The Company of Biologists

RESEARCH ARTICLE

Complex *cis*-regulatory landscape of the insulin receptor gene underlies the broad expression of a central signaling regulator

Yiliang Wei[1,*], Rewatee H. Gokhale[1], Anne Sonnenschein[2], Kelly Mone't Montgomery[3], Andrew Ingersoll[1] and David N. Arnosti[1,2,‡]

ABSTRACT

Insulin signaling plays key roles in development, growth and metabolism through dynamic control of glucose uptake, global protein translation and transcriptional regulation. Altered levels of insulin signaling are known to play key roles in development and disease, yet the molecular basis of such differential signaling remains obscure. Expression of the insulin receptor (*InR*) gene itself appears to play an important role, but the nature of the molecular wiring controlling *InR* transcription has not been elucidated. We characterized the regulatory elements driving *Drosophila InR* expression and found that the generally broad expression of this gene is belied by complex individual switch elements, the dynamic regulation of which reflects direct and indirect contributions of FOXO, EcR, Rbf and additional transcription factors through redundant elements dispersed throughout ∼40 kb of non-coding regions. The control of *InR* transcription in response to nutritional and tissue-specific inputs represents an integration of multiple *cis*-regulatory elements, the structure and function of which may have been sculpted by evolutionary selection to provide a highly tailored set of signaling responses on developmental and tissue-specific levels.

KEY WORDS: dFOXO, Ecdysone, Insulin receptor, Retinoblastoma, Transcriptional enhancer

INTRODUCTION

The insulin signaling pathway plays essential roles in growth and metabolism in metazoans. In mammals, the insulin receptor (INSR) binds to insulin, leading to autophosphorylation, phosphorylation of adaptor proteins, and subsequent activation of the PI3K-Akt and MAPK pathways (Ebina et al., 1985; Ullrich et al., 1985; Oldham and Hafen, 2003). Akt propagates the metabolic effects of the signaling by targeting downstream substrates, including the glucose transporter GLUT4 (also known as SLC2A4) (Bertrand et al., 2008; Gonzalez et al., 2011). The FOXO transcription factor, which is phosphorylated and excluded from the nucleus as a result of insulin signaling, represents an important target (Puig and Tjian, 2005). FOXO regulation is widely conserved in metazoans, serving to mediate the effect of insulin signaling on growth, aging and metabolism in *C. elegans* and *Drosophila* (Taguchi and White,

2008). Interestingly, FOXO directly activates expression of the insulin receptor, representing a negative transcriptional feedback loop (Puig et al., 2003; Puig and Tjian, 2005). The receptor is expressed in most tissues and developmental stages, underscoring the broad physiological relevance of this signaling pathway.

Although insulin levels control pathway activity, varying levels of receptor expression may influence signaling in a tissue-specific manner; the developmental significance of such transcriptional regulation is yet to be understood. Developmental differences in insulin receptor gene expression may be 'hardwired' and subject to evolutionary modification, changing the impact of insulin signaling in the control of body size and morphology. In addition, physiological stimuli influence the expression of *INSR*, although the importance and consequences of this regulation are unknown. Diet, hormone levels and other signals impact *INSR* expression levels, as can viral infection and diabetes (Mamula et al., 1990; Kriauciunas et al., 1993; Chatterjee, 2001; Iritani et al., 2000; Gunton et al., 2005; Kim et al., 2014). Elevated levels of *INSR* expression are observed in numerous cancers, leading to an insulin-dependent growth phenotype (Belfiore and Malaguarnera, 2011). Low levels of *INSR* expression and signaling in the brain are associated with Alzheimer's disease (Frölich et al., 1998; Moloney et al., 2010).

Despite the epidemiological and experimental evidence for changes in expression of *INSR* in cancer and other diseases, we have limited knowledge about the transcriptional controls, precluding a molecular understanding of how receptor expression impacts physiology and disease. The *INSR* gene is very broadly expressed, unlike highly tissue-specific developmental genes, but there is limited data to suggest that the expression is not a product of a simple housekeeping promoter (Lee et al., 1992). For the human *INSR* gene, attention has focused on 2 kb flanking the transcription start site, the activity of which in reporter genes is affected by dexamethasone, glucocorticoids, vitamin D and estrogen (Leal et al., 1992; Lee and Tsai, 1994; García-Arencibia et al., 2005; Calle et al., 2008). This fragment is regulated by Sp1, HMGI, p53 and Rb (Cameron et al., 1992; Shen et al., 1995; Webster et al., 1996; Brunetti et al., 2001). The transcriptional significance of other regions, including extensive introns, remains largely unexplored. Genomic surveys of the mammalian gene reveal functionally uncharacterized chromatin marks and structures consistent with enhancers, and putative intronic enhancers for the gene were identified in mouse T cells (Pasquali et al., 2014; Kundaje et al., 2015; Vanhille et al., 2015).

Drosophila Insulin-like receptor (InR) is activated by a family of insulin-like peptides to control growth and homeostasis (Oldham and Hafen, 2003). The *InR* gene is crucial for embryonic development, function of the nervous system and regulation of organ growth (Petruzzelli et al., 1986; Garofalo and Rosen, 1988; Brogiolo et al., 2001; Song et al., 2003; Wong et al., 2014). *InR* mutations result in pleiotropic recessive phenotypes, leading to embryonic lethality (Fernandez et al., 1995). Its broad expression

[1]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824, USA. [2]Genetics Program, Michigan State University, East Lansing, MI 48824, USA. [3]Pharmaceutical Sciences, School of Pharmacy, University of Wisconsin-Madison, Madison, WI 53705, USA.
*Present address: Cold Spring Harbor Laboratory 1 Bungtown Rd, Cold Spring Harbor, NY 11724, USA.

‡Author for correspondence (arnosti@msu.edu)

D.N.A., 0000-0003-0983-6982

DEVELOPMENT

pattern has been classified as among the more 'transcriptionally stable', similar to other genes of widespread function and fewer extremes in expression (Pérez-Lluch et al., 2015). However, *InR* expression is affected by nutrition and by the steroid hormone ecdysone, which acts through its receptor EcR to control growth and development (Koelle et al., 1991; Riddiford et al., 2000; Hu et al., 2003; Gershman et al., 2007). Ecdysone stimulates expression of *InR* in *Drosophila* Kc cells, and ChIP-seq studies indicate that this regulation is direct; EcR and USP bind to the *InR* locus (Gauhar et al., 2009). More recently, STARR-seq technology in *Drosophila* cell lines identified 20E-responsive elements in the *InR* gene locus (Shlyueva et al., 2014).

An additional regulatory factor for the insulin receptor is the retinoblastoma co-repressor (Rb); this co-factor binds near the *InR* promoter, although the physiological role has yet to be investigated (Acharya et al., 2012; Korenjak et al., 2012; Wei et al., 2015). Rb tumor suppressor proteins are key regulators of the cell cycle; thus, coordinate regulation of the insulin receptor with these genes might provide a link between cell proliferation and growth (Du and Pogoriler, 2006; Giacinti and Giordano, 2006; Acharya et al., 2012). From examination of human ChIP-seq datasets, we note that Rb family proteins also bind the human *INSR* gene (Chicas et al., 2010).

Similar to its mammalian counterpart, the *Drosophila InR* gene is large, with nearly 40 kb of introns (Casas-Tinto et al., 2007). The gene is characterized by various histone modifications and regions exhibiting DNase I hypersensitivity, FAIRE-seq signals, and STARR-seq candidate regulatory elements. However, as for the mammalian gene, we lack an integrated understanding of the direct transcriptional controls of this central player in cell metabolism and development (Kaplan et al., 2011; Li et al., 2011; Nègre et al., 2011; Thomas et al., 2011; Arnold et al., 2013; McKay and Lieb, 2013). Here, we provide a comprehensive identification and characterization of *cis*-regulatory elements associated with the *InR* gene, mapping their dynamic responses to FOXO (also known as dFOXO), ecdysone and Rb. Our detailed mutagenic studies of the active enhancers identify specific elements and motifs required for enhancer activity, providing, in some cases, an incoherent feed-forward regulatory logic. The dynamic regulation of these enhancers by transcriptional inputs indicates that these enhancers play a role in temporal, spatial and critical fine-tuning control of *InR* gene expression. Our study indicates that this gene is subject to a complex transcriptional circuit extending far beyond the previously described simple model of the FOXO-feedback loop mechanism. This gene circuit analysis transforms our understanding of the insulin receptor gene, in that even such a broadly expressed gene requires exquisite controls that are crucial to the roles of this signaling pathway in metabolism, growth control and cancer.

RESULTS
Candidate regulatory regions in *InR* introns
The *Drosophila InR* gene spans ~50 kb, including ~40 kb of introns (Fig. S1A). We found that an 80 kb BAC genomic construct (InR-BAC) covering this locus rescued lethality of an *InR*[E19/GC25] temperature-sensitive mutant (Fig. S1B,C); two copies of the BAC increased *InR* gene expression in these flies 2- to 3-fold (Fig. S1D). Although we cannot rule out possible additional distant *cis*-regulatory elements, it is clear that sequences relevant for *InR* expression are located within this region. We therefore investigated the short 5′ intergenic sequence and introns of *InR* for clues as to transcriptional controls. Data from genome-wide STARR-seq surveys in transfected cells, as well as DNase hypersensitivity data and measurement of open chromatin using FAIRE-seq indicate

that *InR* introns are likely to harbor relevant *cis*-regulatory elements (Fig. 1B, Fig. S2). To evaluate the regulatory potential of intronic regions in the fly, we tested ten GAL4 lines with fragments from the *InR* gene (Pfeiffer et al., 2008). Previous measurements in the embryo indicated that some of these elements drive GFP expression in dynamic patterns (Jenett et al., 2012; Jory et al., 2012; Manning et al., 2012; Li et al., 2014). Three of the fragments also express GFP in larvae and adults, in either ubiquitous or tissue-specific patterns (Fig. 1A, Fig. S3, Table S1). We additionally plotted the results of genome-wide enhancer surveys (from S2 and ovarian stem cells), chromatin accessibility at different developmental stages and tissues as measured by FAIRE-seq, and the enhancer-associated histone modifications H3K27ac, H3K4me1 and presence of the p300 coactivator (also known as Nejire or CBP). The resultant patterns do not provide a consistent, easily interpretable set of correlations across different developmental times. Enhancers found using STARR-seq do point to redundantly acting enhancers in *InR* introns with either shared or cell type-specific patterns (Fig. 1B). These enhancers overlap some of the fragments tested as GAL4 drivers, but there was not a complete agreement between these different methods. The two types of assays relied on distinct basal promoters, which might have biased detection because of enhancer-promoter specificity (Marinić et al., 2013; Zabidi et al., 2015; Lorberbaum et al., 2016).

