Alternative splicing regulates mouse embryonic stem cell pluripotency and differentiation

Nathan Salomonis[a,b,1], Christopher R. Schlieve[a], Laura Pereira[c], Christine Wahlquist[d], Alexandre Colas[d], Alexander C. Zambon[e], Karen Vranizan[f,2], Matthew J. Spindler[a,b], Alexander R. Pico[a], Melissa S. Cline[g], Tyson A. Clark[g], Alan Williams[g], John E. Blume[g], Eva Samal[a], Mark Mercola[d], Bradley J. Merrill[c], and Bruce R. Conklin[a,b,h,1]

[a]Gladstone Institute of Cardiovascular Disease, San Francisco, CA 94158; [b]Pharmaceutical Sciences and Pharmacogenomics Graduate Program, University of California, San Francisco, CA 94143; [c]Department of Biochemistry and Molecular Genetics, University of Illinois, Chicago, IL 60607; [d]Sanford-Burnham Institute for Medical Research, La Jolla, CA 92037; [e]Department of Pharmacology, University of California, La Jolla, CA 92093; [f]Functional Genomics Laboratory, University of California, Berkeley, CA 94720; [g]Affymetrix, Inc., Santa Clara, CA 95051; and [h]Departments of Medicine and Cellular and Molecular Pharmacology, University of California, San Francisco, CA 94143

Edited by Elaine Fuchs, The Rockefeller University, New York, NY, and approved April 27, 2010 (received for review October 27, 2009)

Two major goals of regenerative medicine are to reproducibly transform adult somatic cells into a pluripotent state and to control their differentiation into specific cell fates. Progress toward these goals would be greatly helped by obtaining a complete picture of the RNA isoforms produced by these cells due to alternative splicing (AS) and alternative promoter selection (APS). To investigate the roles of AS and APS, reciprocal exon–exon junctions were interrogated on a genome-wide scale in differentiating mouse embryonic stem (ES) cells with a prototype Affymetrix microarray. Using a recently released open-source software package named AltAnalyze, we identified 144 genes for 170 putative isoform variants, the majority (67%) of which were predicted to alter protein sequence and domain composition. Verified alternative exons were largely associated with pathways of Wnt signaling and cell-cycle control, and most were conserved between mouse and human. To examine the functional impact of AS, we characterized isoforms for two genes. As predicted by AltAnalyze, we found that alternative isoforms of the gene Serca2 were targeted by distinct microRNAs (miRNA-200b, miRNA-214), suggesting a critical role for AS in cardiac development. Analysis of the Wnt transcription factor Tcf3, using selective knockdown of an ES cell-enriched and characterized isoform, revealed several distinct targets for transcriptional repression (Stmn2, Ccnd2, Atf3, Klf4, Nodal, and Jun) as well as distinct differentiation outcomes in ES cells. The findings herein illustrate a critical role for AS in the specification of ES cells with differentiation, and highlight the utility of global functional analyses of AS.

AltAnalyze | microRNA | splice isoforms | Atp2a2 | Tcf7l1

Embryonic stem (ES) cells are a vital tool for studying the regulation of early embryonic propagation and cell-fate decisions. Research in this area has led to the development of new technologies for adult somatic cell reprogramming and insights into the steps required for lineage commitment (1, 2). Several factors critical for the self-renewal of pluripotent cells have been identified in both conventional biochemical screens and genome-wide expression analyses. These include the transcription factors Oct4, Sox2, and Nanog, which interact with a common set of promoters to promote self-renewal and pluripotency (3). Recently, Tcf3, a β-catenin-responsive transcription factor, was implicated in this core transcriptional network as a direct transcriptional repressor of Oct4 and Nanog (4, 5), and is itself a target of these factors (6, 7).

In higher eukaryotes, alternative splicing (AS) and alternative promoter selection (APS) contribute to proteomic diversity by increasing the number of distinct mRNAs from a single gene locus. In different tissues and cellular states, transcript variation can alter protein interaction networks by removing or inserting protein domains, changing subcellular localization, or regulating gene expression (8). AS can also remove binding sites for translational repression by microRNAs (miRNAs) (9). AS and APS appear to increase transcript diversity in ES cells and thus are likely to affect differentiation to distinct tissue lineages (10–13). A better understanding of how AS regulates protein diversity and translational repression during ES cell differentiation may provide critical insights into this process.

Results

AS and APS Are Prominent Features of Mouse ES Cell Differentiation. ES cells and embryoid bodies (EBs) were profiled with a prototype Affymetrix exon–exon junction microarray, which interrogates the mRNAs of ~7,500 genes and more than 40,000 putative exon–exon junctions. For this study, we developed additional analytical methods for a free open-source program named AltAnalyze (http://www.altanalyze.org/) (13) (*SI Materials and Methods*). To identify alternative exons (AEs) that might indicate AS or APS, we added a previously described linear regression-based method (14) to AltAnalyze to compare the expression of reciprocal exon–exon junctions (pairs of exon–exon junctions measuring the inclusion of one or more exons). In addition to scoring AEs, this software determines the likelihood of an AE score occurring by chance, assigns protein associations to regulated exon–exon junctions, and identifies functional sequence elements (domains, motifs, and miRNA binding sites) differing between aligning alternate mRNAs and their corresponding proteins (Fig. 1*A*) (13).

Using AltAnalyze to analyze ES cell differentiation, we identified 170 unique AEs corresponding to 144 genes of 4,269 expressed genes (Dataset S1). Pathway analysis of these genes with the program GO-Elite (15), from within AltAnalyze, showed enrichment in Wnt and TGF-β receptor signaling pathways, actin cytoskeleton, lipid transport, muscle contraction, mRNA metabolism, and embryonic development, among others (Table S1).

Of the 170 AEs, 108 had evidence of AS, and 20 had evidence of APS; the remainder did not associate with a known AS or APS event. The majority of these annotated AS events aligned to pre-

Author contributions: N.S., C.R.S., L.P., C.W., A.C., A.C.Z., K.V., M.J.S., M.S.C., T.A.C., A.W., J.E.B., E.S., M.M., B.J.M., and B.R.C. designed research; N.S., C.R.S., L.P., C.W., A.C., A.C.Z., M.J.S., and E.S. performed research; N.S., C.W., A.C., M.J.S., A.R.P., M.S.C., T.A.C., A.W., J.E.B., and M.M. contributed new reagents/analytic tools; N.S., C.R.S., L.P., C.W., A.C., K.V., and E.S. analyzed data; and N.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo/ (accession no. GSE17021).

