The Hox transcription factor Ultrabithorax binds RNA and regulates co-transcriptional splicing through an interplay with RNA polymerase II

Julie Carnesecchi1,4*, Panagiotis Boumpas1,†, Patrick van Nierop y Sanchez1,†, Katrin Domsch1,2,‡, Hugo Daniel Pinto3,‡, Pedro Borges Pinto1,4 and Ingrid Lohmann1

1Heidelberg University, Centre for Organismal Studies (COS) Heidelberg, Department of Developmental Biology, Heidelberg, Germany, 2Friedrich-Alexander-University Erlangen-Nürnberg, Department Biology, Division of Developmental Biology, Erlangen, Germany, 3Department of Cell Biology, Albert Einstein College of Medicine, New York, NY, USA and 4Institut de Génomique Fonctionnelle de Lyon, Université de Lyon, CNRS UMR5242, Ecole Normale Supérieure de Lyon, Lyon, France

ABSTRACT

Transcription factors (TFs) play a pivotal role in cell fate decision by coordinating gene expression programs. Although most TFs act at the DNA layer, few TFs bind RNA and modulate splicing. Yet, the mechanistic cues underlying TFs activity in splicing remain elusive. Focusing on the Drosophila Hox TF Ultrabithorax (Ubx), our work shed light on a novel layer of Ubx function at the RNA level. Transcriptome and genome-wide binding profiles in embryonic mesoderm and Drosophila cells indicate that Ubx regulates mRNA expression and splicing to promote distinct outcomes in defined cellular contexts. Our results demonstrate a new RNA-binding ability of Ubx. We find that the N51 amino acid of the DNA-binding Homeodomain is non-essential for RNA interaction in vitro, but is required for RNA interaction in vivo and Ubx splicing activity. Moreover, mutation of the N51 amino acid weakens the interaction between Ubx and active RNA Polymerase II (Pol II). Our results reveal that Ubx regulates elongation-coupled splicing, which could be coordinated by a dynamic interplay with active Pol II on chromatin. Overall, our work uncovered a novel role of the Hox TFs at the mRNA regulatory layer. This could be an essential function for other classes of TFs to control cell diversity.

INTRODUCTION

Eukaryotic gene expression is a fascinating process: it creates various proteomes from limited genetic material, thereby promoting cell and tissue diversity in multicellular organisms. To do so, genes are transcribed into premRNAs that undergo a series of RNA processing events such as 5′capping, splicing and 3′polyadenylation (1). Splicing, the mechanism of exon/intron retention or excision, plays an important role in proteome diversity (2). It relies on the dynamic assembly of a ribonucleoprotein complex termed spliceosome and numerous accessory proteins that impel the RNA fate by selecting the appropriate splice sites (3). Splicing is regulated at several levels in the nucleus (2–6). Notably, it mainly happens co-transcriptionally (3,7,8) and depends on the RNA Polymerase II (Pol II) elongation activity (4,9–13). Remarkably, a single pre-mRNA can be spliced in different ways thereby diversifying transcript isoforms and proteins (7). This unique strategy termed alternative splicing (AS) contributes greatly to the diversity of cell and tissue identities in complex organisms (14). Despite its fundamental contribution to the proteome diversity (1,14), realising how transcriptional and splicing programs are coordinated is still a challenge in Biology.

Transcription factors (TFs) are key players of gene expression. They coordinate specific yet flexible transcriptional programs thereby orchestrating developmental diversity and tissue maintenance of living organisms (15,16). They do so by recognising DNA-binding sites and regulating target genes in a precise spatial and temporal manner. Intriguingly, most TFs act at the DNA regulatory layer, however few can bind RNA and/or modulate mRNA splicing (17–20). Notably, the TF Sox9 regulates gene expression via distinct functions on transcription and splicing (18). The homeodomain (HD)-containing TF Bicoid (Bcd) has been shown to interact with caudal RNA in Drosophila thereby inhibiting translation (21). Other TF families such as the C2H2-zinc-finger and the hormone nuclear receptors...
can modulate splicing, the latter via cofactors interaction (19,22). Despite the strong evidences supporting that TFs could be major regulators of mRNA splicing, the mechanistic cues underlying TFs function in AS remain elusive. More importantly, how the mRNA-regulatory function of TFs impacts on cell fate decisions is still enigmatic.

How a given TF drives various transcriptional programs in different cell and tissue types is a longstanding question in Biology. It is assumed that in vivo, TFs establish dynamic protein–protein interaction (PPI) networks (15). We tackled this issue using the Hox TFs, which are key players of animal development and cell homeostasis in adult. The Hox proteins belong to the conserved class of HD TFs (23). They are expressed along the longitudinal axis in cnidarians (24) and bilaterians (25), and orchestrate the development and homeostasis of a diversity of cell and tissue types (25,26). Understanding their molecular function has been a central aim in Developmental Biology. We previously used a tissue-specific proximity-labelling proteomic method to capture interactors of the Hox TF Ultrabithorax (Ubx) in Drosophila embryo (27). Our work revealed that the Ubx protein-networks are assembled at several layers of gene expression. Notably, many partners were regulators of mRNA splicing, revealing an unexpected aspect of the Hox operating mode in vivo. Based on our findings, we propose that Hox TFs coordinate gene expression programs by modulating both transcription and splicing.

In this study, we combined transcriptomic profiling of differentially spliced genes in the embryonic mesoderm and in Drosophila S2R+ cells to uncover a novel function of Ubx in splicing. Notably, Ubx regulates mRNA expression and splicing to promote distinct outcomes in different cell and tissue contexts. Comparison of transcriptome and genomewide chromatin binding profiles indicate that Ubx largely binds its target genes in exons and introns. In line, transcriptome from Drosophila S2R+ cells expressing the UbxN51A mutant, which is no longer able to bind DNA, reveals that the Ubx HD is required for its splicing function. We also uncover a novel RNA-binding ability of Ubx both in cells and in vitro. Intriguingly, the N51 amino acid of the HD, though non-essential for the interaction of Ubx with RNA in vitro, is essential for in vivo Ubx–RNA interaction. The N51A mutation of the HD impacts additionally on the interaction between Ubx and active Pol II (S2Phos). In sum, Ubx binds RNA and regulates co-transcriptional splicing. Altogether, our results strongly suggest that this mechanism is orchestrated by a dynamic interplay between Ubx and Pol II on chromatin. By extending the molecular repertoire of the Hox TFs, our work provides pivotal entry points to understand the Hox function in development and diseases, and unique view point on the role of TFs at the mRNA-regulatory level for governing cell fate decision.

**MATERIALS AND METHODS**

**Biological resources**

The following fly lines were used for the study: **GFP-Ubx** (28), w+18. **Drosophila S2R+** cells were obtained from the Drosophila Genomic Research Centre (DGRC). The PC4 plasmid was kindly obtained from the Drosophila Genomic Research Centre (DGRC). The Ubx plasmids were generated previously for (27). The vector constructs used in the study are pActin-Gal4, PC4 (from DGRC), UAS-GFPnls, UAS-myc-UbxWT, UAS-myc-UbxN51A, UAS-GFP-UBxWT, UAS-GFP-UbxN51A, pET-his-UbxWT, pET-his-UBxN51A, pET-his-MBP-UBxWT, pET-his-MBP-HDWT, pET-his-MBP-HDN51A, pGEX-Ubx-FL, pGEX-Ubx-Nter, pGEX-Ubx-Cter, pGEX-Ubx-HD, pGEX-GST, pActin-BetaGalactosidase, UAS-empty vector, UAS-RFPnls, pActin-GFPnls for plasmid quantity normalisation and quantification. For protein-RNA assay, probes listed in Supplementary Table S4 were cloned in pGEM®-T Easy (Promega) for T7 polymerase transcription.

**Cell culture and transfection**

*Drosophila S2R+* cells were maintained at 25°C in Schneider medium supplemented with 10% FCS, 10 U/ml penicillin and 10 μg/ml streptomycin. Cells were simultaneously seeded and transfected with Effectene (Qiagen) according to the manufacturer’s protocol. For all constructs the Gal4–UAS system was used for inducible protein expression driven by the Actin promoter. For whole cell protein extracts, cells were harvested in phosphate buffered saline (PBS) and pellets were resuspended in RIPA buffer supplemented with protease inhibitor cocktail (Sigma-Aldrich). For RNA analysis, cells were seeded in six-well plates and transfected as described with 750 ng UAS-GFPnls, UAS-myc-UbxWT or UAS-myc-UbxN51A, and 750 ng pActin-Gal4. For interaction, ChIP and RIP assay, 10 × 10⁶ cells were seeded in 100 mm dishes and transfected as described with UAS-GFPnls, UAS-GFP-UbxWT or UAS-GFP-UbxN51A, and pActin-Gal4. Cells were harvested in PBS after 48 hours of transfection and pellets were resuspended in lysis buffer supplemented with protease inhibitor cocktail (Sigma-Aldrich). For RNA analysis, cells were seeded in six-well plates and transfected as described with 750 ng UAS-GFPnls, UAS-myc-UbxWT or UAS-myc-UbxN51A, and 750 ng pActin-Gal4. For interaction, ChIP and RIP assay, 10 × 10⁶ cells were seeded in 100 mm dishes and transfected as described with UAS-GFPnls, UAS-GFP-UbxWT or UAS-GFP-UbxN51A, and pActin-Gal4. Cells were harvested in PBS after 48 hours of transfection and pellets were resuspended in lysis buffer supplemented with protease inhibitor cocktail (Sigma-Aldrich) and 1 mM DTT. For ChIP and interaction experiments, 0.25 or 0.5 μg/ml of Actinomycin D treatment (Sigma-Aldrich) was applied for 20 h. For Pol II C4 (PC4) expression experiments, 5 μg/ml α-amanitin (Sigma-Aldrich) treatment was applied for 25–30 h before RNA extraction or FRAP analysis. 10 μM Triptolide (Sigma-Aldrich) and 10 μM Flavopiridol (Sigma-Aldrich) treatment were applied for 10 and 20 min. For FRAP analysis, cells were seeded and transfected (50 ng of each plasmid) in 12-well plates and transferred with fresh supplemented media (supplemented with α-amanitin for PC4 experiment) in glass bottom dishes coated with Polylysine (Sigma), at least 2 h before image acquisition.

**RNA-Seq from Drosophila S2R+ cells**

Total RNAs were extracted from four independent replicates from *Drosophila S2R+* cells expressing GFPnls, myc-UbxWT or myc-UbxN51A (Gal4-UAS, actin promoter) using Qiagen RNA extraction kit (RNeasy). RNA quality and integrity was assessed using BioAnalyzer 2100TM (Agilent Technologies). Material handling and mRNA-Seq directional libraries were performed by the Deep-Sequencing facility in Heidelberg (Cell Networks) with TruSeq kit and poly(A) selection according to the manufacturer’s protocol (Illumina). Sequencing was performed with NextSeq500
High-Output with a read-length of 75 bp and single-end strands. Replicates were validated by FastQC report and three replicates for each sample were further selected according to Principal Component Analysis (PCA) for the study. The data were generated and uploaded on Gene Expression Omnibus according to the ENCODE guideline (see Data Availability section).