Identification of active enhancers within *InR* introns
To delineate the exact structure of *InR* regulatory regions and identify regulation by FOXO and ecdysone, we divided the *InR* introns into 25 fragments of ~1.5 kb each (Fig. 2A). As noted above, genome-wide assays for *cis*-regulatory elements employed synthetic basal promoters, which may lack functional compatibility with the endogenous enhancers. Therefore, we tested the endogenous basal promoter regions. *InR* has three annotated transcription start sites: T1, T2 and T3 (Casas-Tinto et al., 2007). Genome-wide RNA polymerase II occupancy and the H3K4me3 histone modification, which is linked with transcriptional start sites, showed strong association with the T1 promoter throughout different developmental stages (Fig. S4). This basal promoter appears to be the predominant site of initiation (Casas-Tinto et al., 2007; Graveley et al., 2011; Brown et al., 2014; Attrill et al., 2016). We compared the promoter activities of T1, T2 and T3 in luciferase reporter constructs, using a similarly sized intronic fragment (PT) as a negative control (Fig. 2A). The promoters and the negative control were assayed in S2 and Kc cells; T1 promoter activity was much higher than that of T2 and T3 (Fig. 2B,C, Table S2). T1 was therefore used to assay the 25 intron fragments in reporter constructs. To test for cell type specificity, we used both S2 and Kc cells. Intron fragments 2, 3, 20 and 22 were found to be active in both cell types, and the levels of activity varied. Fragments 4, 12 and 15 were active in one of the two cell types, indicating the presence of cell type-specific enhancers (Fig. 2B,C).

We find active elements in S2 cells in regions 2, 3 and 12, similar to findings from STARR-seq in S2 cells. However, we find no activity in regions 6 or 23-25, where possible enhancers were detected in some STARR-seq assays, while regions 20 and 22 were robust activators, but not consistently identified in STARR-seq (Fig. 2A-C). These differences between the assays might reflect our testing of longer fragments (1.5 kb versus 600 bp), which were thus less likely to divide and inactivate an enhancer element, and our assays relied on enhancers acting on the endogenous basal promoter region, which may provide compatibility that is lacking in the genome-wide approach.
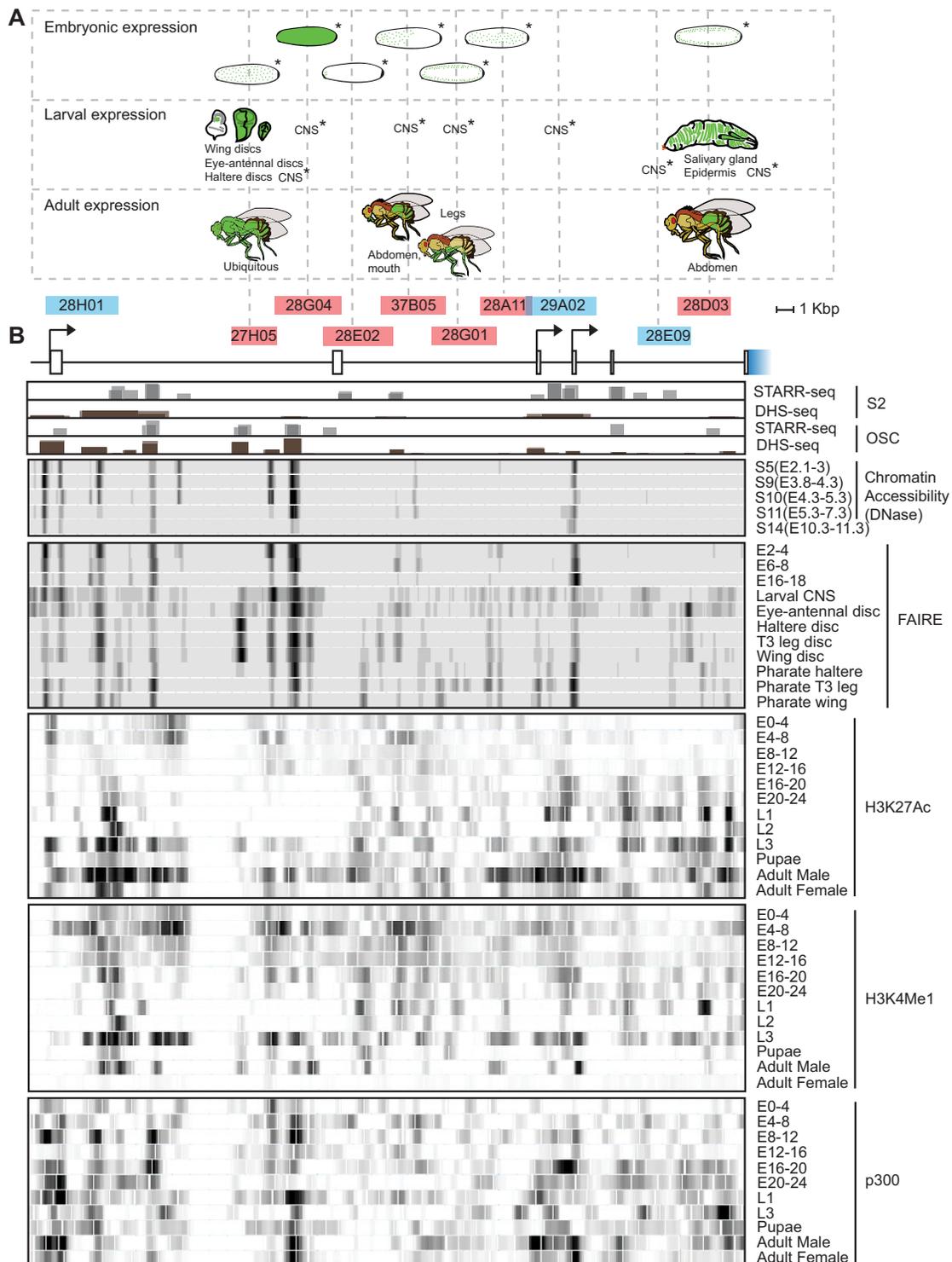
Fig. 1. Regulatory landscape of the *Drosophila InR* locus. (A) Transcriptional output of genomic fragments associated with the *InR* locus. Janelia GAL4 lines that contain genomic fragments in the *InR* locus were crossed to a UAS-GFP line. Extant information collected from the FlyLight database for embryonic and larval activity is also shown (indicated by asterisks). Fragments in red show GFP signal in embryos, larvae or adult flies. Fragments in blue show limited or no expression. Representative images from larvae and adult flies from this study are shown in Fig. S3. Details of each fragment are provided in Table S1. (B) Alignment of the *InR* gene locus with previously identified features. STARR-seq elements are from S2 cells and ovarian stem cells (OSCs) (Arnold et al., 2013), chromatin accessibility was identified by DHS-seq in cell lines (Arnold et al., 2013) or developing embryos (Kaplan et al., 2011; Li et al., 2011; Nègre et al., 2011; Thomas et al., 2011), FAIRE in different developmental stages and tissues (McKay and Lieb, 2013), and the enhancer-associated modifications H3K27ac, H3K4me1 and p300 binding at different developmental stages (modENCODE). For STARR-seq and DHS-seq, the data from two experiments are shown, with darker areas indicating reproducible results and lighter areas found in only one experiment. Bar height indicates enhancer activities or reads for the peaks. Width of STARR-seq signals is set at 600 bp, the average length of the fragments (Arnold et al., 2013). Chromatin accessibility data are for the different developmental stages indicated at the right (S, developmental stage; E, hours of embryonic development). FAIRE, H3K27ac, H3K4me1 and p300 data are presented as peak intensities, with darker shades indicating higher peaks. All data scales are normalized to local maxima. Genome version is *Drosophila* genome dm3/R5.
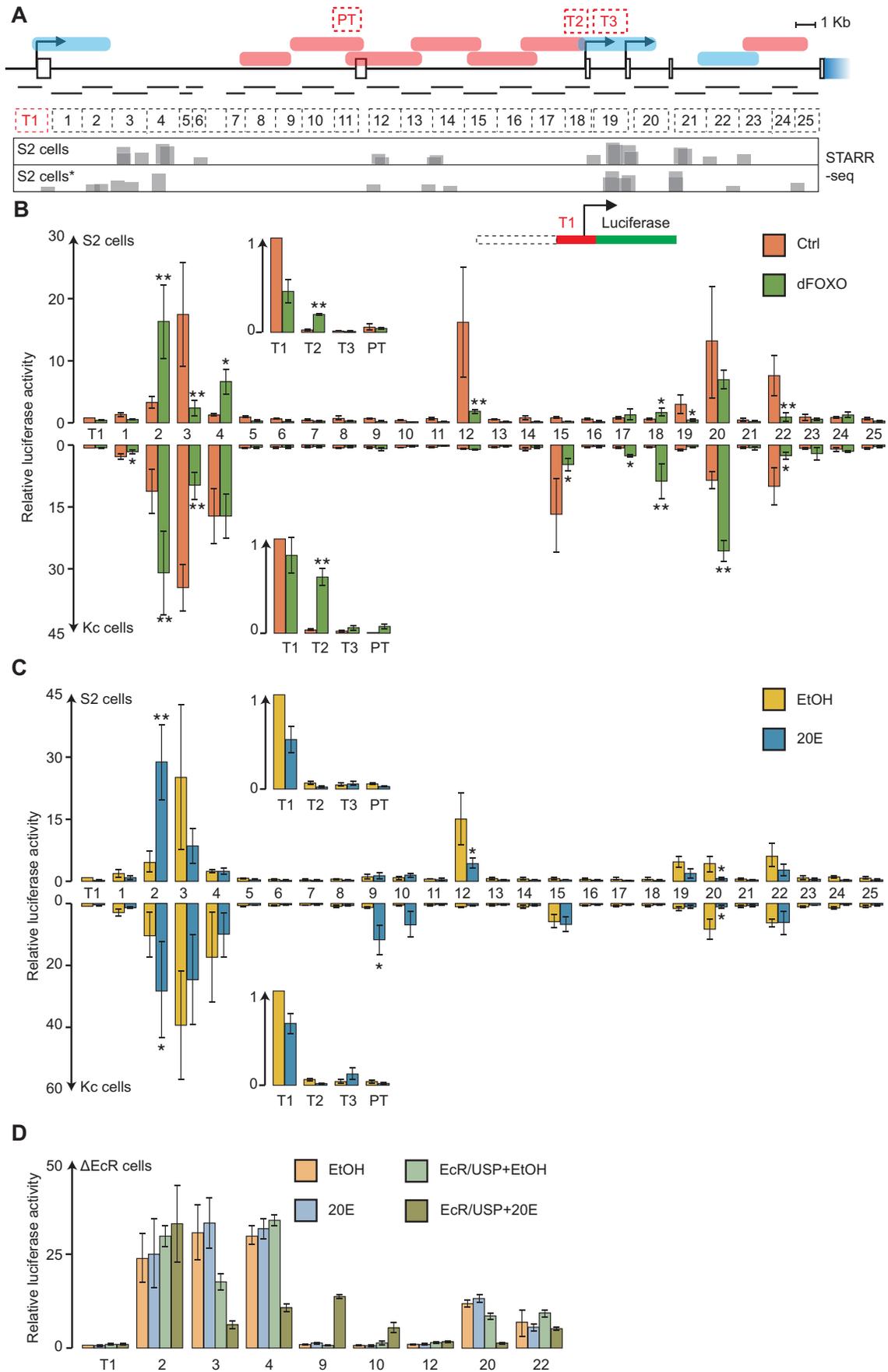
Fig. 2. See next page for legend.

