Control of embryonic stem cell self-renewal and differentiation via coordinated alternative splicing and translation of YY2

Soroush Tahmasebi,1a,b, Seyed Mehdi Jafarnejad,1a,b, Ingrid S. Tam,1a,b, Thomas Gonatopoulos-Pournatzis,1c, Edna Matta-Camacho,1a,b, Yoshihiro Tsukumo,1a,b,2, Akiko Yanagiyama,1a,b, Wencheng Li,3 Yaser Atlantis,4 Maxime Caron,5a,g, Ulrich Braunschweig,4 Dana Pearl,4a,b, Arkady Khoutorsky,4 Christos G. Gkogkas,3 Robert Nadao,5a,g, Guillaume Bourque,5a,g, Xiang-Jiao Yang,1a,b,j, Bin Tian4, Hendrik G. Stunnenberg,1a,b,j, Yojiro Yamanaka,1a,f, Benjamin J. Blencowe,4c,k, Edna Matta-Camachoa,b, Yoshinori Tsukumaa,b,2, Akiko Yanagiyama,1a,b, Wencheng Lid, Yaser Atlasie, Maxime Caronf,g, Vincent Giguère,a,b,1, and Nahum Sonenberg,c,3

Contributed by Nahum Sonenberg, September 22, 2016 (sent for review August 15, 2016; reviewed by George Q. Daley and Maurice S. Swanson)

Translational control of gene expression plays a key role during the early phases of embryonic development. Here we describe a transcriptional regulator of mouse embryonic stem cells (mESCs), Yin-yang 2 (YY2), that is controlled by the translation inhibitors, Eukaryotic initiation factor 4E-binding proteins (4E-BPs). YY2 plays a critical role in regulating mESC functions through control of key pluripotency factors, including Octamer-binding protein 4 (Oct4), Yamanaka factors (OSKM), and transcriptional repressors (Sousa, 2016). YY2 associates with eIF4G to control the activity of eIF4F, the translation initiation complex, and thereby regulates the expression of genes that are critical for mESC self-renewal and differentiation. This study identifies YY2 as a key regulator of mESC self-renewal and differentiation and highlights the importance of translational control in maintaining the pluripotency of mESCs.


Reviews: G.Q.D., Children’s Hospital Boston; and M.S.S., University of Florida. The authors declare no conflict of interest.

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected in 2015.

Significance

Embryonic stem cells (ESCs) maintain a low translation rate; therefore control of mRNA translation is critical for preserving their stemness. We identified a hitherto unstudied transcription factor, Yin-yang 2 (YY2), which is translationally regulated and controls self-renewal and differentiation of mouse ESCs (mESCs). Although YY2 is essential for mESC self-renewal, increased YY2 expression directs differentiation of mESCs toward cardiovascular lineages. Examination of the YY2 5′-UTR revealed a multi-layer regulatory mechanism through which YY2 expression is dictated by the combined actions of the splicing regulator, Polypyrimidine tract-binding protein 1 (PTBP1), and the translation inhibitors, Eukaryotic initiation factor 4E-binding proteins (4E-BPs). YY2 directly controls the expression of several pluripotency and development-related genes. This study describes a synchronized network of alternative splicing and mRNA translation in controlling self-renewal and differentiation.

12360–12367 | PNAS | November 1, 2016 | vol. 113 | no. 44

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1615540113/-/DCSupplemental.
encode proteins that control fundamental cellular processes such as cell proliferation and survival (16).

We showed that 4E-BPs are required for reprogramming mouse embryonic fibroblasts (MEFs) to induced pluripotent stem cells (iPSCs) (17). In the current study, we describe a tightly coordinated network in mESCs whereby the expression of the Yin-yang 2 (YY2) transcription factor is controlled by the splicing regulator Poly(pyrimidine) tract-binding protein 1 (PTBp1) and the 4E-BP translation inhibitors. Our data reveal that stringent regulation of YY2 expression by this network is critical for mESC self-renewal and lineage commitment.

Results

Transcriptome and Translatome Profiling of WT and 4E-BP1/2-Null mESCs. To investigate the role of 4E-BPs in mESCs, we first derived mESCs from WT and Eukaryotic translation initiation factor 4E-binding protein 1 (Eif4ebp1) and Eukaryotic translation initiation factor 4E-binding protein 2 (Eif4ebp2) double-knockout (DKO) mice and then examined the eIF4F complex using a 7-methyl-GTP (m7GTP)-agarose pull-down assay. The eIF4F amount was elevated, as demonstrated by increased (~10.6-fold) eIF4G1 pull-down in DKO mESCs (Fig. S1A). However, polysome profiling (Fig. S1B) and [35S]methionine/cysteine-labeling assays (Fig. S1C) did not detect a substantial difference in global mRNA translation between WT and DKO mESCs. These data are consistent with previous findings that the lack of 4E-BPs affects the translation of a subset of mRNAs rather than affecting global translation (17, 18).