[1]To whom correspondence may be addressed. E-mail: nsalomonis@gladstone.ucsf.edu or bconklin@gladstone.ucsf.edu.

[2]Deceased June 8, 2009.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.0912260107/-/DCSupplemental.

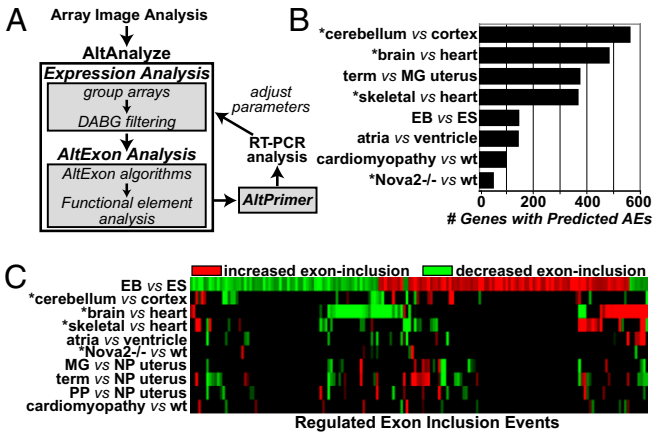

Fig. 1. Unique alternative exon profiles with ES cell differentiation. (*A*) AltAnalyze was used to process array intensities after normalization (expression analysis step), calculate alternative exon scores for reciprocal exon-junction probe sets, annotate these events based on existing mRNA structure annotations (e.g., AS events), and align probe sets to mRNA, protein, and putative functional elements (alternative exon analysis step). These predictions were used for primer design (AltPrimer), RT-PCR validation, and subsequently to reoptimize filtering parameters. (*B*) Resulting data for unique genes associated with AEs were compared for datasets indicative of AS. These comparisons included adult tissue conditions from previous studies (asterisks) or collected in parallel with array data for ES cell differentiation. (*C*) Comparison of profiles for AEs unique to ES cell differentiation. Black indicates the absence of the predicted AE event for a comparison (no scoring thresholds). Alt, alternative; DAGB, detection above background *P* value; $^{-/-}$, knockout; NP, nonpregnant; MG, midgestation; PP, postpartum.

dicted cassette exons (68%). When annotated AS and APS events were linked to protein sequences (e.g., by aligning junction pairs to mRNAs), 80 (74%) of the AEs were predicted to affect the protein domain or motif (InterPro/UniProt) composition. For three genes with evidence of AS, multiple putative miRNA binding sites were identified within retained or excluded exons (Serca2, Pdlim7, Eomes), indicating that AS may selectively alter the ability of miRNAs to regulate these proteins. Thus, an analysis of AEs for protein sequence and putative miRNA binding sites highlights a diverse set of putative functional differences.

To determine the relative extent and overall diversity in alternative exon usage, AEs from ES cells during differentiation were compared with a panel of tissue- and cell-remodeling paradigms, using public and in-house collected data. Datasets with previous evidence of large-scale AEs include cardiac vs. brain or skeletal muscle (14), cardiac atria vs. cardiac ventricle (16, 17), and Nova2 splicing factor knockout in brain vs. wild-type brain (18). Because AS and APS are also implicated in distinct cell-remodeling paradigms, we analyzed a time course of myometrial gestational- and cardiomyopathy-induced remodeling (19–22). Combined analysis of these datasets with the same AltAnalyze parameters showed a wide range in the number of predicted AEs (Fig. 1*B* and Dataset S1). Because the AltAnalyze method is limited to genes expressed in both experimental groups, it is not surprising that same or similar tissue comparisons yielded the largest number of AEs (e.g., cerebellum vs. cortex, skeletal vs. heart, term vs. quiescent uterus). Although fewer predicted AEs were found with ES cell differentiation than with other datasets, most of these AEs were unique to ES cell differentiation (Fig. 1*C*).

Verified Alternative Exons Are Associated with Pathways of Wnt Signaling and Cell-Cycle Control. To verify the expression of ES cell differentiation-predicted AEs, we tested 19 AS and four APS genes by RT-PCR (*SI Materials and Methods*). Except for three AS events, all predicted alternative isoforms were detected (Fig. S1). For 11 AS and three APS events, RT-PCR readily identified the predicted pattern of isoform expression. Analysis of these genes in the context of known interactions revealed a putative signaling network that connects many of these genes to the regulation of self-renewal and differentiation pathways (Fig. 2) (4, 23–33). These genes include the regulation of the Wnt signaling components Tcf3, Ctnnd1, and Map3k7 and the cell-cycle regulators Mark3, Smarcb1, Mxi1, and Epb4.1.

To further characterize these verified isoform changes, we examined predicted protein domain and motif changes along with isoform expression across multiple adult tissues (Fig. 3 and Fig. S2). Several of the verified AEs modified or eliminated characterized domains/motifs (e.g., Tcf3) (4), Mxi1 (34), Epb4.1 (35), thus possibly affecting protein function. Although most isoforms enriched in ES cells were also expressed in other tissues, expression in ES cells was highest for Tcf3, Smarcb1, Map3k7, and Cttnd1 for one of the two regulated isoforms, suggesting these isoforms have a unique role in pluripotency.

Conservation of Detected AS Events Between Mouse and Human. For those splicing events with the largest change in exon inclusion, conservation in human ES cell differentiation was examined using orthologous primer sets (Fig. 4). Orthologous alternative isoforms were identified for all genes, except for Tcf3, with the most similar changes in isoform abundance occurring for Serca2, Ctnnd1, Dnm1l, Kif13a, Mark3, and Smarcb1.