Fig. 2. Mapping of FOXO- and ecdysone-responsive elements. (A) Overall structure of the *InR* gene and scheme for the reporter library. The intronic regions of *InR* were divided into 25 fragments. Each fragment (~1.5 kb) was fused 5′ of the promoter proximal region of the first transcription start site (T1) to drive luciferase reporter gene expression. The promoter proximal regions of the second (T2) and third (T3) transcription start sites were also assayed. An internal region that overlaps fragment 11 and contains a portion of the second exon served as a negative control (PT). The activities of Janelia GAL4 fragments (red, active; blue, less or not active) are indicated above the gene. STARR-seq data from two independent reports are presented beneath the gene (Arnold et al., 2013; *Shlyueva et al., 2014). A small repeat-rich region between fragments 6 and 7 was unclonable. (B) Transcriptional activity of reporters in S2 or Kc cells, and FOXO regulation. Reporters were co-transfected into S2 (upper bars) or Kc (lower bars) cells with or without a FOXO expression construct. Both activation and repression by FOXO was observed. Activity of the T1 promoter region was set to 1. In this and following figures, at least three biological replicates were performed. (C) Transcriptional effects of 20-hydroxyecdysone (20E) treatment. Constructs tested in B were tested for 20E sensitivity by treating cells 24 h after transfection of reporters with ethanol (vehicle control) or 20E (10^{-5} M, in ethanol). (D) Transcriptional responses to 20E treatment require functional EcR. Reporters that respond to 20E treatment in S2/Kc cells fail to respond in an EcR-deficient cell line (Kc-derived ΔEcR cells) (see also Fig. S6). However, these responses were restored by transfecting ΔEcR cells with EcR/USP expression vectors. (B-D) Error bars indicate s.d. *$P<0.05$, **$P<0.01$ (two-tailed *t*-test). Details of reporter activities are provided in Table S2.

We assessed the responsiveness of the T2 and T3 promoters to the regulatory elements identified above. Fusion constructs containing regions 2 or 3 robustly activated transcription from T1, but not T2 or T3, suggesting that these basal elements are unlikely to generate much of the overall transcriptional output, a conclusion supported by RNA-seq analysis (Fig. S5) (Graveley et al., 2011; Brown et al., 2014; Attrill et al., 2016). Previous studies have focused largely on the regulatory potential of T2 (Puig et al., 2003; Casas-Tinto et al., 2007); our analysis indicates that much of the regulatory activity of this locus is likely to be channeled through the distal T1 promoter.

Positive and negative regulation of *InR* enhancers by FOXO
In addition to its role as a downstream effector of insulin signaling, the FOXO transcription factor also impacts *InR* expression (Jünger et al., 2003; Puig et al., 2003). Binding sites for FOXO are present at T2, and reporter genes containing these sequences are activated by FOXO, leading to the notion that T2 allows for FOXO activation of *InR* (Puig et al., 2003; Casas-Tinto et al., 2007). We assayed each element for activation by FOXO. Consistent with previous reports, the weak T2 promoter was modestly activated by FOXO in both S2 and Kc cells (Fig. 2B). By contrast, robust activation by FOXO was observed with region 2, and with region 4 in S2 cells. The T1 promoter was itself slightly repressed by FOXO expression in S2 cells. Strikingly, expression of FOXO had a strong and significant negative effect on other elements, including regions 3 and 22, which were repressed in both cell types. Fragment 20 was repressed in S2 cells, whereas it was activated by FOXO expression in Kc cells (Fig. 2B, Table S2).

Most FOXO response fragments may be indirectly regulated
To determine if the transcriptional effects mediated by FOXO were a consequence of direct interaction of the protein with these regulatory elements, we performed ChIP analysis using anti-FOXO serum. A previously characterized direct target of FOXO, the *Thor* (*4EBP*) promoter, was used as a positive control, which showed strong endogenous FOXO binding (Fig. 3A,B) (Teleman et al., 2008; Alic et al., 2011; Bai et al., 2013). T2 (included in region 18) also showed lower, but significant, FOXO enrichment (Fig. 3A,B)

(Puig et al., 2003; Casas-Tinto et al., 2007). Surprisingly, none of the other elements that were transcriptionally regulated by FOXO expression exhibited strong binding by the factor (Fig. 3A,B). A prominent peak was observed on fragment 10, an element not activated or repressed by FOXO (Fig. 3A,B).

To further assess whether the signals that we observed represented FOXO binding, we treated cells with insulin to activate the signaling pathway, which should result in phosphorylation and exclusion of endogenous FOXO from the nucleus, or subjected cells to serum starvation, which should reduce signaling and increase FOXO activity (Puig et al., 2003). Insulin treatment resulted in modestly reduced FOXO ChIP signals on *Thor*, as well as regions 10 and 18, whereas starvation appeared to increase the ChIP signal (Fig. 3C).

Similarity of ecdysone and FOXO responses
Ecdysone treatment increases the expression of *InR*; however, the molecular mechanism of this regulation has not been elucidated (Gauhar et al., 2009). To determine how the hormone may affect the transcriptional elements of *InR*, we treated cells with 20-hydroxyecdysone (20E). The T1 promoter was slightly repressed by 20E in both cell types, whereas T2 and T3 were unaffected (Fig. 2C). Fragment 2 was robustly activated, whereas fragments 3 and 20 were significantly repressed by 20E in both cell types. In Kc cells, we observed cell type-specific activation of elements 9 and 10, which alone had not shown significant transcriptional potential. A greater number of elements showed reduction in activity after 20E treatment, although some of these effects were modest (Fig. 2C, Table S2). To determine whether these 20E responses required EcR, we assayed reporters in an *EcR*-deficient cell line (ΔEcR) that was derived from Kc cells (Swevers et al., 1996). No constructs responded to 20E treatment in these cells, indicating the requirement for EcR (Fig. 2D, Fig. S6). We transfected the ΔEcR cells with EcR and its heterodimeric partner USP and confirmed that 20E responsiveness was restored, confirming the role of EcR in this regulation (Fig. 2D).

Interestingly, many of the elements tested, including 2, 3, 12 and 20, showed similar responses to 20E and FOXO, suggesting the involvement of linked pathways (Fig. 2B,C). Significantly, 20E signaling has been shown to affect FOXO localization by regulating PI3K activity, suggesting that some of the 20E effects might be mediated by FOXO activity (Colombani et al., 2005). In addition, FOXO has been reported to bind directly to the USP co-factor of EcR (Koyama et al., 2014). Thus, 20E might regulate some enhancers in the *InR* gene via FOXO activity. FOXO overexpression was able to regulate these elements in ΔEcR cells just as in wild-type Kc cells; thus, EcR is not required for FOXO activity (Fig. S7).

To gain further insight into 20E regulation, we compared our data in S2 cells with STARR-seq analysis in the same cell line treated with 20E (Fig. S8) (Shlyueva et al., 2014). Essential features were confirmed in both studies. In contrast to FOXO-responsive enhancers, the 20E-activated enhancers were directly bound by EcR and its binding partner USP, suggesting a mode of regulation involving EcR derepression in the presence of 20E (Fig. S8). None of the 20E-repressed areas correlated with directly bound EcR peaks, and thus these elements might be subject to indirect regulation. One gene induced by EcR is *Eip74EF*, which functions as a repressor (Shlyueva et al., 2014). ChIP-seq analysis in embryos indicates that this protein may interact with repressed regions 19 and 20, but that other repressed elements may be repressed by a different factor. Eip74EF is also found to bind to
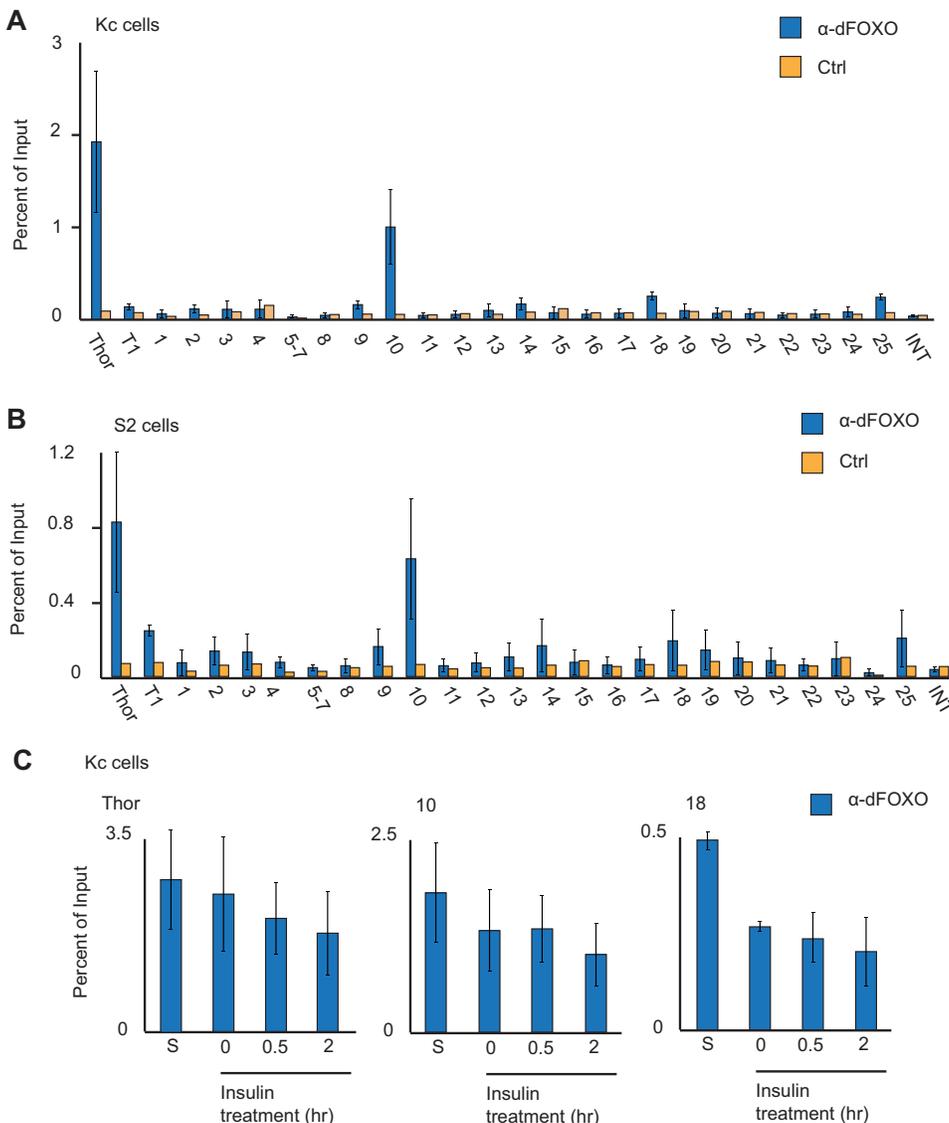
Fig. 3. *In vivo* occupancy of FOXO provides evidence for direct and indirect regulation of *InR*. (A) Binding of FOXO to *InR* genomic regions in Kc cells measured by ChIP. The *Thor* promoter served as a positive control. Regions 10, 18 and 25 showed the highest signals. Error bars indicate s.d. for five biological replicates. (B) Binding of FOXO to *InR* genomic regions in S2 cells measured by ChIP. Error bars indicate s.d. for four biological replicates. (C) Effect of starvation (S) on FOXO binding to the *Thor* promoter and *InR* intron fragments 10 and 18 (which overlaps with T2), and effect of insulin treatment (0.5 and 2 h). Error bars show s.d. for three replicates. With the exception of starvation effects on region 18, changes were not statistically significant, but trends were consistent with expected regulatory effects on FOXO.

region 2, which was activated by 20E. This binding might represent a progressive gene switch, in which initial derepression after loss of EcR binding is later followed by repression, as EcR-driven Eip74EF repressor levels increase.

