

Figures and figure supplements

Conserved functional antagonism of CELF and MBNL proteins controls stem cell-specific alternative splicing in planarians

Jordi Solana et al

eLIFE Research article



Figure 1. Alternative splicing is differentially regulated between planarian stem cells and differentiated cells. (A) Planarian Stem cells ('Neoblasts'), their differentiating progeny, and differentiated cells are purified with FACS ('X1', 'X2', 'Xins', respectively). RNA-seq and computational analyses were *Figure 1 continued on next page*



Figure 1 continued

subsequently used to identify X1-differential AS at a genome-wide scale. (B) Distribution of AS events with increased/decreased inclusion of the alternative sequence in X1. Alt3/5, alternative splice site acceptor/donor selection; IR, intron retention; AltEx, cassette exons. (C) Heatmap of inclusion level values for 293 representative X1-differential AS events. Bars in the dendrogram correspond to AS types in **B**. (D) Representative RT-PCR assays monitoring AS patterns in FACS isolated cell fractions. Red and blue exons indicate those exons with higher and lower inclusion in X1 compared to Xins fractions, respectively. Scatter plot shows correspondence between PSI estimates by RNA-Seq and RT-PCR in whole worms and X1, X2 and Xins fractions for 22 events ($R^2 = 0.92$, n = 88). (E) Proportion of alternatively spliced exons by length class with increased inclusion in X1 ('X1-inc') or Xins ('X1-exc') fractions, or not differentially regulated between X1 and Xins fractions ('Non-X1 reg') DOI: 10.7554/eLife.16797.003

The following source data is available for figure 1:

Source data 1. Schmidtea mediterranea RNA-seq samples used in this study. DOI: 10.7554/eLife.16797.004

Source data 2. List of neoblast-differential AS events. DOI: 10.7554/eLife.16797.005



Figure 1—figure supplement 1. Genome annotation pipeline and summary statistics. (A) Genome annotation pipeline used in this study. (B) Number and percentage of canonical introns (with GY-AG splice sites) in all genes, *Figure 1—figure supplement 1 continued on next page*

Figure 1—figure supplement 1 continued

coding genes, and introns within coding regions. (C) Whole genome intron length distribution shows a marked bimodal distribution. (D) Number of protein-coding genes according to their intron density. DOI: 10.7554/eLife.16797.006



Developmental Biology and Stem Cells | Genomics and Evolutionary Biology

Figure 1—figure supplement 2. Identification and analysis of neoblast differential AS in planarians. (A) Number of alternative splicing events identified by type. Left: number of AS events based on RNA-seq data by type. Alt3/5, Alternative acceptor /donor splice site choice; IR, intron retention; AltEx, alternative exon skipping events (cassette exons). Middle: number of exon skipping events according to the number of neighboring alternative exons that form the event. Right: number of alternative exons according to their length. Microexons are defined as those exons with lengths between 3 and 27 nucleotides. (B) Distribution by type of planarian AS events with increased/ decreased neoblast inclusion of the alternative sequence. 'Alt3/5', alternative splice site acceptor/donor selection; 'IR', intron retention; 'MIC', microexons (length between 3 and 27 nucleotides); 'Single', single cassette long (>27 nucleotides) exon; 'Multi', long exon that is part of an array of multiple neighbouring cassette exons. (C) Percent of neoblast-differential, alternative or constitutive exons with high (average disorder rate >0.67), mid (between 0.33 and 0.67) and low (<0.33) disorder calculated using Disopred2. p-values correspond to 3-way Fisher tests. (D) Percent of residues that overlap a PFAM protein domain. p-values correspond to proportion tests.



Figure 1—figure supplement 3. RT-PCR validation of neoblast-differential AS events. RT-PCR validation assays for 22 and 14 representative AltEx and IR events that are predicted to be differentially regulated between neoblasts and differentiated cells by RNA-Seq analyses. Red and blue exons/introns indicate those exons/introns with higher and lower inclusion in X1 compared to Xins fractions, respectively. Scatter plots show correspondence between inclusion level estimates by RNA-Seq and RT-PCR in whole worms and X1, X2 and Xins fractions (AltEx, R² = 0.92, n = 88; IR, R² = 0.68, n = 56). DOI: 10.7554/eLife.16797.008



Figure 1—figure supplement 4. Gene Ontology analysis of X1-differential AS events. (A-C) Gene Ontology analysis of neoblast-differential AS events: Gene ontology enrichment analysis for 'Biological Process', 'Molecular Function' and 'Cellular Component' terms was performed for neoblast-Figure 1—figure supplement 4 continued on next page



Figure 1—figure supplement 4 continued

differential events that are predicted to (I) generate distinct protein isoforms in X1 and Xins (black), (J) disrupt ORF in X1 (dark grey), and (K) disrupt ORF in Xins (light grey).