Serca2a Is Enriched in EBs and Lacks Serca2b Functional miRNA Binding Sites. Among some of the most intriguing changes predicted by our analysis were the modification of protein domains and the loss of predicted miRNA binding sites due to AS. Among genes with a predicted loss of miRNA binding sites due to AS, Serca2 showed the largest difference in predicted expression between alternative isoforms. Serca2 is a Ca^{2+} pump that hydrolyzes ATP during the translocation of calcium from the cytosol to the sarco/endoplasmic reticulum (SR) in muscle cells. RT-PCR and quantitative (q) PCR analysis of this gene confirmed that an isoform of Serca2, with an additional 44 amino acids and a large 3′UTR region (Serca2b), is expressed in both ES cells and EBs, whereas an isoform with a smaller alternative 3′UTR (Serca2a) was predominantly expressed in EBs (Fig. 5*A*). Serca2a is the predominant cardiac isoform and is essential for cardiac function (36). Interestingly, SERCA2b mRNA is more

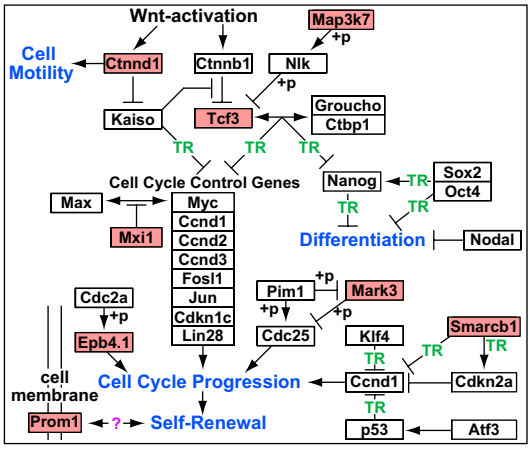

Fig. 2. AS and APS proteins intersect with pathways of pluripotency and differentiation. Proposed model of interaction between Wnt signaling, cell-cycle control, and the regulation of differentiation and pluripotency for validated AS and APS genes (red filled boxes). Arrows indicate promotion, and T-bars indicate inhibition. TR, transcriptional regulation.

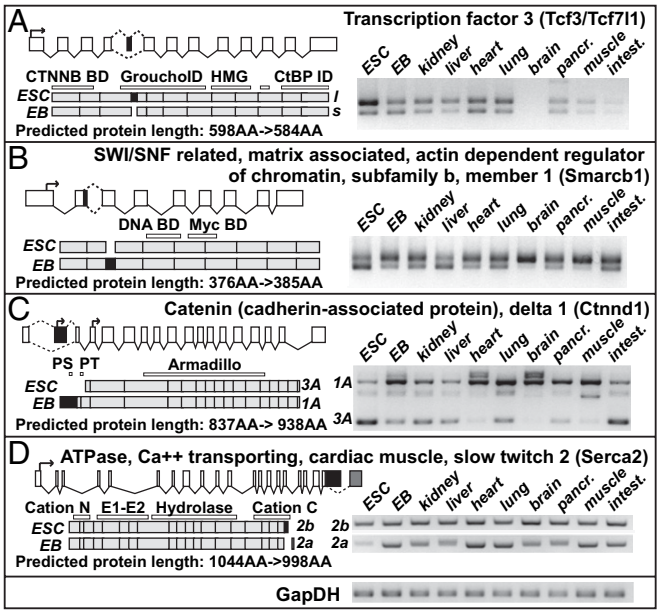

Fig. 3. Validated AS events have diverse expression patterns and alter protein composition. (*A–D*) Exon and protein structure graphs are shown for genes with large differences in isoform expression between ES cells and EBs. For each gene, the detected alternative exon is indicated as a black filled box in the exon structure graph of each panel. Dashed lines between exons indicate AS events. Below each graph are protein segments corresponding to each exon (not to scale) for aligning protein isoforms. Functional protein sequences (e.g., domains and motifs) are annotated above the two protein representations. Predicted changes in protein length are indicated for the ES cell– and EB-enriched isoforms (ES→EB). ES cell, EB, and cross-tissue expression profiles from select adult mouse tissues are shown adjacent to the exon structures for each of the probed mRNA isoforms. AA, amino acids; PS, phosphoserine; PT, phosphotyrosine (UniProt); intest., intestine; pancr., pancreas.

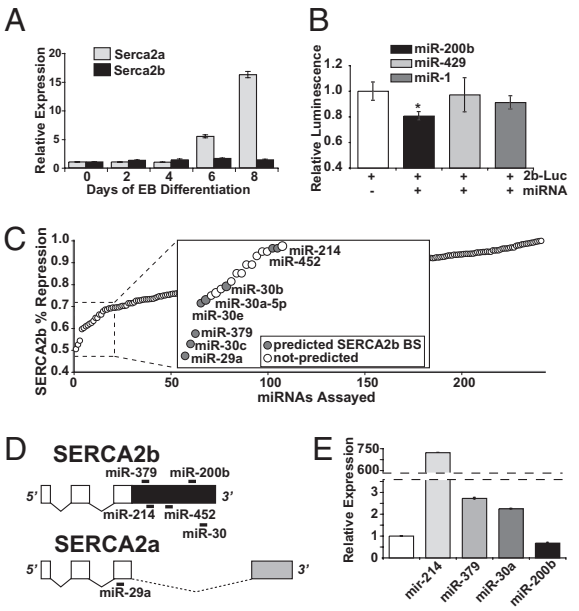

Fig. 5. Serca2 isoforms are targeted by distinct miRNAs during differentiation. (*A*) Relative expression by qPCR of Serca2a and Serca2b at 0, 2, 6, and 8 of differentiation, compared with day 0. (*B*) Relative luciferase activity is shown for mouse cardiac HL-1 cells transfected with a pMiR-luciferase reporter vector containing the Serca2b 3′UTR sequence of the predicted miR-200b binding site (2b-Luc), treated with or without mimics for miR-200b, -429, or -1. (*C*) miRNAs that produced the greatest down-regulation of SERCA2b when screened against a miRNA library in HEK-293 cells. Relative florescence is adjusted to short-hairpin knockdown of SERCA2b (positive control) and knockdown of phospholamban (negative control) florescence. (*D*) Relative position of predicted miRNAs in SERCA2a and SERCA2b exons. (*E*) miRNA relative gene expression by TaqMan for day 8 EBs relative to ES cells (white bar). Values are mean ± SEM of biological (*B*) or technical replicates (*A* and *E*). An asterisk indicates *t* test $P < 0.05$. BS, binding site.

highly degraded than that of SERCA2a in vitro (37), suggesting that the 3′UTR of SERCA2b can inhibit protein expression.

Many miRNA binding-site predictions were found only in the Serca2b 3′UTR and not Serca2a in both mouse and human, with miR-200b, -214, -429, and -30 (a, b, c, e) predicted in both species. Given that miR-200b is highly induced upon cardiac differentiation, as is Serca2a (36, 38, 39), we attempted to verify direct repression of the Serca2b miR-200b binding site in a miRNA-luciferase reporter assay (Fig. 5*B*). Addition of miR-200b specifically repressed the luciferase activity of this reporter, whereas miR-429 (predicted to overlap with the miR-200b binding site) and miR-1 (control) had no effect.