Impact of Rb binding site on the *InR* promoter and enhancers

The direct actions of FOXO and EcR in *InR* expression had previously been supported by genetic and biochemical evidence. More recently, we noted that the T1 proximal promoter region of the *InR* gene is occupied *in vivo* by the Rbf1 (Rbf – FlyBase) tumor suppressor protein, the homolog of mammalian Rb (Acharya et al., 2012; Korenjak et al., 2012; Wei et al., 2015). Binding of Rbf1 appears to be of functional significance, as a T1 promoter construct is repressed by Rbf1 expression in S2 cells (Raj et al., 2012). To further explore the significance of Rbf1 protein interaction with the *InR* T1 promoter, we removed a 100 bp fragment centered on the Rbf1 binding peak (ΔRbf1) (Fig. 4A). This ΔRbf1 promoter showed modest but reproducibly higher activity than the wild-type T1 promoter, indicating that this Rbf1 binding region downregulates expression (Fig. 4B). The T1 promoter proximal region has a relatively modest transcriptional output, so we explored the significance of Rbf1 in the context of more active reporters with

elements 2, 3 or 12. Particularly for the fusion containing region 3, the transcriptional impact of the small T1 deletion was much larger in absolute terms than that observed for just the basal promoter itself, suggesting that Rbf1 might not only reduce the functionality of local activators within T1, but also compromise the utility of the basal promoter for element 3 (Fig. 4C). Similar 'booster' roles for basal elements have been noted in developmentally active genes (Yuh and Davidson, 1996). The removal of the Rbf1 binding region did not dramatically alter the effects of FOXO expression, which activated element 2 and repressed 3 and 12 (Fig. 4C).

As a co-repressor, Rbf1 binds to E2F1 (also known as dE2F1) to block its activation function (Du and Pogoriler, 2006). Removal of the Rbf1 binding element, which includes E2F motifs, does not abrogate the function of T1, suggesting that other regulatory sites contribute to the activity of this promoter. We tested whether E2F1 activates the T1 promoter, and whether this requires the region involved in Rbf1 recruitment. Cotransfection of E2F1 significantly upregulated reporters containing the wild-type promoter as well as the ΔRbf1 T1 promoter; the fold stimulation was similar to that observed for the control *PCNA* promoter (Fig. 4D). *PCNA* was maximally stimulated at lower concentrations of transfected E2F1, suggesting that it might have different affinities for E2F1 binding
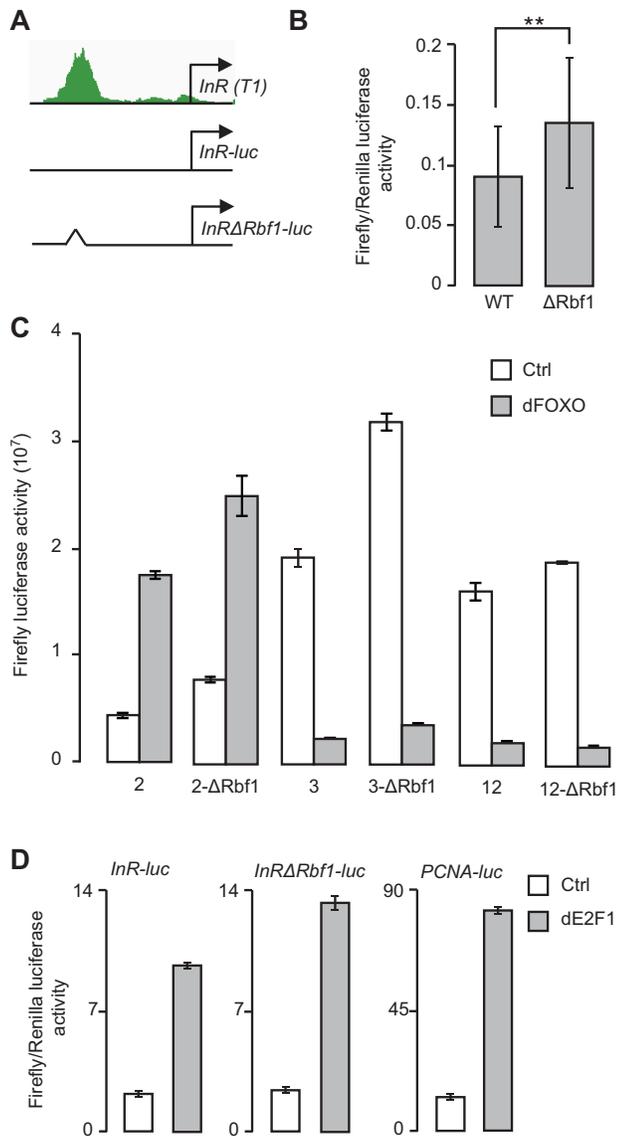
DEVELOPMENT

Fig. 4. Rbf1 role in transcriptional regulation. (A) *In vivo* Rbf1 occupancy of the *InR* T1 promoter proximal region (Acharya et al., 2012). A mutant T1 promoter-luciferase reporter was generated by deleting a 100 bp region associated with the Rbf1 binding peak (*InRΔRbf1-luc*). (B) Deletion of the Rbf1-associated region modestly increases T1 basal promoter activity. **$P<0.01$. Error bars indicate s.d. from four biological replicates. (C) Transcriptional activation by three *InR* intronic enhancers is potentiated by removal of the Rbf1 binding region. Deletion of the Rbf1 binding site had a greater impact on proximal enhancers 2 and 3 than distal enhancer 12, but all showed higher levels of activity on the *InRΔRbf1-luc* construct (white bars). FOXO-mediated activation was slightly enhanced, whereas repression was unaffected by loss of the Rbf1 binding region (gray bars). (D) E2F1 activation of the T1 *InR* promoter. *InRΔRbf1-luc* was activated by E2F1 to higher levels than the wild-type promoter, indicating that there might be E2F1 binding sites outside the Rbf1 peak, or indirect E2F1 regulation. A canonical Rbf1/E2F1 target, the *PCNA* promoter, was used as a positive control. (C,D) Error bars indicate s.d. for three technical replicates.

(Fig. S9). The ΔRbf1 T1 promoter was activated by E2F1 to a higher level than the wild-type T1 promoter, consistent with the removal of the repressive function of Rbf1, and indicating that there might be additional E2F1 binding sites that are not suitable for Rbf1 recruitment, or that the activation occurs through an indirect effect via other transcription factors. Thus, Rbf1 has a repressive function

on the T1 promoter and, more than merely interfering with local activators, this co-repressor might generally influence the ability of linked regulatory regions to fully engage and stimulate transcription from the T1 start site. This mode of regulation contrasts with the all-or-nothing effect observed for Rbf1 and Rb family proteins in general on cell cycle target promoters (Raj et al., 2012), and suggests that Rbf proteins might instead invoke a 'soft' regulatory function on certain target genes.

Transcriptional circuitry of the *InR* gene revealed by precise mapping of CREs

To obtain a more precise understanding of the transcriptional circuitry regulating the *InR* gene, we further analyzed each of the active enhancers and FOXO/20E response enhancers by making serial deletions (∼300 bp each, M1-M5) in each active fragment, testing all in S2 and Kc cells for their response to FOXO or 20E (Fig. 5A). The deletion series revealed the portions of each enhancer necessary for baseline activation, as well as regions that possess inherent repressive potential. Some deletions attenuated or abrogated the response to FOXO or 20E. With enhancer 2, a FOXO- and 20E-activated element, removal of region M1 reduced basal activity to the same level as T1, suggesting that the region contains an essential activator binding site(s) (Fig. 5B). Removal of M3 greatly induced the basal activity, indicating the presence of a repressor binding site(s). FOXO induction was somewhat attenuated by removal of either M3 or M5, indicating potential FOXO-dependent activator binding sites. The effects of these mutations are summarized using symbols for constitutive activators or repressors and for FOXO- or 20E-dependent activator or repressor effects in Fig. 5C.

For enhancer 2, removal of M3 produced a complex effect with 20E treatment: baseline expression increases, but the ability of 20E to activate is lost and instead causes repression (Fig. 5B). Repression on this element is almost certainly due to the direct binding of EcR, as this protein has been found to bind within this region (Fig. S8) and the removal of M3 has no derepressive effect in cells lacking EcR (Fig. 5B). We propose that the M3 mutant is repressed rather than activated by 20E treatment because the region contains activator sites, in addition to EcR binding sites, and these activator sites are important for overall enhancer activity. 20E treatment removes EcR and simultaneously triggers the expression of repressors (e.g. Eip74EF, which might act on 20E-repressed enhancers such as elements 3 and 12). The weaker complement of activators left on this version of enhancer 2 might be dominantly suppressed by the action of these 20E-induced repressors, whereas a wild-type enhancer would not.

We analyzed all of the regulatory fragments using the same deletion analysis in both cell types, as well as ΔEcR cells, and tested for responses to FOXO and 20E (Fig. S10). The results are summarized for all elements using symbols to indicate the presence of activator or repressor activities in subregions M1-M5 (Fig. 5C for results in Kc cells; Fig. S11 for results in S2 cells; Fig. S10 and Table S3).

