysis of the 7.2-kb region between the *dmyc* translation unit and the 3′ end of the JD transgene (Fig. 5A) revealed that when most of the conserved sequences were removed by successive distal to proximal deletions of the J2.1 fragment, a loss of reporter expression in both the brain and imaginal discs was observed (Fig. 5B and C). Interestingly, by staining embryonic and ovarian tissues taken from the largest construct (J2.1) and the smallest deletion (J7), we identified a proximal promoter region within the 5′-UTR, which recapitulated almost all aspects of early embryonic and ovarian *dmyc* expression (Fig. 5D). This observation suggests that the 2-kb 5′-UTR and a further 100 bp upstream of the 5′-UTR region contain regulatory elements important for embryogenesis and

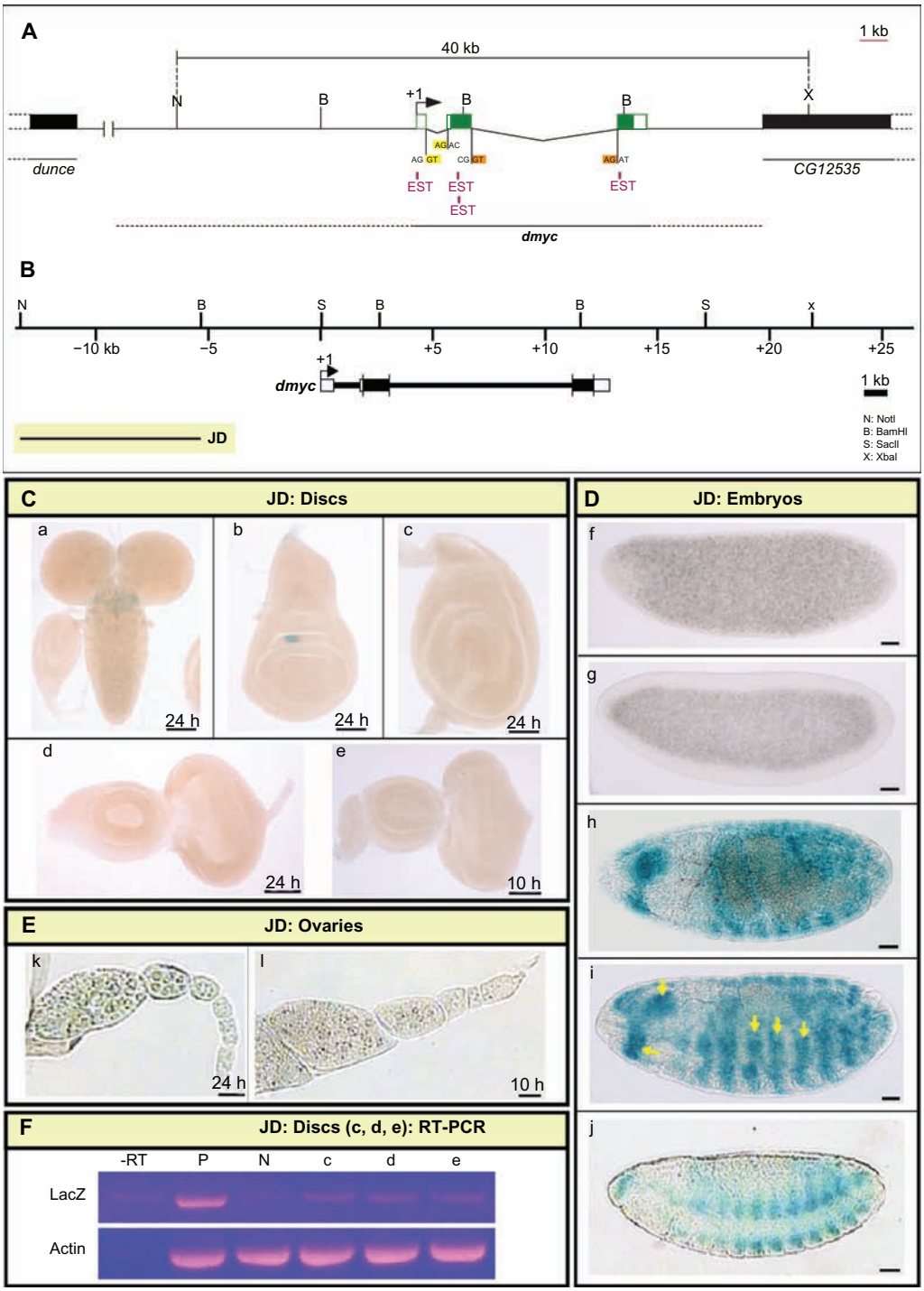


Figure 4. Distal upstream regulatory region of *dmyc* is active in late embryogenesis. (**A**) 40-kb genomic fragment on the chromosome X of *D. melanogaster* between nucleotides −3253633 to +3293632, harboring the *dmyc* locus (nucleotides −3267216 to +3280049), is shown above. The ESTs (pink color) represent portions of the gene transcripts (not all ESTs indicated). The 5´ flanking gene *dunce*, also annotated as *CG*32498, is located ~ 33-kb upstream of the examined 40-kb sequences. The 3´ flanking gene, *CG*12535, is located ~ 6.5-kb to the 3´-end of the *dmyc* gene (nucleotide +3280049). The restriction sites at the genomic locus used in this study, splice donor sites GT (GU on mRNA) and splice acceptor sites AG (highlighted in yellow and orange) are indicated. (**B**) Locations of the *dmyc* and the transgene JD (8-kb in size) relative to the 40-kb genomic locus are shown. (**C**) Activity of the JD transgene in the brain and imaginal tissues. The fragment was neither active in the brain (**a**) nor in imaginal discs (**b**: wing, **c**: leg, and **d**, **e**: eye discs). (**D**) JD fragment was not active in early embryos (**f, g**), however, the fragment recapitulated expression of *dmyc* in late embryogenesis (**h–j**: stage 13–16). (**E**) The staining in ovaries (**k, l**) showed no activity for JD fragment. (**F**) RT-PCR on JD discs (**c, d, e** shown in C), detects no lacZ transcripts beyond background level. Controls used are as follows: –RT: negative control with no transcriptase (brains from Fig. 5B, b); P: Fig. 5B, b, positive control; N: *y[1] w[*1118*]* leg discs, negative control; Actin: *Drosophila* actin as internal control.
Note: Yellow arrow indicates lacZ expression. Staining time for discs and ovaries were either 10 or 24 hours, as indicated (**C, e** one representative for 10 hours discs). The embryos were stained over-night. Scale bar in (**a–j**) indicates 50 µm.

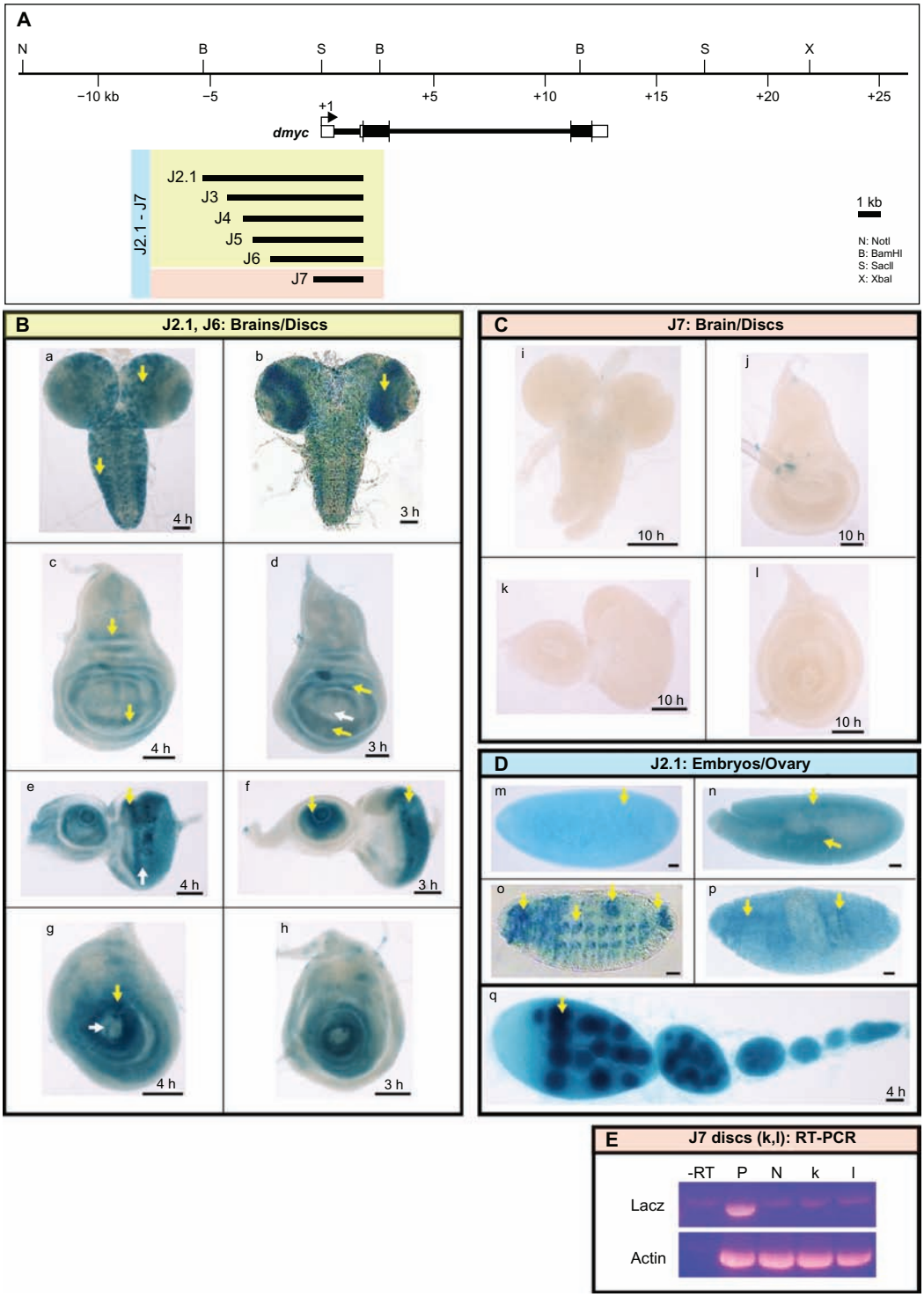


Figure 5. *dmyc* 5′regulatory region contains separable elements. (**A**) The 40-kb genomic locus and the location of the *dmyc* is shown at the top. A fragment of 7.159-kb in size (J2.1) (lacking ORF sequences) from the *dmyc* 5′ region and its deletion restriction fragments were tested for their ability to drive reporter activity in transgenic flies. Colors of the deletion fragments refer to their expression patterns shown in panels **B**, **C** and **D**. (**B**) J2.1 transgene (**a**, **c**, **e**, **g**) and all of its successive deletions up to 2.523-kb in the J6 construct (**b**, **d**, **f**, **h**) were able to express the reporter in a *dmyc* manner in the brain and discs (**a**, **b**: brain; **c**, **d**: wing discs; **e**, **f**: eye discs; **g**, **h**: leg discs). (**C**) Further truncation down to 1.923-kb of J7 resulted in a loss of expression in the brain and discs (**i**–**l**). (**D**) Shown are embryos and ovary taken from transgenic animals carrying J2.1 transgene. J2.1 retained the expression in the embryos (**m**–**p**; embryo stages: **m**: 2–5; **n**: 11–12; **o**, **p**: 13–16) and ovary (**q**). Except for **4D** o, the smallest construct J7 recapitulated the *dmyc* like expression in all the other embryonic stages and in ovary. (**E**) RT-PCR on J7 discs (k, l shown in C), detects no lacZ transcripts beyond background level. Controls used are as follows: –RT: negative control with no transcriptase (brains from Fig. **5B**, b); P: Fig. **5B**, b, positive control; N: *y[1] w[*1118*]* leg discs, negative control; Actin: *Drosophila* actin as internal control.
Note: Yellow arrow indicates lacZ expression and white arrow indicates lack of lacZ activity. Staining time for discs and ovaries is indicated, and embryos were stained over-night. Scale bar in (**a–q**) indicates 50 μm.

oogenesis, but insufficient for patterning in the larval brain or imaginal tissues. It further suggests that the sequences 5.1-kb upstream of 5′-UTR contain tissue-specific enhancer regions responsible for the endogenous pattern of *dmyc* expression in the brain and larval imaginal tissues (Fig. 5B).