To identify other miRNAs that target Serca2b, we developed an unbiased screen to assess miRNA targeting of this isoform in vitro in human cells. To evaluate SERCA2b repression, cDNAs for GFP and SERCA2b were expressed as single transcript and tested against a library of miRNA mimics (Fig. 5*C* and Fig. S3*A*). SERCA2b relative florescence was down-regulated (>30%) for 23 miRNAs, eight of which were predicted in our analysis. Among those eight, seven aligned to the 3′UTR region unique to SERCA2b (Fig. 5*D*). qPCR analysis of differentiating mouse ES cells revealed that miR-379, -30a, and -214 were up-regulated in EBs, whereas miR-200b was highly expressed in both ES and EBs (Fig. 5*E* and Fig. S3*B*). Given the relevance of SERCA2 splicing to cardiac development, we compiled a list of genes with cardiac inhibitory or enhancing ability from the literature and determined which miRNAs were predicted to target these genes (Fig. S3*C* and *SI Materials and Methods*). This analysis found that binding sites for miRNAs of the miR-30 family (-30a, -30b, -30c, -30d, -30e) were enriched among inhibitors of cardiac differentiation. Although not overrepresented, binding sites for miR-214 and -200b were also found within several of these inhibitor genes (Fig. S3*D*). These data suggest that expression of Serca2 escapes direct repression by miRNAs important for promoting cardiac differentiation, through exon exclusion in differentiated ES cells.

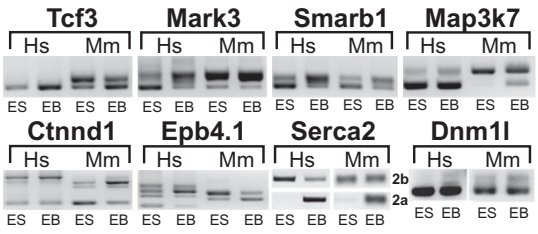

Fig. 4. Conservation of mouse splicing events in human. RT-PCR of verified splicing events in mouse (Mm) ESC differentiation and their human (Hs) orthologs in H9 ES cells (ES) and derived EBs.

Tcf3(l) Is Enriched in ES Cells and Retains Tcf3(s) Transcriptional Repression Activity. One of the most intriguing findings from our analysis was the AS of the Wnt signaling transcription factor Tcf3 (TCF7L1 in human). Though Tcf3 regulates the expression of many downstream targets, it is thought to help maintain pluripotency by repressing the transcription of both Nanog and Oct4 (4, 5). Indeed, ES cells that lack Tcf3 self-renew in the absence of leukemia inhibitory factor (LIF) and cannot give rise to fully differentiated EBs (5, 40). Though several studies focused on the

short form of Tcf3 [Tcf3(s)], our analysis highlights a longer isoform of Tcf3 [Tcf3(l)], enriched in ES cells and down-regulated upon differentiation.

The two Tcf3 isoforms detected by our exon–exon junction microarray analysis differ in the inclusion of a 42-bp cassette exon, which encodes an additional 14 amino acids [i.e., Tcf3(l)] that overlap with the Groucho binding domain (41, 42). The Groucho binding domain is necessary for Tcf3 to repress a Nanog reporter (4). As shown by RT-PCR and qPCR, Tcf3(l) is up-regulated 2-fold in ES cells vs. EBs, and Tcf3(s) is expressed at roughly equivalent levels (Fig. S4 *A* and *B*). Proteins for both isoforms were detected in ES cells, as were cDNAs encoding each isoform in a Tcf3-null ($^{-/-}$) ES cell line (4) (Fig. 6*A*).

To assess their ability to regulate the transcription of known targets, the Tcf3 isoforms were transiently expressed in Tcf3$^{-/-}$ ES cells and assayed for expression of Tcf/Lef-β-catenin transcriptional (TOPFlash) (43) or Nanog promoter-driven luciferase reporter plasmids (4). The isoforms equally inhibited the TOPFlash and Nanog reporters in a concentration-dependent manner (Fig. 6 *B* and *C*). Thus, Tcf3(l) retains transcriptional repressive activity for Tcf/Lef targets and the Nanog promoter-driven luciferase.

Tcf3(l) Knockdown Preferentially Increases Known and Novel Target Gene Expression. Given that Tcf3 isoforms show similar activity on episomal promoters, we used RNAi to examine the effects of these isoforms on endogenous promoters. Stable knockdown of Tcf3(s), Tcf3(l), and all Tcf3 isoforms [Tcf3(a)] was achieved in ES cells with the pSicoR-Ef1α-mCh-puro lentiviral construct with isoforms-specific short hairpin RNAs (shRNAs; Fig. 6*D*). This strategy yielded up to 90% knockdown (KD) of the targeted isoforms, with minimal or no reduction in the nontargeted isoform in ES cell and EB RNA (Fig. S4*C*). When analyzed by Western blot, no Tcf3 protein was detected in Tcf3(a) KD ES cells (Fig. S4*D*).

KD of any Tcf3 isoform increased Nanog expression, with maximal up-regulation by Tcf3(a) KD (4-fold), followed by Tcf3 (l) (3.5-fold), and Tcf3(s) (1.8-fold) KD (Fig. 6*E*). Interestingly, whereas Oct4 was up-regulated with Tcf3(l) KD (2-fold), Oct4 expression levels were unchanged with either Tcf3(s) or Tcf3(a) KD relative to wild-type ES cells (*t* test $P < 0.05$; Fig. 6*F*). Although Tcf3(l) KD produced a greater up-regulation of Oct4 and Nanog than Tcf3(s) KD, this effect might be due to the higher expression of Tcf3(l) in undifferentiated ES cells.