Figure 2. Abundant IR in planarian neoblasts. (A) Representative RT-PCR assays monitoring IR patterns in FACS isolated cell fractions. Red and blue introns indicate those introns with higher and lower inclusion in X1 compared to Xins fractions, respectively. Scatter plot shows correspondence between PSI estimates by RNA-Seq and RT-PCR in whole worms and X1, X2 and Xins fractions for 14 events (R² = 0.68, n = 56). (B) Intron length distributions for all introns (black), all retained introns (grey), X1-included introns (red) and X1-excluded introns (blue). Median length (m) is indicated for each set. (C–D) Gene expression measured using the cRPKM metric for genes containing X1-included (C) or excluded (D) introns in X1, X2 and Xins cell *Figure 2 continued on next page*



Figure 2 continued

fractions. P-value corresponds to a Wilcoxon Sum Rank test between X1 and Xins expression values. (E) Percent of X1-differential AS events by type that are predicted to generate alternative ORF-preserving isoforms (black), disrupt the ORF in neoblasts or differentiated cells (dark/light grey), or overlap non-coding sequences (white).



Figure 3. The neoblast-specific AS program is extensively conserved in *D. japonica*. (A) Left: representative RT-PCR assays monitoring AS patterns for five representative neoblast-differential AS events in FACS isolated cell fractions from *D. japonica*. Scatter plot shows correspondence between ΔPSI (X1-Xins) estimates by RT-PCR in *S. mediterranea* and *D. japonica* for 21 cassette exons (circles) and 6 retained introns (diamonds). Conservation of regulation is observed for both alternative sequences with higher ('X1-inc', red) and lower ('X1-exc', blue) inclusion in neoblasts. (B) Schematic examples of the occurrence of neoblast-differential AS (blue bars) with respect to protein domain organization in two pairs of gene homologues in human and planarian. The examples show that AS events fall in similar protein regions in both human and planarian orthologs.

DOI: 10.7554/eLife.16797.011

The following source data is available for figure 3:

Source data 1. Orthologous gene groups with neoblast/ESC-differential exons in both planarian and humans. DOI: 10.7554/eLife.16797.012

Source data 2. Conservation of stem cell-differential AS events. DOI: 10.7554/eLife.16797.013



Figure 4. Identification of *bruli* and *mbnl* as major regulators of neoblast-specific AS. (A) Scattered plot highlighting differential gene expression for selected tissue-specific AS factors in planarians. X-axis: differential gene expression in X1 vs Xins cell fractions for selected AS factors (color dots) and core spliceosomal components (grey dots). The red line is the median enrichment value for the core spliceosomal components. An AS factor is thus considered to be enriched in X1 or Xins fractions if it is located to the right or to the left of this line, respectively. Y-axis: gene expression levels in whole worms (WW), using the cRPKM metric. (B) Test of regeneration speed after head ablation upon AS factor knockdown. Identification of normally looking eyespots was used as proxy for complete regeneration for 10 individuals per experimental condition. (C) Δ PSI estimates in X1 cell fractions by RT-PCR for two X1-included exons (red), one X1-included retained intron (*c17orf39*, orange) and five X1-excluded exons (blue) 10 days after RNAi treatment with dsRNA coding for AS factor combinations. Δ PSI values for each event and experiment are calculated respect to the average PSIs in three wild type samples. (D) Percent of neoblast-differential exons with sufficient read coverage that change their inclusion levels (Δ PSI \geq 15) in whole worms towards the X1 pattern (as expected for a negative regulator of neoblast-differential AS) upon knockdown of each AS factor.

The following source data is available for figure 4:

Source data 1. Annotation of RBPs in planarians.

DOI: 10.7554/eLife.16797.015

Source data 2. Annotation of spliceosomal components in planarians.