To identify 4E-BP-sensitive mRNAs in mESCs, we performed ribosome profiling (19), which allows precise measurement of the translation of mRNAs on a genome-wide scale, by deep sequencing of ribosome-protected fragments (ribosome footprints; RFPs). We achieved a high degree of reproducibility between the replicates for mRNA sequencing (mRNA-Seq) and RFPs (R² > 0.97, Fig. S1D). Metagene analysis confirmed the enrichment of RFP reads in coding sequences and the expected three-nucleotide periodicity (Fig. S1E). These analyses validated the quality of the mRNA and RFP libraries (Table S1). We used Babel analysis (20) to compute changes in the abundance of RFPs (Fig. L4) independent of changes in the levels of their corresponding mRNAs (Fig. 1B). A significant enhancement in the translation efficiency of a small subset of mRNAs was detected in DKO mESCs [false discovery rate (FDR) < 0.1] (Fig. 1C and Dataset S1), as is consistent with the lack of global change in translation in the DKO mESCs (Fig. S1 B and C). Strikingly, mRNA-Seq data revealed down-regulation of mRNA levels for several pluripotency factors, such as PR domain-containing 14 (Prdm14), ES cell-expressed Ras (Eras), Estrogen-related receptor-β (Esrβ), and Nanog (−1.3, −1, −0.6, and −0.9, respectively; log2 DKO/WT) in DKO mESCs (Dataset S2). Possible reasons for this down-regulation are discussed below.

Ablation of Eif4ebp1 and Eif4ebp2 Results in Reduced Expression of mESC Markers. To validate the mRNA-Seq results, we examined the expression of pluripotency factors in undifferentiated WT and DKO mESCs, and also in WT and DKO mESCs after differentiation, by Western blot and quantitative RT-PCR (RT-qPCR). Although all DKO mESC lines maintained normal morphology under standard mESC culture conditions [in the presence of Leukemia inhibitory factor (LIF) and feeder layer (irradiated MEFs)], expression of the ESC marker NANOG was reduced by 60–70% (Fig. 1D). These changes did not result from unintended consequences of the knockout procedure, because RNAi-mediated depletion of 4E-BP1 and -2 (double knockdown; DKD) resulted in a strong (80%) reduction in NANOG (Fig. S2A). DKO mESCs proliferated more slowly than WT mESCs (Fig. S2B), and when cultured in the absence of feeder layer, they exhibited flattened morphology, which is indicative of cellular differentiation, whereas WT mESCs preserved their normal morphology (Fig. S2C). Similar morphological changes were observed in DKD ESCs (Fig. S2D). Notably, when cultured in the absence of LIF and a feeder layer, NANOG and OCT4 expression was strongly suppressed in DKO mESCs, whereas WT mESCs maintained a higher expression of these proteins (Fig. S2E). In addition to Nanog, mRNA levels of Oct4 and Sox2 were lower in DKO mESCs on day 0 of differentiation (Fig. S2F). Moreover, as determined by RT-qPCR, mRNA levels of several other ESC factors, such as Lin28a, Eras, and ten-eleven translocations (Tets) (Tet1, -2, and -3), were reduced in DKO mESCs (Fig. S2G). Therefore, 4E-BPs are required for the regulation of expression of pluripotency factors. Analysis of embryoid bodies (EBs) derived from WT and DKO ESCs revealed that the lack of 4E-BP1 and 4E-BP2 resulted in the differentiation of mESCs toward mesodermal and endodermal lineages, as indicated by significant up-regulation of Bone morphogenetic protein 4 (Bmp4), an early mesoderm marker, and Gata4 and Gata6, early endoderm markers (Fig. S2H). This up-regulation coincides with the down-regulation of the neuroectoderm markers Microtubule-associated protein 2 (Map2) and Sex-determining region Y (SRY)-box 17 (Sox17) mRNAs in DKO EBs.

The decrease in mRNA levels for several pluripotency factors, such as Nanog, Eras, and Erkb1 in DKO mESCs compared with WT mESCs (Fig. 1D, Fig. S2 F and G, and Dataset S2) suggests that 4E-BP-dependent translational regulation of one or more factor(s) affects these changes in the transcriptome.