Full length intron 2 with downstream core promoter element activated in *dmyc* pattern

In most protein-coding genes of *Drosophila*, there is a downstream core promoter element that functions cooperatively with an initiator to facilitate the binding of transcription factors in the absence of a TATA box.[59] The high throughput expression data for the *D. melanogaster* transcriptome, generated by tilling arrays, has shown that most of introns are transcriptionally active within the early hours of development at specific time points.[60] Scanning of the intron 2 region with *cis*-Decoder initially revealed clusters of multiple conserved sequence blocks common to most *Drosophila* species. A search using the Lasergene GeneQuest module identified a downstream promoter region with no TATA box, comparable with the *Drosophila* consensus sequence (Supplemental Fig. 2).

To analyze the intragenic region of *dmyc*, we generated transgenic animals carrying the 8-kb intron 2 full length fragment (J8), or the subfragment (J8.2) lacking the 2-kb distal sequences (Fig. 6A). The J8 and the J8.2 transgenes both showed *dmyc*-like expression in imaginal discs and brain tissues (Fig. 6B, a–h). During embryogenesis, both transgenes are active in the early and later stages (Fig. 6B, i–n). However, only the J8.2 transgenic animals, which lack the 2-kb upstream sequences, showed strong patterning of *dmyc* in mesodermal tissues (Fig. 6B, o, p). This suggests that the upstream sequences might contain elements that influence *dmyc* expression in presumptive mesodermal tissues. Unlike the difference observed in embryonic activity, the expression of LacZ in ovarian tissues remains unchanged for both transgenes (Fig. 6B, q). The reversed full length intron 2 fragment fused with its 5′ end to the 5′ end of the reporter gene (subfragment J8.1 in Supplemental Fig. 3A), causes the abolition of reporter activity in virtually all tested tissues (Supplemental Fig. 3B, a–g). This observation suggests that the core promoter element only functions unidirectionally. The experimental identification of binding sites for transcriptional regulators in this region remains unresolved.

Analysis of 3′ *dmyc* sequences reveals multiple poly (A) sites

In addition to understanding transcriptional activation of *dmyc*, we are interested in deciphering the mechanism of transcript termination. In animal cells, there is a hexanucleotide-sequence of AAUAAA (occasionally AUUAAA) that is located 10–35 nucleotides upstream of the polyadenylation signal.[61,62] In addition, there is a U/GU-rich region (cleavage stimulation factor binding site) 14–70 nucleotides downstream of the polyadenylation signal.[63] We searched for polyadenylation signals in the dmyc 3′ end comparable with the consensus sequence using the Lasergene GeneQuest module with a threshold of 100% (no errors allowed). The threshold for the T-rich region was set at 80%. Analysis with the AATAAA sequence identified two potential polyadenylation signals, highly similar to the polyadenylation consensus sequence found in animal cells.[61] The first A at the 5′ end of the polyadenylation signal, poly (A)1, corresponds to nucleotide +3761 downstream of the 3′-UTR and the first A at the 5′ end of poly (A)3 corresponds to nucleotide +5952 relative to the 3′-UTR. A search for ATTAAA sequences on the DNA sense strand detected poly (A)2, with a high degree of homology to the consensus sequence. The first A at the 5′ end corresponds to nucleotide +5245 downstream of the 3′-UTR. In order from proximal to distal relative to the transcription start, we named them poly (A)1, poly (A)2, and poly (A)3 (Fig. 7A). To test the detected polyadenylation signals separately, we performed successive deletions from distal towards proximal on the 10.3-kb full length construct (J1, Fig. 7A). The activity of each transgene was invariably high in all the examined tissues (Fig. 7B and C), suggesting that more than the three predicted polyadenylation signals are involved in the formation of the 3′ end of the *dmyc* gene. This finding is consistent with the observation that most developmentally active genes use more than one poly (A) signal to terminate transcription.[64] Direct determination of the efficiency of utilization of polyadenylation signals on the 3′ end of *dmyc* transcripts remains to be resolved.

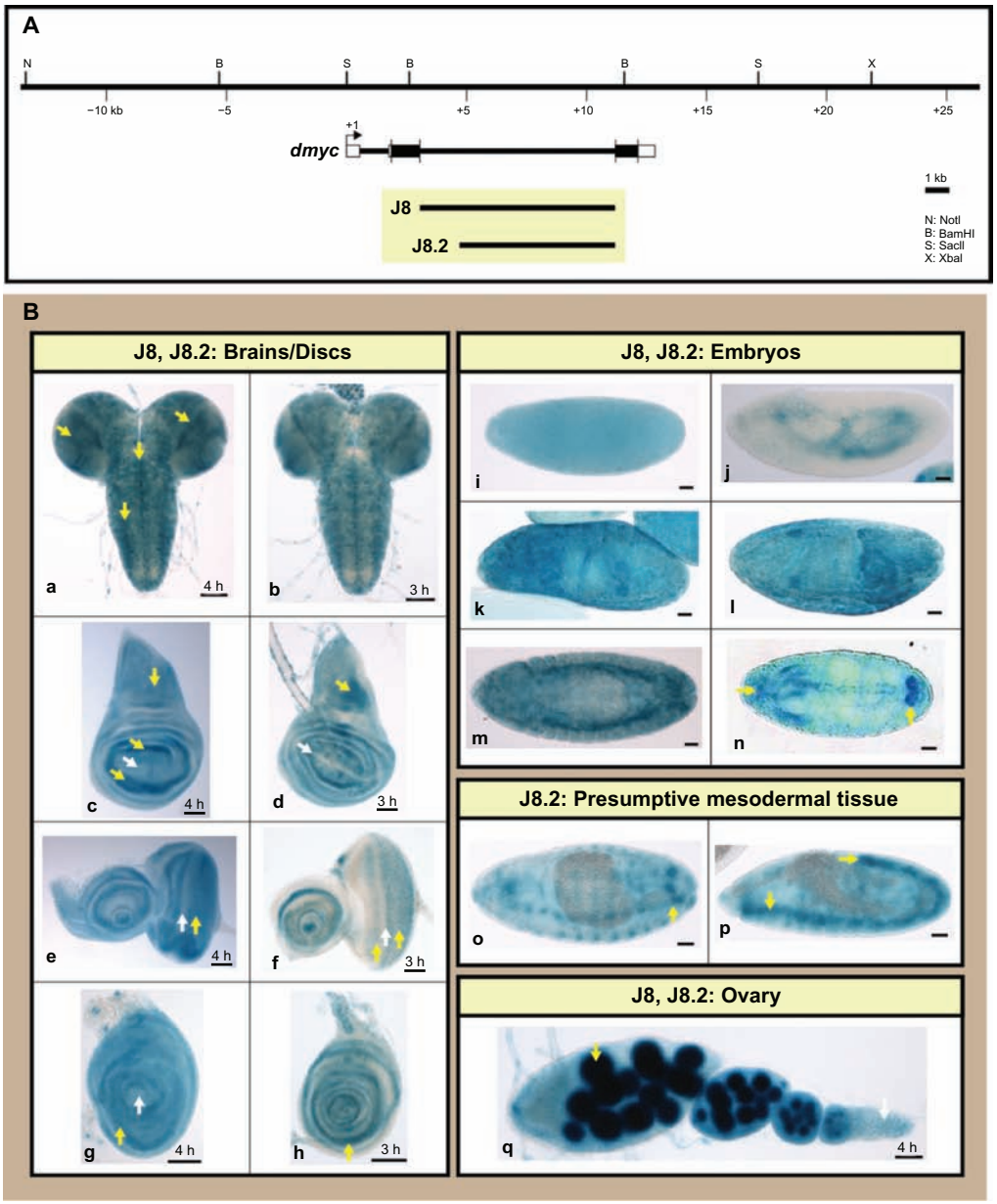


Figure 6. Analysis of the *dmyc* intron 2 identifies a downstream core promoter element, DPE. (**A**) The 8-kb intron 2 full length sequence in the J8 transgene and its truncation, J8.2 (6.1-kb) and their relative location with respect to the *dmyc* locus and the genomic organization are shown. (**B**) The full fragment J8 and its derived sub-fragment were assayed in different tissues during early developmental stages. Both transgenes (J8: **a**, **c**, **e**, **g**; J8.2: **b**, **d**, **f**, **h**) were capable of reflecting *dmyc* expression in the brain (**a**, **b**) and discs (**c**, **d**: wing; **e**, **f**: eye; **g**, **h**: leg). The transgenes J8 and J8.2 both express *lacZ* in the early and late embryogenesis (**i**–**n**) (embryo stages: **i**, **j**: 2–6; **k**–**n**: 9–13). However, only the shorter transgene J8.2 shows expression in mesodermal tissues (**o**, **p**; embryo stages: 12–15). In the ovary (**q**) both transgenes are active.
Note: Yellow arrow indicates lacZ expression and white arrow indicates lack of lacZ activity. Staining time for discs and ovary is indicated, and embryos were stained over-night. Scale bar in (**a**–**q**) indicates 50 μm.

Effect of Upstream Promoter Region on Downstream Core Promoter Element

It has been shown that parameters such as stoichiometry, affinity, spacing, and arrangement of binding sites within the *cis*-regulatory regions influence the output of the transcriptional regulatory sequences.[65] For example, depending on the total number of binding sites for a certain regulator in the same promoter region, stage-specific and tissue-specific expression patterns can be achieved for the same gene during development.[65]

Here we have shown that truncations of conserved sequence blocks and putative regulatory protein binding sites at the *dmyc* proximal 5′ promoter resulted