Combinatorial interactions of *InR* regulatory elements

Our detailed analysis of the *cis*-regulatory landscape of the *InR* gene indicates that multiple, parallel-acting elements contribute to the overall regulation of expression. Early studies emphasized the modularity of multiple enhancers acting on developmental genes, but a number of studies have since shown how some discrete *cis*-regulatory elements function in combinatorial manners (Small et al., 1993; Marinić et al., 2013; Bothma et al., 2015). Are the regulatory units identified in the *InR* gene independently acting units that
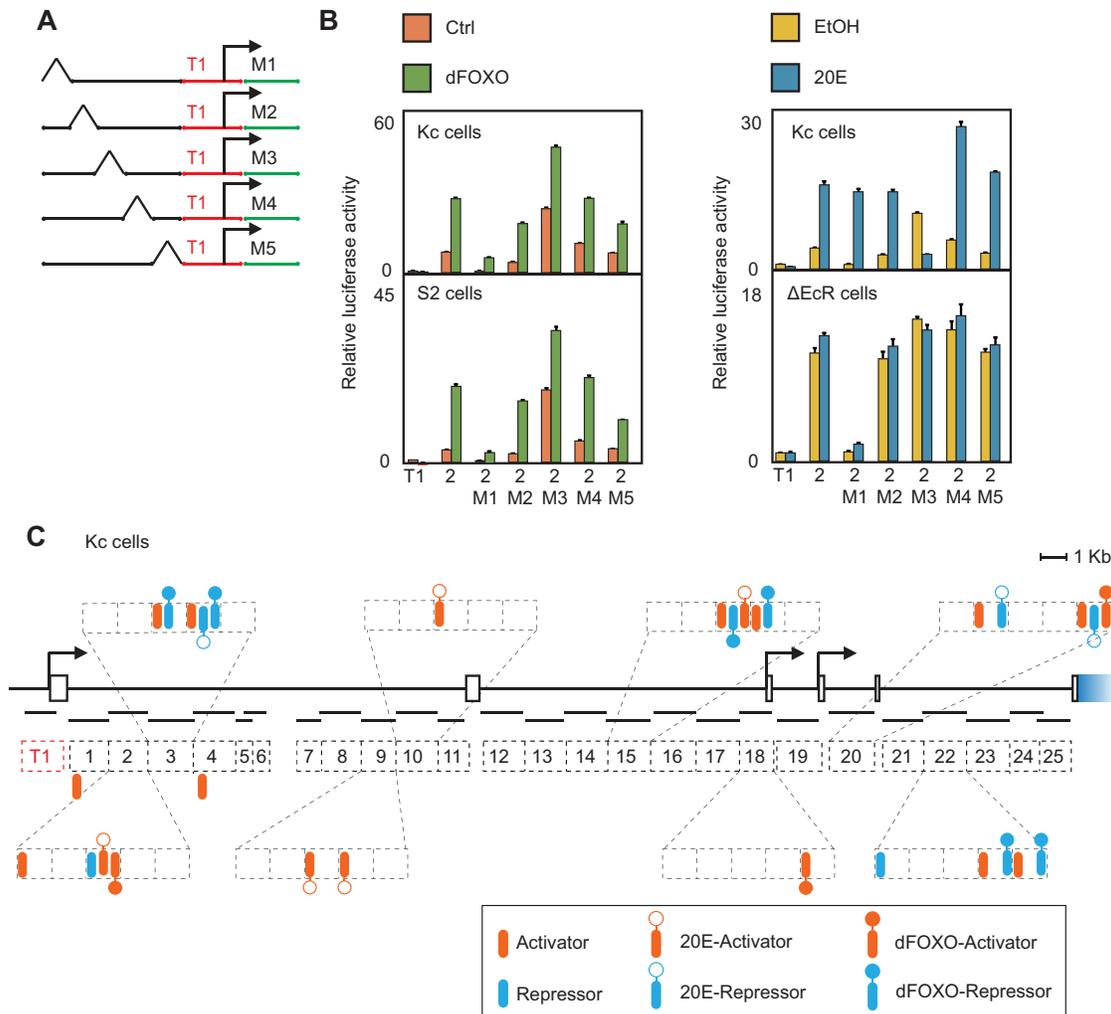
Fig. 5. High-resolution analysis of the *InR* intronic enhancers reveals complex molecular logic circuits for *cis*-regulation. (A) Deletion analysis of active intronic enhancers. Constitutively active or inducible elements were subjected to serial deletions, removing blocks of ~300 bp to yield derivatives M1-M5. The truncated reporters were then tested in S2 and Kc cells (as shown in Fig. 2). (B) Activities of truncated reporters derived from intronic fragment 2. M1 removes a constitutive activator activity, while M3 removes a constitutive repressor as well as 20E and FOXO activation potential. The entire dataset is shown in Fig. S10 and Table S3. Error bars indicate s.d. for three technical replicates. (C) A comprehensive map of enhancers and imputed regulatory factors based on analysis of all mutant reporters in Kc cells. Results for S2 cells are shown in Fig. S11.

function in an additive manner, or might there be higher-order interactions? To test this possibility, we fused *cis*-regulatory regions together and compared their activities with the individual parts (Fig. S12). For regions 2 and 3, the enhancers showed subadditive behavior, meaning that the sum was somewhat less than the individual activities. This effect might simply be a function of distance-dependent activation, a well-known property of *cis*-regulatory elements (despite the generalization that enhancers should work in a distance-independent manner) (Banerji et al., 1981). Although our reductionist analysis of the *cis*-regulatory elements of this gene serves to identify key properties of each of these molecular switches, a quantitative combinatorial understanding will come from re-integrating this information in the intact locus.

DISCUSSION
Gene size and regulatory complexity
With the completion of the first invertebrate metazoan genome sequences, the Carroll laboratory noted a correlation between gene size and functional designation (Nelson et al., 2004). A compact

structure was associated with widely expressed genes, such as those for ribosomal proteins, while transcription factors and signaling molecules were on average encoded by genes with larger intergenic spaces (subsequent annotation of distal 5′ start sites meant that some intergenic spaces are actually large introns). The elaborate expression of very large genes, such as those of the Hox clusters, represents the combined action of multiple tissue-specific regulatory elements active at different developmental stages (Montavon and Soshnikova, 2014). For these genes, the amount of non-coding DNA is clearly linked to this regulatory complexity. On the other hand, signaling molecules, such as the insulin receptor, are broadly expressed, with transcriptional expression that is ranked among the more stable (top third of all genes, rating just below ribosomal protein transcripts) (Pérez-Lluch et al., 2015). That study found that these 'broadly expressed' genes, both compact and large in size, share common histone modification states, in contrast to the different chromatin states associated with dynamically regulated genes, such as those for transcription factors. Thus, broadly expressed genes of both compact and large size share common regulatory and genomic features.

Why do genes, such as the insulin receptor, span substantial genomic regions? One observation relating to developmental gene expression is that genes harboring large introns demonstrate substantially different regulatory kinetics, whereby the traversing of extra genomic distance necessarily introduces a lag in induction and repression (Arnosti, 2011; Bothma et al., 2011). Although this effect might play a role in *InR* expression as well, it is clear that the intronic sequences of the gene serve a role that is far more complex than that of a simple spacer; we find that the transcribed space of the gene contains multiple regulatory elements that provide both redundant as well as contrasting regulatory output, some in a cell-specific manner (Fig. 6A). Other broadly expressed genes, such as those encoding ribosomal proteins, can exhibit significant regulatory responses to environmental signals, yet have relatively compact structures (Teleman et al., 2008). We speculate that there are two selective forces at play that explain the amount of regulatory DNA needed. First, simple promoters, such as those driving ribosomal protein genes, might provide the high levels of activity needed for abundant transcripts, with a certain level of dynamic regulation, but the regulation of these genes may reflect only a few types of signaling input (e.g. TOR, cell cycle) (Powers and Walter, 1999; Martin et al., 2004; Wei et al., 2015). Second, *InR* expression,

by contrast, is wired into many different signaling pathways and might need additional regulatory DNA to provide modules suitable for responses in different developmental settings, including growth control of larval tissues, stem cell niches in the larva and adult, and non-proliferating adult neurons. In addition to providing a wide spectrum of regulatory inputs, the extensive regions of DNA dedicated to control might provide necessary redundancy, so as to achieve a high degree of precision that is buffered from environmental and genetic noise.

In contrast to the highly specific on/off regulation exhibited by many developmental genes, the regulatory elements of *InR* appear to be tuned to maintain moderate responses to signals. For example, both *InR* and the *Eip74EF/Eip75B* genes are regulated by 20E and EcR (Gauhar et al., 2009; Bernardo et al., 2014). The *InR* gene contains elements that are either activated or repressed by 20E, whereas *Eip74EF/Eip75B* contain multiple copies of 20E activator elements. Upon exposure to 20E, *Eip74EF/Eip75B* expression levels increase dramatically, whereas *InR* levels increase much more modestly (Fig. S13) (Bernardo et al., 2014; Mirth et al., 2014). The incoherent feed-forward properties of the *InR* gene, with simultaneous positive and negative effects, might ensure more precise changes in gene expression, preventing pleiotropic impacts
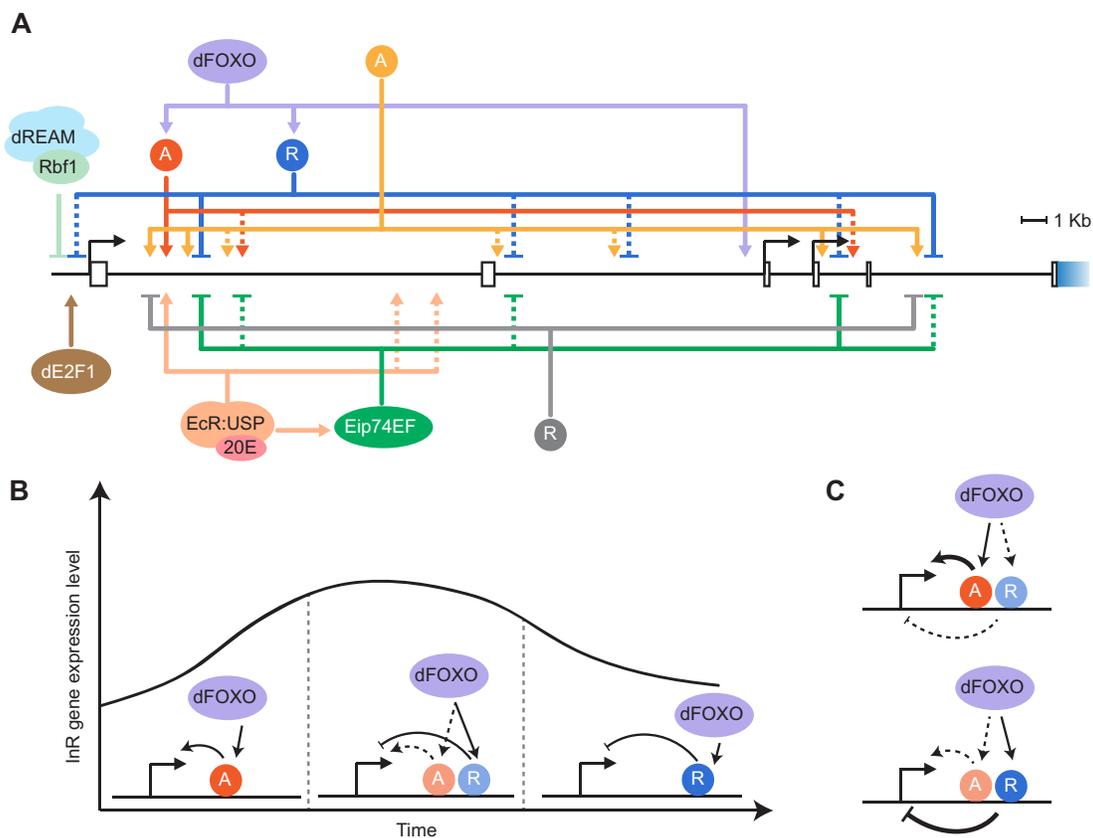

