Genes to Cells

Global transcriptomic analysis of murine embryonic stem cell-derived brachyury$^+$ (T) cells

Michael Xavier Doss[1,a], Vilas Wagh[1], Herbert Schulz[2], Meelis Kull[3], Raivo Kolde[3], Kurt Pfannkuche[1], Tobias Nolden[4], Heinz Himmelbauer[4,b], Jaak Vilo[3], Jürgen Hescheler[1] and Agapios Sachinidis[1]*

[1]*Center of Physiology and Pathophysiology, Institute of Neurophysiology, and Center of Molecular Medicine, University of Cologne (CMMC), Robert-Koch Str. 39, 50931 Cologne, Germany*
[2]*Max-Delbrueck-Center for Molecular Medicine – MDC, Robert-Rössle Str. 10, 13092 Berlin, Germany*
[3]*Institute of Computer Science, University of Tartu, Liivi 2, 50409 Tartu, Estonia and Quretec Ltd, Ulikooli 6a, Tartu, Estonia*
[4]*Department of Vertebrate Genomics, Max-Planck-Institute for Molecular Genetics, Ihnestr.73, D-14195 Berlin, Germany*

Brachyury$^+$ mesodermal cell population with purity over 79% was obtained from differentiating brachyury embryonic stem cells (ESC) generated with brachyury promoter driven enhanced green fluorescent protein and puromycin-*N*-acetyltransferase. A comprehensive transcriptomic analysis of brachyury$^+$ cells enriched with puromycin application from 6-day-old embryoid bodies (EBs), 6-day-old control EBs and undifferentiated ESCs led to identification of 1573 uniquely up-regulated and 1549 uniquely down-regulated transcripts in brachyury$^+$ cells. Furthermore, transcripts up-regulated in brachyury$^+$ cells have overrepresented the Gene Ontology annotations (cell differentiation, blood vessel morphogenesis, striated muscle development, placenta development and cell motility) and Kyoto Encyclopedia of Genes and Genomes pathway annotations (mitogen-activated protein kinase signaling and transforming growth factor beta signaling). Transcripts representing Larp2 and Ankrd34b are notably up-regulated in brachyury$^+$ cells. Knockdown of *Larp2* resulted in a significantly down-regulation BMP-2 expression, and knockdown of *Ankrd34b* resulted in alteration of NF-H, PPARγ and PECAM1 expression. The elucidation of transcriptomic signatures of ESCs-derived brachyury$^+$ cells will contribute toward defining the genetic and cellular identities of presumptive mesodermal cells. Furthermore, there is a possible involvement of *Larp2* in the regulation of the late mesodermal marker BMP-2. *Ankrd34b* might be a positive regulator of neurogenesis and a negative regulator of adipogenesis.

Introduction

During early embryonic development, the mesodermal cells become sequentially determined to more precisely defined cell fates and give rise to a diverse array of functional somatic cell types including muscles, heart, vasculature, blood, kidney, gonads, dermis and cartilage. Understanding of the differentiation process by which mesodermal cells are patterned into diverse somatic cells demands an in-depth analysis of the endogenous signaling cascades and transcription factor networks (Willems & Leyns 2008). Till now, this fascinating puzzle is not yet solved, because isolation of pure mesodermal cells from early-stage embryos in sufficient quantities was not feasible because of technical limitations, and hence the functional genomics of the mesodermal cells has not yet been successfully explored. This has caused a hindrance in our understanding on the patterning of mesoderm into its specific lineages and an impedance to the embryonic/induced pluripotent stem cell-based cell replacement therapy to treat several degenerative diseases. Embryonic stem cells (ESCs) have been proven to be a promising tool to study the early

Communicated by: Moshe Yaniv
**Correspondence*: a.sachinidis@uni-koeln.de
[a]*Present address*: Stem Cell Center, Masonic Medical Research Laboratory, 2150 Bleecker Street, Utica, NY, 13501, USA.
[b]*Present address*: Centre for Genomic Regulation (CRG), UPF, C/Dr. Aiguader 88, E-08003 Barcelona, Spain.

DOI: 10.1111/j.1365-2443.2010.01390.x

embryonic development in a more physiological context *in vitro* because of their versatility over other conventional systems (Doss *et al.* 2004; Gadue *et al.* 2005; Winkler *et al.* 2005).

In mouse embryo, mesoderm is generated from the epiblast through the process of gastrulation that is initiated at approximately day 6.5 of gestation (Tam & Behringer 1997; Baron 2005). Brachyury (also known as T) is considered to be one of the best markers of early pan-mesoderm and is used widely in the field of developmental biology to track the development of this germ layer (Kispert & Hermann 1993; Lacaud *et al.* 2004). It is expressed in all nascent mesoderm and down-regulated as these cells undergo patterning and specification into the derivative tissues including skeletal muscle, cardiac muscle and connective tissues in addition to blood and endothelium (Herrmann 1991; Kispert & Herrmann 1994).

To circumvent the practical difficulties in isolating *in vivo* brachyury-expressing cells in sufficient quantities during embryonic development, we used ESC model as an alternative approach. We established an ESC-derived transgenic brachyury cell line expressing bicistronically both puromycin-*N*-acetyltransferase and enhanced green fluorescent protein (EGFP) under the control of brachyury promoter with the help of an IRES element. We further identified the transcriptomic signatures of ESC-derived brachyury$^+$ cells. Finally, we investigated the functional relevance of two transcripts of unknown functions (TUFs), *La ribonucleoprotein domain family*, *member 2* (*Larp2*) and *ankyrin repeat domain 34B* (*Ankrd34b*) in the differentiating embryoid bodies (EBs) by siRNA-mediated knockdown approaches.

Results

Isolation of brachyury$^+$ cells from the transgenic brachyury ESC line

The 4.5-kb brachyury promoter was isolated from the BAC clone containing brachyury gene using the recombineering approach as outlined in the scheme in the Fig. S1. It was then subcloned in the ESC targeting construct pPuro IRES2 EGFP to get pTbra^P Puro IRES2 EGFP. The brachyury ESC line was generated with the linearized pTbra^P Puro IRES2 EGFP construct by stable transfection. EBs, when generated from the brachyury ESC clone applying the hanging drop protocol and analyzed at regular time intervals, showed the transient expression of brachyury transcripts as shown by RT-PCR (see Fig. 1a). The ESCs and EBs from this transgenic clone behaved faithfully with the wild-type ES and EBs in terms of the brachyury expression and other germ layer-specific transcripts expression (data not shown). The brachyury expression is distributed uniformly throughout the EBs as evident from the EGFP expression in every EB as analyzed using confocal microscopy (see Fig. 1b). Fluorescence-activated cell sorting (FACS) analysis of the EBs showed the same pattern of EGFP expression (Fig. 1c) as that of brachyury mRNA expression (Fig. 1a). Application of puromycin with the experimental protocol outlined in Fig. 1d led to the enrichment of brachyury$^+$ cells in 6-day-old EBs up to 79% (hereafter called 'brachyury$^+$ cells'), compared to 1.4% in the puromycin-untreated control 6-day-old EBs from the same preparation (hereafter called 'control EBs') as analyzed by FACS (see Fig. 1e). Under the same treatment regime, the wild-type EBs did not survive in puromycin selection implying the remaining 21% of the cells in the puromycin-treated EBs are brachyury positive at the time of puromycin selection which later down-regulated brachyury and hence EGFP expression at the end of puromycin selection protocol because of ongoing differentiation process. This finding is in association with the *in vivo* findings that the brachyury expression is very transient, and these cells undergo differentiation very dynamically. Furthermore, the enrichment of brachyury$^+$ cells was confirmed by immunohistochemistry for the coexpression of EGFP and brachyury (see Fig. 1f).