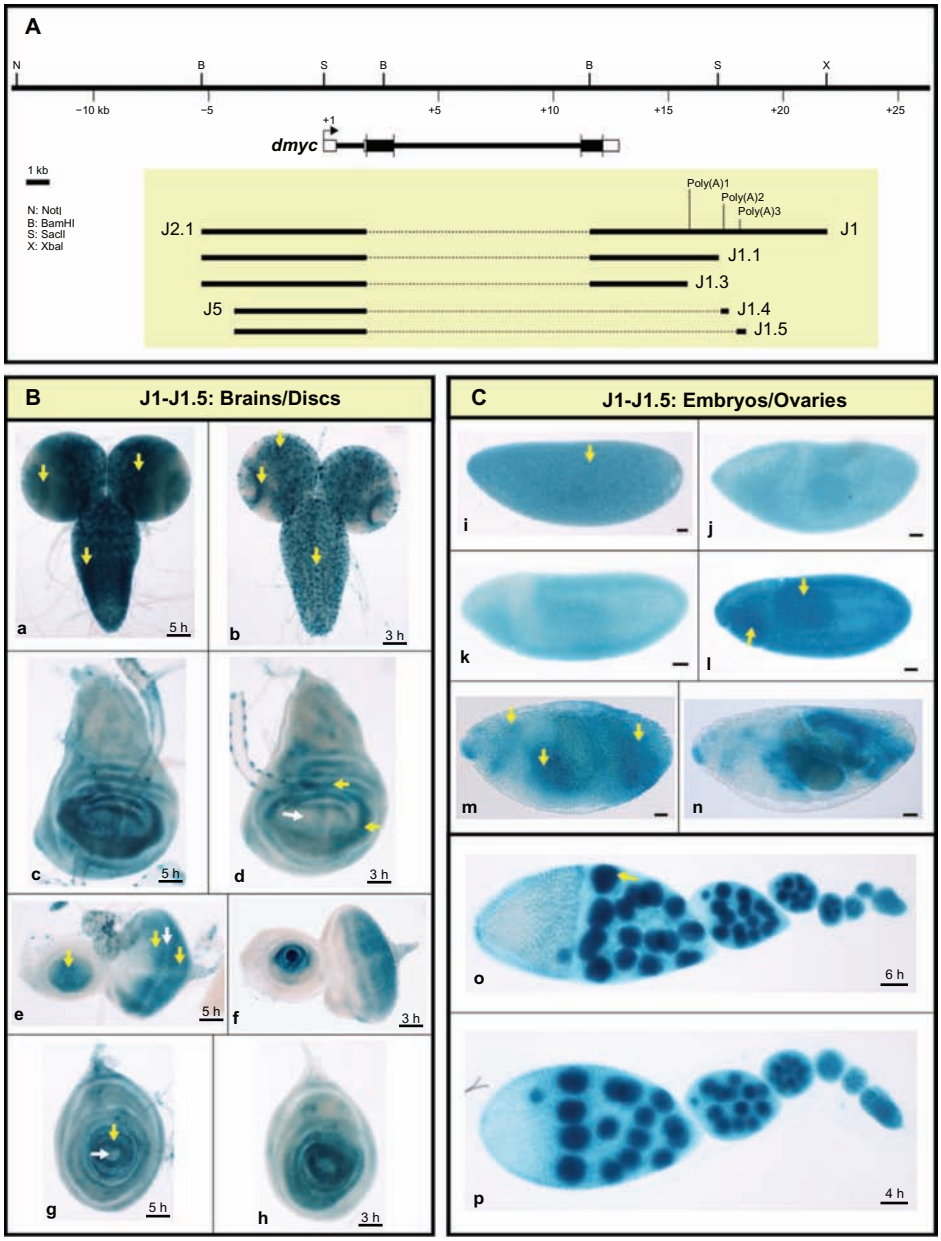


Figure 7. Analysis of the *dmyc* 3′ region reveals multiple functional polyadenylation sites. (**A**) The *dmyc* gene and its location in the genomic region, the 10.3-kb full length construct (J1) and different deletions are shown. In the J1 fragment three potential Poly (A) signals with a high degree of homology to the consensus polyadenylation signal were detected computationally. The transgenes J1, J1.1 and J1.3 are under the control of the full length 5′-promoter (see also Fig. 5 part A). The constructs J1.4 and J1.5 contain the promoter of J5 in Figure 5. J1.1 only contain Poly (A)1, J1.3 does not include the predicted Poly (A) sites, J1.4 contains Poly (A)2 and J1.5 contains Poly (A)3. (**B**) All the constructs J1-J1.5 are capable of mediating regulated expression of the reporter in a manner similar to *dmyc* expression in the third instar larval brain (**a**, **b**) and discs (**c**, **d**: wing; **e, f**: eye; **g**, **h**: leg). (**C**) All the transgenes are active during different stages of embryogenesis (**i**–**n**; embryo stages: i: 2–5; **j**–**l**: 9–11; **m**, **n**: 12–15) and in ovaries (**o**, **p**).
Notes: Yellow arrow indicates lacZ expression and white arrow indicates lack of lacZ activity. Staining times for discs and ovaries are indicated above the scale bar, and embryos were stained over-night. Scale bar in (**a**–**p**) indicates 50 μm.

in a loss of expression in larval brain and different imaginal discs (Fig. 5C). Conversely, removal of upstream sequences in the intron 2 region intensified expression in embryonic mesodermal tissues (Fig. 6B, o, p).

We tested the effect of the combination of these two regulatory regions on reporter activity. Fusion of the fully functional 5′ deletion (J5) inframe to the full length intron 2 transgene (J8) to generate the fusion promoter (J9, Fig. 8A), had no effect on the activity of the reporter in the tested tissues (Fig. 8B). However, the reversed (J5) transgene, fused with its 5′ end to the 5′ end of the (J8) fragment (Supplemental Fig. 4A), caused partial attenuation

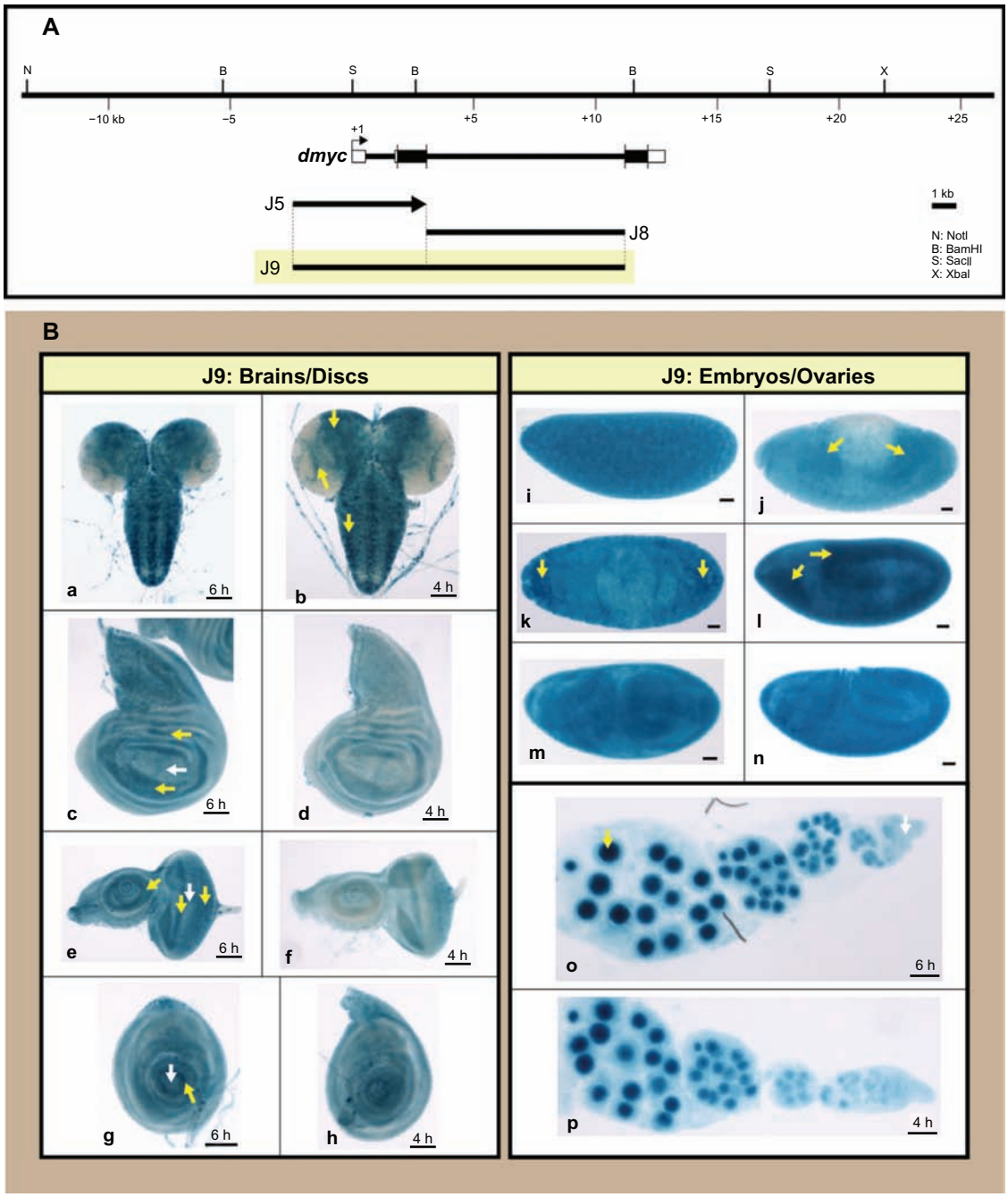
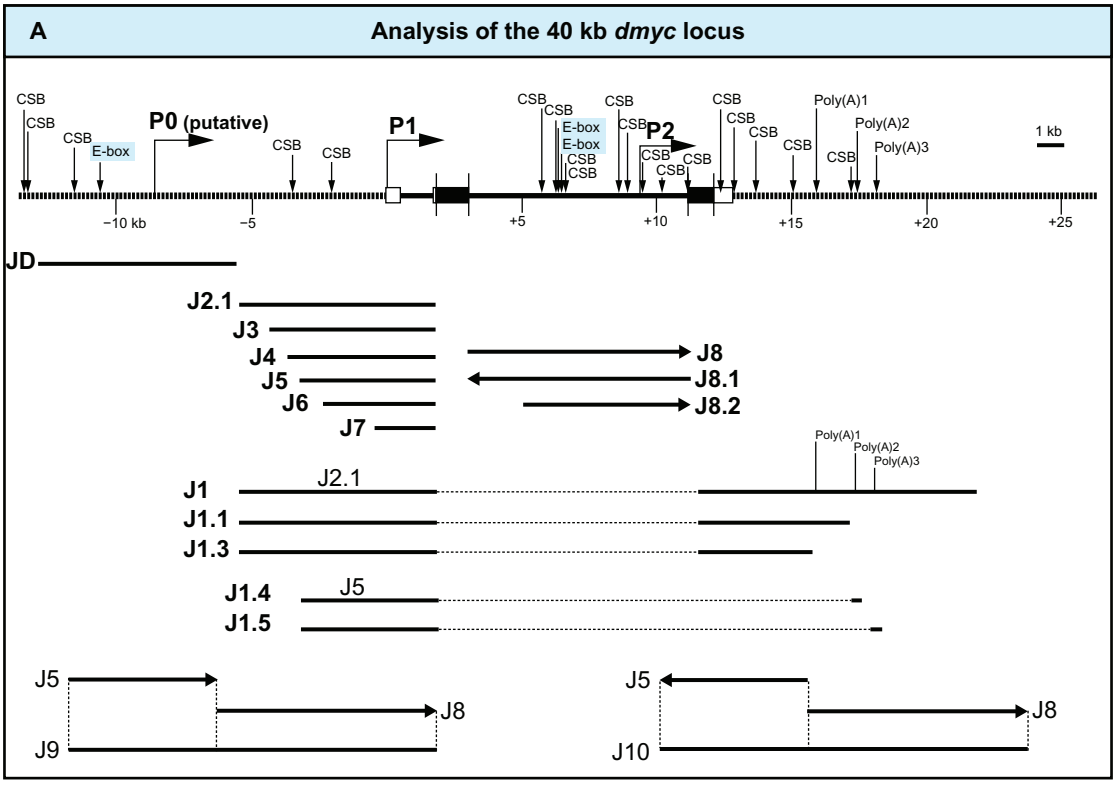


Figure 8. J9 fusion construct (5′ J5 promoter fused to the intronic J8 promoter) has no effect on reporter expression. (**A**) The ~ 3.8-kb J5 promoter, the 8-kb intron 2 in J8 and their fusion product J9 are shown in respect of their relative location to the *dmyc*. (**B**) The ~ 11.1-kb chimeric restriction fragment J9 did not show any alteration in the expression of the *lacZ* reporter in larval tissues (**a**, **b**: brain; **c**, **d**: wing discs; **e**, **f**: eye discs; **g**, **h**: leg discs), embryos (**i**–**n**; embryo stages: **i**: 1–4; **j**–**l**: 9–11; **m**, **n**: 12–15) or ovaries (**o**, **p**).
Notes: Yellow arrow indicates lacZ expression and white arrow indicates lack of lacZ activity. Staining times for discs and ovaries are indicated above the scale bar, and embryos were stained over-night. Scale bar in (**a**–**p**) indicates 50 μm.

of *dmyc* embryonic activity (Supplemental Fig. 4B). The mode of action of enhancer elements responsible for embryonic development was dependent on the spacing and arrangement of binding sites, with the activity in imaginal tissues and ovaries remaining unaffected.