### Expression and differential splicing analysis of RNA-Seq

RNA-Seq analysis was performed by using genome 6 (dm6) according to the ENCODE practices. The quality of the RNA-Seq reads was measured via FastQC. Trimming was performed with java script Trimmomatic Version 0.36 (29), for the removal of the adapter (TruSeq-SE sequencing adapter), with the following command line:

```
java -jar /path/to/java_script/Trimmomatic-0.36/trimmomatic-0.36.jar SE -phred33 <file.txt> <outputfile.txt>
```

ILLUMINA CLIP: TruSeq-SE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36. The reads were aligned using STAR (genomic generated: STAR --runThreadN 14 --runMode genomeGenerate --genomeDir Drosophila_melanogaster.BDGP6.dna.toplevel.fa --genomeSAindexNbases 12; running STAR: STAR --runThreadN 14 --genomeDir /STAR/GenomeDir/ --readFilesIn /STAR/ <file> --outSAMtype BAM Unsorted (30).

Aligned reads were counted by HTSeq (htseq-count -f bam -r pos -m union -s no -t exon acceptedFlanks.bam Drosophila_melanogaster.BDGP6.86.gtf > count.txt) for further analysis in DESeq2 (31). Differential expression analysis was performed with DESeq2 under standard conditions (32).

For the analysis in JunctionSeq the aligned reads were counted and the genome flattened with QoRTs by using the available java scripts (counting: java -jar /path/ToJavaScript/Trimmomatic-0.36/trimmomatic-0.36.jar SE -phred33 <file.txt>

```
<outputfile.txt> ILLUMINA CLIP: TruSeq-SE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36.
```

The reads were aligned using STAR (genomic generated: STAR --runThreadN 14 --runMode genomeGenerate --genomeDir Drosophila_melanogaster.BDGP6.dna.toplevel.fa sgdGTFfile Drosophila_melanogaster.BDGP6.86.gtf --genomeSAindexNbases 12; running STAR: STAR --runThreadN 14 --genomeDir /STAR/GenomeDir/ --readFilesIn /STAR/<file> --outSAMtype BAM Unsorted (30).

Aligned reads were counted by HTSeq (htseq-count -f bam -r pos -m union -s no -t exon acceptedFlaks.bam Drosophila_melanogaster.BDGP6.86.gtf > count.txt) for further analysis in DESeq2 (31). Differential expression analysis was performed with DESeq2 under standard conditions (32).

### ChIP-Seq genome wide distribution

Analysis of Ubx DNA-binding profile in *Drosophila* S2 cells (35) and in the mesoderm (28) were performed as in Domsch et al. (28). The aligned read files (BAM) of the Ubx ChIP-Seq in *Drosophila* S2 cells (35) and in the differentiating mesoderm (28) were downloaded from GEO (accession GSE101556 and GSE121754 respectively). Peak calling was performed using MACS2 with standard narrow peak settings (macs2 callpeak -t ubx_aligned_reads.bam -c input_aligned_reads.bam -q 1.28e8 -f BAM --outdir output -n ubx -q 0.01). Statistics and annotation of the called peaks was performed using HOMER annotatePeaks.pl (annotatePeaks.pl ubx_peaks.narrowPeak dm6 -annStats ubx_stats.txt > annotated.mlUbx_chr_peaks.txt).

The ‘-annStats’ option in annotatePeaks.pl provides a detailed set of annotation statistics, namely the number of called peaks that fall within defined genomic features such as promoters (defined as −1 kb, +100 bp of TSS), gene bodies (exons and introns), intergenic and others (TTS, 3’ and 5’ UTR, ncRNA, miRNA, miscRNA). The generated datasets are available in Supplementary Table S2 and were used to overlap with the list of genes expressed (RPKM value ≥ 1), misexpressed or differentially spliced upon ectopic expression in *Drosophila* S2R+ cells or depletion of Ubx in the embryonic mesoderm.

### RNA extraction, retrotranscription (RT) and quantitative PCR

Total RNAs of three to four independent experiments performed in triplicate were extracted by Trizol (Ambion) according to the manufacturer’s protocol and quantified on NanoDrop. The samples were digested with DNase (Thermo Fisher Scientific) and 1 µg of RNA was converted to first strand cDNA using Reversaid kit (Thermo Fisher Scientific) and random hexamers in 20 µl final volume, according to the manufacturer’s protocol. Real-time PCR experiments have been performed according to MIQE guidelines (36) in qTower3 (Analytik Jena) and CFX96 Real time systems (BioRad). In detail, qPCR was performed in technical duplicate for each sample, in a 96-well plate using the Platinum™ SYBR™ Green (Invitrogen) according to the manufacturer’s protocol in a final volume of 10 µl. Primers were designed with Primer3 for amplions ranging between 70 and 150 bp, verified with nucleotide blast (NCBI), generated by Eurofins and tested by serial dilution of cDNA and melt curve analysis. qPCR cycles were: 95°C for 2 min, 40 cycles of: 95°C, 1 s; 60°C, 30 s, followed by temperature gradient. Data were quantified by the ΔΔ-Ct method and normalised to Actin 5C expression or internal region of constitutively expressed exons of the related gene as indicated. Sequences of the primers used in this study are provided in Supplementary Table S5.

### RNA-immunoprecipitation and quantitative PCR

Confluent *Drosophila* S2R+ cells expressing GFPNLs, GFP-Ubx<sup>Wt</sup> or GFP-Ubx<sup>N51A</sup> plated in 100 mm dishes were collected in cold PBS 48 h after transfection. After several PBS washes, cell pellets were resuspended in buffer A (10 mM HEPES pH 7.9, 10 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.34 M sucrose, 10% glycerol). Lysates were incubated with 0.1% Triton and centrifugated. Nuclear pellets were then resuspended with IP buffer (150 mM NaCl, 50 mM Tris pH 7.5, 1 mM EDTA, 1% Triton), incubated on ice with vigorous regular vortex and sonicated (8 × 30 s on/off, Picorruptor, Diagenode). All buffers were supplemented with protease inhibitor cocktail (Sigma), 1 mM DTT, 0.1 mM PMSF and RNasin (Promega). Input fractions were collected, both for protein and RT-qPCR control. 1.5–2 mg of nuclear lysates were diluted in IP buffer and incubated for 4 h with 40 µl of GFP-Trap beads (Chromotech). Beads were washed 5 times for 5 min with IP buffer. 10% were collected for protein analysis, the remaining beads were resuspended in Trizol (Ambion) and RNAs were extracted according to the
manufacturer’s protocol. Notably, RNAs were precipitated for 1 h at −80°C in isopropanol to increase efficiency. Retro-transcriptions were processed as described, on 2 μg of RNA by doubling the total volume (40 μl final). cDNAs were then diluted by two and 2 μl were used for quantitative PCR. Enrichment was calculated relative to input, which referred to the total RNA present within each sample, and GFP values.

Chromatin-immunoprecipitation coupled with quantitative PCR

Forty-eight hours post-transfection, confluent Drosophila S2R+ cells expressing GFP-UbxWT plated in 100 mm dishes were crosslinked with 1% formaldehyde and quenched them for 5 min in 0.125 M glycine. After several PBS washes, cell pellets were resuspended in lysis buffer (1% SDS, 50 mM Tris-HCl, pH 8, 10 mM EDTA). Sonication (10 cycles, 30 s on/off) was performed with Picoruptor (Diagenode) according to the Diagenode recommendation. Lysates were incubated with 5 μl of antibody against Ubx (Guinea-pig, home-made) or IgG (Santa Cruz) overnight at 4°C on rotation and for one additional hour with mixed Dynabead protein G and A (20:10 μl, Life Technologies). Beads were washed with TSE-150 (0.1% SDS, 1% Triton, 2 mM EDTA, 20 mM Tris, pH 8.1, 150 mM NaCl), TSE-500 (as TSE-150 with 500 mM NaCl), LiCl detergent (0.25 M LiCl, 1% NP40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris, pH 8.1) and Tris-EDTA (5:1 mM). Combined elution and decrosslinking were performed by adding RNAse for 30 min at 37°C, then 0.1% SDS with protease K for 1 h at 37°C followed by additional incubation with NaCl under 900 rpm shaking for 7 h at 65°C. DNA fragments were purified using Qiaquick miniElute (Qiagen) according to the manufacturer protocol and diluted to 1/10 for input and to the half for immunoprecipitated fractions. qPCRs were performed using 2 μl of DNA, and enrichment was calculated relative to input and IgG values.

Co-immunoprecipitation of whole cell lysate and embryos nuclear extract

For co-immunoprecipitation assays, Drosophila S2R+ cells expressing GFPnl5, GFP-UbxWT or GFP-UbxNS1A were harvested in PBS. Pellets were resuspended in NP40 buffer (20 mM Tris pH 7.5, 150 mM NaCl, 2 mM EDTA, 1% NP40) and incubated overnight with 40 μl of GFP-Trap beads (Chromotek). Beads were then washed 5 times with NP40 buffer and all samples were resuspended in Laemmli buffer for immunoblotting analysis. All buffers were supplemented with protease inhibitor cocktail (Sigma), 1 mM DTT and 0.1 mM PMSF. Input fractions represent 1–10% of the immunoprecipitated fraction.

SDS-PAGE and immunoblotting

For western blot analysis, proteins were resolved on 8–15% SDS-PAGE, blotted onto PVDF membrane (Biorad) and probed with specific antibodies after saturation. The antibodies (and their dilution) used in this study were: Ubx (home-made, 1/200e), Histone 3 (Abcam, 1791, 1/10 000e), GFP (Life Technologies, A11122, 1/3000e), Tubulin (Serotec, MCA77G, 1/200e), GST (Cell signalling, 2624, 1/5000e), Pol II total (Bethyl, A300-653A, 1/200e), Pol II S5Phos (Bethyl, A304-408A, 1/1000e), Pol II S2Phos (Bethyl, A300-654A, 1/1000e).

Protein purification and GST pull-down

His-tagged and GST-tagged proteins were cloned for this study or from our previous work (27) in pET or pGEX-6P plasmids respectively. His- and GST-tagged proteins were produced from BL-21 (RIPL) bacterial strain, purified on Ni-NTA agarose beads (Qiagen) or Gluthathione-Sepharose beads (GE-healthcare) respectively and quantified by Coomassie staining. His-tagged proteins were specifically eluted from the beads with imidazole. In vitro interaction assays were performed with equal amounts of GST or GST fusion proteins in affinity buffer (20 mM HEPES, 10 μM ZnCl2, 0.1% Triton, 2 mM EDTA) supplemented with NaCl, 1 mM DTT, 0.1 mM PMSF and protease inhibitor cocktail (Sigma). Proteins produced in vitro were subjected to interaction assays for 2 h at 4°C under mild rotation with Drosophila nuclear extracts or recombinant human CTD (Active Motif). Bound proteins were washed 4 times and resuspended in Laemmli buffer for western-blot analysis. Input fraction was loaded as indicated.