Stringent Control of YY2 Expression in mESCs. One of the mRNAs exhibiting the most significant increase in translation efficiency in DKO mESCs versus WT cells is the YY2 transcription factor (Babel P value = 0.0001) (Fig. 1C and Dataset S1). YY2 exhibits considerable sequence homology (56% identity) with the YY1 transcription factor (21). Like reduced expression 1 (Rex1) [Zinc finger protein 42 (Zfp42)], a well-known ESC marker (22, 23),
YY2 is a retroposed copy of the YY1 gene, which evolved only in placental mammals (24). YY1 is a pleiotropic transcription factor that regulates diverse cellular processes and plays a critical role in early embryonic development, ESC biology, and reprogramming (25–27). Although the N-terminal domains of YY1 and YY2 differ significantly, the C-terminal DNA-binding domain of YY1 is highly conserved in Rex1 and YY2 (22), suggesting that YY2 may play an important role in the regulation of gene expression in mESCs. Although there is no significant change in the level of Y2 mRNA in DKO mESCs compared with WT cells (Fig. S3A), YY2 protein, but not YY1, is elevated 1.9- to 2.4-fold in DKO mESCs (Fig. IE). Importantly, expression of a phosphorilation-resistant 4E-BP1-4A mutant (28) in DKO mESCs reduced YY2 protein levels (Fig. 1F). These data demonstrate that translation of YY2 mRNA is controlled by 4E-BPs.

To study the functional consequence of YY2 up-regulation in mESCs, we generated an mESC line carrying a doxycycline-inducible YY2 construct (dox-YY2). Overexpression of YY2 caused a reduction in the expression of pluripotency factors Nanog, c-Myc, and Oct4 mRNAs (Fig. 2A and B), indicating a negative role for YY2 in mESC self-renewal. However, constitutively expressed shRNA against YY2 in mESCs caused depletion of mESCs in culture (Fig. S3B). This depletion coincided with increased levels (8.4-±2.4-fold) of the apoptosis marker, cleaved caspase-3, in YY2-knockdown cells (Fig. S3C), suggesting that mESCs require a basal level of YY2 expression for survival. To study the impact of a moderate knockdown, we generated an mESC line carrying a doxycycline-inducible shRNA construct against YY2 (shYY2). Although slight induction of shYY2 (0.2 μg/mL doxycycline) enhanced expression of the pluripotency factors Nanog and Sox2 (Fig. 2C), higher doses of doxycycline failed to do so (Fig. 2C), indicating a dose-sensitive effect of YY2 on the expression of mESC pluripotency factors. Importantly, the deleterious effect of complete depletion of YY2 was not limited to mESCs, because CRISPR/Cas9-mediated YY2-knockout blastocysts were unable to maintain their inner cell mass, as demonstrated by blastocyst outgrowth assays (Fig. 2D and Fig. S3 D and E). A similar defective outgrowth has been described previously for Yy1−/− blastocysts (29), suggesting that the lack of YY2 may cause peri-implantation lethality, as described for Yy1−/− mice (26), and demonstrating that YY1 and YY2 fulfill nonredundant functions in blastocyst growth.

Recent studies showing that YY1 directly regulates Nkx2.5 expression and promotes cardiogenesis uncovered a novel function for YY1 in cardiomyocyte differentiation and cardiac morphogenesis (30, 31). To examine the effect of YY2 on mESC differentiation toward cardiovascular lineages, EBs derived from dox-YY2 mESCs were exposed to doxycycline. One week after induction of YY2 expression, foci of beating cardiomyocytes began to appear in the plates with the highest level of YY2 induction (0.2 μg/mL doxycycline) (Movie S1), and the number of foci continued to increase during the following 2 weeks. No beating foci appeared in noninduced EBs up to the third week of differentiation. Consistently, expression of several cardiovascular-specific markers, such as Nkx2.5, Bone natriuretic peptide (Bnp), alpha-Myosin heavy chain (αMHC), Myosin light chain 2a (MLC2a), and MLC2v mRNAs was increased in the YY2-overexpressing EBs in a dose-dependent manner (Fig. 2E). These data suggest that YY1 and YY2 have overlapping functions in directing the differentiation of mESC toward cardiovascular lineages.

**YY2 Binds to the Regulatory Regions of Key Genes for ESC Pluripotency and Differentiation.** Genomic targets of YY1, but not YY2, in mESCs have been documented (25, 32). We determined the genomewide binding sites of YY2 in mESCs by ChIP-Seq in WT mESC cells overexpressing YY2. Because of the high degree of similarity between the C-terminal domains of YY1 and YY2, we used a monoclonal antibody that specifically recognizes the N-terminal domain of YY2 (Fig. 2F and Fig. S3F; also see SI Materials and Methods). YY2-binding sites (Dataset S3) exhibit enrichment for the genomic loci of coding genes (exons or introns, indicated by gene, 43%) (Fig. 3A) and a preference for the promoter regions surrounding the transcription start sites (TSS) (Fig. 3B). Nearly half of the peaks contained the consensus YY1-binding motif (Fig. 3C), as is consistent with a similar sequence preference for YY1 and YY2 (21, 22). Motif distribution across binding peaks revealed enrichment for the known consensus YY1-binding motif directly at YY2 peak centers (Fig. 3D), indicating specific recognition of these binding sites by YY2.