Discussion

The dynamic expression of *dmyc* is initiated from multiple transcription start sites, as summarized in Figures 9 and 10. Tight regulation of dMyc is crucial for cell growth and division during the early phases of development and cell fate specification. In third instar



B

Tissue / Construct	Brain	Wing	Eye	Leg	Early embryo	Late embryo	Ovary
JD	No	No	No	No	No	Yes	No
J2.1	Yes	Yes	Yes	Yes	Yes	Yes	Yes
J3	Yes	Yes	Yes	Yes	Yes	Yes	Yes
J4	Yes	Yes	Yes	Yes	Yes	Yes	Yes
J5	Yes	Yes	Yes	Yes	Yes	Yes	Yes
J6	Yes	Yes	Yes	Yes	Yes	Yes	Yes
J7	No	No	No	No	Yes	Yes	Yes
J8	Yes	Yes	Yes	Yes	Yes	Yes	Yes
J8.1	No	No	No	No	No	No	No
J8.2	Yes	Yes	Yes	Yes	Yes	Yes	Yes
J9	Yes	Yes	Yes	Yes	Yes	Yes	Yes
J10	Yes	Yes	Yes	Yes	Yes	Yes	Yes
J1-J1.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Figure 9. Summary of the computational analyses of the *dmyc* locus and reporter activity studies in the tested tissues. (**A**) Top: The 40-kb *dmyc* locus was scanned with *EvoPrinter* and *cis*-Decoder to detect multiple conserved sequence blocks (CSB) and E-boxes within the noncoding DNA sequences. The neural network genetic algorithm PROMOTER 2.0 predicted three potential promoter regions, P0 (putative), P1 and P2. The scanning with Lasergene module GeneQuest detected three potential polyadenylation signals (Poly(A)1, Poly(A)2, Poly(A)3). Bottom: Deletion constructs derived from the 40-kb locus and their names are given. (**B**) The table shows the summary of the activity of l*acZ* reporter constructs in the tested tissues (For details see Fig. 4–Fig. 8).

larvae, *dmyc* mRNA is detected around the wing pouch and in the notum.[66] *dmyc* activity is absent from the cell cycle-arrested nonproliferating cells that surround the dorsoventral boundary in the wing pouch.[66] In the eye disc, *dmyc* is synthesized in the proliferating cells posterior and anterior to the morphogenetic furrow, but not in the cell cycle-delayed cells of the morphogenetic furrow.[7,35] In antennal discs, *dmyc* mRNA is mainly detected in the central ring of proliferating cells.[7] In the leg disc, endogenous *dmyc* is expressed around the middle of the disc, with the central cells lacking *dmyc* activity.[35] Maternal transcripts

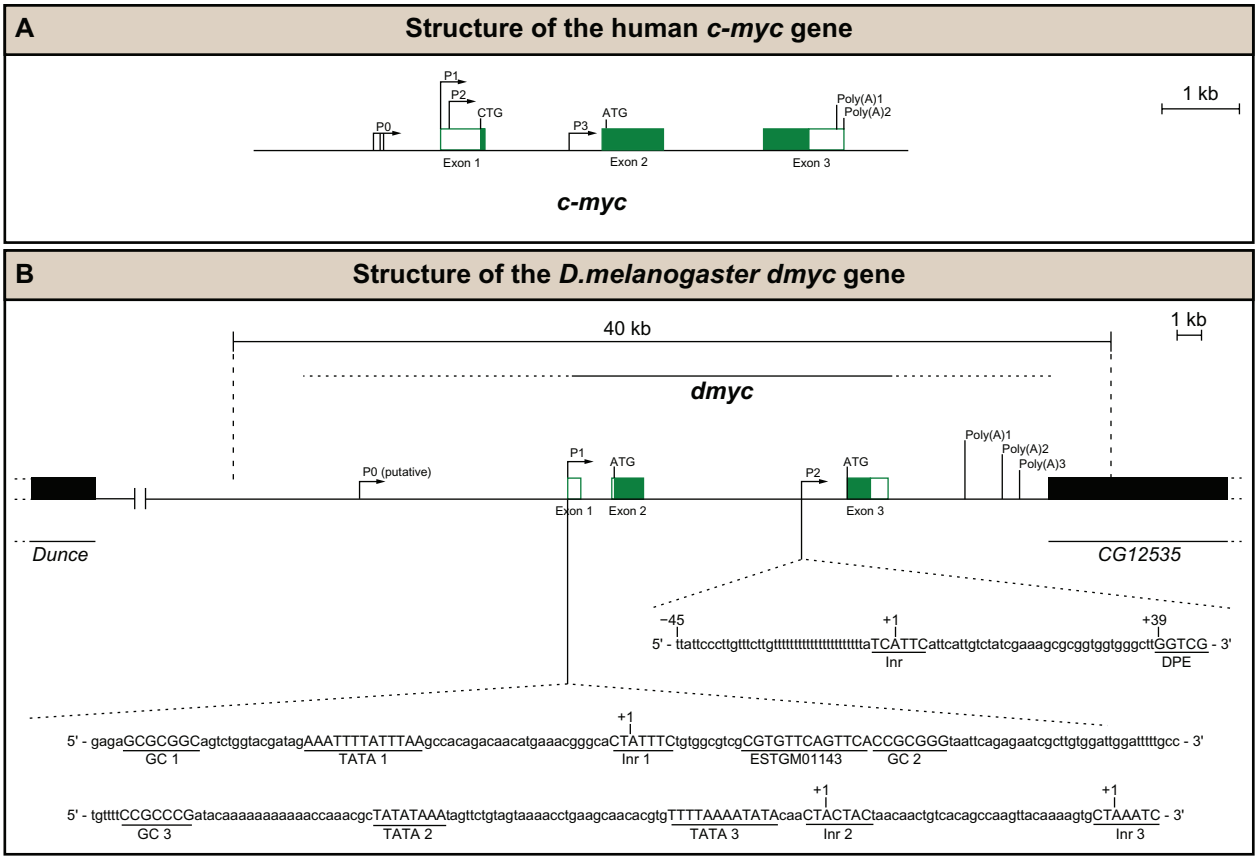


Figure 10. During development, the activity of the *dmyc* regulatory region is integrated by multiple transcription start sites. (**A**) Human *c-myc* gene contains three exons (white boxes indicate noncoding regions, the green boxes represent coding regions), four promoters (P0, P1, P2, P3), two major translation start sites (CTG, ATG) and two polyadenylation signals (Poly(A)1, Poly(A)2). (B) *Drosophila myc* gene contains three exons (white boxes noncoding, the green boxes coding regions), two main promoters (P1, P2), one putative promoter (P0), two major translation start sites (ATG, ATG) and three potential polyadenylation signals (Poly(A)1, Poly(A)2, Poly(A)3). The specified regulatory elements are plotted on the basis of our experimental data (see Results section) and the computational analyses of the *dmyc locus*. The sequences of the major elements for the promoters P1 and P2, shown at the bottom, were identified by search with Lasergene GeneQuest module.

are detected in the nurse cell cytoplasm of the adult ovary, and are subsequently dumped into the oocyte, and can be detected in early embryos.[7] Zygotically derived transcripts can be detected during preblastoderm, with the highest level of *dmyc* mRNA in the anterior and posterior termini. At later stages, *dmyc* mRNA can be detected in the presumptive mesoderm along the ventral midline. During germ band extension, *dmyc* expression intensifies in the mesoderm and endoreplicating cells of the midgut, where expression continues until mid-embryogenesis.[7] The majority of developmental genes achieve patterning via large noncoding regulatory regions containing numerous *cis*-regulatory elements and other diverse regulatory sequences.[67,68] Thus, we set out to determine whether the pattern of *dmyc* expression might be similarly regulated. First, using computational comparative searches of the 40-kb region spanning the *dmyc* gene, we detected multiple conserved sequence blocks. Our subsequent analysis of the *dmyc-lacZ* reporter constructs, containing all of the conserved sequence blocks, suggested that they were transcriptionally active and generated similar patterns of reporter activity as that described for the endogenous *dmyc-lacZ* enhancer trap.[6,39,45,46]

Dissection of the *dmyc* promoter using constructs spanning defined domains of the *dmyc* gene revealed the regions likely required for tissue-specific patterning of *dmyc* transcription. For instance, the *lacZ* reporter for the far upstream 8-kb fragment produced an expression pattern restricted to late embryogenesis in body segments and in presumptive neuromuscular tissues (Fig. 4D). In silico analysis revealed possible regulatory motifs in this region, including core promoter elements and conserved sequence blocks. Regulation by these *cis* elements may be required during embryogenesis,

where *dMyc* is required to specify neuronal fate and facilitate neuroblast proliferation[69] and in control of mesodermal fate determination.[70] In light of this finding, further analysis of *dmyc* transcriptional regulation in this region in response to developmental signals will be of great interest.

Analysis of the 5′ d*myc-lacZ* deletion construct, containing intron 1, the 5′-UTR, and 100 bp upstream of the predicted transcription start site (J7), revealed that this minimal region was sufficient to give reporter activity in a *dmyc*-like pattern in both ovarian nurse cells and in the embryo, but not in larval tissues (Fig. 5C and D). Therefore, we inspected the region extending from nucleotide 100 upstream of the 5′-UTR to nucleotide +187 for initiator consensus sequences. In most mammalian protein-coding genes, there is a TATA box located 25–30 bp upstream of the transcription start site, an initiator element (Inr) overlapping the start site,[71–74] and/or a GC-box (SP1 binding site) 60–100 nucleotides upstream of the transcription start site.[75–77] Experiments with vertebrate cell lines and *Drosophila* embryonic extracts have revealed strict conservation of the Inr consensus sequence, Py Py A_{+1} N T/A Py Py among vertebrates and invertebrates.[73,78]

Analysis of J7 (the region 1914 bp upstream of the predicted translation start site) and the expressed sequence tag (ESTGM01143; beginning at 1812 bp upstream of the translation start) revealed that the 102 bp sequence between the 5′ end of the expressed sequence tag and the 5′ end of the J7 genomic sequence contains a perfect Inr consensus sequence, a TATA box and a GC box (Fig. 10B). We named it the GC box/TATA box/Inr promoter 1 (P1, Fig. 10B). The TATA box is located 39 bp upstream of A_{+1} in the initiator element. Previous reports have shown that occurrence of a TATA box 25–30 bp upstream of the Inr in the same core promoter leads to cooperation between the two elements to enhance promoter strength.[73] Although the distance of 39 bp in the *dmyc* promoter is on the edge of this optimum, cooperation between the TATA box and the Inr element has been shown to extend up to 90 bp in yeast promoters (W Schaffner, personal communication). Together, this suggests that the predicted regulatory elements (TATA box, Inr, and GC box) correspond to the *cis*-regulatory elements that may be responsible for correct developmental *dmyc* expression in larval tissue, the embryo, and the ovary. In addition to the TATA box, Inr, and GC box, two putative TATA boxes (TATA2, TATA3), two GC boxes (GC2, GC3), and two Inr elements (Inr2, Inr3), as shown in Figure 10B, were identified 180 bp downstream of the expressed sequence tag start site. The putative Inr2 element shows one deviation from the Inr consensus sequence at position +4, but the critical positions for Inr activity are +1 and +3, and single bp substitutions at the −2, −1, +4, and +5 positions can still produce an active Inr,[79] suggesting the second Inr element might be functional. The putative Inr3 element shows no deviation from the Inr consensus sequence, suggesting that Inr3 might be as functional as Inr1. TATA box 2 is located 57 bp upstream of A_{+1} in the Inr2 element and TATA box 3 is located 55 bp upstream of A_{+1} in the Inr 3 element, both distances less than 90 bp in yeast promoters.

Most developmentally expressed genes contain separable *cis* regulatory units, which allow patterned expression for tissue-specific roles.[68,80] Indeed, previous work has suggested that both *Drosophila*[36] and mammalian[51] *myc* transcription is also regulated via intronic promoter sequences. In support of these findings, we demonstrated that the J8 transgene, which contains just the intron 2 sequence of the *dmyc* gene, results in *lacZ* reporter activity in all tissues examined (Fig. 6B). Thus we searched for Inr and downstream promoter elements, a sequence motif common to all *Drosophila* downstream promoters in this region. Most protein-coding genes of *Drosophila* contain a downstream core promoter element that functions cooperatively with an initiator to facilitate the binding of transcription factors in the absence of a TATA box.[59,81] A search for consensus *Drosophila* Inr (T-C-A_{+1}-G/T-T-T/C) and downstream core promoter elements (G-A/T-C-G) using DNASTAR Lasergene 9.1, GeneQuest module, revealed the presence of downstream promoter sequence motifs comparable with the *Drosophila* consensus Inr/downstream core promoter element.[59] Sequence motifs were typed in GeneQuest "type in pattern function" and searched for with a threshold of 100% (no errors allowed). The Inr and downstream promoter element motifs at the 3′ end of the intron 2 DNA sequence (Fig. 10B) met all the strict criteria for such elements, in that the Inr sequence motif (T-C-A_{+1}-T-T-C) does not deviate from the consensus, the downstream core promoter element

(G-G-T-C-G) is identical to the core consensus, the spacing between the downstream core promoter element and the Inr (34 nucleotides) is appropriate, and a G nucleotide is correctly positioned between the Inr and the downstream core promoter element (Fig. 10B).

Post-transcriptional 3′ end formation or polyadenylation of the mRNA precursor is a crucial step in mRNA maturation, in which most eukaryotic mRNAs acquire a poly (A) tail at their 3′ ends to promote transcription termination,[62] transport of the mature mRNA from the nucleus,[82] and to enhance the translation and stability of mRNA.[83] Analysis of the entire *dmyc* 3′-UTR for polyadenylation signals and polyadenylation sites revealed three potential consensus sequences, i.e., poly (A)1, poly (A)2, and poly (A)3, which are all capable of terminating transgene transcription. In addition, the 4.362-kb DNA sequence upstream of poly (A)1 at the *dmyc* 3′ end leads to reporter activity in the pattern predicted for *dmyc* expression. Therefore, the *dmyc* gene would be predicted to produce different transcripts with shorter and longer lengths, consistent with the previous analysis of *dmyc* mRNA in which genomic probes derived from this region revealed three alternative transcripts.[7] Comparative analysis of the 3′ end of *dmyc* across 12 sequenced *Drosophila* species revealed multiple conserved sequence blocks in this region. Given that *c-myc* is regulated at the level of mRNA stability[84] via conserved sequences in its 3′-UTR,[85] it will be of interest in the future to determine whether the stability of *dmyc* transcripts depends upon the presence of regulatory domains in its 3′-UTR. The conserved sequence blocks may contain potential microRNA target sites to serve for post-transcriptional modifications, as is the case for the majority of developmental control genes.[86,87] Indeed, *dMyc* has an active role in microRNA biology,[87,88] although regulation of *c-myc* by microRNAs has been reported,[89–92] the evidence for direct regulation of *dmyc* requires investigation. This work provides a starting point for investigating the putative microRNA binding sites and the mechanisms for the interactions between these motifs and their targets.