In vitro transcription with Cy3-UTP labelling

For in vitro transcription, selected fragments of alternatively spliced exons (Supplementary Table S4) were cloned into the pGEM®-T Easy Vector (Promega). To generate the DNA templates for transcription, plasmids were amplified in DH5α bacterial strain, purified and linearised using the ProbeQuant™ G-50 Micro Columns (GE healthcare) and eluted in 50 μl. For RNA 3′-Cy3 labelling, see supplementary methods.
Protein-RNA UV-crosslinking assay

To prepare the protein-RNA complexes for UV-crosslinking, 2 pmol of internally labelled RNA probes were mixed with approximately 0.5–1 µg of his-purified proteins. The binding reaction was performed in a pre-cooled 96-well plate in a volume of 30 µl containing 1× binding buffer (20 mM HEPES pH 7.9, 1.4 mM MgCl₂, 1 mM ZnSO₄, 40 mM KCl, 0.1 mM EDTA, 5% glycerol), 2 µg tRNA (Thermo Fisher Scientific), 3 µg BSA, 10 mM DTT and 0.1% NP40. After 20 min on ice, the samples were irradiated with UV light in a Stratalinker® UV Crosslinker Model 1800 (Stratagene) for 10 min on ice and subsequently transferred in Eppendorf Tubes. 1.5 µl of RNase A (10 mg/ml) (Thermo Fisher Scientific) were added and the samples were incubated for 20 min at 37°C. Cy3-labelled protein–RNA complexes were resolved on 10–12% SDS-PAGE for 1 h at 200 V and detected by fluorescence using INTAS Imager. Following the detection, the gels were stained with Coomassie overnight and imaged using a conventional image scanner (Epson).

Electrophoretic mobility shift assay (EMSA)

The 5′-Cy5 labelled complementary oligonucleotides (Eurofin) commercially produced were annealed before reaction. The sequences used for this study were the following: Ubx sites, Cy5-5′-TTCAAGGCTAATGTTCGAGCTGTCAG-3′. The binding reaction was performed for 20 min in a volume of 30 µl containing 1× binding buffer (20 mM HEPES pH 7.9, 1.4 mM MgCl₂, 1 mM ZnSO₄, 40 mM KCl, 0.1 mM EDTA, 5% glycerol), 0.2 µg Poly(dI-dC), 0.1 µg BSA, 10 mM DTT and 0.1% NP40. For each reaction his-purified proteins were used. Separation was carried out (50 min, 200 V) at 4°C on a 4–6% acrylamide gel in 0.5× Tris–borate–EDTA buffer to visualise complex formation by retardation. Cy5-labelled DNA-protein complexes were detected by fluorescence using INTAS Imager. For RNA EMSA, see Supplementary methods.

Fluorescence recovery after photobleaching acquisition and modelling

Timeseries were acquired on a NIKON A1R CSLM equipped with a 63×, NA 1.27, WI objective. Detection was done on a Galvano scanner at 64 × 64 px. The first 10 pre-bleach frames and 10 post-bleach frames were recorded at 33 frames per second (fps). Subsequent recovery was measured at 30 fps, followed by 0.2 fps for 80 frames, totalling about 4.5 min. Half nuclei were bleached for 250 ms at 100% laser power.

Acquired time series were analysed in ImageJ. Stacks were aligned using the Template Matching and Slice Alignment plugin. Total nuclear intensity (Inuc), half-nucleus bleached intensity (Ibl), and background intensity (Ibg) were used for analysis in R.

FRAP recovery was double normalised as follow:

- FRAP(t) = (I_nuc-I_bg)/(I_bl-I_bg)
- FRAP_norm (t) = (FRAP(t))/I_pre

The double normalised recovery curves were then fitted with the minpack.lm package to single and double exponential diffusion models:

- Single: F(t) = M_mob-M_mob-e^(-k·t)
- t-half: t(1/2) = -(ln(0.5))/k
- Immobile fraction: M_imm = 1-M_mob
- Double: F(t) = (M_fast + M_slow)-M_fast-e^(-k_on-t)-M_slow-e^(-k_off-t)
- t-half: calculated using the investr package
- Immobile fraction: M_imm = 1-M_fast-M_slow

Data analyses and visualisation

For RPKM, gene length referred to the exonic length of each gene while total genomic gene length was used for ChIP-Seq comparative analysis. RPKM calculations were performed with the following formula ‘RPKM = (10⁹ × gene read count)/(total read count × gene length)’. More details are provided in Supplementary methods section.

Gene ontology-term (GO) annotations and overrepresented GO-term analysis were performed with the Web-tool PANTHER using Fisher test and FDR correction.

Integrative Genomics Viewer (IGV Version 2.8.12) was used for Ubx ChIP-Seq data visualisation at logarithm scale using Bedgraph files. Data visualisation was achieved with GraphPad Prism 9, Microsoft Office PowerPoint, Excel and Adobe Illustrator.

Statistical analyses

Gel quantifications were performed with Fiji (is just image j).

Statistical analyses were performed using one-way ANOVA, ch² for distribution and t-test multiparametric using GraphPad Prism 9 software. Each experiment was performed for three to four independent biological replicates for the gels and immunoblots, duplicates for the RIP- and ChIP-qPCR and triplicates for RNA expression.

RESULTS

Ubx modulates mRNA splicing in vivo to coordinate tissue-specific functions

We previously uncovered an interplay between the Drosophila Hox TF Ubx and splicing factors for muscle development (27). Based on these results we hypothesised that Ubx coordinates cell fate decision by modulating splicing.

To test this assumption, we analysed in depth our transcriptome data performed in differentiating muscles (stages 14–17 of embryogenesis) upon mesoderm-specific depletion of Ubx (Figure 1A) (28, 37). The transcriptome profile exhibited a high number of genes differentially expressed upon Ubx knock-down (KD) as previously described (Supplementary Figure S1A, Supplementary Table S1A). These genes are referred to as transcriptionally regulated or misexpressed. We next asked whether mRNA splicing was
affected by Ubx degradation. To this end, we analysed the transcriptomes with JunctionSeq (33), a bioinformatic package that detects differential usage of exons and exon-junction usage, thereby providing information with regards to splice events. We identified 425 differential splicing events upon Ubx degradation compared to control (Figure 1B and C, Supplementary Table S1B). Among these events, 82% were differential exon usage (347/425) with 14% involving the first exon (49/347). The data identified a moderate difference of alternative events upon degradation of Ubx compared to control, within 59% of exons included (high, 205/347) and 41% excluded (low, 142/347). Similarly, we observed 40% (31/78) higher and 60% (47/78) lower usages of splice sites (Figure 1C). These events were related to a substantial number of genes differentially spliced upon Ubx degradation (Figure 1D). We overlapped the list of genes differentially spliced and misexpressed upon Ubx degradation (Figure 1D, 283). We overlapped the list of genes differentially spliced and misexpressed upon Ubx degradation (Figure 1D). We overlapped the list of genes differentially spliced and misexpressed upon Ubx degradation (Figure 1D). We overlapped the list of genes differentially spliced and misexpressed upon Ubx degradation (Figure 1D). We overlapped the list of genes differentially spliced and misexpressed upon Ubx degradation (Figure 1D). We overlapped the list of genes differentially spliced and misexpressed upon Ubx degradation (Figure 1D). We overlapped the list of genes differentially spliced and misexpressed upon Ubx degradation (Figure 1D). We overlapped the list of genes differentially spliced and misexpressed upon Ubx degradation (Figure 1D). We overlapped the list of genes differentially spliced and misexpressed upon Ubx degradation (Figure 1D). We overlapped the list of genes differentially spliced and misexpressed upon Ubx degradation (Figure 1D). We overlapped the list of genes differentially spliced and misexpressed upon Ubx degradation (Figure 1D).

Figure 1. Ubx regulates splicing in vivo in the embryonic mesoderm. (A) Schematic of tissue context for the differential transcriptome performed in embryonic mesoderm (28), upon tissue-specific degradation of endogenous GFP-Ubx with the DegradFP system. (B) MA plot from JunctionSeq shows the fold change of differential splicing events (higher/lower) in Ubx-degradFP (Degrad) or control experiment, plotted with the mean of the normalised coverage (FDR 0.1). (C) Summary of differential splicing events detected upon Ubx degradation (exons, splice sites), higher or lower upon Ubx knock-down (KD). (D) Venn diagram overlapping the misexpressed (2891) and differentially spliced (283) genes upon Ubx depletion in embryonic mesoderm. (E) Fold enrichment of Gene Ontology (GO) term of biological processes enriched for the list of genes exclusively differentially spliced (199/283, P-value < 0.05, fold enrichment). See also Supplementary Figures S1–S3, Supplementary Table S1A and B.

translation (Supplementary Figure S1C). This suggests that Ubx regulates mRNA expression and splicing to coordinate distinct cell functions. The molecular function (GO term) of the differentially spliced genes upon Ubx degradation were enriched for cytoskeletal proteins as well as Rho GTPase activity, which is essential for cell shape and motility (38) (Supplementary Figure S1D). Consistently, cell components of the differentially spliced genes were related to neuromuscular junction, myofibril and sarcomere module (Supplementary Figure S1D). The list of genes included Tropomyosin I, Paramyosin, Troponin T as well as Complexin, Ephexin and Neto for neuromuscular junction (Supplementary Figure S2A and B). This strongly suggests that Ubx regulates muscle development not only by stabilising the lineage-specific transcriptional program but also by coordinating muscle function and signal transduction through the regulation of alternative splicing (AS) and specific isoform production.

One of the various AS mechanisms relies on modulating the expression and/or splicing of mRNA regulatory proteins themselves (39). In this context, we examined the genes related to mRNA processing, which were misexpressed and/or differentially spliced upon Ubx degradation in the mesoderm. The GO term analysis revealed a low enrichment of events related to mRNA processing only detected among the genes repressed by Ubx (i.e. up-regulated upon Ubx degradation) (Supplementary Figure S1C). Ubx depletion was also associated with the differential splicing of a small set of mRNA processing factors, such as Squid (Sqd), a hnRNP for which isoform-specific functions have been described (40), RbFox1, involved in cardiac hypertrophy and heart failure (41) or the serine threonine kinase DOA, producing multiple isoforms for which non-redundant func-
tions have been suggested (42) (Supplementary Figure S3). Thus, Ubx could modulate splicing by both direct and indirect effect.

Overall, these data revealed that Ubx regulates tissue-specific functions through the regulation of mRNA expression and splicing. We had shown that Ubx regulates alternative fate gene expression at the transcriptional level (28). Our results now suggest a novel regulatory function of Ubx in muscle development by coordinating gene programs at the splicing level.