Pathway analysis of genes associated with YY2-binding peaks using the Ingenuity Pathway Analysis (IPA) program revealed significant enrichment for genes related to ESC pluripotency (Fig. 3E), such as Oct4, Teratocarcinoma-derived growth factor 1 (Tdgf1), Ezh2, and Forkhead box protein D3 (FoxD3). We also found enrichment for genes involved in the activation of the retinoic acid receptor (RAR)-signaling pathway. Previous studies showed that activation of the RAR pathway promotes differentiation of ESCs to cardiomyocytes, particularly MLC2v+ ventricular cardiomyocytes, and that the RAR-signaling pathway plays a critical role in cardiogenesis (33, 34). These findings indicate that activation of this pathway, along with other cardiovascular-related YY2 targets [e.g., Bnp, Mesoderm posterior protein 2 (Mesp2), and MKL (megakaryoblastic leukemia)/myocardin-like 1 (Mkl1)], is responsible for engendering the differentiation of mESCs toward the cardiovascular lineage by YY2 (Fig. 2F). We validated the ChIP-Seq results for a selected number of genes with ChIP-qPCR (Fig. 3 F and G and Fig. S3G).
RT-qPCR analysis of a selected number of pluripotency factors in the ChiP-Seq dataset, such as Oct4, Esrmb, Tet1, and Tet2 mRNAs, revealed that YY2 overexpression suppresses their expression in mESCs (Fig. 4A). The ChiP-Seq analysis did not identify Nanog as a target of YY2. Thus, the decrease in Nanog mRNA expression in DKO mESCs (Dataset S2) or upon overexpression of YY2 (Fig. 2A and B) is likely secondary to the down-regulation of other pluripotency factors. In agreement with our previous observation (Fig. 2C), doxycycline-inducible Y2 knockdown in mESCs revealed a dose-sensitive regulation of its target genes (Fig. 4B).

We analyzed the target genes identified by the YY2 ChiP-Seq assay relative to those of YY1 in mESCs (32) and found that 27.7% of YY2 targets are shared with YY1 (Fig. 4C and Dataset S3). Conversely, a large portion of the YY2 targets (~72%) are not present among the YY1 targets (Fig. 4C), a finding that supports a previous study of genome-wide mRNA expression in HeLa cells demonstrating that YY1 and YY2 regulate some shared but mostly unique sets of genes (35). Although some pluripotency-related genes such as Tdfl1 are unique YY2 targets, Oct4, Krüppel-like factor 5 (Klf5), and Foxd3 are common targets of YY1 and YY2 (Fig. 4C). Notably, Yy1 is a target of both YY1 and YY2, and our data showed that YY2 has a dose-sensitive effect on YY1 expression (Fig. 2A and C). The similar consensus-binding motifs of YY1 and YY2 are consistent with these factors exhibiting overlapping or competing effects on common target genes (21, 36–38). One plausible explanation for their distinct activities is the considerable divergence in their N-terminal domains (Fig. S4A). A recent report showed that the binding of YY1 to active promoters/enhancers in ESCs through its C-terminal DNA-binding domain is facilitated by concurrent binding of its N-terminal domain to RNA species transcribed from these regulatory elements (32). We hypothesized that the N-terminal domain of YY2 cannot interact with RNA. We examined this possibility by EMSA using purified mouse YY1 and YY2 proteins (Fig. S4 B and C) and DNA and RNA probes derived from the AT-rich interactive domain-containing protein A1 (ARID1A) promoter, which interact with YY1 (32). Notably, Arid1a is among the YY2 targets in our ChiP-Seq analysis (Dataset S3). Although YY2 interacts with the DNA probe with only slightly less efficiency than YY1 (Fig. 4D), only YY1 binds to the RNA probe (Fig. 4E), because no visible binding to the Arid1a promoter was seen by negative control.
promoter-derived RNA was detected for YY2, even after prolonged exposure (Fig. S4D). In agreement with these results, analysis of the N-terminal domains of mouse Y1 and YY2 proteins by BindN, an RNA-binding prediction server (39), identified two distinct RNA-binding motifs for YY1 that are not conserved in YY2 (Fig. S4E). Our data suggest a mechanism by which YY2 may act differently from YY1 and that their differential affinity for the promoter-derived RNAs underlies the opposing effects of YY1 and YY2 on certain shared promoters.