Differentially expressed transcriptome in brachyury$^+$ cells

Transcriptomic profiling of the undifferentiated transgenic brachyury ESCs (hereafter called 'ESCs'), brachyury$^+$ cells and control EBs was carried out from three biologically independent set of experiments using the Affymetrix mouse MG 430 v2.0 3′ oligonucleotide microarrays (Affymetrix UK Ltd, High Wycombe, UK). Raw expression results were Robust Multi-array Average (RMA)-normalized (Irizarry *et al.* 2003). A parametric ANOVA (*F*-tests) and three pairwise comparisons using the least significant differences (LSD) *post hoc* test have been carried out on the transcriptome results obtained from ESCs, control EBs and brachyury$^+$ cells. The false discovery rate (FDR) of the *F*-test-set has been calculated using the Benjamini–Hochberg procedure (Hochberg & Benjamini 1990). An ANOVA FDR-value $<10^{-3}$ has been used to identify and restrict the number of 10 090 probe sets transcripts that were differentially expressed among the

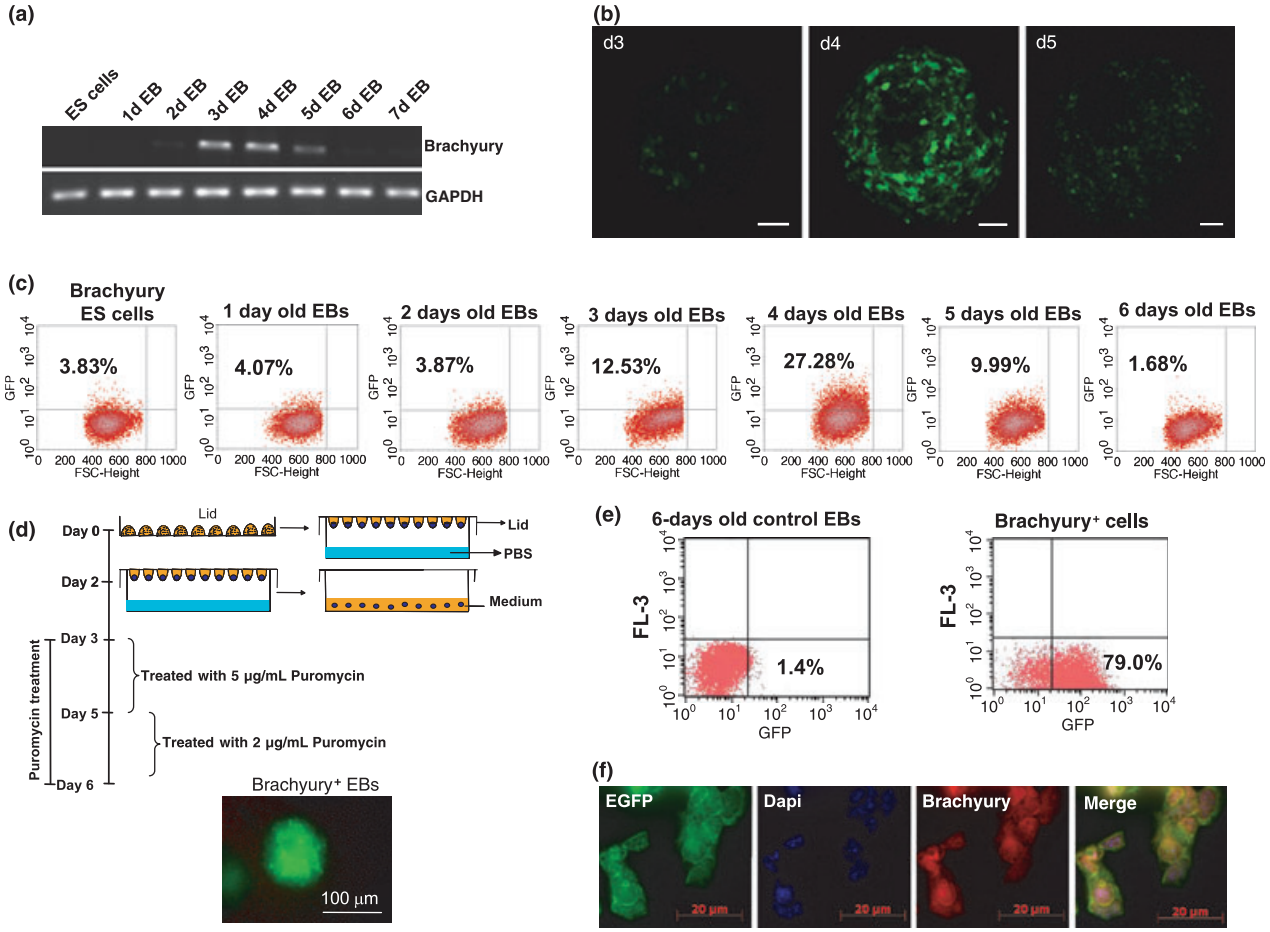
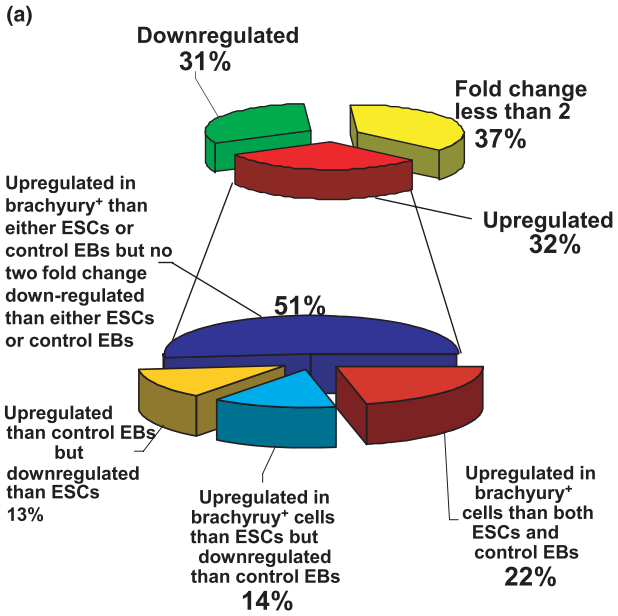
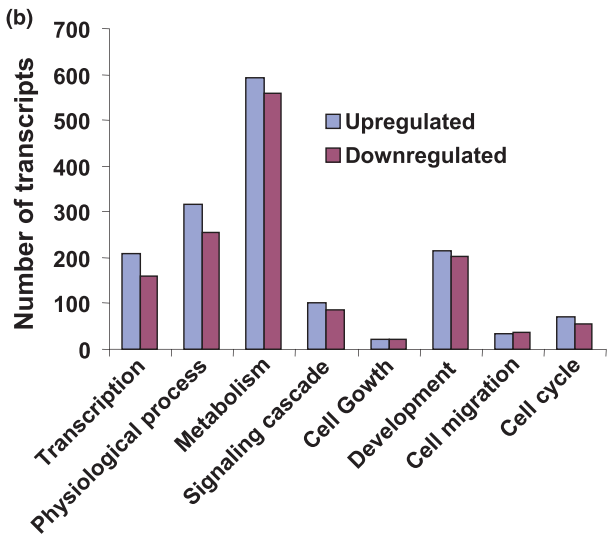


Figure 1 Isolation of brachyury$^+$ cells from transgenic brachyury embryonic stem cells (ESCs). (a) Semi-quantitative RT-PCR analysis of the brachyury transcript expression in the ESCs and the embryoid bodies (EBs) from brachyury ESCs at daily intervals (for primers and conditions see Table S5). (b) Confocal fluorescence microscopy of EGFP-labeled brachyury$^+$ cells within the EBs at day 3 to day 5 of differentiation (bar scale 50 µm). (c) Fluorescence-activated cell sorting (FACS) analysis of the EGFP-labeled brachyury$^+$ cells after dissociation of undifferentiated brachyury ESCs and differentiated EBs at daily intervals. Only the endogenous EGFP$^+$ cells were scored. For precise calculation of percentage of EGFP$^+$ cells, density plot was used. (d) Scheme of the experimental protocol used for the enrichment of brachyury$^+$ cells. (e) FACS analysis of enriched brachyury$^+$ cells in comparison with the day 6 puromycin untreated control EBs. For precise calculation of percentage of EGFP$^+$ cells, dot plot with FL3 channel was used. No staining for FL3 channel was carried out. (f) Immunohistochemical staining of the enriched brachyury$^+$ cells to demonstrate the co-expression of their brachyury expression and EGFP. EGFP, enhanced green fluorescent protein.

conditions day-0 undifferentiated ESCs, day-6 control EBs and day-6 brachyury$^+$ cells. Further filtering of the transcripts with LSD *post hoc* $P < 0.01$ (in each of conditions – in brachyury$^+$ cells vs. ESCs as well as in brachyury$^+$ cells vs. day 6 control EBs), a fold change cutoff of 2 and redundancies removal led to 3022 unique transcripts.