UbxCy distinctly coordinates transcription and splicing

The mesoderm is composed of various cell lineages (somatic, visceral and cardiac). Importantly, the muscle types are specified by defined transcription and splicing programs (43,44). Consequently, the heterogeneity of the mesoderm population might shade the impact of Ubx in splicing. In addition, Ubx regulates the expression of mRNA processing factors thereby challenging the identification of its direct function in splicing. To decouple the role of Ubx on transcription and splicing, we chose to investigate its molecular function in an accessible and homogenous cell context, which does not express any Hox proteins. Thus, we investigated the role of Ubx in splicing using the Hox-free Drosophila S2R+ cell system. Specifically, we performed RNA-Seq experiments upon ectopic expression of Ubx Wild-Type (WT) or GFP fused to a nuclear localisation sequence (nls) as control (Figure 2A). We transiently expressed the constructs under the control of the Gal4-UAS system (45) driven by the ubiquitous actin promoter. Three independent biological replicates were evaluated and Principal Component Analysis (PCA) validated the similarity of replicates (Supplementary Figure S4A). The differential expression profile revealed that ectopic expression of UbxWT induced a global change of the transcriptome compared to the control (GFP), with 932 up-regulated and 985 down-regulated genes (Supplementary Figure S4B and C, Supplementary Table S1C). The transcriptome performed in the differentiating muscles revealed that Ubx represses alternative fate genes at the transcriptional level to stabilise the cell lineage (28). In contrast, the Drosophila S2R+ cells represent a pool of somatic cells, which are not specified. In this context, the genes activated by UbxWT were largely related to various tissue identities and differentiation processes, such as heart development, motor neuron axon guidance, histoblast morphogenesis, stem cell differentiation. This mirrors the pivotal role of Ubx, and of the Hox TFs in general, in the development of numerous tissue types (Supplementary Figure S4D). It also demonstrates the inability of Ubx to control one cell-specific transcriptional program in absence of the proper set of TFs essential for lineage commitment (46). In contrast, UbxWT repressed genes related to general processes such as translation and ribosome biogenesis in this cellular context (Supplementary Figure S4E). Subsequent analysis using JunctionSeq revealed that UbxWT expression induced a significant change of the mRNA splicing profile in the cells, with 133 events differentially regulated compared to control (Figure 2B and C, Supplementary Table S1D). Similar to the mesoderm transcriptome (Figure 1), we observed 85.7% of differential exons usage (114/133) within 19% involving the first exon (Figure 2C, 22/114). The data indicated a moderate difference of events upon Ubx expression for the splice sites usage within 58% higher (11/19) and 42% lower (8/19) events compared to control. However, the difference was noticeable for the exon usage, as 68% of exon inclusion (78/114, higher) and 32% of exon exclusion (lower 36/114) were observed upon UbxWT expression (Figure 2C). This strongly suggests that Ubx largely promotes the retention of exon cassettes in Drosophila cells. These events related to 81 genes differentially spliced upon UbxWT expression (Figure 2D). We confirmed the role of Ubx in splicing by analysing the differential exon retention of several target genes upon UbxWT expression, namely Chascon (Chas, Figure 2F, I–J), the poly(A) binding protein encoding gene pAbp (Figure 2G, K–L), the small GTPase Rgl1 (Figure 2H, M–N), the RhoGEF Paratrophin-I-like (Pura, Supplementary Figure S5A, C–D), the histone H3.3B (Supplementary Figure S5B, E–F), the cAMP phosphodiesterase Dunce (Dnc, Supplementary Figure S5G–H), CG34417 (Supplementary Figure S5I–J) and the ribosomal protein Rps13 (Supplementary Figure S5K–L). This is exemplified with Chas, a gene involved in hair bristle development and muscle-tendon junction (47), which is regulated by Ubx both at the transcriptional and splicing level (Figure 2F, I–J). Chas contains an exon cassette retained upon UbxWT expression (Figure 2F, exon E5). We confirmed the result by assessing the expression of constitutive and alternative exons related to the housekeeping gene Actin5C. This showed a change at the expression level for all exons tested (E1, E3, E5, Figure 2I). In contrast, pAbp was regulated at the splicing level exclusively, and only exon cassette E1 was significantly modulated at the expression level (Figure 2K). We further examined the differential splicing of the exon cassette E5 of Chas compared to its constitutive exon E3 showing an increase of exon E5 retention (Figure 2J). In contrast, the constitutive exon E1 of Chas was not differentially spliced upon UbxWT expression.

We next overlapped the lists of genes misexpressed and differentially spliced upon UbxWT expression (Figure 2D), 54% of the genes (44/81) were exclusively differentially spliced without their expression being affected, indicating that Ubx regulates both transcription and splicing of mRNA. Interestingly, these genes related to cell function similar to the genes differentially spliced in the mesoderm, notably cell communication, signal transduction and cell migration (Figure 2E, Supplementary Figure S4F). As observed for the mesoderm transcriptome, these functions were distinct from the one related to the misexpressed genes (Supplementary Figure S4D–E). This reinforces our observation that Ubx coordinates transcription and splicing to control distinct cellular functions in defined cell and tissue contexts. Notably, its regulatory role in splicing is essential for cell-cell communication and cell shape. Consistently, we identified 16 genes differentially spliced both in the embryonic mesoderm and Drosophila S2R+ cells, highlighting common as well as cell-type specific target genes related to specialised cell behaviour (Supplementary Figure S4G).
Figure 2. Ubx regulates transcription and splicing in Drosophila S2R+ cells. (A) Schematic of cell context for the differential transcriptome performed in Hox free Drosophila S2R+ cells, upon ectopic expression of myc-Ubx^{WT} or GFP fused to a nuclear localisation sequence (nls) control, using the Gal4–UAS system driven by the actin promoter. RNA-Seq data were further analysed for three independent biological replicates. (B) MA plot from JunctionSeq showed the fold change of differential splicing events (higher/lower) upon ectopic expression of Ubx^{WT} or GFP control in Drosophila S2R+ cells, plotted with the mean of the normalised coverage (FDR = 0.1). (C) Summary of differential splicing events detected upon Ubx^{WT} expression (exons, splice sites) higher or lower compared to control. (D) Venn diagram overlapping the misexpressed (1917) and differentially spliced (81) genes upon Ubx^{WT} expression in Drosophila S2R+ cells. (E) Gene ontology (GO) term of biological processes enriched for the list of genes exclusively differentially spliced (44 genes) and D. melanogaster (Drosophila melanogaster) extended panel), and the gene level normalised read count (right Y-axis, narrow panel) of Ubx^{WT} expression (wt, blue line) compared to control (CTRL, red line). Significant differential splicing events are highlighted in purple. Isoforms including differentially spliced exon or junction usages (purple) upon Ubx^{WT} expression are displayed below each read count. (I–N) RT-qPCR experiments showing the differential expression (I, K, M) over Actin5C and (J, L, N) differential retention of exon cassettes over constitutive exons for Chas (I–J), pAbp (K–L) and Rgk1 (M–N) in Drosophila S2R+ cells expressing GFP control (white) or Ubx^{WT} (blue). (E + number) = exon related to JunctionSeq annotation. Differentially spliced exons are underlined (purple), n = 4 independent biological triplicates. Bars represent mean ± SEM. Statistical test by one-way ANOVA (*p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, ns = non-significant). See also Supplementary Figure S4-5, Supplementary Table S1C and D.
UbxB chromatin-binding events are enriched in the gene body of target genes

Ubx regulates its direct transcriptional targets by specific recognition and binding of cis-regulatory elements (28,48,49). We thus reasoned that Ubx could modulate mRNA splicing through direct binding of its target genes. To study this question, Ubx chromatin binding profiles performed in Drosophila S2 cells (35) (Figure 3A, Supplementary Table S2A) were compared to the transcriptome and splicing profiles upon Ubx WT expression. Both cell lines derived from the same origin (50) and their similarity was further confirmed by PCA (Supplementary Figure S6A) and Pearson coefficient showing high correlation between the S2 and S2R+ cells transcriptome datasets ($r = 0.93$, Supplementary Figure S6B). Additional overlap of the transcriptome profiles (RPKM $\geq 1$) validated the high similarity of the datasets (Supplementary Figure S6C). Subsequently, we overlapped our transcriptome with the genome-wide Ubx binding profile (35) showing that 71.5% (1372/1917 genes) of the misexpressed and 85% (69/81 genes) of the differentially spliced genes upon Ubx expression were bound by Ubx (Figure 3B). In contrast, only 67.6% (5100/7545 genes) of the expressed genes (RPKM $\geq 1$) had at least one Ubx binding event (Figure 3B, Supplementary Table S2A). This indicates that these genes are direct targets of Ubx. Previous analysis showed that Ubx binding is enriched at promoters (35). Consistently, we detected 41% of the binding events at promoters (promoter and TSS) (Figure 3A). Importantly, Ubx binding events were highly enriched within the gene body (48%), in particular in introns (73%), at Transcription Termination Sites (TTS 16%) and in exons (10%). In this context, we asked whether Ubx could be enriched in the gene body of its differentially spliced target genes. The genomic distribution analysis revealed that the genes differentially spliced upon Ubx WT expression had an enrichment of binding events in introns (29% of the total binding events) and exons (16% of the events) (Figure 3C, Supplementary Figure S6). In sharp contrast, the genes globally expressed were characterised by 18% of binding events detected in introns and 8% of binding events in exons (Figure 3C, chi$^2 P = 0.0017$). This difference was moderate compared to the genes misexpressed upon Ubx WT expression (chi$^2 P = 0.1$). Moreover, the low Pearson coefficient correlating the number of peaks detected in the gene body to the gene length indicated that the gene length does not account for the differential enrichment of binding events detected in the gene body (Supplementary Figure S6H-K). Remarkably, Ubx binding events were not enriched in alternative exon cassettes as only 6 over 24 exons having a binding event were also differentially spliced. Instead, we noticed a broad distribution of Ubx along the gene body of various target genes, as exemplified for Chas, Rgk1 and Pura (Figure 3D, Supplementary Figure S7A–C).

Similarly, the analysis of Ubx binding profile in the differentiating muscles (stages 14–17) revealed that 56% of the binding events were localised in the gene body, including 60% in introns, 16% at TTS and 23% in exons (Supplementary Figure S7D, Supplementary Table S2B). Notably, 50% of the differentially spliced genes upon Ubx degradation had an Ubx binding event (141/283 genes, Supplementary Figure S7E), whereas only 32% of the expressed genes were bound by Ubx (3424/10 635 genes, Supplementary Figure S7E, right panel, Supplementary Table S2B). Moreover, we observed an enrichment of Ubx events in the gene body of the genes differentially spliced upon Ubx depletion compared to the expressed genes (chi$^2 P = 0.019$, Supplementary Figure S7F). Although the distribution of events was unchanged in exons, more events were detected in introns and less in intergenic regions compared to the expressed genes. This indicated an enrichment of Ubx binding events within the gene body (Supplementary Figures S6L–O, S7F).

Overall, the analysis reveals that Ubx regulates transcription as well as splicing via direct chromatin binding of its target genes. Notably, Ubx binding is enriched along the gene body of its spliced target genes.

Ubx splicing activity requires the Homeodomain

The Hox DNA-binding domain, the Homeodomain (HD), plays a pivotal role in the DNA recognition, binding and regulatory function of Hox TFs (49,51). Thus, the HD could be an essential platform for Ubx transcriptional and splicing functions. To evaluate the role of Ubx binding ability in splicing, we performed transcriptome profiling upon transient expression of Ubx N51A, a mutant version of Ubx which is no longer able to bind DNA (27). N51A (asparagine to alanine) is a single mutation of the amino-acid N51 of the HD that is necessary for the interaction with the major groove of DNA (52). The Ubx N51A mutant has been generated and validated in our previous study, showing its inability to induce homeotic transformation in embryos or activate Ubx synthetic enhancer in Drosophila cells (27).