**A Retained 5′-UTR Intron Renders Yy2 mRNA Sensitive to 4E-BP-Dependent Translational Repression.** To search for eIF4E-sensitive elements in the 5′-UTR of Yy2 mRNA, we first used 5′-RACE to annotate the sequence in mESCs. In addition to the annotated mRNA sequence of the 217-nt 5′-UTR (RefSeq: NM_001098723.1), henceforth referred to as “variant A,” we uncovered two additional variants possessing 290-nt and 100-nt 5′-UTRs, respectively (Fig. 5A and B). The 290-nt variant (designated “variant B”) is a 5′ extension of variant A; the 100-nt variant is a spliced version of variant A, lacking 117 nucleotides. The spliced region harbor all the features of a canonical intron, including a GU dinucleotide at the 5′ splice site (SS), an AG dinucleotide at the 3′ SS, and a polypyrimidine tract and putative branch site A nucleotide within 20 nt of the 3′ SS (Fig. 5A). Considering that the intronic sequence is shared between variants A and B, we termed their corresponding spliced variants ΔA and ΔB, respectively (Fig. 5B). Retroposed genes are generally intronless, because they are generated through the reverse transcription of mature mRNAs (40, 41). Therefore, the intron acquisition by Yy2 is most likely a recent evolutionary event that occurred after the retroposition of Yy2 from Yy1 in placental mammals.

To measure intron retention in Yy2 mRNA during differentiation of mESCs by RT-PCR, we used two forward primers (Fw1 and Fw2) and a common reverse primer (Rv) to target the flanking exons (Table S2). Primers Fw1 and Rv detect only variants B and ΔB (297-bp and 180-bp PCR products, respectively), whereas Fw2 and Rv amplify a 209-bp PCR product for variants A and B and a 92-bp PCR product for variants ΔA and ΔB. We examined Yy2 intron retention events in mESCs and EBs at days 4 and 6 postdifferentiation and measured the degree of intron retention using the percent intron retention (PIR) as a metric (42). The percentage of nonspliced variants (A and B) is higher in mESCs than in EBs, demonstrating that differentiation coincided with a marked reduction in intron retention (Figs. 5C and Fig. S5A). Notably, this alternative splicing event is not restricted to mESCs, because various degrees of intron retention were detected at different embryonic stages and across different tissues (Fig. 5D and Fig. S5B), demonstrating developmental and tissue-specific regulation of Yy2 alternative splicing.

To identify the trans-acting factor(s) responsible for Yy2 alternative splicing, we used the RBPremap web server to predict consensus motifs for RNA-binding proteins (43). We identified two canonical PTBP recognition motifs (CUCUCUC) flanking the Yy2 5′-UTR intron (Fig. 5A and B and Fig. S5C). PTBP1 and PTBP2 (a neural- and testis-enriched paralog) are RNA-binding proteins implicated in several aspects of mRNA metabolism, including alternative splicing, stability, localization, and polyadenylation (44). PTBP1 is essential for embryonic growth before gastrulation, because Ptbp1−/− mESCs have severe proliferation defects (45, 46). To explore its role in Yy2 5′-UTR intron retention, PTBP1 was depleted in mESCs using shRNA (Fig. S5D). PTBP1 knockdown resulted in reduced Yy2 intron retention, demonstrating that it acts as an inhibitor of Yy2 splicing in mESCs (Fig. 5E). This knockdown was associated with reduced growth and smaller colonies of mESCs (Fig. S5E). To demonstrate that PTBP1 directly affects Yy2 5′-UTR splicing, we

![Fig. 5](https://www.pnas.org/doi/10.1073/pnas.1615540113)