We verified the Affymetrix data by examining the expression levels of four randomly chosen representative genes namely *Nanog*, *bone morphogenetic protein 2* (*Bmp2*), *Sox17 and neurofilament H* (*Nefh*) by applying the quantitative real-time PCR (qPCR) method. As indicated in Fig. S2, results from the Affymetrix analyses clearly correspond to the results obtained from qPCR. Among them, 1573 transcripts were found to be up-regulated in brachyury$^+$ cells in comparison with at least one of the two control population, brachyury$^+$ ESCs, control EBs or both, and 1549 transcripts were down-regulated compared to both control populations as summarized in





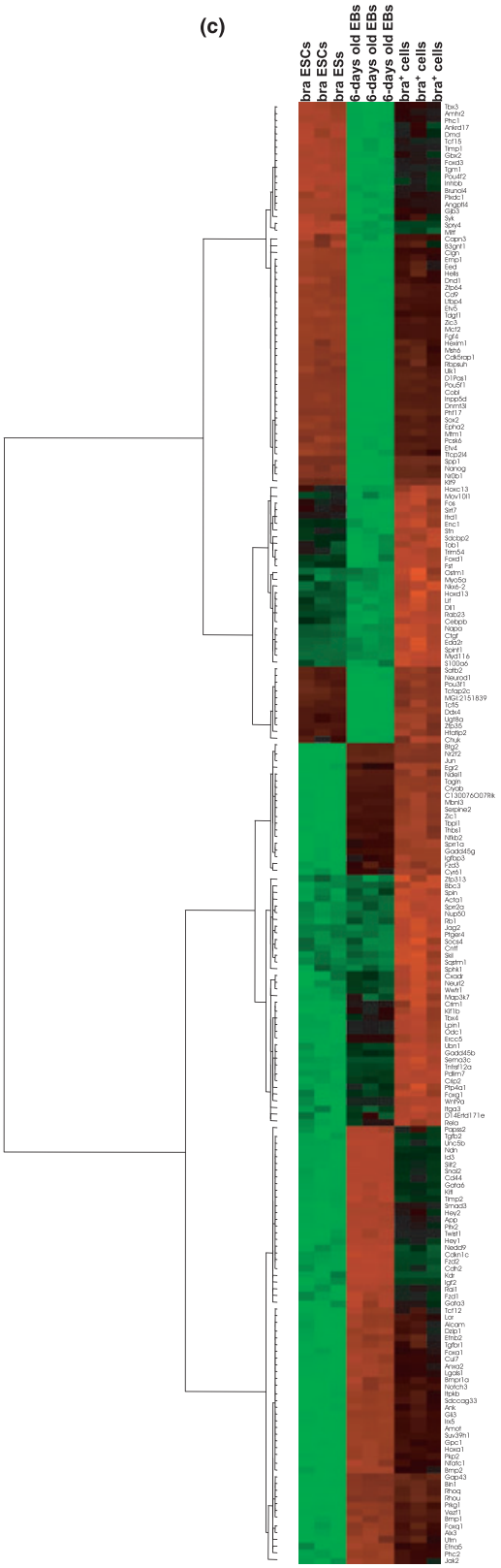


Fig. 2a. To be more specific, the transcripts up-regulated in brachyury$^+$ cells were further categorized on their differential expression level relative to either of the control populations. There are 341 transcripts up-regulated in brachyury$^+$ cells in comparison with both ESCs and control EBs, 215 transcripts up-regulated in comparison with ESCs but down-regulated in comparison with control EBs and 203 transcripts up-regulated in comparison with control EBs but down-regulated to ESCs. The remaining 814 transcripts of the brachyury$^+$ cells show up-regulation in either one of the ESCs or control population but do not show a twofold down-regulation in either of these, as summarized in Fig. 2a. The basic gene ontology's (GO) overrepresented by up- and down-regulated transcripts in brachyury$^+$ cells are shown in Fig. 2b. These transcripts were then further analyzed applying the Database for Annotation, Visualization and Integrated Discovery (DAVID) tool (Dennis *et al.* 2003). As demonstrated, over 400 transcripts were identified to be involved in transcriptional processes (see Fig. 2b). Interestingly, over 400 transcripts seem to be involved in the regulation of the developmental processes. Among them, 216 transcripts were identified to be up-regulated in brachyury$^+$ cells. Hierarchical clustering of these 216 up-regulated genes associated with development is shown in Fig. 2c.

Analysis of the genes differentially regulated in brachyury$^+$ cells

The transcripts up- or down-regulated in brachyury$^+$ cells in comparison with at least one of brachyury ESCs and control EBs are shown in Fig. S3a,b, respectively, using heatmap and dendrogram views of hierarchical clustering results. These transcripts were further analyzed, categorized and listed as follows. The functional annotations overrepresented by the transcripts up-regulated in brachyury$^+$ cells are listed in Table 1. Hierarchical clustering of transcripts down-regulated in brachyury$^+$ cells in comparison with at least one of ESCs and control 6-day-old control EBs is shown in Fig. S3b. Interestingly, GOs containing genes regulating apoptosis, development, cell differentiation, cell death, transcription and blood vessel morphogenesis are highly enriched. The functional annotations overrepresented by the transcripts down-regulated in brachyury$^+$ cells are listed in the Table 2. Similar to the up-regulated genes, several down-regulated genes have been identified to be involved in developmental processes such as Bmp4, Hand1, Hand2, Dkk1, Fgf15, Gata4, Tbx2 and Gli2. The transcripts up-regulated in brachyury$^+$ cells in comparison with both ESCs and control EBs were categorized with their functional role as transcription factors and the transcripts involved in development and are listed in Table S1, respectively. The development-related transcription factors and signal transducers up-regulated in brachyury$^+$ cells compared to ESCs but down-regulated in comparison with control EBs are listed in Table S2. The development-related transcription factors and signal transducers up-regulated in brachyury$^+$ cells when compared to 6-day-old control EBs but down-regulated in comparison with brachyury ESCs are listed in Table S3. The transcripts of the brachyury$^+$ cells that show up-regulation in comparison with one of the ESCs or control population but in turn do not show a twofold down-regulation among either of the control populations (i.e. ESCs or control EBs) are further categorized and listed in Table S4 on the basis of their role in development and as transcription factors.

Figure 2 Analysis of the differentially expressed genes in brachyury+ cells when compared to undifferentiated and 6-day-old embryoid bodies (EBs) using DAVID analysis. (a) Pie diagram showing the distribution of genes with more than twofold up- or down-regulation. A total of 3022 unique transcripts, among them, 1573 transcripts are up-regulated in brachyury$^+$ cells in comparison with at least one of the two control population, brachyury$^+$ embryonic stem cells (ESCs), control EBs or both and 1549 transcripts are down-regulated compared to both control populations (b) DAVID was used to obtain gene ontology biological processes annotations. The bar chart shows the number of genes in the categories adhesion, cell cycle, cell death, cell–cell signaling, cellular metabolism, development, stress response, signal transduction, transcription and transport separately for up-regulated and down-regulated transcripts. (c) Hierarchical clustering of probe sets identified as up-regulated and associated with 'developmental processes' in brachyury$^+$ at least twofold higher compared to either ESCs, control EBs or both. Each transcript is represented by a single row of colored boxes; each array is represented by a single column. Rectangles corresponding to intermediately expressed transcripts are colored black, up-regulated transcripts are indicated with red of increasing intensity, and weakly expressed transcripts with green of increasing intensity.