DOI: 10.7554/eLife.16797.016

Source data 3. Single and multiple splicing factor knockdown RNAi groups. DOI: 10.7554/eLife.16797.017



Figure 4—figure supplement 1. Identification of *bruli* and *mbnl* as major regulators of neoblast-specific AS. (A) FACS scatter plots for *bruli* and control RNAi treated animals 10 days after RNAi inyection. (B) Proportion of cells in X1, X2 and Xins fractions 10 days after *bruli* and control RNAi treatment. DOI: 10.7554/eLife.16797.018



Figure 4—figure supplement 2. Enrichment of RBP binding motifs associated with neoblast-differential exons. Motif enrichment analyses for all RNAcompete-derived motifs (*Ray et al., 2013*) for neoblast-differential alternatively spliced exons with increased or decreased inclusion in X1 fraction. Each intronic box corresponds to a 20-nucleotide bin at the indicated location with respect to the AS exon. The whole exonic sequence was evaluated as a single bin. In each box, counts of the top 10% most strongly bound 7-mers were compared between the AS set of interest and a non-regulated background set. Colors encode significance of enrichment: log10 P-value Fisher's exact test.







Figure 5—figure supplement 1. BRULI and MBNL factors domain architecture and effects on regeneration and alternative splicing. (A-B) Amino acid sequence (A) and schematization of protein domains (B) of BRULI, MBNL-1, MBNL-LIKE-1, MBNL-LIKE-2 and MBNL-LIKE-3. RRM1 domains are depicted Figure 5—figure supplement 1 continued on next page



Figure 5—figure supplement 1 continued

in green and Zinc-finger domains in magenta, and their CCCH aminoacids indicated in red. The X1-inc alternative exon of *mbnl-1* is depicted in cyan. (C) Sequence identity matrix between planarian and human CELF RRM1 (top) and MBNL ZnF (bottom) domains. DOI: 10.7554/eLife.16797.021



Figure 5—figure supplement 2. *bruli* and *mbnl* factors effects on regeneration and alternative splicing. (A) The effects of *bruli* and *mbnl* knockdown and *mbnl-1* single knock down are reproducible and dependent on the level of cut. Top: Workflow of the experiment: 10 animals per experimental *Figure 5—figure supplement 2 continued on next page*



Figure 5—figure supplement 2 continued

group and time point were injected with dsRNA coding for *bruli, gfp* as a control and a mix of dsRNAs against the four *mbnl* genes and cut 22 days after RNAi; trunk and tail pieces were monitored independently. Two independent dsRNAs coding for different regions of *mbnl-1* were used (*mbnl-1* and *mbnl-1**). The experiment was replicated twice independently. Bottom: Results of the scoring for trunk and tail regeneration experiments; color key is indicated below. (B) The effects of *bruli* and *mbnl* knockdown on AS events are detectable 5 days after RNAi and are permanent for at least 30 days: quantifications for two X1-inc exons (red), one X1-inc retained intron (c17orf39, orange) and five X1-exc exons (blue) 5, 10, 15, 20, 25 and 30 days after RNAi treatment with dsRNA coding for *bruli, gfp* as a control and a mix of dsRNAs against the four *mbnl* genes. DOI: 10.7554/eLife.16797.022

eLIFE Research article



Figure 6. *bruli* and *mbnl* antagonistically regulate neoblast-specific AS. (A) High negative association ($p<1.4 \times 10^{-19}$, one-sided binomial test) between differences in inclusion levels (Δ PSI) of X1-differential AS events in X1 versus Xins fractions, and differences in *bruli* and control RNAi treated X1 fractions. Red/Blue dots correspond to X1-included/excluded AS events. (B) High positive association ($p<2.1 \times 10^{-11}$, one-sided binomial test) between differences in inclusion levels (Δ PSI) of X1-differential AS events in X1 versus Xins fractions, and differences in *mbnl* and control RNAi treated Xins fractions. Only AS events with sufficient read coverage in the control and KD samples and an absolute Δ PSI>15 are plotted. (C) RT-PCR assays monitoring AS patterns for 5 representative X1-differential AS events in FACS isolated cell fractions treated with *bruli*, control or *mbnl* RNAi. (D) Most X1-differential AS events that are affected by both *bruli* and *mbnl* knockdown are regulated in an antagonistic manner ($p<2.2 \times 10^{-7}$, one-sided binomial test). (E) 2-dimensional histogram of RNA-compete 7-mer Z-scores, comparing the sequence-specific binding of planarian and human proteins. Unspecific 7-mers with Z-score < 0 for both RBPs were excluded. ρ , Spearman rank correlation; R, Pearson correlation. Top 10 planarian motifs are highlighted. (F) Motif-enrichment analysis. Each intronic box corresponds to a 20-nucleotide bin at the indicated location relative to the AS exon (middle box). Color encodes the significance of enrichment (Fisher's exact test, Bonferroni corrected for the number of tested bins) of high affinity 7-mers for BRULI or MBNL, comparing each differentially spliced exon set with an unaffected background set. Exon sets (I-IV) correspond to those in quadrants indicated in panels B and C.