Because c-Myc is a potent mitogen, the level of *c-myc* transcription must be tightly regulated. Myc transcription responds to developmental signaling molecules,[33,93,94] which are likely to modulate the complement of a wide variety of transcription factors at the *myc* promoter.[95,96] The evidence presented in this work reinforces the idea that *dmyc* represents a tightly and dynamically regulated gene. Further genetic studies combined with genomic approaches will be required to identify the molecular mechanism controlling *dmyc* transcription via the regulatory elements identified here.

Acknowledgements

We thank W McGinnis and D Bopp for critical reading of the manuscript, and E Kubli and E Künnemann for the original editing of the manuscript. The plasmid BlueScript was a generous gift from O Georgiev at the W Schaffner laboratory, and the pCaSpeR4 vector containing lacZ sequences was obtained from Drosophila Genomic Resources. The enhancer trap line *dpp-lacZ* used as the positive control was a gift from D Gligorov at the F Karch laboratory, University of Geneva. All the other fly strains were received from Bloomington Stock Center and Kevin Cook assisted JK with the fly genetics. The injections were done at Rainbow Transgenic Flies Inc, CA. The photographs were taken at the Center for Microscopy, University of Zurich. This study was partly supported by a scholarship from the "Bildungsdirektion Kanton Zürich, Abteilung Stipendien" to JK.

Disclosures

Author(s) have provided signed confirmations to the publisher of their compliance with all applicable legal and ethical obligations in respect to declaration of conflicts of interest, funding, authorship and contributorship, and compliance with ethical requirements in respect to treatment of human and animal test subjects. If this article contains identifiable human subject(s) author(s) were required to supply signed patient consent prior to publication. Author(s) have confirmed that the published article is unique and not under consideration nor published by any other publication and that they have consent to reproduce any copyrighted material. The peer reviewers declared no conflicts of interest.

References

1. Kulkarni MM, Arnosti DN. Information display by transcriptional enhancers. *Development.* Dec 2003;130(26):6569–75.
2. Clyde DE, Corado MS, Wu X, Pare A, Papatsenko D, Small S. A self-organizing system of repressor gradients establishes segmental complexity in Drosophila. *Nature.* Dec 18, 2003;426(6968):849–53.

3. Li L, Edgar BA, Grewal SS. Nutritional control of gene expression in Drosophila larvae via TOR, Myc and a novel cis-regulatory element. *BMC Cell Biol.* 2010;11:7.
4. Orian A, Grewal SS, Knoepfler PS, Edgar BA, Parkhurst SM, Eisenman RN. Genomic binding and transcriptional regulation by the Drosophila Myc and Mnt transcription factors. *Cold Spring Harb Symp Quant Biol.* 2005;70:299–307.
5. van Riggelen J, Yetil A, Felsher DW. MYC as a regulator of ribosome biogenesis and protein synthesis. *Nat Rev Cancer.* Apr 2010;10(4):301–9.
6. Siddall NA, Lin JI, Hime GR, Quinn LM. Myc—what we have learned from flies. *Curr Drug Targets.* Jul 2009;10(7):590–601.
7. Gallant P, Shiio Y, Cheng PF, Parkhurst SM, Eisenman RN. Myc and Max homologs in Drosophila. *Science.* Nov 29, 1996;274(5292):1523–7.
8. de la Cova C, Johnston LA. Myc in model organisms: a view from the flyroom. *Semin Cancer Biol.* Aug 2006;16(4):303–12.
9. Blackwood EM, Eisenman RN. Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science.* Mar 8, 1991;251(4998):1211–7.
10. Grandori C, Cowley SM, James LP, Eisenman RN. The Myc/Max/Mad network and the transcriptional control of cell behavior. *Annu Rev Cell Dev Biol.* 2000;16:653–99.
11. Herkert B, Eilers M. Transcriptional repression: the dark side of myc. *Genes Cancer.* Jun 2011;1(6):580–6.
12. Adhikary S, Eilers M. Transcriptional regulation and transformation by Myc proteins. *Nat Rev Mol Cell Biol.* Aug 2005;6(8):635–45.
13. Eilers M, Eisenman RN. Myc's broad reach. *Genes Dev.* Oct 15, 2008; 22(20):2755–66.
14. Dominguez-Sola D, Ying CY, Grandori C, et al. Non-transcriptional control of DNA replication by c-Myc. *Nature.* Jul 26, 2007;448(7152):445–51.
15. Boxer LM, Dang CV. Translocations involving c-myc and c-myc function. *Oncogene.* Sep 10, 2001;20(40):5595–610.
16. Dang CV. c-Myc target genes involved in cell growth, apoptosis, and metabolism. *Mol Cell Biol.* Jan 1999;19(1):1–11.
17. Eisenman RN. Deconstructing myc. *Genes Dev.* Aug 15 2001;15(16):2023–30.
18. Gomez-Roman N, Felton-Edkins ZA, Kenneth NS, et al. Activation by c-Myc of transcription by RNA polymerases I, II and III. *Biochem Soc Symp.* 2006;(73):141–54.
19. Lee FS, Hagler J, Chen ZJ, Maniatis T. Activation of the IkappaB alpha kinase complex by MEKK1, a kinase of the JNK pathway. *Cell.* Jan 24, 1997;88(2):213–22.
20. Levens D. Disentangling the MYC web. *Proc Natl Acad Sci U S A.* Apr 30, 2002;99(9):5757–9.
21. Lutz W, Leon J, Eilers M. Contributions of Myc to tumorigenesis. *Biochim Biophys Acta.* Mar 14, 2002;1602(1):61–71.
22. Mai S, Mushinski JF. c-Myc-induced genomic instability. *J Environ Pathol Toxicol Oncol.* 2003;22(3):179–99.
23. Nasi S, Ciarapica R, Jucker R, Rosati J, Soucek L. Making decisions through Myc. *FEBS Lett.* Feb 16, 2001;490(3):153–62.
24. Oskarsson T, Trumpp A. The Myc trilogy: lord of RNA polymerases. *Nat Cell Biol.* Mar 2005;7(3):215–7.
25. Oster SK, Ho CS, Soucie EL, Penn LZ. The myc oncogene: MarvelouslY Complex. *Adv Cancer Res.* 2002;84:81–154.
26. Pelengaris S, Khan M. The c-MYC oncoprotein as a treatment target in cancer and other disorders of cell growth. *Expert Opin Ther Targets.* Oct 2003;7(5):623–42.
27. Schmidt EV. The role of c-myc in cellular growth control. *Oncogene.* May 13, 1999;18(19):2988–96.
28. Soucek L, Evan G. Myc-Is this the oncogene from Hell? *Cancer Cell.* Jun 2002;1(5):406–8.
29. Wade M, Wahl GM. c-Myc, genome instability, and tumorigenesis: the devil is in the details. *Curr Top Microbiol Immunol.* 2006;302:169–203.
30. Schreiber-Agus N, Stein D, Chen K, Goltz JS, Stevens L, DePinho RA. Drosophila Myc is oncogenic in mammalian cells and plays a role in the diminutive phenotype. *Proc Natl Acad Sci U S A.* Feb 18, 1997;94(4):1235–40.
31. Trumpp A, Refaeli Y, Oskarsson T, et al. c-Myc regulates mammalian body size by controlling cell number but not cell size. *Nature.* Dec 13, 2001; 414(6865):768–73.
32. Benassayag C, Montero L, Colombie N, Gallant P, Cribbs D, Morello D. Human c-Myc isoforms differentially regulate cell growth and apoptosis in Drosophila melanogaster. *Mol Cell Biol.* Nov 2005;25(22): 9897–909.
33. Duman-Scheel M, Johnston LA, Du W. Repression of dMyc expression by Wingless promotes Rbf-induced G1 arrest in the presumptive Drosophila wing margin. *Proc Natl Acad Sci USA.* Mar 16, 2004;101(11):3857–62.
34. Johnston LA, Gallant P. Control of growth and organ size in Drosophila. *Bioessays.* Jan 2002;24(1):54–64.
35. Johnston LA, Prober DA, Edgar BA, Eisenman RN, Gallant P. Drosophila myc regulates cellular growth during development. *Cell.* Sep 17, 1999; 98(6):779–90.
36. Neto-Silva RM, de Beco S, Johnston LA. Evidence for a growth-stabilizing regulatory feedback mechanism between Myc and Yorkie, the Drosophila homolog of Yap. *Dev Cell.* Oct 19, 2010;19(4):507–20.
37. Herranz H, Perez L, Martin FA, Milan M. A Wingless and Notch double-repression mechanism regulates G1-S transition in the Drosophila wing. *EMBO J.* Jun 4, 2008;27(11):1633–45.
38. Smith-Bolton RK, Worley MI, Kanda H, Hariharan IK. Regenerative growth in Drosophila imaginal discs is regulated by Wingless and Myc. *Dev Cell.* Jun 2009;16(6):797–809.
39. Mitchell NC, Johanson TM, Cranna NJ, et al. Hfp inhibits Drosophila myc transcription and cell growth in a TFIIH/Hay-dependent manner. *Development.* Jul 28, 2010.
40. Froldi F, Ziosi M, Garoia F, et al. The lethal giant larvae tumour suppressor mutation requires dMyc oncoprotein to promote clonal malignancy. *BMC Biol.* 2010;8:33.
41. Song H, Hasson P, Paroush Z, Courey AJ. Groucho oligomerization is required for repression in vivo. *Mol Cell Biol.* May 2004;24(10):4341–50.
42. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* Nov 11, 1994;22(22):4673–80.
43. Yavatkar AS, Lin Y, Ross J, Fann Y, Brody T, Odenwald WF. Rapid detection and curation of conserved DNA via enhanced-BLAT and EvoPrinterHD analysis. *BMC Genomics.* 2008;9:106.
44. Brody T, Rasband W, Baler K, Kuzin A, Kundu M, Odenwald WF. cis-Decoder discovers constellations of conserved DNA sequences shared among tissue-specific enhancers. *Genome Biol.* 2007;8(5):R75.
45. Cranna N, Quinn L. Impact of steroid hormone signals on Drosophila cell cycle during development. *Cell Div.* 2009;4:3.
46. Cranna NJ, Mitchell NC, Hannan RD, Quinn LM. Hfp, the Drosophila homolog of the mammalian c-myc transcriptional-repressor and tumour suppressor FIR, inhibits dmyc transcription and cell growth. *Fly (Austin).* Apr 1, 2011;5(2).
47. Jones RM, Branda J, Johnston KA, et al. An essential E box in the promoter of the gene encoding the mRNA cap-binding protein (eukaryotic initiation factor 4E) is a target for activation by c-myc. *Mol Cell Biol.* Sep 1996;16(9):4754–64.
48. Blackwell TK, Huang J, Ma A, et al. Binding of myc proteins to canonical and noncanonical DNA sequences. *Mol Cell Biol.* Sep 1993;13(9):5216–24.
49. Blackwell TK, Kretzner L, Blackwood EM, Eisenman RN, Weintraub H. Sequence-specific DNA binding by the c-Myc protein. *Science.* Nov 23, 1990;250(4984):1149–51.
50. Cleveland JL, Huleihel M, Bressler P, et al. Negative regulation of c-myc transcription involves myc family proteins. *Oncogene Res.* 1988;3(4):357–75.
51. Wierstra I, Alves J. The c-myc promoter: still MysterY and challenge. *Adv Cancer Res.* 2008;99:113–333.
52. Goodliffe JM, Wieschaus E, Cole MD. Polycomb mediates Myc autorepression and its transcriptional control of many loci in Drosophila. *Genes Dev.* Dec 15, 2005;19(24):2941–6.
53. Khan A, Shover W, Goodliffe JM. Su(z)2 antagonizes auto-repression of Myc in Drosophila, increasing Myc levels and subsequent trans-activation. *PLoS One.* 2009;4(3):e5076.
54. Bentley DL, Groudine M. Novel promoter upstream of the human c-myc gene and regulation of c-myc expression in B-cell lymphomas. *Mol Cell Biol.* Oct 1986;6(10):3481–9.