To expand the search for differentially responsive Tcf3 transcriptional targets, we examined the expression of a panel of 17 known and 17 novel candidate TCF target genes. Novel Tcf3 target genes were identified as being affected by Tcf3 knockout, KD, induction, and promoter occupancy microarray studies (6, 40, 44). Of the 34 genes examined, 20 genes were differentially expressed with KD of at least one of Tcf3 isoform ($P < 0.05$; Fig. S5). Among these differentially expressed genes, six (Atf3, Ccnd2, Jun, Klf4, Nodal, and Stmn2) exhibited a 2-fold or greater difference in expression between Tcf3(s) and Tcf3(l) KD (Fig. 6*G*). These patterns of expression persisted throughout ES cell differentiation (Fig. S6*A*). These factors are implicated in both the regulation cell-cycle progression and a broad range of developmental pathways (Fig. 2) (30, 32, 33, 45). Tcf3 promoter occupancy was previously observed for Atf3, Ccnd2, Jun, Nodal, and Stmn2, as was opposing regulation by Tcf3 knockout and induction for Atf3, Jun, Klf4, and Stmn2 (6, 40, 44). These data suggest that the two Tcf3 isoforms regulate both an overlapping set as well as distinct sets of target genes.

Delayed Differentiation of Tcf3 Knockdown Lines. Isoform-specific Tcf3 KD lines displayed unique cell-colony morphologies in the presence and absence of LIF (Fig. 7 and Fig. S7). With LIF, wild-type E14 ES cells typically distribute into a monolayer when grown on gelatinized culture plates (without mouse feeder cells). Tcf3(a), Tcf3(s), and Tcf3(l) KD ES cells resulted in clustered, rounded colonies when maintained in LIF. When wild-type ES cells were removed from LIF, they displayed a differentiated morphology (cobblestone-like) by day 3 and could not be passaged further, whereas all of the Tcf3-KD lines could be further passaged to day 6 and did not have a predominantly differentiated morphology. In

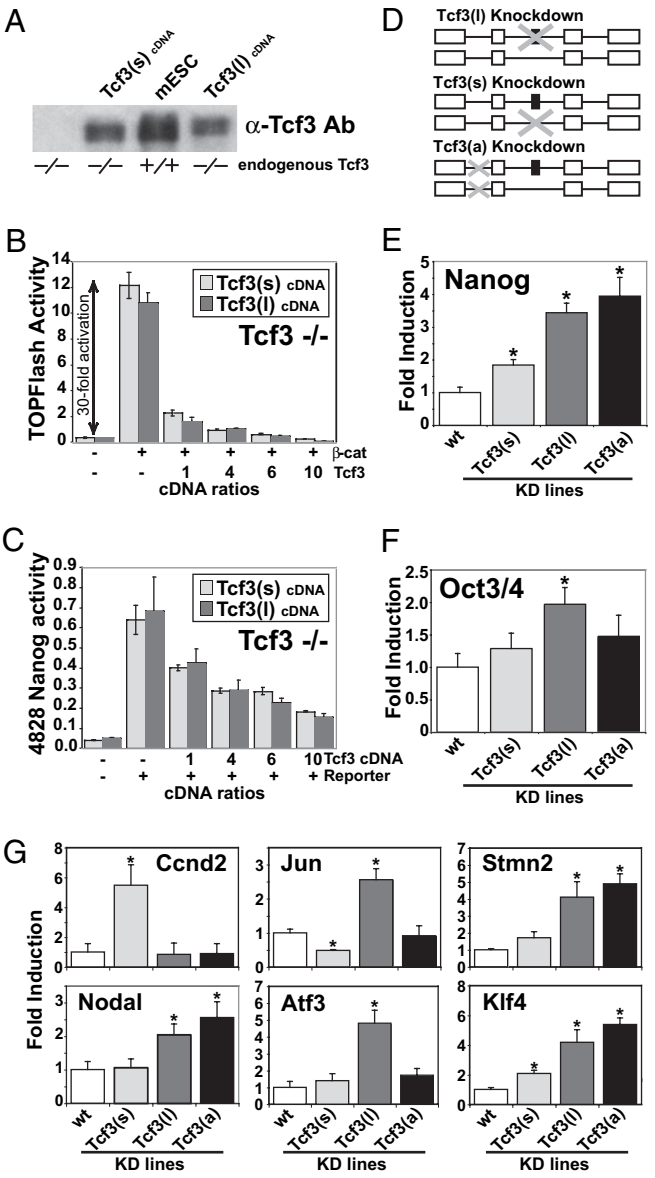

Fig. 6. Selective Tcf3 isoform expression and transcriptional activities in ES cells. (*A*) Tcf3$^{-/-}$ ES cells were transfected with full-length cDNAs for both isoforms expressed under the control of a cytomegalovirus promoter and compared with wild-type ES cells. Expression of Tcf3(s) and Tcf3(l) isoforms resulted in a shift in the detection of Tcf3 protein on a polyacrylamide gel with a Tcf3 antibody common to both isoforms. To assess activity of these isoforms, Tcf3$^{-/-}$ ES cells were transfected with (*B*) a Tcf/Lef transcriptional reporter construct (TOPFlash) and a stable form of β-catenin (β-cat) or (*C*) Nanog promoter reporter, driving luciferase expression, with or without increasing amounts of transfected cDNA for the Tcf3 isoforms. (*D*) To achieve isoform-specific KD, Tcf3 regions unique to or in common with each isoform were targeted using shRNAs directed against either exon junctions [E3–E4 for all or E4–E5 for Tcf3(s)] or exons [exon E4a for Tcf3(l)]. Quantitative expression levels for (*E*) Oct4, (*F*) Nanog, and (*G*) putative Tcf3 transcriptional targets are reported for undifferentiated wild-type (wt) and isoform-specific Tcf3 KD lines cultured with LIF. Fold induction is relative to wild-type ES cells. Values are mean ± SEM of biological triplicates. An asterisk indicates *t* test $P < 0.05$.

DEVELOPMENTAL BIOLOGY

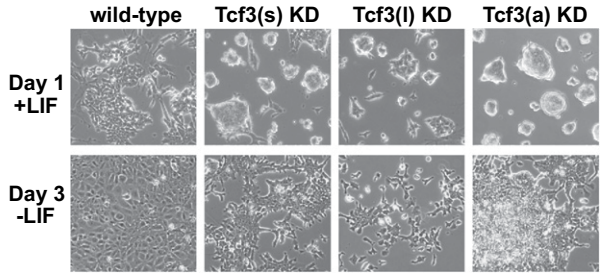

Fig. 7. Altered morphology of ES cells and differentiated EBs with Tcf3 isoform knockdown. Cell morphology, clustering, and differentiation were observed for wild-type, Tcf3(s), Tcf3(l), and Tcf3(a) KD ES cells before and after removal of LIF (–LIF). See Fig. S7 for additional time points.