RNA-Seq experiments were performed in Drosophila S2R+ cells expressing Ubx N51A and PCA validated the high similarity of the three biological replicates (Supplementary Figure S8A). The differential expression profile revealed that ectopic expression of Ubx N51A induced a global change of gene expression compared to control, with 436 up-regulated and 631 down-regulated genes (Supplementary Figure S8B, Figure 3E–F, Supplementary Table S1E). Intriguingly, 22.8% (213/932) of the up-regulated and 50.2% (495/985) of the down-regulated genes upon Ubx WT expression were equally regulated by Ubx N51A mutant (Figure 3E, F). This could be due to an indirect effect, residual chromatin loading (53) via protein-protein interactions or a so far uncovered repressive ability of the HD (54).

Analysis using JunctionSeq revealed that Ubx N51A expression induced rare differential splicing events with only 24 events detected (Figure 3G, H, Supplementary Figure S8C) in 20 genes (Figure 3I, Supplementary Table S1F). This demonstrates that Ubx N51A has a minor effect on splicing. Comparison of differential splicing profiles revealed that 89% (119/133) of the splicing events and 85% (69/81) of differentially spliced genes were exclusively regulated by Ubx WT (Figure 3H, I). This difference did not account for Ubx expression level as both Ubx WT and Ubx N51A constructs were expressed at comparable level (Supplementary Figure S8D). We confirmed this finding by analysing the differential exon retention of selected target genes upon Ubx N51A expression (Figure 3J–L, Supplementary Figure S9). As expected, Ubx N51A did not promote the differen-
Figure 3. Ubx regulates splicing through its Homeodomain. (A) Genomic distribution (plotted in percentage %) of the regions bound by Ubx, namely promoter and transcription start site (TSS), intergenic, gene body, using the ChIP-Seq data of Ubx from *Drosophila* S2 cells generated by Zouaz *et al.* (35). The distribution in the gene body is further detailed for introns, transcription termination sites (TTS) and exons (including 5′ and 3′ UTRs). See also Supplementary Table S2A for the overall dataset used for the percentage of distribution. (B) Venn diagram representing the overlap of genes bound, misexpressed and differentially spliced upon UbxWT expression (left panel) as well as bound and expressed (RPKM ≥ 1) (right panel) in *Drosophila* cells. (C) Graphical view of the peak to gene distribution (plotted in percentage %) for the overall binding events (all), the binding events for the genes globally expressed, misexpressed and differentially spliced upon UbxWT expression. Distribution was done according to intergenic, promoter + TSS, exon, intron and TTS regions. Chi² tests were performed to estimate the statistical differences between each distribution profile: ‘all versus misexp.’: \( P = 0.596 \), ‘all versus diff.spliced’: \( P = 0.0017 \), ‘misexp. versus diff.spliced’: \( P = 0.12 \), ‘exp. versus misexp.’: \( P = 0.53 \), ‘exp. versus diff.spliced’: \( P = 0.0016 \). (D) Visualisation of Ubx binding events in the gene body of *Chas* by logarithmic scale. Transcription directionality is indicated with a horizontal blue arrow. Differentially spliced exons are indicated by a vertical blue arrow. (E, F) Venn diagrams from the RNA-Seq data performed in *Drosophila* S2R+ cells, representing the overlap of genes up-regulated (E) and down-regulated (F) upon UbxWT and UbxN51A expression compared to control GFP (three independent biological replicates for each sample). (G) MA plot from JunctionSeq showed the fold change of differential splicing events (higher/lower) upon ectopic expression of UbxN51A or GFP control in *Drosophila* S2R+ cells, plotted with the mean of the normalised coverage (FDR = 0.1). (H, I) Venn diagram representing the overlap of the differentially spliced events (H) and differentially spliced genes (I) upon UbxWT and UbxN51A expression in *Drosophila* S2R+ cells. (J–L) Relative RNA expression (RT-qPCR) revealed the differential retention of exon cassettes over constitutive exons for *Chas* (J), *pA4p* (K) and *Rgk1* (L) in *Drosophila* S2R+ cells expressing GFP control (white), UbxWT (blue) or UbxN51A (purple), but not for constitutive exons. (E + number) \( = \) exon relates to JunctionSeq annotation, differentially spliced exons are underlined (purple). \( n = 4 \) independent biological triplicates. Bars represent mean ± SEM. Statistical test by one-way ANOVA (** ** \( P < 0.0001 \), ns = non-significant). See also Supplementary Figure S6–S10, Supplementary Table S1E–F, S2, S3.
tial splicing of Chas, Rgk1, pAbp, Pura and H3.3B. Only Dnc, one of the common spliced target genes of Ubx<sup>WT</sup> and Ubx<sup>N51A</sup> exhibited a differential retention of exon E11 upon expression of each Ubx version (Supplementary Figure S9H, I). Importantly, Ubx splicing function in Drosophila cells is not a bias of ectopic Ubx expression as differential splicing was observed at level comparable to physiological expression of Ubx in the mesoderm (Supplementary Figure S10). Interestingly, some of the genes repressed by both Ubx<sup>WT</sup> and Ubx<sup>N51A</sup> are splicing factors (Rm62, SF3B3). It indicates once more that while Ubx regulates the expression of splicing factors, its activity on alternative splicing is most likely mediated by direct effects on mRNA targets. Taken together, the data show that Ubx splicing activity requires a functional HD.

**Ubx associates with RNA in vivo and is enriched on target alternative exon cassettes**

Our data indicated that N51 of the HD is necessary for Ubx splicing activity, yet we did not observe an enrichment of Ubx DNA-binding events on specific exon cassette (Figure 3D, Supplementary Figure S7A–C). Therefore, we wondered if the splicing activity of Ubx could be mediated by a so far uncovered RNA-binding ability thereby providing specificity at the exon level.

To this end, we performed RNA-immunoprecipitation experiments of GFP fused proteins (RIP-RT-qPCR) using Drosophila nuclear S2R+ cells extracts. We reasoned that nuclear extracts exhibit the RNA-binding function linked to transcription and mRNA processing, while interactions occurring in the cytoplasm are most likely related to mRNA transport and translation. We assessed the enrichment of constitutive and alternative exons in GFP, GFP-Ubx<sup>WT</sup> and GFP-Ubx<sup>N51A</sup> fractions of the Ubx spliced genes Chas, pAbp, Rgk1 (Figure 4A–C) and Pura, H3.3B (Supplementary Figure S11A, B). The enrichment of Actin5C mRNA was measured as a negative control (Figure 4A). For all target genes, we observed an enrichment of exon cassettes (Chas E5, pAbp E1, Rgk1 E13-E19, Pura E11, H3.3B E4) in the Ubx<sup>WT</sup> fraction compared to control GFP. This result demonstrates a novel ability of Ubx to associate with RNA in vivo. Interestingly, the constitutive exons of Chas, pAbp, Rgk1 and Pura were weakly or to a lesser extent enriched in the Ubx<sup>WT</sup> fraction compared to exon cassettes (Figure 4A–C, Supplementary Figure S11A). This indicated a binding specificity of Ubx toward alternatively spliced exons. Only H3.3B presented a similar enrichment for both constitutive E2 and cassette E4 exons in Ubx<sup>WT</sup> fraction (Supplementary Figure S11B). The smaller mRNA size of H3.3B compared to the other mRNA targets (mRNA H3.3B = 1.3 kb while other mRNAs > 2.4 kb) could account for this observation. Importantly, we observed a significant decrease of RNA pull-down in the Ubx<sup>N51A</sup> fraction compared to Ubx<sup>WT</sup> for all RNA targets studied (Figure 4A–C, Supplementary Figure S11A, B). As control, we analysed the pull-down efficiency by immunoblotting, showing a similar enrichment of GFP-fused proteins (Supplementary Figure S11C).

In sum, the results highlight that Ubx associates with RNA and that the N51A mutation impacts on its RNA-binding ability in vivo. More importantly, the results show that Ubx is specifically enriched in mRNA exonic regions regulated at the splicing regulatory level by Ubx, in contrast to its DNA-binding profiles which spread over the gene body.

**Ubx binds RNA directly in vitro and HD-N51 is a non-essential amino acid**

We subsequently asked whether Ubx could directly bind RNA. To this end, we performed in vitro protein-RNA interaction assays by Ultra Violet (UV)-crosslinking using purified his-tagged proteins and fluorescent labelled RNA probes. The probes corresponded to the exact or nearest exonic sequences (if Uracil content was too low) identified in cells, and were labelled using Cy3-UTP nucleotides (Supplementary Table S4). After crosslinking of protein-RNA complexes, unprotected RNAs were visualised on denaturing gels for fluorescent RNA detection and protein content by Coomassie staining (Figure 4D–K, Supplementary Figures S11 and S12). Remarkably, the assays revealed a direct binding of purified Ubx protein on the RNA probes for the exon cassettes of Chas (E5), pAbp (E1), Rgk1 (E19) and Pura (E11) compared to GFP control (Figure 4D, F, Supplementary Figure S11D, F). In contrast, Ubx binding was weaker for the probes of constitutive exon for Chas (E3), Rgk1 (E1), Pura (E19) (Figure 4E, Supplementary Figure S11E, G) and not detected for pAbp (E6) (Figure 4G). As negative control we found that Ubx does not bind a control RNA sequence, both by crosslinking assay and by native RNA Electrophoretic Mobility Shift Assay (EMSA, Supplementary Figure S11Q–S). Although we cannot account for protein-RNA interaction in UTP-free exonic sequences, this strongly suggests that Ubx exhibits binding specificity toward RNA sequences.