Table 1 Functional annotations overrepresented by transcripts upregulated in brachyury^{+} cells in comparison with at least one of either embryonic stem cells or control embryoid bodies

Category	Term	Count	*P*-value
SP_PIR_KEYWORDS	transcription	133	1.80E-19
SP_PIR_KEYWORDS	zinc	157	2.30E-19
SP_PIR_KEYWORDS	alternative splicing	187	2.70E-18
SP_PIR_KEYWORDS	zinc-finger	132	3.10E-18
GOTERM_MF_ALL	protein binding	395	4.20E-18
SP_PIR_KEYWORDS	nucleotide-binding	134	5.10E-16
GOTERM_MF_ALL	nucleic acid binding	315	1.40E-15
GOTERM_MF_ALL	DNA binding	206	2.20E-13
GOTERM_BP_ALL	cell differentiation	119	6.20E-13
GOTERM_BP_ALL	development	216	8.50E-13
GOTERM_MF_ALL	transcription regulator activity	140	5.70E-12
GOTERM_BP_ALL	transcription, DNA-dependent	209	8.30E-11
GOTERM_BP_ALL	transcription	218	2.00E-10
SP_PIR_KEYWORDS	developmental protein	61	5.20E-10
GOTERM_BP_ALL	cell death	77	7.00E-10
GOTERM_BP_ALL	regulation of transcription	210	8.90E-10
GOTERM_BP_ALL	apoptosis	70	1.30E-08
GOTERM_BP_ALL	organ development	106	1.60E-07
SP_PIR_KEYWORDS	differentiation	32	1.20E-06
GOTERM_BP_ALL	intracellular signaling cascade	101	2.10E-06
GOTERM_MF_ALL	sequence-specific DNA binding	49	2.60E-06
GOTERM_BP_ALL	morphogenesis	98	5.20E-06
GOTERM_BP_ALL	cell cycle arrest	13	1.10E-05
GOTERM_BP_ALL	neuron differentiation	36	1.30E-05
GOTERM_BP_ALL	cell cycle	71	1.90E-05
GOTERM_BP_ALL	muscle development	21	2.30E-05
GOTERM_BP_ALL	organ morphogenesis	58	3.80E-05
GOTERM_BP_ALL	nervous system development	58	5.70E-05
GOTERM_BP_ALL	regulation of cell differentiation	24	6.00E-05
GOTERM_MF_ALL	protein kinase activity	67	7.30E-05
SP_PIR_KEYWORDS	transmembrane	198	7.60E-05
INTERPRO_NAME	Ras GTPase	27	9.00E-05
GOTERM_MF_ALL	kinase activity	93	9.50E-05
SP_PIR_KEYWORDS	rna-binding	37	1.70E-04
INTERPRO_NAME	Small GTP-binding protein domain	26	1.70E-04
GOTERM_MF_ALL	protein serine/threonine kinase activity	55	1.80E-04
INTERPRO_NAME	Ras small GTPase, Rab type	24	1.80E-04
GOTERM_BP_ALL	cell development	38	4.00E-04
GOTERM_BP_ALL	regulation of development	26	5.90E-04
INTERPRO_NAME	Zinc finger, C2H2-type	73	6.30E-04
GOTERM_BP_ALL	regulation of cell proliferation	32	6.90E-04
GOTERM_BP_ALL	cell motility	34	7.40E-04
GOTERM_BP_ALL	blood vessel morphogenesis	21	8.60E-04
GOTERM_BP_ALL	neuron development	26	9.60E-04
INTERPRO_NAME	Basic leucine zipper	7	1.10E-03
GOTERM_BP_ALL	vasculature development	23	1.20E-03
INTERPRO_NAME	Ras small GTPase, Ras type	20	1.30E-03
GOTERM_BP_ALL	ectoderm development	14	1.50E-03
SP_PIR_KEYWORDS	cell adhesion	27	1.60E-03
SP_PIR_KEYWORDS	cytoskeleton	26	1.70E-03

Table 1 (*Continued*)

Category	Term	Count	*P*-value
GOTERM_BP_ALL	tissue development	27	1.90E-03
GOTERM_BP_ALL	blood vessel development	22	2.20E-03
GOTERM_BP_ALL	epidermis development	13	2.20E-03
SP_PIR_KEYWORDS	leucine zipper	6	2.40E-03
GOTERM_BP_ALL	axon guidance	13	2.50E-03
GOTERM_BP_ALL	cell proliferation	41	2.60E-03
GOTERM_BP_ALL	positive regulation of transcription	27	2.70E-03
GOTERM_BP_ALL	embryonic development	33	4.80E-03
GOTERM_BP_ALL	cell fate commitment	14	5.70E-03
GOTERM_BP_ALL	striated muscle development	11	9.50E-03
INTERPRO_NAME	Zinc finger, RING-type	32	1.10E-02
INTERPRO_NAME	Basic helix-loop-helix dimerization region bHLH	16	1.10E-02
GOTERM_BP_ALL	interphase of mitotic cell cycle	9	1.10E-02
KEGG_PATHWAY	MAPK SIGNALING PATHWAY	34	1.20E-02
GOTERM_BP_ALL	axonogenesis	17	1.20E-02
GOTERM_BP_ALL	muscle cell differentiation	7	1.20E-02
GOTERM_BP_ALL	growth	21	1.40E-02
GOTERM_BP_ALL	neuron morphogenesis during differentiation	19	1.40E-02
GOTERM_BP_ALL	regulation of bone remodeling	6	1.40E-02
GOTERM_BP_ALL	sexual reproduction	27	1.50E-02
GOTERM_BP_ALL	negative regulation of programmed cell death	16	1.60E-02
KEGG_PATHWAY	NEURODEGENERATIVE DISORDERS	8	1.70E-02
GOTERM_BP_ALL	negative regulation of development	11	1.80E-02
GOTERM_BP_ALL	positive regulation of apoptosis	18	1.90E-02
SP_PIR_KEYWORDS	mitogen	6	1.90E-02
GOTERM_BP_ALL	skeletal development	17	2.00E-02
GOTERM_BP_ALL	cellular morphogenesis during differentiation	20	2.30E-02
KEGG_PATHWAY	TGF-BETA SIGNALING PATHWAY	14	2.40E-02
GOTERM_MF_ALL	MAP kinase phosphatase activity	4	2.40E-02
GOTERM_BP_ALL	neuron fate commitment	5	2.50E-02
GOTERM_BP_ALL	determination of adult life span	3	2.50E-02
GOTERM_BP_ALL	reproduction	30	3.00E-02
GOTERM_BP_ALL	embryonic morphogenesis	16	3.10E-02
GOTERM_BP_ALL	DNA packaging	23	3.30E-02
GOTERM_BP_ALL	regulation of ossification	5	4.30E-02
GOTERM_BP_ALL	vasculogenesis	5	4.30E-02
GOTERM_BP_ALL	myoblast differentiation	5	4.30E-02
GOTERM_BP_ALL	ER-nuclear signaling pathway	4	4.30E-02
BIOCARTA	p38 MAPK Signaling Pathway	8	4.40E-02
GOTERM_BP_ALL	embryonic limb morphogenesis	8	4.60E-02
GOTERM_BP_ALL	MAPKKK cascade	13	5.30E-02
BIOCARTA	p53 Signaling Pathway	5	5.30E-02
GOTERM_BP_ALL	skeletal muscle fiber development	6	5.60E-02
GOTERM_BP_ALL	neuron migration	7	5.70E-02
GOTERM_BP_ALL	tube development	13	5.90E-02
GOTERM_BP_ALL	epidermis morphogenesis	7	7.50E-02
GOTERM_BP_ALL	placenta development	4	7.70E-02
GOTERM_BP_ALL	fibroblast proliferation	3	7.70E-02
INTERPRO_NAME	POU homeobox	4	8.40E-02

Table 1 (*Continued*)

Category	Term	Count	*P*-value
GOTERM_BP_ALL	enzyme linked receptor protein signaling pathway	22	9.30E-02
KEGG_PATHWAY	CELL CYCLE	14	9.80E-02
BIOCARTA	Melanocyte Development and Pigmentation Pathway	3	9.90E-02

MAPK, mitogen-activated protein kinase.