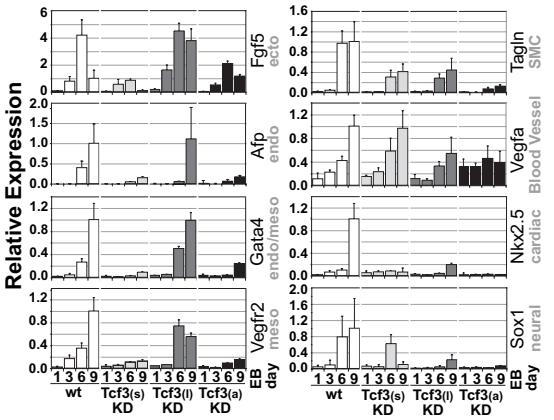

Fig. 8. Distinct patterns of lineage marker expression for all Tcf3 knockdown lines. qPCR was performed on RNA extracted from multiple days of EB differentiation (days 1–9) for wild-type and Tcf3 isoform KD lines. Relative gene expression changes for all cell lines and time points are compared with gene expression of wild-type ESCs at day 9 of differentiation, where gene expression was typically highest. The lineages for which each marker is associated are listed in gray under the gene name for each. Values are mean ± SEM of technical triplicates of pooled EB plates (n = 96 EBs or greater).

addition to confirming that complete loss of Tcf3 allows for LIF-independent ES cell propagation (5, 40), these data suggest that loss of either Tcf3 isoform enhances self-renewal.

Tcf3 Knockdowns Inhibit Distinct Differentiation Pathways. To assess differentiation outcomes with loss of individual Tcf3 isoforms, wild-type and Tcf3 KD lines were derived for EBs at 1, 3, 6, and 9 d of differentiation. qPCR analysis of the time-course data revealed that Tcf3(s) KD had the most profound overall effect on differentiation marker expression, blunting the expression of most lineage (ectoderm, endoderm, mesoderm) and tissue-specific markers examined (Fig. 8 and Fig. S6*B*). This effect was most similar to Tcf3(a) KD, except for the ectodermal markers Fgf5 and Nestin, the endodermal markers FoxA2 and Sox17, and the osteoblast marker Sp7, which were expressed at similar or greater levels than in wild-type ES cells. Tcf3(l) KD largely did not affect expression of lineage markers, but blunted expression of the examined cardiac and neural markers. These data suggest that Tcf3 isoforms have distinct roles in certain, but not all, lineage pathways.

Discussion

Our study uncovered over 100 putative AS events predicted to differ substantially in ES cells and EBs. These events appear to be unique to ES cell differentiation and occur in genes essential to the maintenance of pluripotency, lineage specification, and cell-cycle progression. In many cases, the predicted AS events appear to affect the putative domain structure of the resulting proteins, suggesting such proteins have a functional role in differentiating ES cells. Analysis of Serca2 confirmed removal of targeted miRNA binding sites upon ES cell differentiation in both a candidate approach and an unbiased screen. One targeted miRNA binding site, miR-200b, is restricted to the cardiac progenitor lineage, where the nontargeted isoform, Serca2a, is also most highly expressed in this cell type (38, 39). This observation has broad implications for the ability of AS to regulate protein expression, without the need to regulate target gene or miRNA transcription. In the case of Tcf3, although selective expression of Tcf3 isoforms similarly limited self-renewal of ES cells, a handful of transcriptional targets were regulated by one but not both isoforms. Interestingly, these target genes are associated with a broad range of self-renewal and differentiation pathways, raising the interesting hypothesis that isoform-specific regulation of these targets affects distinct lineage commitment decisions. Further analysis of Tcf3 isoforms will be critical to determine how these isoforms interact differently with target promoters and binding partners, and in which cells they are restricted to during lineage commitment.

Our analysis shows multiple strategies for identifying and validating functionally relevant splicing changes in vitro. Although only a minority of AS events may impact self-renewal or pluripotency, analysis with AltAnalyze should highlight changes in isoform expression most likely to affect protein composition and expression. As whole-genome exon, junction, and RNA sequencing data become the standard method for expression analysis, we anticipate that these methods will become increasingly important to score for AS and predict the potential impact on protein composition and expression.

Materials and Methods

Tissue Isolation and Sample Preparation. Mouse E14 ES cells were grown in monolayer on gelatin-coated culture plates, maintained in medium supplemented with 10% FBS, sodium pyruvate, nonessential amino acids, β-mercaptoethanol, and LIF, and passaged with trypsin. EBs were derived by the hanging-drop method as described (39) in 20% FBS to enrich for the mesoderm lineage. For the microarray analyses, EBs were allowed to differentiate for 14 d. Isolated mouse (FVB/N) tissues consisted of age-matched adult cardiac atria and ventricle, myometrium from adult virgin, 14.5 d (quiescent), 18.5 d (term) gestation, and 6 h postpartum (46), and mouse adult cardiac ventricles from a model of dilated cardiomyopathy (47) and single-transgenic α-myosin heavy chain promoter–tetracycline transactivator controls. Affymetrix CEL files from a published tissue compendium (14) and Nova2 knockout (18) dataset were also included. Total RNA was isolated from cell cultures or snap-frozen tissues, in biological triplicates or greater, using TRIzol extraction and purified with the Qiagen RNA Purification Kit. The AltMouse A microarray design, sample preparation, and hybridization protocol have been described (14). Microarray data were deposited at the NCBI Gene Expression Omnibus (GEO) database (accession no. GSE17021). Additional details on alternative exon analyses are provided in *SI Materials and Methods*.

miRNA Targeting and Expression of Serca2 Isoforms. To assess miRNA repression of SERCA2b, we prepared a DNA construct that transcribes SERCA2b and GFP in a single transcript. This construct was derived by substituting the CMV promoter of the Clontech pDsRed-N1 vector with the PGK promoter driving the expression of eGFP and replacing the DsRed sequence with human SERCA2b full-length cDNA (BC035588). The SERCA2b construct was transfected into HEK-293 cells (300 ng per well) in the presence of siRNAs for phospholamban (negative control), SERCA2 (positive control), or mimics for each miRNA of the Pre-miR miRNA Precursor Library–Human V2 (Ambion) at a concentration of 10 nM. Transfections were performed in duplicates, and fluorescence was quantified using an automated fluorescent microscope (InCell Analyzer 1000; Amersham Biosciences) and a custom algorithm (48). miRNA binding sites identified in the screen were bioinformatically predicted using annotations from AltAnalyze, sequence alignment, and RNA22 (http://cbcsrv.watson.ibm.com/rna22_download_content.html). To assess miRNA binding-site targeting of miRNA-200b, an oligonucleotide duplex for the predicted Serca2b 3′UTR miRNA-200b binding site was obtained from Integrated DNA Technologies and cloned into the pMIR-reporter vector

(Ambion) immediately downstream of the luciferase protein. The miRNA-200b pMIR-reporter was transfected into HL-1 cells (0.4 μg) with Lipofectamine 2000 (Invitrogen) with or without miR-200b, -429, or -1 mimics (50 pmoles) from Dharmacon, and luciferase activity was measured after a 24-h incubation. Repression of the pMIR-reporters was measured by normalizing luciferase to Renilla activity.