Concurrently, we analysed the RNA-binding ability of Ubx<sup>N51A</sup> mutant. Surprisingly, Ubx<sup>N51A</sup> exhibited similar RNA-binding ability as Ubx<sup>WT</sup>, which was supported by the quantification of protein-RNA interactions (Figure 4D–K, Supplementary Figure S11D–K, M–P). In contrast to the Ubx–RNA binding profile in Drosophila cells, the in vitro interaction assay revealed that the N51 amino-acid is non-essential for Ubx RNA-binding in sharp contrast to its DNA-binding ability (Supplementary Figure S11L). To evaluate if the HD domain mediates the interaction, we performed UV-crosslinking assay with purified HD<sup>WT</sup> and HD<sup>N51A</sup> both bind RNA (Figure 4H–K, Supplementary Figure S11H–K). Further quantifications revealed that the Ubx HD did not recapitulate the full binding of the full-length protein for Chas exons E5 and E3, pAbp exon E1 (Figure 4L, N, Supplementary Figure S12A–C), Rgk1 exons E19, E1 and Pura exons E11, E19 (Supplementary Figure S12E–L). Unexpectedly, the HD alone provided a weak binding affinity toward the RNA probes of pAbp constitutive exon E6 which was not detected with the full-length Ubx protein (Figure 4O, Supplementary Figure S12D). As control, we verified that the
Figure 4. Ubx binds RNA in vivo and in vitro, partly via its Homeodomain. (A–C) RNA-immunoprecipitation (RIP-RT-qPCR) experiments of Drosophila S2R+ cells expressing GFP control (white), GFP-UbxWT (blue) and GFP-UbxN51A (purple) showing an enrichment of targeted exonic regions of Chas (A), pAbp (B) and Rgk1 (C). Values are RNA relative enrichment over GFP calculated as percentage of input. The input represents the total RNA detected in each sample and thus, normalizes the enrichment to the total RNA expression level. (E + number) = exon related to JunctionSeq annotation, differentially spliced exons are underlined (purple). n = 3 independent biological duplicates. Bars represent mean ± SEM. The results exhibited a specific enrichment of differentially spliced exonic sequences in UbxWT fraction compared to GFP and UbxN51A. (D–K) Fluorescent protein–RNA interaction assay followed by UV-crosslinking and RNase digestion, performed in vitro with purified proteins namely his-MBP-GFP as control, his-Ubx (WT and N51A) full-length (FL) (D–G) and the HD alone his-MBP-UbxHD (WT and N51A) (H–K). Interactions were detected on denaturing gels by Cy3-UTP signal (upper panel), and gels were stained by Coomassie to reveal the protein content (lower panel). Each probe is indicated relative to the genes and exons. Molecular marker is indicated showing the size of each protein. MBP fused proteins are named his-MBP-X, his-fused proteins are named his-X. (L–O) Graphical view showing the quantification of relative RNA-binding of Ubx HD compared to full-length (FL) MBP fused proteins for Chas exon cassette E5 (L), constitutive E3 (M), pAbp exon cassette E1 (N), constitutive E6 (O) normalised to Coomassie staining. n = 3 independent biological replicates. Statistical test by one-way ANOVA (*P < 0.05, **P < 0.01, ***P < 0.001, ns = non-significant). See also Supplementary Figures S11 and S12, Supplementary Table S4.
MBP tag does not affect Ubx-RNA binding affinity (Supplementary Figure S12M–T). This indicates that while the HD mediates RNA-binding, it is not sufficient to recapitulate the RNA-binding profile of the full-length Ubx protein.

In sum, these results reveal a novel ability of Ubx to directly bind RNA in vitro. Importantly, the N51 amino-acid of the HD, although essential for Ubx interaction with DNA, is not essential for its direct RNA-binding. This strongly suggests that Ubx recognises and binds RNA and DNA via its HD yet, using a different protein-RNA interaction mode.

**Ubx interacts with active RNA polymerase II and requires its functional HD**

Our results indicated that the N51 amino-acid of the HD is necessary for Ubx RNA-binding and splicing function in vivo, but is not essential for its RNA interaction in vitro. Thus, Ubx loading onto the chromatin seems to mediate its RNA interaction in vivo and to regulate splicing. In this context, we explored in depth the molecular mechanism by which Ubx controls splicing.

Splicing happens mainly co-transcriptionally and depends on the RNA polymerase II (Pol II) activity (4). Moreover, Ubx interacts with several components of the basal transcriptional machinery (55,56) and can modulate transcription events, such as Pol II pausing (35). Based on these evidences, we reasoned that Ubx and Pol II could collaborate for regulating splicing during transcription elongation. Co-immunoprecipitation experiments were performed on GFP-fused proteins in *Drosophila* S2R+ cells, revealing that UbxWT interacts with both paused (initiation, S5Phos) and active (elongation, S2Phos) Pol II (Figure 5A, B). In contrast, UbxN51A mutant interacted similarly with the paused Pol II (S5Phos), while exhibiting a significantly weaker interaction with active Pol II (S2Phos) (Figure 5C). Notably, co-immunoprecipitation performed on nuclear embryo extracts revealed that Ubx interacts equally well with the two phosphorylated forms of Pol II in embryos (Supplementary Figure S13).

We subsequently asked whether Ubx interaction with Pol II required a specific domain. To this end, GST pull-down experiments were performed using Ubx full-length (FL), truncated N-terminal, C-terminal or HD purified proteins with *Drosophila* S2R+ cell extracts (Figure 5D–F). The assay revealed that Ubx full-length protein efficiently pulled down Pol II from nuclear extract, while each fragment pulled down Pol II to a lesser extent (50–80% of the full-length interaction). Interestingly, we noticed a stronger interaction of Pol II with the N-terminal domain of Ubx compared to the HD-containing C-terminal domain (80/50% Figure 5F). We subsequently examined if Ubx could directly interact with Pol II, a so far unknown molecular feature of Hox TFs. We performed GST pull-down assays with Ubx derivatives and the purified human carboxy terminal domain (CTD) of Pol II (Figure 5G, H). We observed a direct interaction between the CTD of Pol II and Ubx full-length protein as well as the N-terminal, C-terminal and HD domain. The pull-down was ten times stronger with Ubx full-length than with the Ubx truncated fragments (Figure 5H). We noticed a stronger interaction with the N-terminal domain compared to the C-terminal, however, this was not significant after quantification (n = 4).

In sum, the data show that Ubx interacts with Pol II in vitro and in vivo and revealed that the integrity of the HD is essential for the interplay between Ubx and active Pol II. This strongly suggests that Ubx binding to the chromatin is essential to mediate Ubx/Pol II interaction during active transcription.

**Ubx regulates splicing via an elongation-mediated process**

We next asked whether the Pol II activity can impact on Ubx splicing function. To do so, we first assessed the interaction between Ubx and Pol II upon treatment with the transcription inhibitor Actinomycin D, a DNA intercalator that accumulates hyperphosphorylated Pol II (Figure 6A–D (57)). Actinomycin D treatment reduced the interactions between UbxWT or UbxN51A with paused Pol II (S5Phos, 20%), and even more between UbxWT and active Pol II (S2Phos) for which the interaction was reduced by 70% (Figure 6D). As a control, the input fraction confirmed that this effect was not due to a decrease of Pol II phosphorylation (Figure 6B).

This indicates that Ubx/Pol II interaction depends on active transcription as well as on the integrity of the HD at transcription. To test how active transcription impacts on Ubx splicing activity, we analysed the effect of transcriptional drugs on Ubx splicing activity. The cells were treated with Flavopiridol (FP), an inhibitor of the Pol II kinase CDK9 which blocks the release of Pol II, thereby impairing elongation (57,58). In parallel, cells were treated with Triptolide (TP), an inhibitor of TFIH that prevents the assembly of the Pol II pre-initiation complex, while the assembled and engaged Pol II complex can still achieve transcription cycle (57,58) (Figure 6A). Upon treatment with Flavopiridol, the differential splicing induced by UbxWT was significantly reduced for most of its target exon cassettes (Figure 6E–H, Supplementary Figure S14A, D). Flavopiridol specifically impairs the elongation process, indicating that Ubx splicing activity occurs in tandem with elongation. In contrast, the transcription inhibitor Triptolide had no effect on Ubx splicing activity on the selected target genes (Figure 6E–H, Supplementary Figure S14A, D). Of note, the differential splicing of the *H3.3B* exon E4 was reduced, most likely due to its small gene size (Supplementary Figure S14A). This indicates that Ubx splicing activity depends on transcriptional elongation but is decoupled from initiation process.

We next reasoned that the rate of Pol II should impact on Ubx spliced target genes if Ubx regulates splicing co-transcriptionally. To test this hypothesis, we used a mutant of the biggest subunit of Pol II (rpII215), termed Pol II C4 (PC4) (11,59). Pol II C4 is associated with a slower elongation rate than the Pol II WT thereby impacting on splicing in a gene-specific manner (13). This mutant is resistant to α-amanitin, a drug inducing degradation of the Pol II WT form (57). Upon expression of Ubx in combination with Pol II C4 (PC4) and α-amanitin treatment in *Drosophila* S2R+ cells, we observed a gene-specific effect of Pol II C4 on Ubx spliced targets (Figure 6I–L, Supplementary Figure S14B, E). *Chas* exon E5, *p.Abp* exon E1 and *Rgk1* exon E19 retentions were reduced while *Rgk1* exon E13, *Pura* exon E11 splicing was...
Figure 5. Ubx interacts with active RNA Polymerase II via its functional Homeodomain. (A) Co-immunoprecipitation of endogenous Pol II total, paused (SSPhos) and active (elongating, S2Phos) forms with GFP fusion proteins (GFPnls, GFP-UbxWT, GFP-UbxN51A), ectopically expressed in Drosophila S2R+ cells. Western blots were probed with the indicated antibodies. The input is shown as a control of expression levels (lanes 1–3), Histone 3 (H3) is used as a loading control. Low and high exposure (exp.) are presented. (B) Schematic of paused Pol II (SSPhos) in the initiation complex, loaded onto the promoter and elongating Pol II (S2Phos) actively transcribing the gene. (C) Quantification of relative enrichment of Pol II, Pol II SSPhos and Pol II S2Phos relative to UbxWT and UbxN51A pulldown proteins, showing a specific enrichment of active Pol II S2Phos for UbxWT compared to UbxN51A. n = 3 independent biological replicates. Statistical test by one-way ANOVA (**P < 0.01, ns = non-significant). (D) Schematic of GST-fused Ubx derivatives. The HD is underlined in gray. The amino acids are indicated for each truncated construct. FL = full-length, Nter = N-terminal, Cter = C-terminal, HD = Homeodomain. (E, H) Pull down assay using the indicated GST-fused Ubx derivatives and Drosophila S2R+ cells nuclear extracts (E) or in vitro purified human Carboxy Terminal Domain (CTD) of Pol II (G). Input is loaded as indicated. (F, H) Quantification of interactions relative to GST-UbxFL (full-length) signal is indicated in (F) for Pol II from nuclear extract and in (H) for purified CTD. n = 3 independent biological replicates. Bars represent mean ± SEM. See also Supplementary Figure S13.
Figure 6. Ubx mediates co-transcriptional splicing via Pol II cooperation. (A) Schematic of the inhibitory effect of transcription drugs, Flavopiridol (FP), Triptolide (TP) and Actinomycin D (Act.D) on Pol II. (B) Co-immunoprecipitation of Pol II total, paused (SSPhos) and active (elongating S2Phos) with GFP fusion proteins (GFPnls, GFP-UbxWT, GFP-UbxN51A), expressed in *Drosophila* S2R+ cells treated for 20 h with Actinomycin D (Act.D) as indicated or DMSO as control (−). Western blots were probed with the indicated antibodies. The input is shown as a control of expression levels (lanes 1–9). Histone 3 (H3) is used as a loading control. (C–D) Quantification of relative enrichment of Pol II SSPhos (C) and S2Phos (D) relative to Ubx WT and Ubx N51A pull down showing once more a specific enrichment of active Pol II S2Phos for Ubx WT compared to Ubx N51A, and a significant decrease of Pol II SSPhos (30%) and S2Phos enrichment (70%) upon Actinomycin D treatment. n = 4 independent biological replicates. Statistical test by one-way ANOVA (*P* < 0.05, **P* < 0.01, ns = non-significant). (E–L) RT-qPCR experiments showing the differential retention of exon cassettes over constitutive exons for Chas (E, I), pAbp (F, J) and Rgk1 (G–H, K–L) in *Drosophila* S2R+ cells expressing GFP control (white), Ubx WT (blue). (E–H) 10 and 20 min treatments with Flavopiridol (FP, elongation repressed) or Triptolide (TP, elongation proceeds until termination) showed that Ubx activity on splicing depends on active transcription; (I–L) ectopic expression of the slow Pol II C4 (PC4) mutant characterised by a slow transcription rate and α-amanitin resistance, combined with α-amanitin treatment (α-am) showed that Ubx effect on splicing depends on the Pol II rate. (M–O) Chromatin-immunoprecipitation (ChIP) experiments of Ubx coupled with Actinomycin D treatment are presented as percent of enrichment relative to input and IgG control (horizontal bar set to 1). Binding of Ubx on the proximal and distal exons to the Transcription Start Site (TSS) of Chas (M), pAbp (N) and Rgk1 (O) are displayed relative to a schematic of the genes architecture. Differentially spliced exons are highlighted in purple. The exon number follows the JunctionSeq annotation file. The transcription directionality is represented by the black arrow. The alternative TSS and TTS are represented by the black brackets. n = 3 independent biological in triplicates (RNA) or duplicates (ChIP). Bars represent mean ± SEM. Statistical test by one-way ANOVA (*P* < 0.05, **P* < 0.01, ***P* < 0.001, ns = non-significant). See also Supplementary Figure S14.
not affected and H3.3B exon E4 retention was significantly increased (Figure 6I–L, Supplementary Figure S14B, E). Notably, the effect was observed for the majority of the target genes already without α-amanitin treatment. This could be due to competition between the Pol IIWT and mutant forms as previously observed (59–61). These results indicate that Pol II elongation rate impacts on Ubx splicing activity in an exon-specific manner.