Figure 6—figure supplement 1 continued

introns indicate those exons/introns with higher and lower inclusion in X1 compared to Xins fractions, respectively. Scatter plot shows correspondence between inclusion level estimates by RNA-Seq and RT-PCR ($R^2 = 0.87$, n = 225). DOI: 10.7554/eLife.16797.024



Figure 6—figure supplement 2. *bruli* and *mbnl* antagonistically regulate neoblast-specific AS. (A-B) *bruli* and *mbnl* have a minor impact on neoblast-differential AS events in Xins and X1, respectively: (A) Lack of association (p = 0.206, one-sided binomial test) between differences in inclusion levels (ΔPSI) of neoblast-differential AS events in X1 versus Xins fractions, and differences in *bruli* and control RNAi treated Xins fractions. (B) Lack of association (p = 0.714, one-sided binomial test) between differences in inclusion levels (ΔPSI) of neoblast-differential AS events in X1 versus Xins fractions, and differences in *mbnl* and control RNAi treated X1 fractions. (C) RNAcompete-derived motifs for BRULI and MBNL proteins. Top-scoring RNA-binding 7mers and corresponding logos identified by RNAcompete analysis are shown for MBNL-1 (excluding a neoblast-differential exon; X1 isoform), MBNL-1IKE-1, MBNL-1IKE-2, and BRULI. Scatterplots show the correlation between 7-mers Z-scores derived from microarray data splits, SetA and SetB. Logos are created after aligning the top 10 highest scoring 7-mers. Both isoforms of MBNL-1 display similar binding preferences. RNAcompete retrieved simpler 4-mer or 3-mer containing GCU motifs for MBNL-LIKE-1 and MBNL-LIKE-2, which contain only one pair of zinc fingers. (D) Conserved binding specificities for planarian and human CELF and MBNL splicing factors: Venn diagrams showing the overlap of RNA-compete 7-mers, comparing the top 100 (top) and the top 1000 (bottom) 7-mers of planarian and human proteins. P-values correspond to hypergeometric tests.

eLIFE Research article



Figure 7. *bruli* and *mbnl* antagonistically regulate neoblast biology. (A) Schematic representation of phenotypic analysis experiments. (B) Gene expression changes estimated by RNA-Seq of several progeny markers 25 days after *bruli* (green) or *mbnl* (purple) RNAi treatment compared to controls. (C) Quantification of gene expression levels by qPCR of *prog-1*, *prog-2-1*, and *prog-1-1* in whole worms 25 days after treatment with *bruli*, *control* or *mbnl* RNAi. (D) Proportion of genes belonging to each of the 6 clusters defined in (*Onal et al., 2012*) based on their X1vs Xins enrichment that are decreased 20, 25 and 30 days after *bruli* and *mbnl* RNAi treatment. P-values correspond to hypergeometric tests for Clusters 1 (P = 5.6 $\times 10^{-56}$, P = 4.0 $\times 10^{-113}$, P = 2.7 $\times 10^{-178}$, for *bruli(RNAi*) days 20, 25 and 30, respectively) and 6 (P = 1.2×10^{-3} , P = 1.0×10^{-14} , P = 1.2×10^{-16} , for *mbnl(RNAi*) days 20, 25 and 30, respectively). All other tests were not significant. Black histograms on the right side indicate schematically, for each cluster, the relative gene expression levels in each cell fraction. (E) *In situ* hybridization of three representative Cluster 1 (red bars) and Cluster 6 (blue bars) genes in whole worms 25 days after *bruli*, *control* or *mbnl* RNAi. (G) Principal Component 1 separates transcriptomes from *bruli* RNAi treated samples together with *Smedwi-2(RNAi*)(which impairs neoblast differentiation).