Isoform-Specific Expression/shRNA in Mouse ES Cells. Stable isoform-specific KD of Tcf3 alternatively spliced isoforms were obtained in E14 ES cells with sequence-specific shRNAs, delivered by lentiviral infection. Three 19-mer shRNAs were designed to target Tcf3(l) (GGATGGTGCCTCCCACATT), Tcf3(s) (CCAGCACACTTGTCCAACA), and a constitutive region of Tcf3 (GCACCTACCTACAGATGAA) by selecting overlapping predictions from the program PSICOLIGOMAKER 1.5 (http://web.mit.edu/jacks-lab/protocols/pSico.html) and the Broad Institute mouse hairpin library (http://www.broadinstitute.org/rnai/trc/lib). The pSicoR-Ef1α-mCh-Puro lentiviral construct was created by replacing the CMV promoter of pSicoR-mCherry with Ef1α and adding a T2A-puromycin-resistance gene cassette following mCherry (49). Isoform-specific shRNAs were ligated into the pSicoR-Ef1α-mCh-puro construct and cotransfected with the viral packaging plasmids pMDLgpRRE, pRSV_Rev, and pVSV-G (Clontech) into HEK293 cells with FuGENE6 (Roche) as described (49). Harvested supernatant from viral-producing HEK293 cells was filtered through a 0.45-μm filter, and 100 μL incubated with 200,000 ES cells on rotator for 3 h. Cells were plated onto gelatinized tissue-culture plates, grown under feeder-free conditions in the presence of LIF, and selected for puromycin-resistant colonies for at least 5 d. Clonal populations of mCherry-expressing ES cells were screened with isoform-specific qPCR primers to select for clones with optimal isoform-specific KD. cDNAs for both the two Tcf3 isoforms were also expressed in Tcf3$^{-/-}$ ES cells on 129/Sv background (GS1) by electroporation or transfection of a linearized pCDNA3-Tcf3(s) (pBM58) (4) or pCDNA3-Tcf3(l) constructs. The pCDNA3-Tcf3(l) construct was obtained by removal of a 670-bp fragment by Kpn1-Pml1 (NEB) digestion of pBM58 and insertion of the corresponding 712-bp ES cell RT-PCR fragment from Tcf3(l).

ACKNOWLEDGMENTS. We thank Gary Howard and Stephen Ordway, Drs. Chuck Sugnet, Jernej Ule, and Kathryn Ivey for data analysis advice, and Jo Dee Fish and Chris Barker from the Gladstone Histology and Genomic Cores. This work was supported by National Institutes of Health Grants GM080223, HG003053, HL66621 (to B.R.C. and N.S.), HL057181, HL092851 (to E.S.), R01CA128571 (to B.J.M.), R37HL059502, R33HL088266 and California Institute for Regenerative Medicine RC1-000132-1 (to M.M.).