**Ubx dynamically binds chromatin during transcription**

Our data indicate that Ubx binds and regulates the differential splicing of its target genes. Moreover, Ubx regulates splicing co-transcriptionally, in cooperation with Pol II and its interaction with active Pol II depends on its functional HD. Altogether, this strongly supports the hypothesis that Ubx regulates splicing via a dynamic interplay with Pol II during active transcription, thereby binding or travelling along the gene body to regulate splicing. We sought to test this model by examining how active transcription affects Ubx binding dynamics within gene bodies. To this end, we performed Chromatin-immunoprecipitation (ChIP) experiments of Ubx after Actinomycin D treatment in *Drosophila* S2R+ cells (Figure 6M–O, Supplementary Figure S14F). We observed a significant enrichment of Ubx binding in exons proximal and distal to the Transcription Start Site (TSS—first exon) of *Chas, pAbp, Rgk1* and *Pura*. Interestingly, Actinomycin D treatment significantly reduced the binding of Ubx in the distal exons of *Chas* (E3, E1), *pAbp* (E5, E6), *Rgk1* (E13, E11, E2) and *Pura* (E5, E11, E14, E19), while proximal exons to the transcription start sites were still bound by Ubx similarly to control condition. This result, we propose that i), the immobile fraction reflects Ubx stable binding on enhancer and promoter regions, ii), the slow mobile fraction refers to Ubx strong interactions or scanning of the chromatin, iii), the fast mobile fraction is a combination of diffuse, and transiently bound molecules involved in the turn-over of transcription (i.e. initiation-termination). In this context, we propose that the slow polymerase reduces the half-time recovery of Ubx on the chromatin and the immobile fraction as Ubx molecules are still bouncing or travelling along the gene body due to the slower transcription cycle (Figure 7). This model is further supported by co-immunoprecipitation experiments showing that UbxWT (but not UbxNS1A) interacted stronger with the active Pol IIWT (PC4) when Pol IIWT is degraded upon α-amanitin treatment (Figure 7D–F).

Altogether, the data indicate that Ubx dynamically binds to the chromatin during active transcription. This suggests that Ubx regulates co-transcriptional splicing via a dynamic interplay with Pol II along the gene body, which is coupled with a specific recognition and binding of differentially spliced exonic RNA (Figure 7G–I).

**DISCUSSION**

Our studies uncovered a novel molecular function of the Hox TF Ubx in mRNA splicing. Ubx coordinates mRNA expression and splicing via common and distinct modes of action, thereby promoting cell-type specific function in defined cell and tissue context. This expands the repertoire of Hox molecular function and its impact on cell fate. Our previous work revealed that Ubx interacts with a distinct set of regulatory cofactors in different embryonic tissues (27). The results further exhibit that Ubx interacts with several players of gene expression including mRNA splicing factors to coordinate tissue development (27). Altogether, this provides the molecular basis to explore the role of Ubx in
Figure 7. Models for Ubx regulation of co-transcriptional splicing and protein dynamics. (A) Normalised curves of fluorescence recovery (t) after photobleaching (FRAP) related to time (s, second) from data acquired using Drosophila S2R+ cells expressing GFP-UbxWT or GFP-UbxNS1A co-expressed or not with slow Pol IIIC4 (PC4) coupled with /H9251-amanitin treatment. First post-bleach acquisition data point is set to t(0) = 0. Modelling and fitting are described in Materials and Methods. (B) Distribution of Ubx populations is plotted for the different samples: immobile represents the fraction stably loaded onto the chromatin, slow mobile, the intermediate interactions and fast mobile, the transient interaction as well as diffusion molecules of Ubx. Mean ± SEM are shown. Statistical significance evaluated by t-test (*P < 0.05). (C) Value of the calculated half-time (1/2time) recovery and statistical differences compared to UbxWT condition are displayed. It shows that Ubx half time recovery is significantly faster for UbxNS1A than UbxWT. The slow Pol IIIC4 (PC4) reduces Ubx immobile population and UbxWT half time recovery, mirroring a potential decrease of stable binding of Ubx molecules on chromatin. n = 17 for GFP-UbxNS1A, n = 24 for GFP-UbxNS1A + PC4, n = 21 for GFP-UbxWT and n = 24 for GFP-UbxWT + PC4 nuclei included in the analysis. (D) Co-immunoprecipitation of total, paused (S5Phos) and active (elongating S2Phos) Pol II with GFP fusion proteins (GFPnls, GFP-UbxWT, GFP-UbxNS1A), co-expressed (or not) with slow Pol IIIC4 (PC4), coupled (or not) with /H9251-amanitin treatment (PC4/H9251). Western blots were probed with the indicated antibodies. The input is shown as a control of expression levels (lanes 1–9), histone 3 (H3) is used as a loading control. (E, F) Quantification of relative enrichment of Pol II S5Phos (E) and S2Phos (F) relative to UbxWT and UbxNS1A pull down showing a specific enrichment of active Pol II S2Phos for UbxWT compared to UbxNS1A, and a significant increase of Pol II S2Phos enrichment upon /H9251-amanitin treatment (PC4/H9251) compared to PC4 alone. n = 4 independent biological replicates. Statistical test by one-way ANOVA (*P < 0.05, ns = non-significant). (G–I) Schematic of the proposed model for co-transcriptional splicing mediated by Ubx. (G) Ubx regulates co-transcriptional splicing via a dynamic interplay with Pol II, the splicing machinery and a dual requirement of DNA-RNA interface. Two non-exclusive models for Ubx molecular mode of action are proposed: (H) Ubx is dynamically loaded along the gene body relying on its DNA-binding abilities. Like a bouncing/scanning behaviour, Ubx recruits the spliceosome de novo on spliced exons which are recognised and bound via Ubx RNA-binding abilities. (I) Ubx, paused Pol II and splicing factors are loaded together onto the promoter. Upon transcription activation, Ubx travels with Pol II (S2Phos) and the splicing factors thereby scanning both the chromatin and RNA through different HD interface to allow the specific recognition and regulation of targeted exons. See also Supplementary Figure S15.
mRNA splicing in vivo and a new perspective of Hox function for development and tissue homeostasis.

Our results indicate that the N51 amino-acid of the HD is essential for Ubx splicing activity while still partially mediating transcriptional function by repressing gene expression. First, this could be due to residual chromatin binding via protein interactions. ChiP-Seq experiments performed in *Drosophila* Kc167 cells with a combinatorial mutated version of Ubx HD showed that the Ubx mutant retained few binding to the chromatin (53). Nonetheless, FRAP experiments indicated that the punctual mutation N51A is essential for the stable binding of Ubx to the chromatin (Fig. 7). Second, the effect could be due to a loss of interaction with splicing factors. However, our previous study revealed that Ubx<sup>N51A</sup> protein is still able to interact with splicing factors in cells and in embryos while the interaction with TFs and chromatin binding proteins is specifically reduced or impaired (27). Third, it has been shown that a version of Ubx with a point mutation in the HD (Q50K) retained 65% of its repressive activity on the DME-reporter in embryos, mediating Distal-less abdominal repression (54).

Thus, the transcriptional repression observed upon expression of Ubx<sup>N51A</sup> might not be a bias of the ectopic system but an uncovered molecular function of Ubx. This opens a new perspective to study uncharacterised activity of the HD and its versatility for transcriptional repression. Importantly, the N51A mutation impairs Ubx DNA-binding ability, but is not essential for RNA-binding of Ubx in vitro. The effect mediated by Ubx<sup>N51A</sup> on Ubx target genes expression could be due to its ability to bind mature mRNA in the cytoplasm. In this context, Ubx could have additional RNA-binding specific functions, such as regulating RNA transport, decay or translation as suggested for other Hox proteins (64,65). We previously showed that ectopic expression of Ubx<sup>N51A</sup> is not able to induce homeotic transformation in embryos (27). However, Ubx<sup>N51A</sup> partly shapes the transcriptional (but not the splicing) program in the Hox-free *Drosophila* S2R+ cells. Analysing the impact of Ubx<sup>N51A</sup> on homeosis and morphogenesis may provide key information about the co- and post-transcriptional functions of Ubx in development (66).

Our study shed light on a novel ability of Ubx to bind RNA in vivo and in vitro. This defines Ubx as member of the DNA- and RNA-binding proteins (DRBPs), together with the Sox TFs or the HD-TF Bcd (67–69). Ubx binds several RNA exonic sequences in vivo and in vitro. In contrast, the RNA-binding ability of Bcd is restricted to) caul) for which Bcd binds a putative RNA sequence in the 3′UTR (70). This depends on the amino acid R54 of Bcd and Pol II, our results reveal that Ubx mediates elongation-dependent splicing. Interestingly, the rate of the Pol II affects Ubx–RNA binding affinity compared to the full-length protein. Sox2 also mediates RNA-binding via an additional 60 amino acid RNA-binding motif (RBM) (76). Even more important, the RBM domain of Sox2 provides RNA sequence specificity as shown by SELEX experiments. Altogether, it suggests that Ubx–RNA binding affinity relies on multiple interaction interfaces.

In this study, we provide the first evidence for a dynamic interplay between Ubx and Pol II to regulate splicing. First, our data indicate that Ubx interacts with Pol II, both with the paused (S5Phos) and active (S2Phos) forms. This is consistent with a large amount of studies showing that Ubx and the Hox TFs in general regulate transcription activation (77–80), promoter pausing, Pol II release (35,81,82) and interact with various components of the pre-initiation complex (56,83). Yet, no direct interaction between Hox TFs and Pol II were identified. Our results show for the first time a direct interaction between Ubx and the conserved CTD of Pol II. How this interaction is mediated in vivo and impacts on Ubx transcriptional and splicing activity is yet to be explored.