Table 2 Functional annotations overrepresented by transcripts downregulated in brachyury^{+} cells in comparison with both embryonic stem cells and control embryoid bodies

Category	Term	Count	*P*
GOTERM_BP_ALL	development	203	8.50E-13
GOTERM_MF_ALL	protein binding	340	1.50E-12
GOTERM_CC_ALL	intracellular	586	2.10E-12
GOTERM_BP_ALL	metabolism	559	1.10E-11
GOTERM_CC_ALL	intracellular organelle	503	3.00E-10
GOTERM_CC_ALL	organelle	503	4.00E-10
GOTERM_BP_ALL	organ morphogenesis	68	9.00E-10
GOTERM_BP_ALL	organ development	105	5.20E-09
GOTERM_MF_ALL	metal ion binding	263	7.90E-09
GOTERM_MF_ALL	transferase activity	166	2.80E-08
GOTERM_MF_ALL	binding	740	2.90E-08
GOTERM_BP_ALL	regulation of biological process	283	3.50E-08
GOTERM_BP_ALL	cell motility	41	3.40E-07
GOTERM_BP_ALL	regulation of physiological process	256	7.00E-07
GOTERM_BP_ALL	cellular morphogenesis	51	8.30E-07
GOTERM_BP_ALL	regulation of cellular process	258	8.40E-07
GOTERM_MF_ALL	phosphotransferase activity, alcohol group as acceptor	76	1.10E-06
GOTERM_BP_ALL	regulation of cellular physiological process	246	1.70E-06
GOTERM_BP_ALL	protein modification	138	2.10E-06
GOTERM_MF_ALL	kinase activity	92	2.50E-06
GOTERM_MF_ALL	transferase activity, transferring phosphorus-containing groups	101	4.60E-06
GOTERM_BP_ALL	system development	62	4.90E-06
GOTERM_MF_ALL	cation binding	232	5.10E-06
GOTERM_BP_ALL	enzyme-linked receptor protein signaling pathway	34	5.40E-06
GOTERM_BP_ALL	embryonic development	40	5.90E-06
GOTERM_MF_ALL	adenyl nucleotide binding	117	7.40E-06
GOTERM_MF_ALL	guanyl-nucleotide exchange factor activity	21	1.40E-05
GOTERM_MF_ALL	small GTPase regulator activity	27	1.70E-05
GOTERM_BP_ALL	negative regulation of biological process	76	1.90E-05
GOTERM_MF_ALL	ATP binding	112	2.00E-05
GOTERM_MF_ALL	protein serine/threonine kinase activity	53	4.20E-05
GOTERM_BP_ALL	negative regulation of cellular process	70	4.50E-05
GOTERM_BP_ALL	heart development	19	4.50E-05
UP_SEQ_FEATURE	active site:Proton acceptor	59	6.70E-05
GOTERM_MF_ALL	transition metal ion binding	167	7.20E-05

Table 2 (*Continued*)

Category	Term	Count	*P*
GOTERM_BP_ALL	angiogenesis	20	7.30E-05
GOTERM_BP_ALL	cell proliferation	44	7.90E-05
GOTERM_BP_ALL	biosynthesis	105	9.30E-05
GOTERM_MF_ALL	protein kinase activity	61	1.10E-04
GOTERM_BP_ALL	cell differentiation	84	1.50E-04
GOTERM_MF_ALL	oxidoreductase activity	73	1.50E-04
GOTERM_BP_ALL	regulation of cell proliferation	32	1.80E-04
GOTERM_MF_ALL	enzyme regulator activity	59	2.40E-04
GOTERM_BP_ALL	cell adhesion	58	2.40E-04
GOTERM_BP_ALL	cellular process	836	3.00E-04
UP_SEQ_FEATURE	splice variant	185	3.00E-04
UP_SEQ_FEATURE	binding site:ATP	48	3.00E-04
GOTERM_BP_ALL	blood vessel morphogenesis	21	3.20E-04
GOTERM_BP_ALL	blood vessel development	23	3.30E-04
GOTERM_BP_ALL	neurogenesis	33	3.50E-04
KEGG_PATHWAY	ECM-RECEPTOR INTERACTION	20	3.70E-04
GOTERM_BP_ALL	cell development	36	3.80E-04
GOTERM_MF_ALL	zinc ion binding	135	3.90E-04
GOTERM_BP_ALL	intracellular signaling cascade	85	3.90E-04
GOTERM_BP_ALL	vasculature development	23	4.30E-04
GOTERM_BP_ALL	cellular morphogenesis during differentiation	24	4.40E-04
GOTERM_BP_ALL	embryonic development (sensu Metazoa)	20	4.80E-04
GOTERM_BP_ALL	transmembrane receptor protein tyrosine kinase signaling pathway	20	5.30E-04
GOTERM_BP_ALL	cell growth	17	5.30E-04
UP_SEQ_FEATURE	nucleotide phosphate-binding region:ATP	70	6.10E-04
GOTERM_BP_ALL	regulation of growth	18	8.10E-04
GOTERM_BP_ALL	regulation of cell size	17	8.20E-04
GOTERM_BP_ALL	protein amino acid phosphorylation	58	8.30E-04
GOTERM_BP_ALL	tube development	17	1.00E-03
GOTERM_MF_ALL	DNA binding	144	1.30E-03
GOTERM_MF_ALL	protein-tyrosine kinase activity	35	1.40E-03
GOTERM_BP_ALL	neuron development	24	1.60E-03
KEGG_PATHWAY	FOCAL ADHESION	34	1.70E-03
GOTERM_BP_ALL	skeletal development	19	1.80E-03
GOTERM_BP_ALL	axonogenesis	18	2.50E-03
GOTERM_BP_ALL	neuron morphogenesis during differentiation	20	2.90E-03
GOTERM_BP_ALL	neurite morphogenesis	20	2.90E-03
GOTERM_BP_ALL	endocytosis	20	3.40E-03
GOTERM_MF_ALL	phosphoric monoester hydrolase activity	27	3.50E-03
UP_SEQ_FEATURE	active site:Proton donor	17	3.80E-03
GOTERM_BP_ALL	transcription	163	5.30E-03
GOTERM_BP_ALL	cellular biosynthesis	85	5.70E-03
GOTERM_BP_ALL	positive regulation of biological process	60	7.00E-03
GOTERM_BP_ALL	monosaccharide catabolism	11	7.00E-03
UP_SEQ_FEATURE	mutagenesis site	41	7.40E-03
GOTERM_BP_ALL	organelle organization and biogenesis	79	9.00E-03
GOTERM_BP_ALL	cell cycle	55	9.60E-03
GOTERM_CC_ALL	envelope	38	1.10E-02
GOTERM_BP_ALL	regulation of progression through cell cycle	34	1.20E-02
GOTERM_BP_ALL	regulation of cell cycle	34	1.30E-02

Table 2 (*Continued*)