**Fig. 5.** Retention of the 5′-UTR intron renders Yy2 sensitive to 4E-BP-mediated translation suppression. (A) Sequence of the promoter region of the mouse Yy2 gene. The two conserved TSS are marked by arrowheads; the boxed sequence shows the retained intron; the two exons highlighted in red are the consensus PTBP-binding motifs; and the underlined ATG is the translation start codon for Yy2 mRNA. (B) A cartoon depicting the four variants of Yy2 5′-UTR. vB and vΔA represent a long variant with and without intron retention, respectively; vA and vΔA represent a short variant with and without intron retention, respectively. CD, coding DNA sequence. (C) RT-PCR using the primer pair Fw2 and Rv designed to recognize all four possible variants to estimate the splicing efficiency of the 5′-UTR intron in mESCs and mEBs on days 4 and 6 postdifferentiation. Gapdh mRNA was used as the control. (D) RT-PCR analysis of intron retention (IR) in the Yy2 5′-UTR in different mouse embryonic stages and adult tissues using the primers described in C. Gapdh mRNA was used as the control. (E) RT-PCR analysis of intron retention in the Yy2 5′-UTR using the primers described in C upon depletion of PTBP1 expression in mESCs by two independent shRNAs. Yy2-ORF primers amplifying a segment of the coding region of Yy2 were used to demonstrate the change in overall expression of Yy2 mRNA. β-actin mRNA was used as the internal control. (F) RT-PCR amplification (primers Fw2 and Rv) of the in vitro splicing products of the A, B, and ΔA variants in WERI retinoblastoma cell extracts with different amounts of recombinant PTBP1 protein. Recombinant BSA was used as a negative control. (G) Luciferase reporter assay with Firefly (Fluc) and Renilla (Rluc) luciferase reporter mRNAs, as described in Fig. S5J. The in vitro-transcribed mRNAs were purified and transfected into WT and DKO mESCs. (Left) The normalized luciferase activity of each construct in WT mESCs. (Right) Comparison of the luciferase activity of each construct in DKO and WT mESCs. Fluc mRNA was cotransfected with Rluc mRNA as a transfection control. Data are mean ± SD (n = 3). **P < 0.01, ***P < 0.001; ns, nonsignificant. CTR, control; RLiu, relative luminescence units.
performed an in vitro splicing assay using a WERI retinoblastoma (WERI-Rb1) cell extract, purified recombinant PTBP1 protein, and the three Y2 5′-UTR variants (A, B, and ΔA) transcribed in vitro. Incubation in the cell extract resulted in splicing of the introns from A and B variants but had no effect on the ΔA variant (Fig. S5 F and G). Importantly, the addition of recombinant PTBP1 protein (Fig. S5H) dramatically suppressed splicing and resulted in the complete retention of the intron in a dose-dependent manner (Fig. 5F and Fig. S5J).

The length and complexity of the 5′-UTR play a critical role in mRNA translation, because mRNAs with long and structured 5′-UTRs are more sensitive to eIF4E activity (15, 47). The combined use of alternative TSS and alternative splicing determine the complexity of the Y2 5′-UTR. To examine the effect of 5′-UTR variants on mRNA translation, we constructed luciferase reporters containing the 5′-UTRs of the A, B, and ΔA variants (Fig. S5J) and used the reporters to generate mRNAs that were transfected into WT and DKO mESCs. The B, and to a lesser extent A, variant mRNAs were poorly translated, whereas translation of the intron-less ΔA variant was markedly more efficient (more than sixfold higher than control) in WT mESCs (Fig. 5G, Left). These data demonstrate that the 5′-UTR containing the retained intron sequence inhibits translation. The translation of A and B variant mRNAs was significantly elevated in the DKO mESCs, whereas the ΔA variant remained insensitive to 4E-BPs levels (Fig. 5G, Right). These results demonstrate that intron retention in the Y2 5′-UTR, in combination with the activity of 4E-BPs, determines the outcome of Y2 mRNA translation. Intron retention adds an extra 117 nt to the 5′-UTR, increasing the complexity of its secondary structure (Fig. S6A) and rendering it sensitive to 4E-BP–mediated translation repression. As is consistent with the increased translation of the Y2 spliced variant (Fig. 5G), the highest expression of Y2 protein was detected in the heart and muscle tissues (Fig. S6B), in which the lowest degree of Y2 5′-UTR intron retention is observed (Fig. 5D and Fig. S5J).

This double-layered control mechanism, consisting of the retention of the Y2 5′-UTR intron by PTBP1 and suppression of translation of the resulting mRNA variant by 4E-BPs, allows the modulation of Y2 mRNA translation.

Discussion
Many mammalian mRNAs contain heterogeneous 5′-UTRs (48). Length, structure, and sequence elements in the 5′-UTRs strongly impact translation (49, 50). We found that the combination of alternative TSS and splicing produces four Y2 mRNA variants with different translation efficiencies. Thus, the relative level of each variant, in combination with 4E-BP activity, dictates the rate of YY2 protein synthesis. ESC differentiation, which is a known regulator of mRNA splicing in these tissues (53, 54). Intron retention has emerged as a widespread mechanism to regulate gene expression in different cell and tissue types as well as during stem cell differentiation (42, 55, 56). The finding that a retained intron at the 5′-UTR of Y2 mRNA controls its translation underscores the intricate interplay between alternative splicing and translational control (57, 58).