Fig. 6. Transcriptional circuitry indicates complex controls of the *InR* gene for temporal or spatial regulation. (A) Transcriptional circuitry of the *InR* gene. Regulatory inputs relating to FOXO, ecdysone, E2F1, Rbf1 and additional activators and repressors are indicated. Arrows indicate activation, T-lines repression. Dashed lines indicate elements that are active in a cell type-specific manner. The dREAM complex proteins are associated with Rbf1 at the promoter, but their functional relevance for *InR* expression is unknown (Acharya et al., 2012; Georlette et al., 2007; Wei et al., 2015). Uncharacterized activator(s) indicated by A and repressor(s) by R. Eip74EF-mediated repression role is speculative. (B) Model for temporal control of *InR* gene expression by FOXO – an incoherent feed-forward mechanism involving activation and repression. In this model, FOXO first activates *InR* through direct binding to the gene and by inducing the expression of transcriptional activators that bind *InR*. In a later phase, FOXO induces the expression of transcriptional repressors to reduce gene expression, resulting in a temporal pulse of *InR* expression. (C) Model for spatial control of *InR* gene expression by FOXO. In this model, differential action of FOXO on *InR*-binding activators and repressors leads to greater or lesser induction of *InR* expression. Both processes may be partially active to achieve an intermediate level of expression, allowing different equilibrium expression levels to be reached.

on the downstream signaling pathway. Similarly, the presence of FOXO-activated and FOXO-repressed enhancers within the *InR* locus might allow FOXO to achieve precise temporal and spatial control of *InR*. In one model, the incoherent signaling may enable a temporally complex expression pattern, whereby the direct action of FOXO first transiently upregulates *InR* gene expression, followed by a delayed downregulation via the indirectly repressed enhancers (Fig. 6A,B). In addition, the multiple layers of regulation by FOXO might also provide tissue specificity, whereby FOXO-driven activation and repression signals may be weighted differently in different cellular contexts (Fig. 6A,C). Interestingly, a recent study of the broadly expressed *ptc* gene in *Drosophila* reveals a similar genetic architecture, with large amounts of regulatory DNA devoted to fine-tuning of gene expression via multiple independent elements (Lorberbaum et al., 2016). Our results suggest that the complex regulatory control found for *InR* might be representative of certain broadly expressed genes that have functions that necessitate complex developmental and physiological inputs, and thus a high degree of regulatory precision. Whether gene size generally reflects regulatory sophistication in other metazoan genomes remains to be explored; the selection on non-coding DNA varies in different lineages (Hartl, 2000). It is interesting that the mammalian *INSR* gene spans ~200 kb, with many features consistent with enhancer complexity, including a predicted 'super-enhancer' state within the transcription unit (Wei et al., 2016).

FOXO regulation via direct and indirect pathways

Our study provides new insights into the transcriptional regulation mediated by FOXO, a key player in nutritionally driven developmental plasticity and insulin sensitivity (Tang et al., 2011). Feedback regulation by FOXO transcription factors controls the expression of *InR* (Jünger et al., 2003; Puig et al., 2003; Puig and Tjian, 2005). This aspect of transcriptional regulation of *InR* has been studied in molecular detail; based on genetic perturbation studies and transcriptional reporter assays with short segments of the *InR* gene, a previous model suggested that FOXO regulation consists of direct binding to activate a single internal promoter (Puig et al., 2003; Casas-Tinto et al., 2007). Our data indicate that FOXO regulation is far more complex; we confirm the direct, if modest, activation role for FOXO on an internal promoter, but FOXO indirectly activates or represses at least half a dozen additional enhancers located within introns of the *InR* gene (Fig. 2B, Fig. 6A). The majority of this regulation appears to rely on transcriptional intermediates and multiple regulatory layers (Fig. 6A).

How common is such concerted direct and indirect regulation by FOXO? Hundreds of genes are suggested to be direct targets of FOXO regulation in *Drosophila*, although few have been investigated further for transcriptional regulation (Alic et al., 2011; Bai et al., 2013). For genes with small promoter regions, such as *Thor*, direct activation by FOXO might represent the bulk of the regulation. However, other genes appear to be subject to incoherent feed-forward regulation, in which a factor confers both positive and negative effects. The *RpL24-like* promoter is directly repressed by FOXO and is activated by the transcription factor Myc, which is in turn activated by FOXO, establishing a two-layer regulation of this ribosomal protein gene (Teleman et al., 2008; Alic et al., 2011; Herter et al., 2015). The compact promoter of *RpL24-like* probably does not approach the complexity of regulation seen with *InR*, which might reflect the importance of fine control of the receptor gene at the apex of this signaling cascade. Thus, it remains to be established how often FOXO target genes are regulated via multiple enhancers through complex direct and indirect paths, but

simple direct activation may represent only one class of important FOXO effects.

Regulation of insulin receptor expression in development and disease

Molecular analysis of the transcriptional controls for Hox and other developmental genes has transformed our understanding of the mechanisms of gene expression, development and evolution. The transcriptional control of broadly expressed genes, such as *InR*, has received less attention. However, the sophistication of *InR* transcriptional wiring, which produces fewer extremes of expression, clearly points to strong selection for specific types of expression. Distinct variants in the coding regions of this gene are strongly selected in different *Drosophila* populations, and it is likely that sequence variation within *InR* intronic enhancers will be targets for evolution at a population and species level (Paaby et al., 2014) (A.S., unpublished). Understanding the impact of standing natural variation is also likely to provide insights into pathological states. In a *Drosophila* cancer model system, tumorigenesis associated with a high-sugar diet involves *InR* upregulation via the Wnt signaling pathway, and misregulation of human *INSR* is noted in cancer, type 2 diabetes and Alzheimer's disease (Gunton et al., 2005; Freude et al., 2009; Belfiore and Malaguarnera, 2011; Hirabayashi et al., 2013). Indeed, human sequence variants associated with type 2 diabetes and Alzheimer's disease lie within candidate *INSR* enhancers (Pasquali et al., 2014). The role we find for retinoblastoma protein in the regulation of *InR* points to a coordination of cell cycle and cellular signaling.

Our transcriptional map identifies enhancers that are likely to be targets of functional mutations, although more detailed studies will provide a better measure of candidate transcription factor binding sites. Many genome-wide studies utilize chromatin marks as proxies for active enhancers; however, it is significant that despite the many classes of ENCODE data available for this system, direct tests were essential for the identification and characterization of enhancers, which were not easily identified from general chromatin features. In fact, even dynamic chromatin features may reflect off-target effects of transcription factors, rather than functional interactions (Kok et al., 2015). Although likely to be incomplete, our identification of the *cis*-regulatory circuitry of *InR* represents a first step in enabling the construction of computational models, which can be tested in physiological settings to understand the impact of regulatory sequence variation and signaling (Samee et al., 2015; Sayal et al., 2016).

MATERIALS AND METHODS
Fly strains and transgenic lines
Fly strains were obtained from the Bloomington Stock Center as indicated in the supplementary Material and Methods. Transgenic flies were generated using a BAC containing the entire *InR* locus. mRNA from 3-day-old transgenic flies was analyzed by qPCR. For details, see the supplementary Material and Methods.

Luciferase reporters and assays
Luciferase reporters were constructed as previously described (Zhang et al., 2014). For details of reporters and luciferase assays, see the supplementary Material and Methods, Tables S4 and S5.

Cell culture and transfection
Drosophila S2 cells, Kc cells (Kc167) and Δ*EcR* cells (derived from Kc cells, obtained from Drosophila Genomics Resource Center, ID: L57-3-11) were cultured in Schneider's medium (Gibco) supplied with 10% FBS

(Gibco) and penicillin-streptomycin (100 units/ml penicillin, 100 μg/ml streptomycin, Gibco). Details of transfections are provided in the supplementary Material and Methods.

FOXO ChIP-qPCR

ChIP-qPCR assays in cell culture were conducted as previously described (Carey et al., 2009). For details, see the supplementary Material and Methods.

Ecdysone treatment

Kc or S2 cells were treated with 20E and then subject to qPCR analysis as described in the supplementary Material and Methods.

Genome data

Published STARR-seq, DNase hypersensitive site (DHS)-seq, FAIRE and ChIP-seq data were obtained and analyzed as described in the supplementary Material and Methods.

References

Acharya, P., Negre, N., Johnston, J., Wei, Y., White, K. P., Henry, R. W. and Arnosti, D. N. (2012). Evidence for autoregulation and cell signaling pathway regulation from genome-wide binding of the Drosophila retinoblastoma protein. *G3 (Bethesda)* **2**, 1459-1472.

Alic, N., Andrews, T. D., Giannakou, M. E., Papatheodorou, I., Slack, C., Hoddinott, M. P., Cocheme, H. M., Schuster, E. F., Thornton, J. M. and Partridge, L. (2011). Genome-wide dFOXO targets and topology of the transcriptomic response to stress and insulin signalling. *Mol. Syst. Biol.* **7**, 502.

Arnold, C. D., Gerlach, D., Stelzer, C., Boryn, L. M., Rath, M. and Stark, A. (2013). Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science* **339**, 1074-1077.

Arnosti, D. N. (2011). Transcriptional repressors: shutting off gene expression at the source affects developmental dynamics. *Curr. Biol.* **21**, R859-R860.

Attrill, H., Falls, K., Goodman, J. L., Millburn, G. H., Antonazzo, G., Rey, A. J. and Marygold, S. J. (2016). FlyBase: establishing a Gene Group resource for Drosophila melanogaster. *Nucleic Acids Res.* **44**, D786-D792.

Bai, H., Kang, P., Hernandez, A. M. and Tatar, M. (2013). Activin signaling targeted by insulin/dFOXO regulates aging and muscle proteostasis in Drosophila. *PLoS Genet.* **9**, e1003941.

Banerji, J., Rusconi, S. and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**, 299-308.

Belfiore, A. and Malaguarnera, R. (2011). Insulin receptor and cancer. *Endocr. Relat. Cancer* **18**, R125-R147.

Bernardo, T. J., Dubrovskaya, V. A., Xie, X. and Dubrovsky, E. B. (2014). A view through a chromatin loop: insights into the ecdysone activation of early genes in Drosophila. *Nucleic Acids Res.* **42**, 10409-10424.

Bertrand, L., Horman, S., Beauloye, C. and Vanoverschelde, J.-L. (2008). Insulin signalling in the heart. *Cardiovasc. Res.* **79**, 238-248.

Bothma, J. P., Magliocco, J. and Levine, M. (2011). The snail repressor inhibits release, not elongation, of paused Pol II in the Drosophila embryo. *Curr. Biol.* **21**, 1571-1577.

Bothma, J. P., Garcia, H. G., Ng, S., Perry, M. W., Gregor, T. and Levine, M. (2015). Enhancer additivity and non-additivity are determined by enhancer strength in the Drosophila embryo. *Elife* **4**, e07956.

Brogiolo, W., Stocker, H., Ikeya, T., Rintelen, F., Fernandez, R. and Hafen, E. (2001). An evolutionarily conserved function of the Drosophila insulin receptor and insulin-like peptides in growth control. *Curr. Biol.* **11**, 213-221.

Brown, J. B., Boley, N., Eisman, R., May, G. E., Stoiber, M. H., Duff, M. O., Booth, B. W., Wen, J., Park, S., Suzuki, A. M. et al. (2014). Diversity and dynamics of the Drosophila transcriptome. *Nature* **512**, 393-399.