55. Bentley DL, Groudine M. A block to elongation is largely responsible for decreased transcription of c-myc in differentiated HL60 cells. *Nature.* Jun 12–18, 1986;321(6071):702–6.
56. Bareket-Samish A, Cohen I, Haran TE. Signals for TBP/TATA box recognition. *J Mol Biol.* Jun 16, 2000;299(4):965–77.
57. Yean D, Gralla J. Transcription reinitiation rate: a special role for the TATA box. *Mol Cell Biol.* Jul 1997;17(7):3809–16.
58. Stewart TA, Bellve AR, Leder P. Transcription and promoter usage of the myc gene in normal somatic and spermatogenic cells. *Science.* Nov 9, 1984; 226(4675):707–10.
59. Kutach AK, Kadonaga JT. The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoters. *Mol Cell Biol.* Jul 2000;20(13):4754–64.
60. Daines B, Wang H, Wang L, et al. The Drosophila melanogaster transcriptome by paired-end RNA sequencing. *Genome Res.* Feb 2011;21(2):315–324.
61. Anderson S, Bankier AT, Barrell BG, et al. Sequence and organization of the human mitochondrial genome. *Nature.* Apr 9, 1981;290(5806):457–65.
62. Zhao J, Hyman L, Moore C. Formation of mRNA 3′ ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol Mol Biol Rev.* Jun 1999;63(2):405–45.
63. Lutz CS. Alternative polyadenylation: a twist on mRNA 3′ end formation. *ACS Chem Biol.* Oct 17, 2008;3(10):609–17.
64. Neilson JR, Sandberg R. Heterogeneity in mammalian RNA 3′ end formation. *Exp Cell Res.* May 1, 2010;316(8):1357–64.
65. Kulkarni MM, Arnosti DN. cis-regulatory logic of short-range transcriptional repression in Drosophila melanogaster. *Mol Cell Biol.* May 2005;25(9):3411–20.
66. Wu DC, Johnston LA. Control of wing size and proportions by Drosophila myc. *Genetics.* Jan 2010;184(1):199–211.
67. Zeitlinger J, Stark A. Developmental gene regulation in the era of genomics. *Dev Biol.* Mar 15, 2010;339(2):230–9.
68. Carroll SB. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell.* Jul 11, 2008;134(1):25–36.
69. Orian A, Delrow JJ, Rosales Nieves AE, et al. A Myc-Groucho complex integrates EGF and Notch signaling to regulate neural development. *Proc Natl Acad Sci U S A.* Oct 2, 2007;104(40):15771–6.
70. Demontis F, Perrimon N. Integration of Insulin receptor/Foxo signaling and dMyc activity during muscle growth regulates body size in Drosophila. *Development.* Mar 2009;136(6):983–93.
71. Kollmar R, Farnham PJ. Site-specific initiation of transcription by RNA polymerase II. *Proc Soc Exp Biol Med.* Jun 1993;203(2):127–39.
72. Breathnach R, Chambon P. Organization and expression of eucaryotic split genes coding for proteins. *Annu Rev Biochem.* 1981;50:349–83.
73. Smale ST, Kadonaga JT. The RNA polymerase II core promoter. *Annu Rev Biochem.* 2003;72:449–79.
74. Juven-Gershon T, Hsu JY, Kadonaga JT. Perspectives on the RNA polymerase II core promoter. *Biochem Soc Trans.* Dec 2006;34(Pt 6):1047–50.
75. Liu H, Zhou M, Luo X, et al. Transcriptional regulation of BRD7 expression by Sp1 and c-Myc. *BMC Mol Biol.* 2008;9:111.
76. Black AR, Black JD, Azizkhan-Clifford J. Sp1 and kruppel-like factor family of transcription factors in cell growth regulation and cancer. *J Cell Physiol.* Aug 2001;188(2):143–60.
77. Lang JC, Whitelaw B, Talbot S, Wilkie NM. Transcriptional regulation of the human c-myc gene. *Br J Cancer Suppl.* Dec 1988;9:62–6.
78. Lo K, Smale ST. Generality of a functional initiator consensus sequence. *Gene.* Dec 5, 1996;182(1–2):13–22.
79. Javahery R, Khachi A, Lo K, Zenzie-Gregory B, Smale ST. DNA sequence requirements for transcriptional initiator activity in mammalian cells. *Mol Cell Biol.* Jan 1994;14(1):116–27.
80. Kirchhamer CV, Bogarad LD, Davidson EH. Developmental expression of synthetic cis-regulatory systems composed of spatial control elements from two different genes. *Proc Natl Acad Sci U S A.* Nov 26, 1996; 93(24):13849–54.
81. Burke TW, Willy PJ, Kutach AK, Butler JE, Kadonaga JT. The DPE, a conserved downstream core promoter element that is functionally analogous to the TATA box. *Cold Spring Harb Symp Quant Biol.* 1998;63:75–82.
82. Huang Y, Carmichael GG. A suboptimal 5′ splice site is a cis-acting determinant of nuclear export of polyomavirus late mRNAs. *Mol Cell Biol.* Nov 1996;16(11):6046–54.
83. Preiss T, Hentze MW. Dual function of the messenger RNA cap structure in poly(A)-tail-promoted translation in yeast. *Nature.* Apr 2, 1998;392(6675):516–20.
84. Brewer G. An A + U-rich element RNA-binding factor regulates c-myc mRNA stability in vitro. *Mol Cell Biol.* May 1991;11(5):2460–6.
85. Lemm I, Ross J. Regulation of c-myc mRNA decay by translational pausing in a coding region instability determinant. *Mol Cell Biol.* Jun 2002;22(12):3959–69.
86. Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS. MicroRNA targets in Drosophila. *Genome Biol.* 2003;5(1):R1.
87. Herranz H, Hong X, Perez L, et al. The miRNA machinery targets Mei-P26 and regulates Myc protein levels in the Drosophila wing. *EMBO J.* May 19, 2010;29(10):1688–98.
88. Harris RE, Pargett M, Sutcliffe C, Umulis D, Ashe HL. Brat promotes stem cell differentiation via control of a bistable switch that restricts BMP signaling. *Dev Cell.* Jan 18, 2011;20(1):72–83.
89. Aguda BD, Kim Y, Piper-Hunter MG, Friedman A, Marsh CB. MicroRNA regulation of a cancer network: consequences of the feedback loops involving miR-17-92, E2F, and Myc. *Proc Natl Acad Sci U S A.* Dec 16, 2008;105(50):19678–83.
90. Bueno MJ, Malumbres M. MicroRNAs and the cell cycle. *Biochim Biophys Acta.* Feb 17, 2011.
91. Bueno MJ, Perez de Castro I, Malumbres M. Control of cell proliferation pathways by microRNAs. *Cell Cycle.* Oct 2008;7(20):3143–8.
92. Cannell IG, Bushell M. Regulation of Myc by miR-34c: A mechanism to prevent genomic instability? *Cell Cycle.* Jul 15, 2010;9(14):2726–30.
93. Johnston LA, Edgar BA. Wingless and Notch regulate cell-cycle arrest in the developing Drosophila wing. *Nature.* 1998;394(6688):82–4.
94. de la Cova C, Abril M, Bellosta P, Gallant P, Johnston LA. Drosophila myc regulates organ size by inducing cell competition. *Cell.* Apr 2, 2004;117(1):107–16.
95. Liu J, Kouzine F, Nie Z, et al. The FUSE/FBP/FIR/TFIIH system is a molecular machine programming a pulse of c-myc expression. *Embo J.* May 17, 2006;25(10):2119–30.
96. Hallikas O, Palin K, Sinjushina N, et al. Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. *Cell.* Jan 13, 2006;124(1):47–59.

Supplementary Materials

Table S1. Analytical primer pairs.

No.	Name	Primer sequence (5′ → 3′)	bp	Usage
1	PX1	ACG GTT GGC ATT GTT GAC TC	20	BAC clone RP98-2A13 sequencing primers
2	PN1	AAT TCG TTT TCC ATA GAT A	19	
3	BAC-F	CTT CGC GTC CAA CAG ATG	18	pC-RP27 sequencing primers
4	BAC-R	GCG CCA AAG CAA GAG GGA ATG	21	
5	pZpA-F	CCA AAA AAG GGT TTC ATT AAC TTG TAC ACA TAC	33	Sequence pLacZ-SV40 polyA
6	pZpA-R	TTG GGG ATT TTA ATA GCG GGC CCT GTG TGT	30	
7	SV40	GTT CAG GGG GAG GTG TGG G	19	SV40 P(A) sequencing primer
8	pZR	CGG GCC TCT TCG CTA TTA CG	20	pJ8 reverse sequencing primer
9	pcaF	GAC GGC GAT ATT TCT GTG GAC	21	pCaSpeR4 sequencing primers
10	pcaR	CCT TAG CAT GTC CGT GGG GTT TGA	24	
11	PolyA0-F	CCC CCT TGC TAT AAC CCG TAT AT	23	PCR – pJ1.3
12	PolyA0-R	AAA CGG TTT CTA TAT ATAGGCCTG TGT GTG T (StuI)	31	
13	pA2-F-04	TTT GCT GCC ACT TCT GTGGATCCT CAT ACA TT (BamHI)	32	PCR – pJ1.4
14	pA2-R-05	GGT TTT AGC GGC TTT GAT CGG TCTAGA AAT AAT ATT TT (XbaI)	38	
15	pA3-F	AAG GGG AAA CTG ATT GGATCC AAA AAG CAC A (BamHI)	31	PCR – pJ1.5
16	pA3-R	CGC CTC GCT TTC ATA CAA TCTAGA TAA TGT TC (XbaI)	32	
17	dm5E2F	GCG CCA AAG CAA GAG GGA ATG AAC	24	PCR – pJ2.1
18	dm5E2R	CGG GAG AG GGTACC TGC GAT TAT GTT GTC T GGG TTT TTT TTT C ACC651	43	
19	In2-5F	GAC CCC CTGCGGCCGCGG TGA GTC AAA TTT ATA TAC TTT T (NotI)	40	PCR – In2.5′
20	In2-5R	CCC CGT CGCGGCCGC TCA CGT AAG CTT TTA TAC TAA T (NotI)	37	
21	In2-3F	GCA ACA GGTACC GGC TTG TGT GCA TTT TAT TT (Acc65I)	32	PCR – In2.3′
22	In2-3R	TCG ATT GGTACC TCT GCA CAG CGA TAG TAA AAA AA (Acc65I)	35	
23	P9-F2	TGT AGGGCCCCA AAG CAA GAG GGA ATG AAC AA (PspOMI)	32	PCR pJ9
24	P9-R2	TTG TGT GTGCGGCCGCTG CGA TTA TGT TGT CT (NotI)	32	
25	P1-3-F	GTT GGA CGA CGC GAA GAT GAA AGA GAA	27	pJ1.3 sequencing primer
26	P9-R	TTG TGT GTG AGG GCC CTG CGA TTA TGT TGT CT	32	pJ9 sequencing primer
27	pJ8.1-F	TGA GCG CGC GTGGGCCCA CTC ACT ATA (PspOMI)	27	PCR pJ8.1
28	pJ8.1-R	CATGTACAT GGCGGCCGC TCA AGA CAC C (BsrGI, NotI)	28	
29	CDS3-F	TTA CTA TCG CGG TCTAGA TGA GGA AAT CG (XbaI)	29	PCR *dmyc* CDS3′

(*Continued*)

Table S1. (*Continued*)

No.	Name	Primer sequence (5′ → 3′)	bp	Usage
30	CDS3-R	TAA GCC GGGCCC ATG AGA CAA CAC TAT (PspOMI)	27	
31	CDS5-F	ACA ACA TGGTACCAA TGG CCC TTT ACC GCT CTG (Acc65I)	33	PCR *dmyc* CDS5′
32	CDS5-R	TAA ATTCTAGAC ACC GGA ATC TGA GGG GGT CTC CA (XbaI)	35	
33	ACT-F01	CGC TCC CCG TGC TGT CTT CC	20	JD, J7 discs RT-PCR:
34	ACT-R01	GCG GTG GCC ATC TCC TGC TC	20	detection of actin transcripts as internal control
35	lacZ-pcr F	TGG TTA TGC CGA TCG CGT CAC ACT AC	26	JD, J7: RT-PCR on discs
36	lacZ-pcr R	CGG ATA AAC GGA ACT GGA AAA ACT GC	26	for detection of lacZ transcripts

Notes: The oligonucleotides were designed with Bioinformatics tool Laser Gene 9.1 module Primer Select and synthesized at Microsynth AG, Switzerland. As indicated the primer sets were used for sequencing and amplification of *dmyc* noncoding regions as well as for sequencing of cloning vectors and injection plasmids. Orientation of the indicated primer sequences is 5′ → 3′. Length of primers is in base pair (bp).
Abbreviations: bp, base pair; CDS, coding sequence; RT, reverse transcriptase; PCR, polymerase chain reaction.