We established Y2 mRNA as a target of 4E-BP–dependent translation suppression in mESCs and demonstrated that a basal level of YY2 expression is essential for mESC survival. Similar to Yl1-knockout embryos, CRISPR-mediated Y2-knockout embryos survived the preimplantation period, but the growth of their inner cell mass was impaired, as revealed by blastocyst outgrowth assays (Fig. 2D). These observations exclude the possibility of redundant functions for YY1 and YY2 in the early stages of embryonic development. The critical function of YY2 in mESCs is most likely mediated by its direct transcriptional regulation of ESC genes such as Oct4, Eras, Taf1, Tet2, and Tfdg1. We demonstrated that mESCs are highly sensitive to YY2 and that a modest manipulation of YY2 levels significantly affects the expression of pluripotency factors, ESC self-renewal, and differentiation (Figs. 2 A–C and 4 A and B). Consistent with the observed phenotype in DKO mESCs, overexpression of YY2 in WT mESCs induces down-regulation of several pluripotency factors and promotes differentiation. Although in the current study we focused on the characterization of YY2, the ribosome-profiling experiment identified several other mRNAs, such as Triggering receptor expressed on myeloid cells 2 (Triml2), Sumo1, Amigo3, Lefty, and Nodal, that display 4E-BP sensitivity. The change in translation of these additional targets in DKO mESCs could contribute to the differentiation-prone phenotype of DKO mESCs and explain the differences that may be observed between DKO mESCs and mESCs overexpressing YY2.

We documented that increased expression of YY2 directly affects translation of mouse embryoid bodies (mEBs) toward the cardiovascular lineage. The importance of YY2 in cardiomyocytes is likely not limited to early differentiation, because the heart expresses a higher level of YY2 than other tissues of adult mouse (Fig. S6B). This function is likely mediated by transcriptional control of the RAR pathway and multiple YY2 targets, such as Bmp4, Nodal, Bnp, Mesp2, and Mkl1, as evident from the ChipSeq analysis. Recent studies showed binding of YY1 to the promoters of highly expressed ribosomal proteins and the nuclear encoded mitochondrial membrane, enzymes, and ribosomal proteins (55) and highlighted the importance of YY1 in cardiomyocyte differentiation and heart morphogenesis (30, 31). Notably, cardiac-specific ablation of Yl1 causes severe abnormalities in the heart, indicating that YY2 in the heart does not compensate for YY1 deletion (31). Comparison of Y2 and Yl1 ChipSeq data (Fig. 4C and Dataset S3) demonstrated overlapping sets of nuclear encoded mitochondrial proteins, such as the aminocyl RNA synthetases Turs2, and Lrs2, as well as mitochondrial ribosomal proteins Mrps18c, Mrpl1, Mrpl44, and Mrpl53 (Fig. 4C and Dataset S3). That YY2 and YY1 have both unique and overlapping target genes explains their convergent but nonredundant functions in development.

The reduced expression of several pluripotency markers (e.g., Oct4, Nanog, PRDM14, and Esrrb) in DKO mESCs and their greater tendency for differentiation (e.g., slower proliferation and a flattened morphology) resembles a “primed” state of pluripotency (59). Although the transcriptional regulatory network safeguarding the naive/primed state of pluripotency has been studied extensively, the role of posttranscriptional control mechanisms, particularly translational control, remained largely unexplored. The importance of postranscriptional regulatory processes has been highlighted recently by the identification of the RNA-binding proteins Pum1 and Lin28 as regulators of naive/primed pluripotency (60, 61). Detailed analysis of mRNA translation therefore should provide valuable information about the role of translational control in these pluripotent states.

In summary, we have described a mechanism by which mESC self-renewal and lineage commitment are controlled via stringent regulation of YY2 expression at two stages: (i) by repressing the splicing of the Y2 5′-UTR intron via PTBP1 and (ii) by suppressing the translation of the resulting mRNA variant by 4E-BP1 and -2.
Fig. 6. Regulation of YY2 expression by splicing and mRNA translational control. A proposed model depicting the regulation of YY2 expression at two different steps: (i) alternative splicing (via PTBP1) and (ii) mRNA translation (via 4E-BP1/2). A basal level of YY2 expression is required for maintenance of mESC self-renewal, but increased translation of YY2 mRNA directs cardiovascular lineage commitment.

These two layers of control (Fig. 6) coalesce to limit the expression of YY2 protein to a low basal level in mESCs, thereby maintaining their self-renewal and pluripotency.