Brunetti, A., Manfioletti, G., Chiefari, E., Goldfine, I. D. and Foti, D. (2001). Transcriptional regulation of human insulin receptor gene by the high-mobility group protein HMGI(Y). *FASEB J.* **15**, 492-500.

Calle, C., Maestro, B. and García-Arencibia, M. (2008). Genomic actions of 1,25-dihydroxyvitamin D3 on insulin receptor gene expression, insulin receptor number and insulin activity in the kidney, liver and adipose tissue of streptozotocin-induced diabetic rats. *BMC Mol. Biol.* **9**, 65.

Cameron, K. E., Resnik, J. and Webster, N. J. (1992). Transcriptional regulation of the human insulin receptor promoter. *J. Biol. Chem.* **267**, 17375-17383.

Carey, M. F., Peterson, C. L. and Smale, S. T. (2009). Chromatin immunoprecipitation (ChIP). *Cold Spring Harb. Protoc.* **2009**, pdb.prot5279.

Casas-Tinto, S., Marr, M. T., Andreu, P. and Puig, O. (2007). Characterization of the Drosophila insulin receptor promoter. *Biochim. Biophys. Acta* **1769**, 236-243.

Chatterjee, M. (2001). Vitamin D and genomic stability. *Mutat. Res.* **475**, 69-87.

Chicas, A., Wang, X., Zhang, C., McCurrach, M., Zhao, Z., Mert, O., Dickins, R. A., Narita, M., Zhang, M. and Lowe, S. W. (2010). Dissecting the unique role of the retinoblastoma tumor suppressor during cellular senescence. *Cancer Cell* **17**, 376-387.

Colombani, J., Bianchini, L., Layalle, S., Pondeville, E., Dauphin-Villemant, C., Antoniewski, C., Carre, C., Noselli, S. and Leopold, P. (2005). Antagonistic actions of ecdysone and insulins determine final size in Drosophila. *Science* **310**, 667-670.

Du, W. and Pogoriler, J. (2006). Retinoblastoma family genes. *Oncogene* **25**, 5190-5200.

Ebina, Y., Ellis, L., Jarnagin, K., Edery, M., Graf, L., Clauser, E., Ou, J.-H., Masiarz, F., Kan, Y. W. and Goldfine, I. D. (1985). The human insulin receptor cDNA: the structural basis for hormone-activated transmembrane signalling. *Cell* **40**, 747-758.

Fernandez, R., Tabarini, D., Azpiazu, N., Frasch, M. and Schlessinger, J. (1995). The Drosophila insulin receptor homolog: a gene essential for embryonic development encodes two receptor isoforms with different signaling potential. *EMBO J.* **14**, 3373-3384.

Freude, S., Schilbach, K. and Schubert, M. (2009). The role of IGF-1 receptor and insulin receptor signaling for the pathogenesis of Alzheimer's disease: from model organisms to human disease. *Curr. Alzheimer Res.* **6**, 213-223.

Frölich, L., Blum-Degen, D., Bernstein, H.-G., Engelsberger, S., Humrich, J., Laufer, S., Muschner, D., Thalheimer, A., Türk, A., Hoyer, S. et al. (1998). Brain insulin and insulin receptors in aging and sporadic Alzheimer's disease. *J. Neural Transm.* **105**, 423-438.

García-Arencibia, M., Davila, N., Campion, J., Carmen Carranza, M. and Calle, C. (2005). Identification of two functional estrogen response elements complexed with AP-1-like sites in the human insulin receptor gene promoter. *J. Steroid Biochem. Mol. Biol.* **94**, 1-14.

Garofalo, R. S. and Rosen, O. M. (1988). Tissue localization of Drosophila melanogaster insulin receptor transcripts during development. *Mol. Cell. Biol.* **8**, 1638-1647.

Gauhar, Z., Sun, L. V., Hua, S., Mason, C. E., Fuchs, F., Li, T.-R., Boutros, M. and White, K. P. (2009). Genomic mapping of binding regions for the Ecdysone receptor protein complex. *Genome Res.* **19**, 1006-1013.

Georlette, D., Ahn, S., MacAlpine, D. M., Cheung, E., Lewis, P. W., Beall, E. L., Bell, S. P., Speed, T., Manak, J. R. and Botchan, M. R. (2007). Genomic profiling and expression studies reveal both positive and negative activities for the Drosophila Myb MuvB/dREAM complex in proliferating cells. *Genes Dev.* **21**, 2880-2896.

Gershman, B., Puig, O., Hang, L., Peitzsch, R. M., Tatar, M. and Garofalo, R. S. (2007). High-resolution dynamics of the transcriptional response to nutrition in Drosophila: a key role for dFOXO. *Physiol. Genomics* **29**, 24-34.

Giacinti, C. and Giordano, A. (2006). RB and cell cycle progression. *Oncogene* **25**, 5220-5227.

Gonzalez, E., Flier, E., Molle, D., Accili, D. and McGraw, T. E. (2011). Hyperinsulinemia leads to uncoupled insulin regulation of the GLUT4 glucose transporter and the FoxO1 transcription factor. *Proc. Natl. Acad. Sci. USA* **108**, 10162-10167.

Graveley, B. R., Brooks, A. N., Carlson, J. W., Duff, M. O., Landolin, J. M., Yang, L., Artieri, C. G., van Baren, M. J., Boley, N., Booth, B. W. et al. (2011). The developmental transcriptome of Drosophila melanogaster. *Nature* **471**, 473-479.

Gunton, J. E., Kulkarni, R. N., Yim, S., Okada, T., Hawthorne, W. J., Tseng, Y.-H., Roberson, R. S., Ricordi, C., O'Connell, P. J., Gonzalez, F. J. et al. (2005). Loss

of ARNT/HIF1beta mediates altered gene expression and pancreatic-islet dysfunction in human type 2 diabetes. *Cell* **122**, 337-349.

Hartl, D. L. (2000). Molecular melodies in high and low C. *Nat. Rev. Genet.* **1**, 145-149.

Herter, E. K., Stauch, M., Gallant, M., Wolf, E., Raabe, T. and Gallant, P. (2015). snoRNAs are a novel class of biologically relevant Myc targets. *BMC Biol.* **13**, 25.

Hirabayashi, S., Baranski, T. J. and Cagan, R. L. (2013). Transformed Drosophila cells evade diet-mediated insulin resistance through wingless signaling. *Cell* **154**, 664-675.

Hu, X., Cherbas, L. and Cherbas, P. (2003). Transcription activation by the ecdysone receptor (EcR/USP): identification of activation functions. *Mol. Endocrinol.* **17**, 716-731.

Iritani, N., Sugimoto, T. and Fukuda, H. (2000). Gene expressions of leptin, insulin receptors and lipogenic enzymes are coordinately regulated by insulin and dietary fat in rats. *J. Nutr.* **130**, 1183-1188.

Jenett, A., Rubin, G. M., Ngo, T.-T. B., Shepherd, D., Murphy, C., Dionne, H., Pfeiffer, B. D., Cavallaro, A., Hall, D., Jeter, J. et al. (2012). A GAL4-driver line resource for Drosophila neurobiology. *Cell Rep.* **2**, 991-1001.

Jory, A., Estella, C., Giorgianni, M. W., Slattery, M., Laverty, T. R., Rubin, G. M. and Mann, R. S. (2012). A survey of 6,300 genomic fragments for cis-regulatory activity in the imaginal discs of Drosophila melanogaster. *Cell Rep.* **2**, 1014-1024.

Jünger, M. A., Rintelen, F., Stocker, H., Wasserman, J. D., Vegh, M., Radimerski, T., Greenberg, M. E. and Hafen, E. (2003). The Drosophila forkhead transcription factor FOXO mediates the reduction in cell number associated with reduced insulin signaling. *J. Biol.* **2**, 20.

Kaplan, T., Li, X.-Y., Sabo, P. J., Thomas, S., Stamatoyannopoulos, J. A., Biggin, M. D. and Eisen, M. B. (2011). Quantitative models of the mechanisms that control genome-wide patterns of transcription factor binding during early Drosophila development. *PLoS Genet.* **7**, e1001290.

Kim, S.-C., Jang, H.-C., Lee, S.-D., Jung, H.-J., Park, J.-C., Lee, S.-H., Kim, T.-H. and Choi, B.-H. (2014). Changes in expression of insulin signaling pathway genes by dietary fat source in growing-finishing pigs. *J. Anim. Sci. Technol.* **56**, 12.

Koelle, M. R., Talbot, W. S., Segraves, W. A., Bender, M. T., Cherbas, P. and Hogness, D. S. (1991). The Drosophila EcR gene encodes an ecdysone receptor, a new member of the steroid receptor superfamily. *Cell* **67**, 59-77.

Kok, K., Ay, A., Li, L. M. and Arnosti, D. N. (2015). Genome-wide errant targeting by Hairy. *Elife* **4**, e06394.

Korenjak, M., Anderssen, E., Ramaswamy, S., Whetstine, J. R. and Dyson, N. J. (2012). RBF binding to both canonical E2F targets and noncanonical targets depends on functional dE2F/dDP complexes. *Mol. Cell. Biol.* **32**, 4375-4387.

Koyama, T., Rodrigues, M. A., Athanasiadis, A., Shingleton, A. W. and Mirth, C. K. (2014). Nutritional control of body size through FoxO-Ultraspiracle mediated ecdysone biosynthesis. *Elife* **3**, e03091.

Kriaucijunas, K. M., Goldstein, B. J., Lipes, M. A. and Kahn, C. R. (1993). Modulation of expression of insulin and IGF-I receptor by Epstein-Barr virus and its gene products LMP and EBNA-2 in lymphocyte cell lines. *J. Cell. Physiol.* **154**, 486-495.

Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M. J. et al. (2015). Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317-330.

Leal, M. A., Cabanas, C., Rius, C., Aller, P. and Calle, C. (1992). Modulation by dexamethasone of insulin binding and insulin receptor mRNA levels in U-937 human promonocytic cells. *Biochimie* **74**, 545-549.

Lee, J. K. and Tsai, S. Y. (1994). Multiple hormone response elements can confer glucocorticoid regulation on the human insulin receptor gene. *Mol. Endocrinol.* **8**, 625-634.

Lee, J. K., Tam, J. W., Tsai, M. J. and Tsai, S. Y. (1992). Identification of cis- and trans-acting factors regulating the expression of the human insulin receptor gene. *J. Biol. Chem.* **267**, 4638-4645.

Li, X.-Y., Thomas, S., Sabo, P. J., Eisen, M. B., Stamatoyannopoulos, J. A. and Biggin, M. D. (2011). The role of chromatin accessibility in directing the widespread, overlapping patterns of Drosophila transcription factor binding. *Genome Biol.* **12**, R34.

Li, H.-H., Kroll, J. R., Lennox, S. M., Ogundeyi, O., Jeter, J., Depasquale, G. and Truman, J. W. (2014). A GAL4 driver resource for developmental and behavioral studies on the larval CNS of Drosophila. *Cell Rep.* **8**, 897-908.