Category	Term	Count	*P*
GOTERM_BP_ALL	pattern specification	20	1.60E-02
GOTERM_BP_ALL	physiological process	793	2.00E-02
GOTERM_CC_ALL	extracellular space	163	2.70E-02
KEGG_PATHWAY	TIGHT JUNCTION	19	3.10E-02
GOTERM_MF_ALL	insulin-like growth factor binding	5	3.20E-02
GOTERM_BP_ALL	positive regulation of cellular process	48	3.40E-02
GOTERM_CC_ALL	nonmembrane-bound organelle	104	3.90E-02
GOTERM_CC_ALL	intracellular nonmembrane-bound organelle	104	3.90E-02
GOTERM_BP_ALL	regulation of signal transduction	19	4.50E-02
GOTERM_MF_ALL	transferase activity, transferring glycosyl groups	21	4.70E-02
GOTERM_CC_ALL	endomembrane system	28	4.90E-02
UP_SEQ_FEATURE	transmembrane region	211	5.20E-02
GOTERM_BP_ALL	cell division	18	6.00E-02
KEGG_PATHWAY	AXON GUIDANCE	19	6.10E-02
GOTERM_BP_ALL	tissue development	20	6.40E-02
GOTERM_BP_ALL	positive regulation of development	9	6.60E-02
GOTERM_CC_ALL	extracellular region	176	7.40E-02
GOTERM_BP_ALL	mitotic cell cycle	18	8.00E-02
GOTERM_BP_ALL	lung development	6	8.00E-02
GOTERM_BP_ALL	mesoderm development	5	8.60E-02
GOTERM_BP_ALL	circulation	8	8.70E-02
GOTERM_BP_ALL	negative regulation of signal transduction	9	9.70E-02

Promoter analysis of the transcripts up-regulated in brachyury$^+$ cells

To study the specific transcription regulatory mechanisms related to brachyury$^+$ cells, we obtained a list of the genes up-regulated in the BMP-2$^+$ cells and myosin heavy chain α (α-MHC)$^+$ cells as reference sets from our previous published work (Doss *et al.* 2007a,b). We used these reference data sets to eliminate the transcripts common to the late mesodermal/somatic lineages. We carried out binding site analysis in the promoters of the genes up-regulated specifically in the brachyury$^+$ cells (see Experimental procedures for the detailed analysis). When scanned through all the brachyury$^+$ cell-specific transcription factors in TRANSFAC® (version 11.4) (BIOBASE GmbH, Wolfenbüttel, Germany) (Wingender *et al.* 1996) and Jaspar (Sandelin *et al.* 2004) databases, two candidates qualifying most of our criteria of analysis were filtered out (see Fig. 3). Transfac motif M00242, peroxisome proliferator-activated receptor α (PPARα), scored over threshold 9.89 in 15 of 109 (13.8%) promoters of the brachyury$^+$-specific list. Transfac motif M00258, interferon-sensitive response element (ISRE), scored over threshold 10.63 in 23.9% of the promoters in the brachyury$^+$-specific genes. Target promoters of PPARα in the brachyury$^+$-specific list were from genes AC158524.2, ACAA1B, ACTA1, ASPA, CDCP1, GPATCH2, HIST1H1C, MPZL3, PRDM8, RB1, SERPINB9B, SLC23A3, SLC7A6OS, SPRR2B and TBX18. Target promoters of ISRE in the brachyury$^+$-specific list were from genes 5630401D24RIK, B230120H23RIK, CDK9, CXCL10, EPS8L2, F3, GBP1, GBP2, GPATCH2, H2-D1, H2-Q7, H2-Q8, H2-T3, HHIPL2, HIST1H1C, IL1RAP, LATS2, MARS, NEFM, PLK2, PPM1D, RIOK3, SERPINB9B, SLC23A3, SMC5 and ZBTB4.

Functional evaluation of *Larp2* and *Ankrd34b*, two TUFs specifically up-regulated in brachyury$^+$ cells

The selection criteria for the functional analysis of the TUFs for further investigation were as follows: (i) their functions should be unknown so far at least for their role in embryonic developmental context. (ii) TUFs should be up-regulated with high fold change compared to the ESCs and control EBs from the Affymetrix data (see Fig. 4). (iii) TUFs should be developmentally regulated in early embryos (see Fig. 4b). (iv) TUFs should not have any information other than their conserved domain contained in their protein, if any.

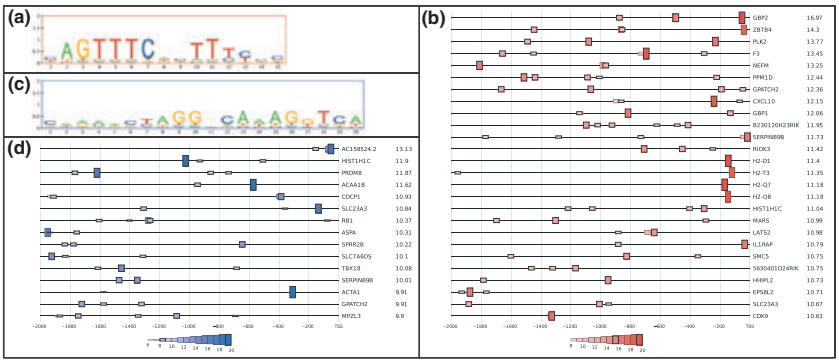


Figure 3 Results of the promoter analysis. (a, c) Sequence logos of PPARA and interferon-sensitive response element (ISRE), respectively. (b, d) The predicted target genes of PPARA and ISRE, respectively, sorted by the average of the three highest scoring hits within the promoter of length 2000 bp. All the hits scoring over 6 in case of PPARA and 8 in case of ISRE are visualized on the promoters.

Accordingly, two TUFs, *La ribonucleoprotein domain family, member 2* (*Larp2*) (also known as *1700108L22Rik*) and *ankyrin repeat domain 34B* (*Ankrd34b*), were chosen after assessing fulfilment of the above criteria (see Fig. 4). As demonstrated in Fig. 4a, both transcripts are specifically up-regulated in brachyury^{+} cells as well as in ES-derived BMP-2^{+} cells which represents late mesodermal cells (Doss *et al.* 2007a) in contrast to functional cardiovascular cell types such as cardiomyocytes (Doss *et al.* 2007b) and contractile smooth muscle cells isolated from ESCs (Potta *et al.* 2009). *Larp2* (http://www.ncbi.nlm.nih.gov/UniGene/clust.cgi?UGID=271837&TAXID=10090&SEARCH=mm.46782) encodes for an RNA-binding protein that is developmentally regulated in early murine embryonic stages (http://www.ncbi.nlm.nih.gov/sites/entrez?db=geo&cmd=search&term=1700108L22Rik). This transcript is up-regulated in preimplantation embryos. The function of this gene has not yet been reported. LA domain codes for a protein, which binds to RNA. Generally, the La protein plays a role in the transcription of RNA polymerase III. It is most probably a transcription termination factor. Normally, it binds to the 3′ termini of virtually all nascent polymerase III transcripts. It is associated with precursor forms of RNA polymerase III transcripts including tRNA and 4.5S, 5S, 7S and 7-2 RNAs (by similarity).

Ankrd34b (UniGene Mm.213623, http://www.ncbi.nlm.nih.gov/UniGene/clust.cgi?ORG=Mm&CID=213623) is normally expressed during embryonic development in spinal cord belonging to the CNS (http://www.ncbi.nlm.nih.gov/geo/gds/profileGraph.cgi?&dataset=sztDAC&dataset=wywCBC$&gmin=0.129683&gmax=2.000224&absc=&gds=2209&idref=1435307_at&annot=Ankrd34b). The protein product of this transcript contains conserved domains for ankyrin-repeat domains. Ankyrin repeats mediate protein–protein interactions thereby modulating several biological processes such as differentiation (Hortsch *et al.* 2009). In this context, it has been demonstrated that interaction of L1-type cell adhesion molecules (CAMs) with ankyrins promotes neurogenesis (Hortsch *et al.* 2009). Ankyrins are also involved in the neurogenesis in zebrafish (Diks *et al.* 2006) and Xenopus neurogenesis (Lahaye *et al.* 2002).

Both transcripts were functionally evaluated by systematic 'loss of function' studies using the siRNA knockdown approach. The relative mRNA expression levels of *Nefh*, *brachyury*, *Bmp2*, *Sox17*, *Nanog*, *Oct4*, *Myh6*, *PPARγ*, *NF-H* (also known as *Nefh*) and *PECAM1* were analyzed using qRT-PCR. *Nefh*, *brachyury*, *Bmp2* and *Sox17* are the canonical germ layer markers. *Nanog* and *Oct4 are* pluripotent ESCs markers. *Myh6*, *PPARγ*, *NF-H* and *PECAM1* are mesodermal-derived somatic cell markers for cardiac, adipogeneic, neurogeneic and endothelial cells, respectively. siRNA-transfected ESCs were monitored for any morphological changes associated with any differentiated cell types.