Table S2. List of the Fly Stocks used in the study.

No.	Bloomington stock #	Name	Genotype with FlyBase links
1	6598	y w	y^1 w^{1118}
2	23648	attp-86F	P{hsp70-flp}1, y^1 w^*; M{3xP3-RFP.attP}ZH-86Fb; M{vas-int.B}ZH-102D
3	24480	attp-2A	y^1 M{3xP3-RFP.attP}ZH-2A w^*; M{vas-int.Dm}ZH-102D
4	24482	attp-51C	y^1 M{vas-int.Dm}ZH-2A w^*; M{3xP3-RFP.attP′}ZH-51C
5	24485	attp-68E	y^1 M{vas-int.Dm}ZH-2A w^*; M{3xP3-RFP.attP′}ZH-68E
6	24871	attp-VK00033	y^1 M{vas-int.Dm}ZH-2A w^*; PBac{y^+-attP-3B}VK00033
7	24872	attp-VK00037	y^1 M{vas-int.Dm}ZH-2A w^*; PBac{y^+-attP-3B}$haf^{VK00037}$
8	2376	Oregon	Oregon-R-P2
9	2475	double balancer	w^*; T(2;3)ap^{Xa}, ap^{Xa}/CyO; TM3, Sb^1
10	11108	blue balancer	Cyo, P{lArB}A66.2F2/b^1 Adh^* cn^* l(2)**; ry^{506}
11	8412	*dpp*-lacZ	y^1 w^{1118}; P{dpp-lacZ.Exel.2}3

Notes: The attp-lines, 86F, 2A, VK00033 and VK000037 show strong activity of the reporter lacZ. The expression of transgene in the attp-strains 51C and 68E is weak. For random P-element mediated transgenesis embryos were taken from the y w (6598) flies. Oregon flies were used for *dmyc in situ* hybridization in different imaginal tissues. The fly stock "blue balancer" that expresses lacZ in embryos and ovaries, was used as a source for positive control in the embryos and ovaries staining. The dpp-lacZ line was used as positive control for the imaginal discs staining.

Figure S1. EvoDifference analysis of the distal 8-kb upper strand of the 40-kb *D. melanogaster dmyc* locus detects conserved sequence blocks. Each block of uppercase bases represents a cluster of conserved sequences. Black capital letter bases are conserved in all or in all but one species including *D. sechellia, D. yakuba, D. erecta, D. ananassae, D. pseudoobscura, D. willistoni or D. virilis*. There is one E-Box, conserved in all species except *D. virilis* (white text with black highlight). The neural network genetic algorithm PROMOTER 2.0 predicts three possible promoter regions using ccaat or bHLH recognition motifs (lower case black letters). The yellow highlighted sequence includes a highly likely promoter region (dubbed as P0 putative). Sequence orientation is 5′ → 3′.

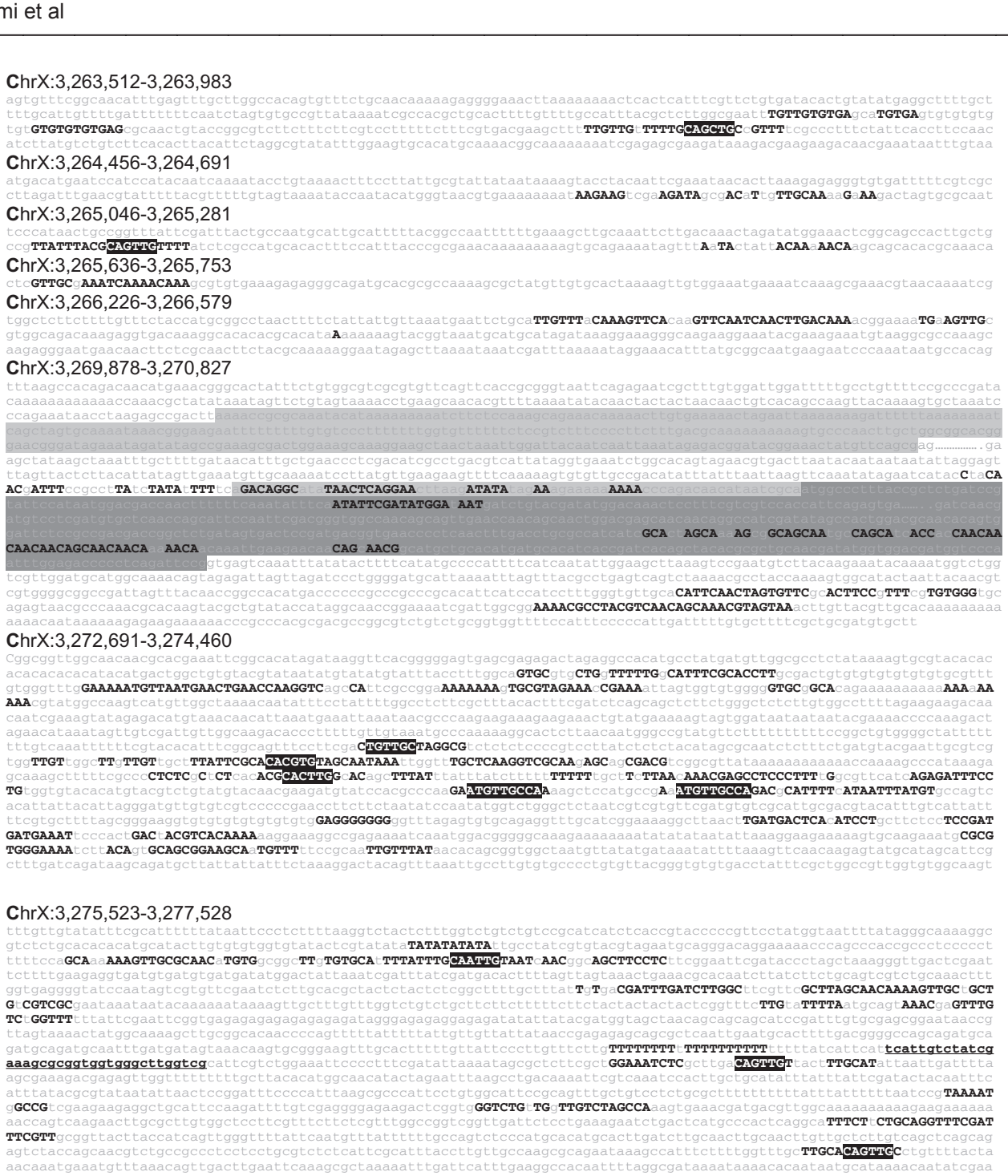


Figure S2. (*Continued*)

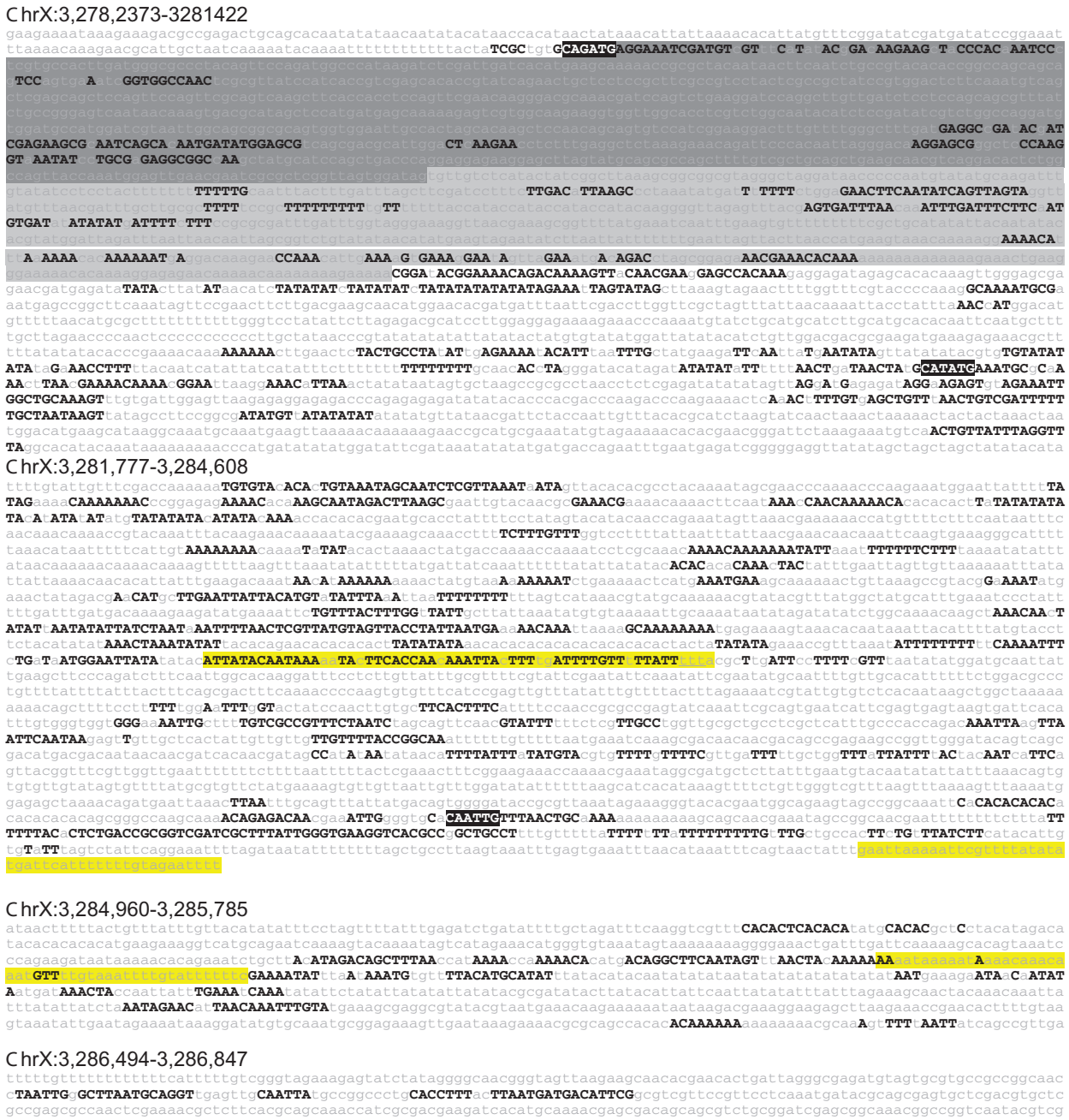


Figure S2. EvoDifference analysis of the 40-kb top strand *D. melanogaster dmyc* locus. The 12.834-kb *dmyc* gene resides within the 40-kb fragment. Each block of uppercase bases represents a cluster of conserved sequences. Uppercase black nucleotides represent bases in the *D. melanogaster dm* EvoP 40-kb reference sequence that are conserved in all or all but one of the other 5 orthologous DNAs, *D. sechellia, D. yakuba, D. erecta, D. willistoni or D. virilis*. The *D. melanogaster dmyc* transcribed sequence is annotated according to FlyBase as follows: 5′-UTR (light grey), protein-coding sequence (dark grey) and 3′-UTR (light grey). The conserved bHLH binding sites and E-boxes are in white text with black highlight, as is the repeat motif ATGT-TGCCA and the nearby sub-repeat TGTTGC. PolyA1, 2, 3 in the 3′ UTR (yellow highlights), detected by Lasergene GeneQuest module, is indicated. The DPE in the intron 2 region is lower case bold and underlined. ORF orientation is 5′ → 3′.