Lorberbaum, D. S., Ramos, A. I., Peterson, K. A., Carpenter, B. S., Parker, D. S., De, S., Hillers, L. E., Blake, V. M., Nishi, Y., McFarlane, M. R. et al. (2016). An ancient yet flexible cis-regulatory architecture allows localized Hedgehog tuning by patched/Ptch1. *Elife* **5**, pii: e13550.

Mamula, P. W., McDonald, A. R., Brunetti, A., Okabayashi, Y., Wong, K. Y., Maddux, B. A., Logsdon, C. and Goldfine, I. D. (1990). Regulating insulin-receptor-gene expression by differentiation and hormones. *Diabetes Care* **13**, 288-301.

Manning, L., Heckscher, E. S., Purice, M. D., Roberts, J., Bennett, A. L., Kroll, J. R., Pollard, J. L., Strader, M. E., Lupton, J. R., Dyukareva, A. V. et al. (2012). A resource for manipulating gene expression and analyzing cis-regulatory modules in the Drosophila CNS. *Cell Rep.* **2**, 1002-1013.

Marinić, M., Aktas, T., Ruf, S. and Spitz, F. (2013). An integrated holo-enhancer unit defines tissue and gene specificity of the Fgf8 regulatory landscape. *Dev. Cell* **24**, 530-542.

Martin, D. E., Soulard, A. and Hall, M. N. (2004). TOR regulates ribosomal protein gene expression via PKA and the Forkhead transcription factor FHL1. *Cell* **119**, 969-979.

McKay, D. J. and Lieb, J. D. (2013). A common set of DNA regulatory elements shapes Drosophila appendages. *Dev. Cell* **27**, 306-318.

Mirth, C. K., Tang, H. Y., Makohon-Moore, S. C., Salhadar, S., Gokhale, R. H., Warner, R. D., Koyama, T., Riddiford, L. M. and Shingleton, A. W. (2014). Juvenile hormone regulates body size and perturbs insulin signaling in Drosophila. *Proc. Natl. Acad. Sci. USA* **111**, 7018-7023.

Moloney, A. M., Griffin, R. J., Timmons, S., O'Connor, R., Ravid, R. and O'Neill, C. (2010). Defects in IGF-1 receptor, insulin receptor and IRS-1/2 in Alzheimer's disease indicate possible resistance to IGF-1 and insulin signalling. *Neurobiol. Aging* **31**, 224-243.

Montavon, T. and Soshnikova, N. (2014). Hox gene regulation and timing in embryogenesis. *Semin. Cell Dev. Biol.* **34**, 76-84.

Nègre, N., Brown, C. D., Ma, L., Bristow, C. A., Miller, S. W., Wagner, U., Kheradpour, P., Eaton, M. L., Loriaux, P., Sealfon, R. et al. (2011). A cis-regulatory map of the Drosophila genome. *Nature* **471**, 527-531.

Nelson, C. E., Hersh, B. M. and Carroll, S. B. (2004). The regulatory content of intergenic DNA shapes genome architecture. *Genome Biol.* **5**, R25.

Oldham, S. and Hafen, E. (2003). Insulin/IGF and target of rapamycin signaling: a TOR de force in growth control. *Trends Cell Biol.* **13**, 79-85.

Paaby, A. B., Bergland, A. O., Behrman, E. L. and Schmidt, P. S. (2014). A highly pleiotropic amino acid polymorphism in the Drosophila insulin receptor contributes to life-history adaptation. *Evolution* **68**, 3395-3409.

Pasquali, L., Gaulton, K. J., Rodriguez-Segui, S. A., Mularoni, L., Miguel-Escalada, I., Akerman, I., Tena, J. J., Moran, I., Gomez-Marin, C., van de Bunt, M. et al. (2014). Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nat. Genet.* **46**, 136-143.

Pérez-Lluch, S., Blanco, E., Tilgner, H., Curado, J., Ruiz-Romero, M., Corominas, M. and Guigo, R. (2015). Absence of canonical marks of active chromatin in developmentally regulated genes. *Nat. Genet.* **47**, 1158-1167.

Petruzzelli, L., Herrera, R., Arenas-Garcia, R., Fernandez, R., Birnbaum, M. J. and Rosen, O. M. (1986). Isolation of a Drosophila genomic sequence homologous to the kinase domain of the human insulin receptor and detection of the phosphorylated Drosophila receptor with an anti-peptide antibody. *Proc. Natl. Acad. Sci. USA* **83**, 4710-4714.

Pfeiffer, B. D., Jenett, A., Hammonds, A. S., Ngo, T.-T. B., Misra, S., Murphy, C., Scully, A., Carlson, J. W., Wan, K. H., Laverty, T. R. et al. (2008). Tools for neuroanatomy and neurogenetics in Drosophila. *Proc. Natl. Acad. Sci. USA* **105**, 9715-9720.

Powers, T. and Walter, P. (1999). Regulation of ribosome biogenesis by the rapamycin-sensitive TOR-signaling pathway in Saccharomyces cerevisiae. *Mol. Biol. Cell* **10**, 987-1000.

Puig, O. and Tjian, R. (2005). Transcriptional feedback control of insulin receptor by dFOXO/FOXO1. *Genes Dev.* **19**, 2435-2446.

Puig, O., Marr, M. T., Ruhf, M. L. and Tjian, R. (2003). Control of cell number by Drosophila FOXO: downstream and feedback regulation of the insulin receptor pathway. *Genes Dev.* **17**, 2006-2020.

Raj, N., Zhang, L., Wei, Y., Arnosti, D. N. and Henry, R. W. (2012). Ubiquitination of retinoblastoma family protein 1 potentiates gene-specific repression function. *J. Biol. Chem.* **287**, 41835-41843.

Riddiford, L. M., Cherbas, P. and Truman, J. W. (2000). Ecdysone receptors and their biological actions. *Vitam. Horm.* **60**, 1-73.

Ryu, J.-R. and Arnosti, D. N. (2003). Functional similarity of Knirps CtBP-dependent and CtBP-independent transcriptional repressor activities. *Nucleic Acids Res.* **31**, 4654-4662.

Samee, M. A. H., Lim, B., Samper, N., Lu, H., Rushlow, C. A., Jimenez, G., Shvartsman, S. Y. and Sinha, S. (2015). A systematic ensemble approach to thermodynamic modeling of gene expression from sequence data. *Cell Syst.* **1**, 396-407.

Sayal, R., Dresch, J. M., Pushel, I., Taylor, B. R. and Arnosti, D. N. (2016). Quantitative perturbation-based analysis of gene expression predicts enhancer activity in early Drosophila embryo. *Elife* **5**, pii: e08445.

Shen, W.-J., Kim, H. S. and Tsai, S. Y. (1995). Stimulation of human insulin receptor gene expression by retinoblastoma gene product. *J. Biol. Chem.* **270**, 20525-20529.

Shlyueva, D., Stelzer, C., Gerlach, D., Yanez-Cuna, J. O., Rath, M., Boryn, L. M., Arnold, C. D. and Stark, A. (2014). Hormone-responsive enhancer-activity maps reveal predictive motifs, indirect repression, and targeting of closed chromatin. *Mol. Cell* **54**, 180-192.

Small, S., Arnosti, D. N. and Levine, M. (1993). Spacing ensures autonomous expression of different stripe enhancers in the even-skipped promoter. *Development* **119**, 762-772.

Song, J., Wu, L., Chen, Z., Kohanski, R. A. and Pick, L. (2003). Axons guided by insulin receptor in Drosophila visual system. *Science* **300**, 502-505.

DEVELOPMENT

Swevers, L., Cherbas, L., Cherbas, P. and Iatrou, K. (1996). Bombyx EcR (BmEcR) and Bombyx USP (BmCF1) combine to form a functional ecdysone receptor. *Insect Biochem. Mol. Biol.* **26**, 217-221.

Taguchi, A. and White, M. F. (2008). Insulin-like signaling, nutrient homeostasis, and life span. *Annu. Rev. Physiol.* **70**, 191-212.

Tang, H. Y., Smith-Caldas, M. S. B., Driscoll, M. V., Salhadar, S. and Shingleton, A. W. (2011). FOXO regulates organ-specific phenotypic plasticity in Drosophila. *PLoS Genet.* **7**, e1002373.

Teleman, A. A., Hietakangas, V., Sayadian, A. C. and Cohen, S. M. (2008). Nutritional control of protein biosynthetic capacity by insulin via Myc in Drosophila. *Cell Metab.* **7**, 21-32.

Thomas, S., Li, X.-Y., Sabo, P. J., Sandstrom, R., Thurman, R. E., Canfield, T. K., Giste, E., Fisher, W., Hammonds, A., Celniker, S. E. et al. (2011). Dynamic reprogramming of chromatin accessibility during Drosophila embryo development. *Genome Biol.* **12**, R43.

Ullrich, A., Bell, J. R., Chen, E. Y., Herrera, R., Petruzzelli, L. M., Dull, T. J., Gray, A., Coussens, L., Liao, Y.-C., Tsubokawa, M. et al. (1985). Human insulin receptor and its relationship to the tyrosine kinase family of oncogenes. *Nature* **313**, 756-761.

Vanhille, L., Griffon, A., Maqbool, M. A., Zacarias-Cabeza, J., Dao, L. T. M., Fernandez, N., Ballester, B., Andrau, J. C. and Spicuglia, S. (2015). High-throughput and quantitative assessment of enhancer activity in mammals by CapStarr-seq. *Nat. Commun.* **6**, 6905.

Webster, N. J., Resnik, J. L., Reichart, D. B., Strauss, B., Haas, M. and Seely, B. L. (1996). Repression of the insulin receptor promoter by the tumor suppressor gene product p53: a possible mechanism for receptor overexpression in breast cancer. *Cancer Res.* **56**, 2781-2788.

Wei, Y., Mondal, S. S., Mouawad, R., Wilczynski, B., Henry, R. W. and Arnosti, D. N. (2015). Genome-wide analysis of Drosophila Rbf2 protein highlights diversity of RB family targets and possible role in regulation of ribosome biosynthesis. *G3 (Bethesda)* **5**, 1503-1515.

Wei, Y., Zhang, S., Shang, S., Zhang, B., Li, S., Wang, X., Wang, F., Su, J., Wu, Q., Liu, H. et al. (2016). SEA: a super-enhancer archive. *Nucleic Acids Res.* **44**, D172-D179.

Wong, D. M., Shen, Z., Owyang, K. E. and Martinez-Agosto, J. A. (2014). Insulin- and warts-dependent regulation of tracheal plasticity modulates systemic larval growth during hypoxia in Drosophila melanogaster. *PLoS ONE* **9**, e115297.

Yuh, C. H. and Davidson, E. H. (1996). Modular cis-regulatory organization of Endo16, a gut-specific gene of the sea urchin embryo. *Development* **122**, 1069-1082.

Zabidi, M. A., Arnold, C. D., Schernhuber, K., Pagani, M., Rath, M., Frank, O. and Stark, A. (2015). Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* **518**, 556-559.

Zhang, L., Wei, Y., Pushel, I., Heinze, K., Elenbaas, J., Henry, R. W. and Arnosti, D. N. (2014). Integrated stability and activity control of the Drosophila Rbf1 retinoblastoma protein. *J. Biol. Chem.* **289**, 24863-24873.

DEVELOPMENT