As shown in Fig. 5a–i, transfection by siRNA targeted against *Larp2* resulted in more than 50% significant knockdown of *Larp2* mRNA expression in undifferentiated ESCs (d0) at 24 h post-transfection. At 48-h post-transfection, the transfected undifferentiated ESCs were used for differentiation applying 'hanging drop' protocol, and then total RNA was isolated at regular intervals from 1 to 10 EBs. As

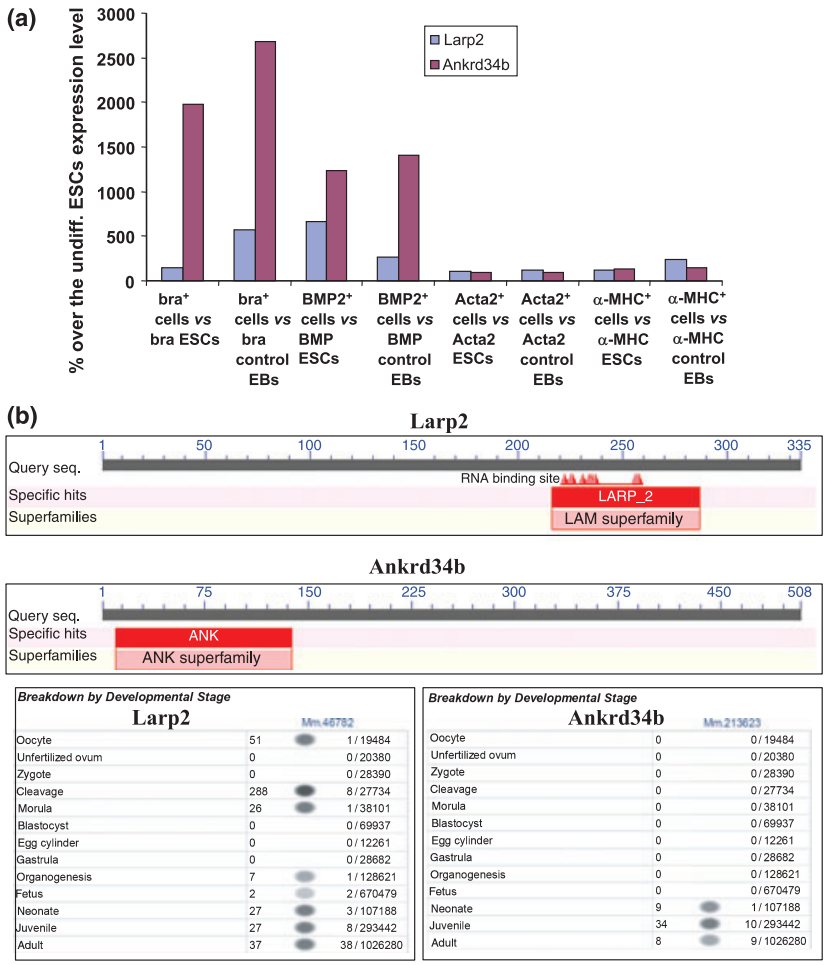


Figure 4 Expression levels of *Larp2* and *Ankrd34b* in embryonic stem cells-derived mesodermal cells and somatic cells as well as during mouse embryo development. (a) Expression levels of *Larp2* and *Ankrd34b* that are specifically up-regulated in the mesodermal cells [brachyury$^+$ and BMP-2$^+$ cells (Doss *et al.* 2007a) when compared to the ES-derived α-MHC$^+$ cardiomyocytes (Doss *et al.* 2007b) and Acta2$^+$ smooth muscle cells (Potta *et al.* 2009). The arrays gene expression level of the undifferentiated ES cell lines (brachyury, BMP-2, Acta2, and the α-MHC ES cell line) was set 100%. (b) Conserved LA domain contained in Larp2 (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi?INPUT_TYPE=precalc&SEQUENCE=94536725&log$=ProtAccnMenu) and ankyrin repeat domain 34B (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi?INPUT_TYPE=precalc&SEQUENCE=172072675) as well as the developmental stage-specific expression of Larp2 (http://www.ncbi.nlm.nih.gov/UniGene/ESTProfileViewer.cgi?uglist=Mm.46782) and *Ankrd34b* (http://www.ncbi.nlm.nih.gov/UniGene/ESTProfileViewer.cgi?uglist=Mm.213623). MHC, myosin heavy chain.

expected, the expression remained still down-regulated in 1–9-day-old EBs (see Fig. 5a–ii) because of siRNA knockdown. The expression of pluripotent and germ layer representative genes was determined using the qRT-PCR. As shown in Fig. 5b, the BMP-2 expression level only at day 3 was significantly down-regulated, compared to vehicle-treated and untreated differentiated ESCs (see Fig. 5b). No significant differences have been observed regarding the expression of NF-H, brachyury and Sox17 (see Fig. 5a) and pluripotent markers Oct4 and Nanog as well as on the expression of the cardiomyocytes markers such as GATA4, Nkx2.5 and α-MHC (data not shown). As shown in Fig. 6a, transfection of the brachyury$^+$ cells with siRNA directed to *Ankrd34b* resulted in a significant knockdown of the *Ankrd34b* first detectable in 2-day-old EBs (80% inhibition of the mRNA expression) and a complete knockdown within the 3-day-old EBs compared to vehicle-treated cells. Normal expression levels like in the