Second, using transcriptional drugs and mutant of the Pol II, our results reveal that Ubx mediates elongation-coupled splicing. Interestingly, the rate of the Pol II affects Ubx spliced target genes differently. It has been shown that the rate of Pol II can impact both on exon retention and exclusion (84). This strongly suggests that co-transcriptional splicing requires fine-tuned transcription rate, which may...
depend on cis- (RNA consensus sequences) and trans- (repressor or activator recruitment) regulations. Thus, Ubx may regulate splicing via various molecular mechanisms, such as by promoting the recruitment of splicing activators (SR proteins) or repressors (hnRNP), by modifying the RNA folding or impacting on the chromatin landscape (9,13,85) or by regulating the expression (and splicing) of splicing factors. Taken together, this further extends the possible combination of Hox molecular mode of actions through direct or indirect mechanisms.

Third, ChIP and FRAP experiments reveal that Ubx chromatin binding in the gene body and protein dynamic are affected by the transcription rate. Interestingly, the fluorescence recovery of Ubx was not extended in presence of slow Pol II C4 (Supplementary Figure S15B). On one hand, this may reflect the increase of Ubx molecules interacting with active Pol II, having a shorter half time recovery (Figure 7C). This is supported by co-immunoprecipitation experiments showing stronger interaction between Ubx and slow Pol II C4 (Figure 7D). On the other hand, if the dynamic of Ubx is impacted by the slow Pol II C4, it is conceivable that other Ubx-protein associations would be perturbed, thereby influencing the global dynamic behaviour of Ubx as well. Thus, Ubx could use different mechanisms to regulate splicing, and one of them would be in close cooperation with Pol II.

Fourth, our data strongly suggest that Ubx chromatin binding is pivotal for its interaction with the active Pol II and its interplay with active elongation process. This is supported by several evidences: i), we showed previously that in vivo, Ubx N51A and Ubx WT interactomes are distinct. Notably, Ubx N51A does not interact with chromatin binding proteins nor TFs whereas it interacts with mRNA processing factors, and splicing factors (27); ii), Ubx N51A interacts with total and paused Pol II but weakly with elongating Pol II (Figures 5A, 6B); iii), Ubx N51A exhibits a distinct behaviour compared to Ubx WT, with a much shorter residence time on the chromatin (Figure 7, Supplementary Figure S15B-C); iv), in contrast to Ubx WT form, Ubx N51A dynamic behaviour is not affected by the Pol II C4 slow mutant. This mutant only relates to the active form of Pol II traveling along the gene body. In this configuration Ubx N51A dynamic does not rely on the Pol II rate as the interaction with active/elongating Pol II is impaired (Figures 5A–C, 6A–D).

Altogether, these data suggest a model in which Ubx regulates co-transcriptional splicing via a dynamic interplay with Pol II, the splicing machinery and a dual requirement of DNA-RNA interface. We propose herein two non-exclusive models for Ubx molecular mode of action (Figure 7G–I). First, Ubx is dynamically loaded along the gene body, with a binding/bouncing behaviour, thereby allowing the recruitment of the spliceosome machinery for its de novo assembly (86). This model is supported by the presence of Ubx in active transcriptional cluster in vivo allowing close-proximity of Ubx molecules and chromatin (87). Second, Ubx is initially loaded onto the chromatin in complex with Pol II and splicing factors at promoter regions. Upon transcriptional activation, they travel together to regulate transcription and splicing.

Ubx-Pol II interaction could as well be driven by chromatin looping between enhancers and promoters (Figure 7G). In this configuration, Ubx would be stably bound onto an enhancer and contacts Pol II loaded on the promoter. While the interface with the chromatin seems at play for Ubx splicing function, we did not observe specific enrichment of Ubx on chromatin region of differentially spliced exons. In contrast, Ubx binds specifically RNA sequences of differentially spliced exons (Figure 4A-C, Supplementary Figure S11A-C). Thus, we propose that during active transcription, Ubx regulates differential splicing by recognising and binding specific RNA sequences thereby promoting the loading of the spliceosome and specific exon retention. In future, single molecule imaging strategy will be crucial to investigate these co-transcriptional regulatory models (88).

All in all, our work lays the groundwork to understand the role of Hox proteins in mRNA splicing, thereby providing new perspectives of Hox function in development and diseases. Beyond the Hox TFs, it broadens our insights into the molecular mechanisms employed by TFs to coordinate the variety of cell and tissue identities.

DATA AVAILABILITY

Raw data of RNA-Seq performed in Drosophila S2R+ cells have been deposited on Gene Expression Omnibus (GSE171547) according to the ENCODE guideline. We used published data from embryonic mesoderm of Ubx ChIP-Seq performed in late mesoderm (stages 14–17) (GSE121752) and mesoderm RNA-Seq of early (10–13), late (14–17) stages as well as upon depletion of Ubx (GSE121670) from (28). We further re-analysed transcriptomes (GSE101557, control) and Ubx ChIP-Seq (GSE101556) performed in Drosophila S2 cells in (35).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the Bloomington Center for fly lines, and DGRC for plasmids. We would like to further thank Alexandra Moreira for sharing PTB plasmid, Jean-Yves Rognant for the RIP protocol, Guido Grossmann and Jan Lohmann for access to critical equipment. We thank the Nikon Imaging Center for access to FRAP facilities as well as the deep sequencing facilities of the Cell Networks. We are grateful to the people who helped improving the manuscript, in particular Samir Merabet who offered significant discussion during the revision process. We deeply thank the Merabet team for helping to establish the scientific environment to perform some of the experiments critical for the revisions. Author contributions: Conceptualisation: J.C., P.B.P. Experimental design: J.C., P.B.P. Experimental procedures: J.C., P.B. FRAP acquisition and modelling: P.N. Bioinformatic data curation: P.B., K.D., H.D.P. Data analysis: J.C.. Revision: J.C., P.B. Supervision: J.C. Writing, editing: J.C. with support of all authors. Funding acquisition: I.L., J.C.

FUNDING

Fondation pour la Recherche Médicale [ARF20200401178 to J.C.]; DFG [LO 844/8-1 to I.L., in part]. Funding for open access charge: Heidelberg University.
Conflict of interest statement. None declared.

REFERENCES

1. Moore, M.J. and Proudfoot, N.J. (2009) Pre-mRNA processing reaches back to transcription and ahead to translation. Cell, 136, 688–700.
2. Chen, M. and Manley, J.L. (2009) Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nat. Rev. Mol. Cell Biol., 10, 741–754.
3. Wahl, M.C., Will, C.L. and Lührmann, R. (2009) The spliceosome: design principles of a dynamic RNP machine. Cell, 136, 701–718.
4. Bentley, D.L. (2014) Coupling mRNA processing with transcription in time and space. Nat. Rev. Genet., 15, 163–175.
5. Spector, D.L. and Lamond, A.I. (2011) Nuclear speckles. Cold Spring Harb. Perspect. Biol., 3, a000646.
6. Galganski, L., Urbanek, M.O. and Krzyzosiak, W.J. (2017) Nuclear speckles: molecular organization, biological function and role in disease. Nucleic Acids Res., 45, 10350–10368.
7. Blencowe, B.J. (2006) Alternative splicing: new insights from global analyses. Cell, 126, 37–47.
8. Hegele, A., Kamburov, A., Grossmann, A., Sourlis, C., Wowro, S., Weimann, M., Will, C.L., Pena, V., Lührmann, R. and Stelzl, U. (2012) Dynamic protein–protein interaction wiring of the human spliceosome. Mol. Cell, 45, 567–580.
9. Carrocci, T.J. and Neugebauer, K.M. (2019) Pre-mRNA splicing in the nuclear landscape. Cold Spring Harb. Quant. Biol., 8, 11–20.
10. Khodor, Y.L., Rodriguez, J., Abruzzi, K.C., Tang, C.-H.A., Marr, M.T. and Weimann, M. (2017) Nuclear landscape. Cell, 168, 10350–10368.
11. de la Mata, M., Alonso, C.R., Kadener, S., Fededa, J.P., Blaustein, M., Rebagliati, M.R., Nishikura, K., Xi, G., Geng, W., Levine, S.S. et al. (2017) The Hox proteins Ubx and AbdA collaborate with the transcription pausing factor M1BP to regulate gene transcription. EMBO J., 36, 2887–2906.
12. Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol., 15, 520–2512.
13. Rhee, D.Y., Cho, D.-Y., Zhai, B., Slattery, M., Ma, L., Mintseris, J., Wong, C.Y., White, K.P., Celniker, S.E., Przytycka, T.M. et al. (2014) Coupling of RNA polymerase II transcription elongation with pre-mRNA splicing. J. Mol. Biol., 428, 2623–2635.
14. Aubeouf, D. (2018) Alternative mRNA processing sites decrease genetic variability while increasing functional diversity. Transcription, 9, 75–87.
15. Junion, G., Spivakov, M., Girardot, C., Braun, M., Gustafson, E.H., Birney, E. and Furlong, E.E.M. (2012) Alternative splicing collectively defines cardiac cell fate and reflects lineage history. Cell, 148, 473–486.
16. Rambout, X., Dequiedt, F. and Maquat, L.E. (2018) Beyond transcription: roles of transcription factors in pre-mRNA splicing. Chem. Rev., 118, 4339–4364.
17. Girardot, M., Bayet, E., Maurin, J., Fort, P., Roux, P. and Raynaud, P. (2018) SOX9 has distinct regulatory roles in alternative splicing and transcription. Nucleic Acids Res., 46, 9106–9118.
18. Aubeouf, D. (2002) Coordinate regulation of transcription and splicing by steroid receptor coregulators. Science, 298, 416–419.
19. Han, H., Braunsweg, U., Konstantinopoulou-Pournazis, T., Weatheritt, R.J., Hirsch, C.L., Ha, K.C.H., Radovani, E., Nabeel-Shah, S., Sterne-Weiler, T., Wang, J. et al. (2017) Multilayered control of alternative splicing regulatory networks by transcription factors. Mol. Cell, 65, 539–553.
20. Riddel, C.J., Gilles, A.F. and Averbouh, M. (2013) MicroRNAs act as cofactors in bicoid-mediated translational repression. Curr. Biol., 23, 1579–1584.
21. Cramer, P., Pesce, C.G., Baralle, F.E. and Kornblitt, A.R. (1997) Functional association between promoter structure and transcript alternative splicing. Proc. Natl. Acad. Sci. U.S.A., 94, 11456–11460.
22. Bürghlin, T.R. and Affolter, M. (2016) Homeodomain proteins: an update. Chromosoma, 125, 497–521.
23. He, S., del Viso, F., Chen, C.-Y., Ikmi, A., Kroesen, A.E. and Gibson, M.C. (2018) An axial Hox code controls tissue segmentation and body patterning in Nematostella vectensis. Science, 361, 1377–1380.
24. Pearson, J.C., Lemons, D. and McGinnis, W. (2005) Modulating Hox gene functions during animal body patterning. Nat. Rev. Genet., 6, 893–904.
25. Castelli-Gair, Hombri, J., Sánchez-Higuera, C. and Sánchez-Herrero, E. (2016) Control of organogenesis by Hox genes. In: Castelli-Gair, Hombri, J. and Bovolenta, P. (eds.) Organogenetic Gene Networks. Springer International Publishing, Cham, pp. 319–373.