transcript id	type of marker			C [bru	li_ctrl1	log2 F	C [mbi	l-ctrl1	
		marker	day 20	day 25	day 30	day 20	day 25	day 30	
comp2592 c0 seq1		bruli	day 20	uuy 20	aay oo	uuy 20	uuy 20	uuy oo	
comp2838_c0_seq1	KD mRNAs	mbnl-1							
comp4169_c0_seq1		mbnl-like-2							
comp7489_c0_seq1		mbnl-like-1							
comp659_c0_seq1	neoblast markers	smedw-1							
comp756 c0 seq1	manhlanti ONC	smedwi-2							
comp1985 c0 seq1	neoplast+CINS	vasa-1							
comp1582_c0_seq1	markers	smedtud1							
comp6020_c0_seq1		fgfr-1							
comp3935_c0_seq1		egr-1							
comp5563_c0_seq1		p53							
comp5942_c0_seq1	zeta-neoblast (van	soxP-3							
comp8720_c0_seq1	Wolfswinkel et al	yala 1/2/3							
comp7033_c0_seq1		hox3							
comp15471 c0 seq1	2014)	six6							
comp6729 c0 seq1		nkx2.4							
comp4998_c1_seq1		tcf15							
comp4078_c0_seq1		foxJ							
comp8483_c0_seq1		soxP-2							
comp/22/_c0_seq1		soxP-1							
comp1024_c0_seq1	sigma-neoblast	1gtr-4							
comp4172_c0_seq1	(van Wolfswinkel	nlk-1							
comp3098_c0_seq1		nhx-1							
comp13371 c0 seq1	et al . 2014)	lhx-2							
comp7324_c0_seq1		pbx4							
comp13215_c0_seq1		soxB-1							
comp13772_c0_seq1	gamma-neoblast	pro x-1							
comp1694_c0_seq1	(van Wolfswinkel	hnf-4							
comp4075_c0_seq1	ot al. 2014)	gata4/5/6							
comp332_c0_seq1	et al . 2014)	nrog-1							
comp69_c0_seq1	early progeny	prog-1							
comp478 c0 seq1		prog-1-1							
comp213_c0_seq1		prog-2-5							
comp3549_c0_seq1		pmp-3							
comp61_c0_seq1		prog-2-3							
comp363_c0_seq1		prog-2-2							
comp3665_c0_seq1		prog-1-3							
comp5958 c0 seq1		5V2-RP							
comp5517 c0 sea1		RPC-2							_
comp2178_c0_seq1		prog-1-8							
comp3226_c0_seq1	Now onidormal	pmp-8							
comp1013_c0_seq1	ivew epidernial	pmp-7							-
comp2147_c0_seq1	progenitor genes	pmp-4							
comp4492_c0_seq1	(Zhu <i>et al</i> . 2015)	prog-1-6							
comp771 c0 seq1		prog-1-10							
comp468 c0 seq1		prog-Pr							-
comp343 c0 seq1		prog-1-9							
comp12027_c0_seq1		SLC7a11							
comp775_c0_seq1		pmp-10							-
comp263_c0_seq1		pmp-6							
comp719_c0_seq1		prog-1-5							
comp6130_c0_seq1		DCOR-1							
comp1925_c0_seq1		prog-1-4							
		prog-F2							
comp899 c0 seq1		agat2							
comp2385 c0 sea1	lata 1947	agat3							
comp5493_c0_seq1	late progeny	odc							
comp2449_c0_seq1		тср							
10 = 1 0 0 1		********							

Figure 7—figure supplement 1. bruli and mbnl knockdown have contrasting effect of specific gene markers. Log2 fold change in gene expression of key markers 20, 25 and 30 days after bruli (left) or mbnl (right) treatment, measured from RNA-Seq. Red/Green intensities indicate different levels of Figure 7—figure supplement 1 continued on next page



Figure 7—figure supplement 1 continued

down/up regulation. bruli and mbnl knockdown have distinct effects on different types of markers. In particular, they show antagonistic effects of 'Early progeny' markers.



Figure 7—figure supplement 2. Gene Ontology analysis of genes downregulated upon *bruli* and *mbnl* knockdown. Gene ontology enrichment analysis for 'Biological Process', 'Molecular Function' and 'Cellular Component' terms of genes downregulated upon *bruli* and *mbnl* knockdown was performed for genes that were significantly downregulated after treatment with *bruli* (green) or *mbnl* (red) RNAi. DOI: 10.7554/eLife.16797.028