Materials and Methods
ESC Cell Culture and Differentiation. Mouse ES cells were maintained in DMEM (Wisent Inc.), 1% nonessential amino acids (Gibco), 1% l-Glutamine (Wisent Inc.), 1% sodium pyruvate (100 μM stock from Invitrogen), 0.1 mM β-mercaptoethanol, 5% (vol/vol) FBS, 1,000 U mouse lLIF (ESGRO; Millipore), penicillin (50 μg/mL), and streptomycin (50 μg/mL) and were expanded on the feeder layer or gelatin. For mESC differentiation (62), 800–2,000 mESCs were cultured for 2 d in hanging drops containing differentiation medium (DMEM, 20% (vol/vol) FBS, 1% nonessential amino acids (Gibco), 1% l-Glutamine (Wisent Inc.), penicillin (50 μg/mL), and streptomycin (50 μg/mL)). The resulting EBs were transferred to bacteriological dishes and cultured in suspension for 3 d. Then they were plated onto gelatin-coated tissue-culture plates for the rest of the differentiation process.

Single-Guide RNA Synthesis. The DNA template for YY2 single-guide RNA (sgRNA) was synthesized by PCR reactions using px330 (Addgene) as a template. Two primers were used: TTAGG-sgRNA forward and T-SCGAAGCTTTAGGATTCGTCAGATCACTTATAAGTTGATAGGGCAGTTTACCAAGTGTTAGTCAGATAATGC-3) and scramble (5′-AAAAGACGCACTGCAATGGC-3). The PCR product was purified with the PCR DNA fragments extraction kit (Geneaid) and was used as a template for sgRNA synthesis with the T7 MAXscript kit (Ambion). The synthesized sgRNA was ETOH-precipitated and dissolved in RNase-free water.

Cytoplasmic Microinjection, Embryo Culture, and Blastocyst Outgrowth. The animal ethics committee of the McGill University approved all the animal procedures. Cas9 mRNA (50 ng/μL) (Sigma) and 50 ng/μL, Y2-sgRNA in 10 mM KCl were injected into CD1 zygotes in M2 medium (Zenith Biotech). Cytoplasmic injection was performed with a Femtojet microinjector (Eppendorf) and a Cyto721 intracellular amplifier (World Precision Instruments) for the procedures. Cas9 mRNA (50 ng/μL) was injected into embryos at the 2 cell stage, and the zona pellucidae of the injected zygotes were removed with acid Tyrode, 15% (vol/vol) FBS, 1,000 U mouse lLIF/mL (ESGRO; Millipore), penicillin (50 μg/mL), and streptomycin (50 μg/mL) and 5% (vol/vol) CO2 incubator at 37 °C. The embryos were cultured for 2 d in hanging drops containing differentiation medium (DMEM, 20% (vol/vol) FBS, 1% nonessential amino acids (Gibco), 1% l-Glutamine (Wisent Inc.), penicillin (50 μg/mL), and streptomycin (50 μg/mL)). The resulting EBs were transferred to bacteriological dishes and cultured in suspension for 3 d. Then they were plated onto gelatin-coated tissue-culture plates for the rest of the differentiation process.

Random RNA Fragmentation and mRNA-Seq. Cytoplasmic RNA (150 μg) was used for mRNA-Seq analysis. Poly(A)+ mRNAs were purified using magnetic oligo-dT Dynabeads (Invitrogen) according to the manufacturer’s instructions. Purified RNA was eluted from the beads and mixed with an equal volume of 2x alkaline fragmentation solution (2 mM EDTA, 10 mM Na2CO3, 90 mM NaHCO3, pH 9.2) and incubated for 20 min at 95 °C. Fragmentation reactions were mixed with stop/precipitation solution [300 mM NaOAc (pH 5.5) and GlycoBlue (Ambion)], followed by isopropanol precipitation. Fragmented mRNA was size-selected on a denaturing 10% (v/vol) polyacrylamide urea gel, and the area corresponding to 35–50 nt containing the RFPs was excised, eluted, and precipitated with isopropanol.

ACKNOWLEDGMENTS. We thank Martin Klar, Tommy Alain, Joshua Dunn, Jelena Popic, and Ola Larsson for discussions; Serge Gueroussov, Bushra Raj, and Vinagolu Rajasekhar for reagents; Joseph Marrotti for the HEK293F cell line; and Annie Sylvester and Annik Larfane for animal care.

This work was supported by Canadian Institutes of Health Research CIHR Grants MOP-142200 (to N.S.), MOP-125888 (to V.G.), MOP-115090 (to G.B.), and MOP-111197 (to Y.Y.), and MOP-14690 (to B.J.B.). S.M.J. is the recipient of a CHRI Postdoctoral fellowship. S.T. is the recipient of a Richard and Edith Strauss fellowship. U.B. is the recipient of a Human Frontier Science Program Long Term Fellowship. T.G.-P. was supported by an EMBO (European Molecular Biology Organization) Fellowship and an Ontario Institute for Regenerative Medicine Fellowship.