1. Yamanaka S (2008) Pluripotency and nuclear reprogramming. *Philos Trans R Soc Lond B Biol Sci* 363:2079–2087.
2. Vallier L, Pedersen RA (2005) Human embryonic stem cells: An in vitro model to study mechanisms controlling pluripotency in early mammalian development. *Stem Cell Rev* 1:119–130.
3. Boyer LA, et al. (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122:947–956.
4. Pereira L, Yi F, Merrill BJ (2006) Repression of Nanog gene transcription by Tcf3 limits embryonic stem cell self-renewal. *Mol Cell Biol* 26:7479–7491.
5. Tam WL, et al. (2008) Tcf3 regulates embryonic stem cell pluripotency and self-renewal by the transcriptional control of multiple lineage pathways. *Stem Cells* 26: 2019–2031.
6. Cole MF, Johnstone SE, Newman JJ, Kagey MH, Young RA (2008) Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. *Genes Dev* 22:746–755.
7. Loh YH, et al. (2006) The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet* 38:431–440.
8. Cooper TA (2005) Alternative splicing regulation impacts heart development. *Cell* 120:1–2.
9. Duursma AM, Kedde M, Schrier M, le Sage C, Agami R (2008) miR-148 targets human DNMT3b protein coding region. *RNA* 14:872–877.
10. Pritsker M, Doniger TT, Kramer LC, Westcot SE, Lemischka IR (2005) Diversification of stem cell molecular repertoire by alternative splicing. *Proc Natl Acad Sci USA* 102: 14290–14295.
11. Yeo GW, et al. (2007) Alternative splicing events identified in human embryonic stem cells and neural progenitors. *PLOS Comput Biol* 3:1951–1967.
12. Kunarso G, Wong KY, Stanton LW, Lipovich L (2008) Detailed characterization of the mouse embryonic stem cell transcriptome reveals novel genes and intergenic splicing associated with pluripotency. *BMC Genomics* 9:155.
13. Salomonis N, et al. (2009) Alternative splicing in the differentiation of human embryonic stem cells into cardiac precursors. *PLOS Comput Biol* 5:e1000553.
14. Sugnet CW, et al. (2006) Unusual intron conservation near tissue-regulated exons found by splicing microarrays. *PLOS Comput Biol* 2:e4.
15. Krishnan K, Salomonis N, Guo S (2008) Identification of Spt5 target genes in zebrafish development reveals its dual activity in vivo. *PLoS ONE* 3:e3621.
16. Chu PJ, Larsen JK, Chen CC, Best PM (2004) Distribution and relative expression levels of calcium channel beta subunits within the chambers of the rat heart. *J Mol Cell Cardiol* 36:423–434.
17. Sato N, et al. (2003) A novel variant of cardiac myosin-binding protein-C that is unable to assemble into sarcomeres is expressed in the aged mouse atrium. *Mol Biol Cell* 14: 3180–3191.
18. Ule J, et al. (2005) Nova regulates brain-specific splicing to shape the synapse. *Nat Genet* 37:844–852.
19. Tyson-Capper AJ (2007) Alternative splicing: An important mechanism for myometrial gene regulation that can be manipulated to target specific genes associated with preterm labour. *BMC Pregnancy Childbirth* 7 (Suppl 1):S13.
20. Curley M, Morrison JJ, Smith TJ (2004) Analysis of Maxi-K alpha subunit splice variants in human myometrium. *Reprod Biol Endocrinol* 2:67.
21. Dabertrand F, Fritz N, Mironneau J, Macrez N, Morel JL (2007) Role of RYR3 splice variants in calcium signaling in mouse nonpregnant and pregnant myometrium. *Am J Physiol Cell Physiol* 293:C848–C854.
22. Biesiadecki BJ, Elder BD, Yu ZB, Jin JP (2002) Cardiac troponin T variants produced by aberrant splicing of multiple exons in animals with high instances of dilated cardiomyopathy. *J Biol Chem* 277:50275–50285.
23. Spring CM, et al. (2005) The catenin p120ctn inhibits Kaiso-mediated transcriptional repression of the beta-catenin/TCF target gene matrilysin. *Exp Cell Res* 305:253–265.
24. Imbalzano AN, Jones SN (2005) Snf5 tumor suppressor couples chromatin remodeling, checkpoint control, and chromosomal stability. *Cancer Cell* 7:294–295.
25. Bachmann M, Hennemann H, Xing PX, Hoffmann I, Möröy T (2004) The oncogenic serine/threonine kinase Pim-1 phosphorylates and inhibits the activity of Cdc25C-associated kinase 1 (C-TAK1): A novel role for Pim-1 at the G2/M cell cycle checkpoint. *J Biol Chem* 279:48319–48328.
26. Dugast-Darzacq C, Grange T, Schreiber-Agus NB (2007) Differential effects of Mxi1-SRalpha and Mxi1-SRbeta in Myc antagonism. *FEBS J* 274:4643–4653.
27. Huang SC, et al. (2005) Mitotic regulation of protein 4.1R involves phosphorylation by cdc2 kinase. *Mol Biol Cell* 16:117–127.
28. Ishitani T, et al. (1999) The TAK1-NLK-MAPK-related pathway antagonizes signalling between beta-catenin and transcription factor TCF. *Nature* 399:798–802.
29. Katoh Y, Katoh M (2007) Comparative genomics on PROM1 gene encoding stem cell marker CD133. *Int J Mol Med* 19:967–970.
30. Wang H, Mo P, Ren S, Yan C (2010) Activating transcription factor 3 activates p53 by preventing E6-associated protein from binding to E6. *J Biol Chem* 285:13201–13210.
31. Rocha S, Martin AM, Meek DW, Perkins ND (2003) p53 represses cyclin D1 transcription through down regulation of Bcl-3 and inducing increased association of the p52 NF-kappaB subunit with histone deacetylase 1. *Mol Cell Biol* 23:4713–4727.
32. Wisdom R, Johnson RS, Moore C (1999) c-Jun regulates cell cycle progression and apoptosis by distinct mechanisms. *EMBO J* 18:188–197.
33. Vallier L, Alexander M, Pedersen RA (2005) Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J Cell Sci* 118: 4495–4509.
34. Dugast-Darzacq C, Pirity M, Blanck JK, Scherl A, Schreiber-Agus N (2004) Mxi1-SRalpha: A novel Mxi1 isoform with enhanced transcriptional repression potential. *Oncogene* 23:8887–8899.
35. Gee SL, et al. (2000) Alternative splicing of protein 4.1R exon 16: Ordered excision of flanking introns ensures proper splice site choice. *Blood* 95:692–699.
36. Greene AL, et al. (2000) Overexpression of SERCA2b in the heart leads to an increase in sarcoplasmic reticulum calcium transport function and increased cardiac contractility. *J Biol Chem* 275:24722–24727.
37. Misquitta CM, Mwanjewe J, Nie L, Grover AK (2002) Sarcoplasmic reticulum Ca(2+) pump mRNA stability in cardiac and smooth muscle: Role of the 3′-untranslated region. *Am J Physiol Cell Physiol* 283:C560–C568.
38. Domian IJ, et al. (2009) Generation of functional ventricular heart muscle from mouse ventricular progenitor cells. *Science* 326:426–429.
39. Ivey KN, et al. (2008) MicroRNA regulation of cell lineages in mouse and human embryonic stem cells. *Cell Stem Cell* 2:219–229.
40. Yi F, Pereira L, Merrill BJ (2008) Tcf3 functions as a steady state limiter of transcriptional programs of mouse embryonic stem cell self renewal. *Stem Cells* 26: 1951–1960.
41. Cavallo RA, et al. (1998) Drosophila Tcf and Groucho interact to repress Wingless signalling activity. *Nature* 395:604–608.
42. Roose J, et al. (1998) The Xenopus Wnt effector XTcf-3 interacts with Groucho-related transcriptional repressors. *Nature* 395:608–612.
43. Veeman MT, Slusarski DC, Kaykas A, Louie SH, Moon RT (2003) Zebrafish prickle, a modulator of noncanonical Wnt/Fz signaling, regulates gastrulation movements. *Curr Biol* 13:680–685.
44. Nishiyama A, et al. (2009) Uncovering early response of gene regulatory networks in ESCs by systematic induction of transcription factors. *Cell Stem Cell* 5:420–433.
45. Chiellini C, et al. (2008) Stathmin-like 2, a developmentally-associated neuronal marker, is expressed and modulated during osteogenesis of human mesenchymal stem cells. *Biochem Biophys Res Commun* 374:64–68.
46. Salomonis N, et al. (2005) Identifying genetic networks underlying myometrial transition to labor. *Genome Biol* 6:R12.
47. Redfern CH, et al. (2000) Conditional expression of a Gi-coupled receptor causes ventricular conduction delay and a lethal cardiomyopathy. *Proc Natl Acad Sci USA* 97: 4826–4831.
48. Bushway PJ, Mercola M (2006) High-throughput screening for modulators of stem cell differentiation. *Methods Enzymol* 414:300–316.
49. Grskovic M, Chaivorapol C, Gaspar-Maia A, Li H, Ramalho-Santos M (2007) Systematic identification of cis-regulatory sequences active in mouse and human embryonic stem cells. *PLoS Genet* 3:e145.