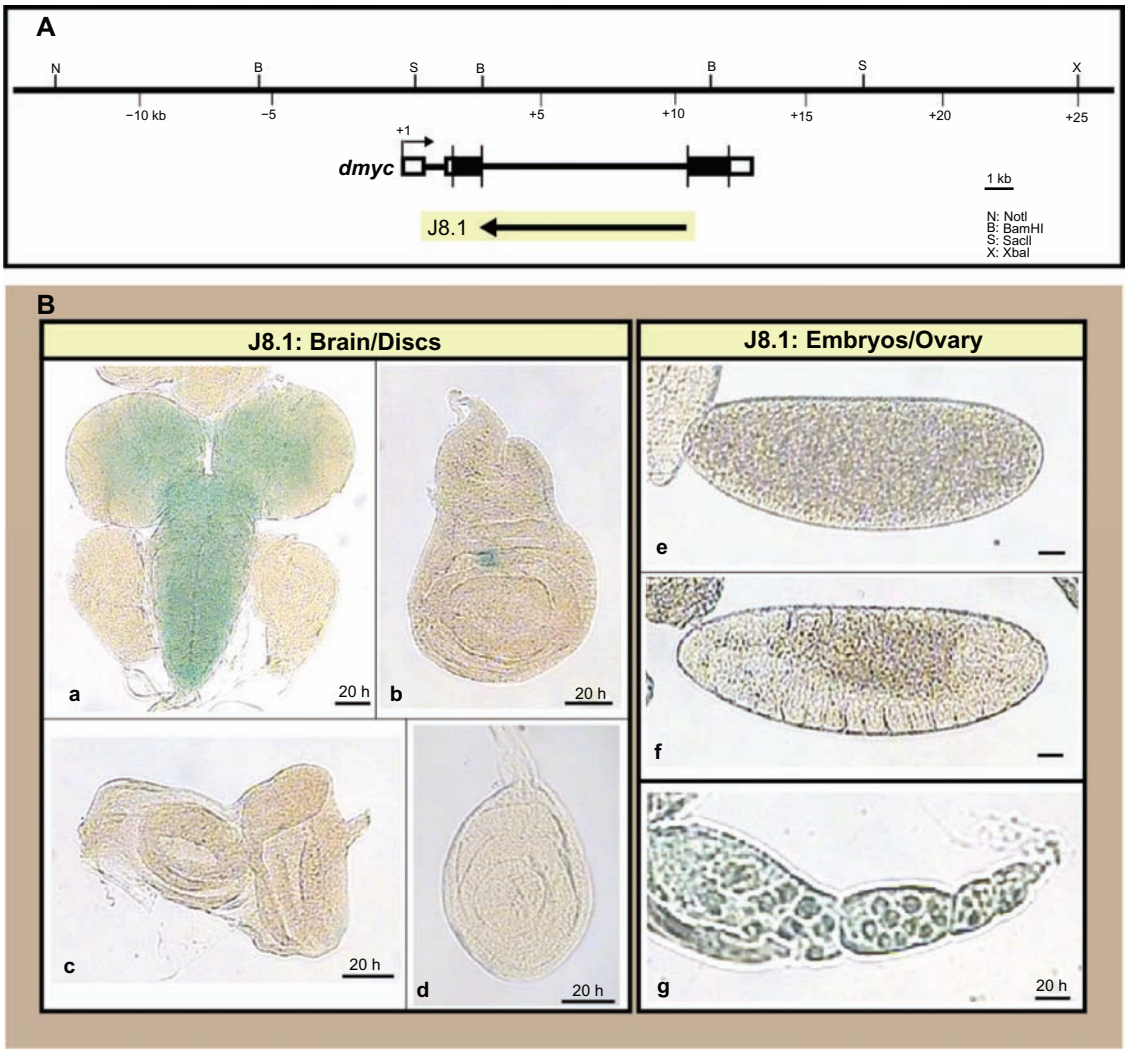


Figure S3. The TATA-less downstream promoter element DPE functions unidirectionally. (**A**) The 8-kb intron 2 full length sequence in the J8.1 transgene and its relative location with respect to the *dmyc* locus and genomic organization is shown. (**B**) The J8.1 fragment was examined for its ability to drive reporter expression in different tissues during early developmental stages. Except for a minimal basal expression in the brain (**a**) the activity of the reporter was abolished in imaginal discs **(b–d)**, embryos (**e**, stage 1–2; **f**, stage 13–16) and ovary (**g**), emphasizing the unidirectionality of the minimal core promoter within the J8.1 fragment responsible for full activity in all the tissues (compared to the J8 activity in Fig. 6).
Note: Staining times for the discs and ovary are indicated above the scale bars. Embryo staining took place over-night. Scale bar in (**a–g**) indicates 50 µm.

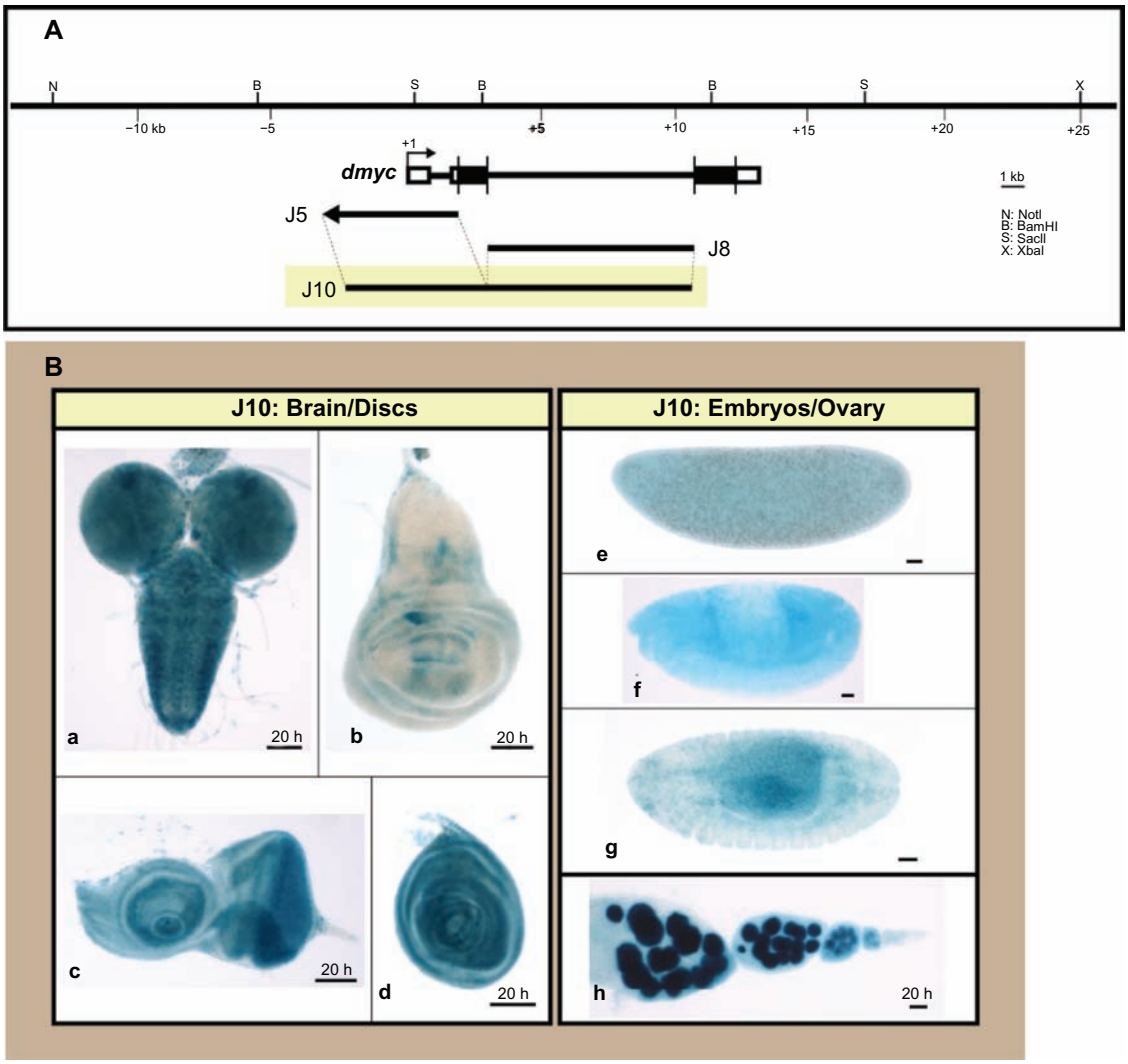


Figure S4. The fusion of J5 in the direction opposite to J8 attenuates expression in embryos. (**A**) The approximately 4.1-kb J5 promoter, the 8-kb intron 2 full length sequence in the J8 transgene, their fusion product J10 and their relative location with respect to the *dmyc* locus and genomic organization are shown. (**B**) The approximately 12.1-kb chimeric restriction fragment in J10 results in reporter activity in the brain (**a**), discs (**b–d**) in a manner similar to fragments J2.1 to J6 in Figure 5 or intron 2 transgenes in Figure 6. However, the expression is weak during different stages of embryogenesis (**e**, stage 1–3; **f**, **g**, stage 9–12). The expression in ovary remained similar to J2.1 and its deletions and J8, J8.2 transgenes.
Note: Staining times for the discs and ovary are indicated above the scale bars. Embryo staining took place over-night. Scale bar in (**a–h**) indicates 50 μm.

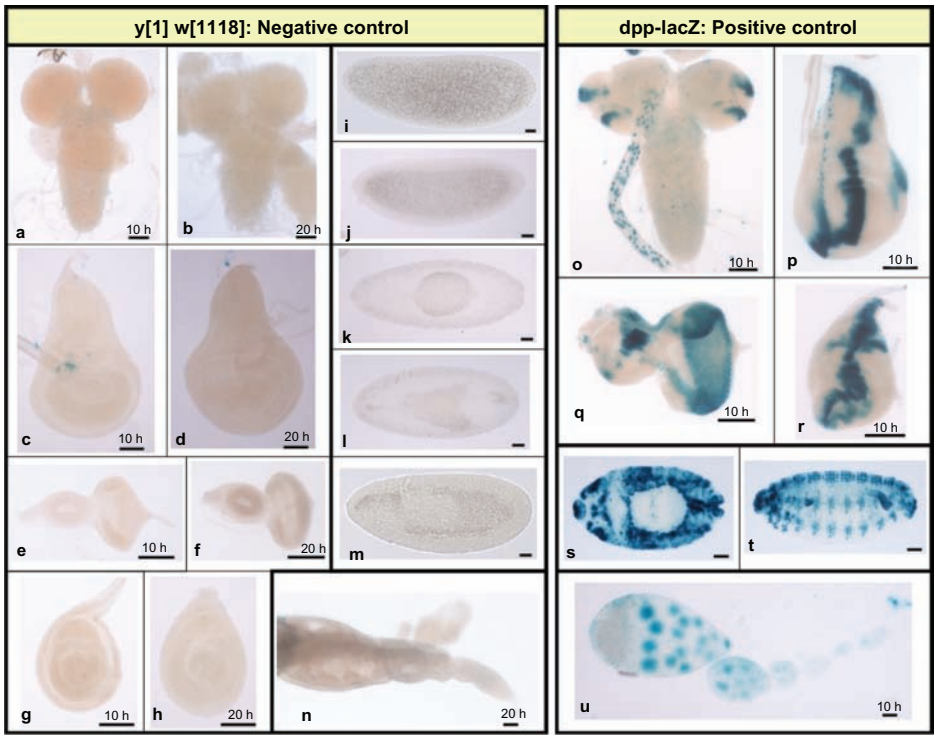


Figure S5. Positive and negative controls for lacZ staining. For each lacZ staining of transgenes negative and positive control stainings were performed. As negative control different tissues were taken from the "y[1] w[1118] " and "attp-flies" listed in Table 2. *dpp*-lacZ flies served as source for positive control. Depicted are negative controls (**a–n**) and positive controls (**o–u**), (**a**, **b**, **o**, brain; **c–h**, **p–r**, discs; **i–m**, **s**, **t**, embryos; and **n**, **u** ovaries).
Note: Staining times for the discs and ovary are indicated above the scale bars. Embryo staining took place over-night. Scale bar in (**a–u**) indicates 50 µm.

Publish with Libertas Academica and every scientist working in your field can read your article

"I would like to say that this is the most author-friendly editing process I have experienced in over 150 publications. Thank you most sincerely."

"The communication between your staff and me has been terrific. Whenever progress is made with the manuscript, I receive notice. Quite honestly, I've never had such complete communication with a journal."

"LA is different, and hopefully represents a kind of scientific publication machinery that removes the hurdles from free flow of scientific thought."

Your paper will be:

- Available to your entire community free of charge
- Fairly and quickly peer reviewed
- Yours! You retain copyright

http://www.la-press.com