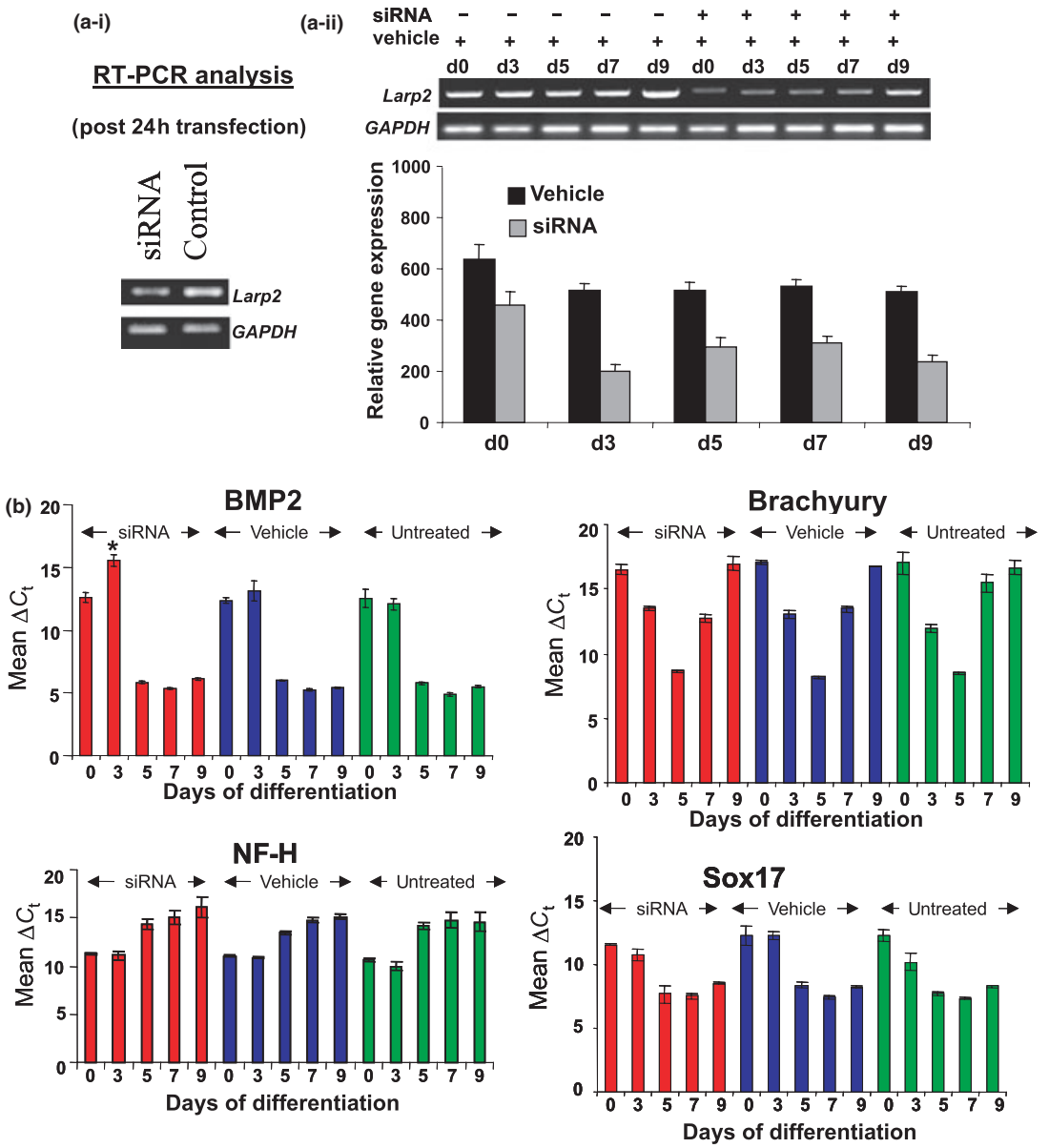


Figure 5 Effect of siRNAs directed to *Larp2* on specific gene expression markers in embryonic stem cells (ESCs). (a) To assess the efficiency of knockdown with siRNA against *Larp2*, transfection of ESCs cultured in culture petri-dishes (approximately 50% confluence) with the appropriate siRNA (100 nM) was carried out. (i) RNA was collected 24 h after transfection and analyzed for the expression level of *Larp2*. (ii) 48 h after transfection (day 0 as per the differentiation protocol), the siRNA-transfected ESCs were used for making 'hanging drop' embryoid bodies (EBs). Upon isolation of total RNA, RT-PCR was carried out using the respective primers shown in Table S5 (see Experimental procedures). As an internal loading control, PCRs were carried out using the GAPDH primers (Table S5). Diagram summarizes densitometric analysis of three independent experiments (mean ± SD, $n = 3$, $\star P < 0.05$ for the siRNA-treated vs. vehicle-treated cells). (b) Effect of siRNA on gene expression of germ layers-specific genes in ESCs and EBs generated by the hanging drop protocol. After RNA isolation, the expression level was determined using qRT-PCR (for more details see Materials). Experiments were carried out in triplicates (mean ± SD, $n = 3$, $\star P < 0.05$ for the siRNA-treated vs. untreated or vehicle-treated cells).

vehicle-treated 4-day-old EBs have been observed in the 4-day-old EBs generated from siRNA-treated ESCs.

As demonstrated in Fig. 6b, the expression of NF-H in EBs generated with siRNA-treated cells was rapidly down-regulated in the 2-day-old EBs,

compared to the high expression level of the NF-H in undifferentiated and in 1-day- to 4-day-old EBs.

Furthermore, siRNA-treated ESCs and EBs showed higher PPARγ expression levels compared to the respective vehicle-treated control ESC and EBs (Fig. 6c). Also, the expression level of PECAM1 in EBs generated from siRNA-treated cells remained significantly higher between days 1 and 2 in comparison with that in vehicle-treated EBs (see Fig. 6d). No differences in mRNA expression levels of the pluripotent marker genes (Oct4 and Nanog), mesodermal marker genes (brachyury and BMP-2), cardiomyocytes and endothelial marker genes between vehicle and siRNA-treated cells could be observed (see Fig. S4).

Discussion

Analysis of the genes differentially regulated in brachyury$^+$ cells

Evidently, GOs associated with differentiation and developmental processes as well as GOs associated with an increased transcriptional activity are highly enriched in genes up-regulated in the brachyury$^+$ cells (Table 1). These results demonstrate the high developmental capacity of brachyury$^+$ cells. Obviously, GOs associated with transcripts regulating apoptosis and cell death are also highly enriched. It is well known that apoptosis is one of the key biological

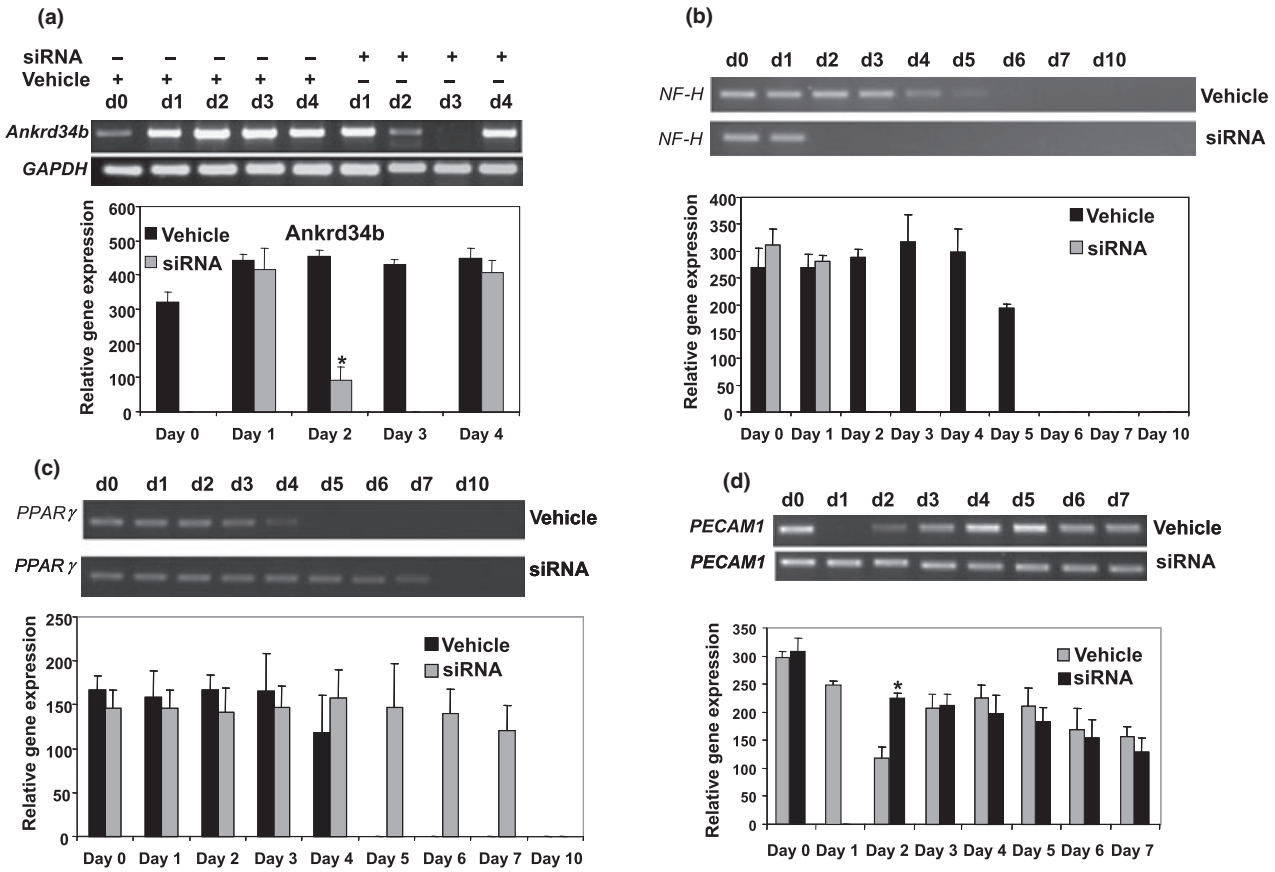


Figure 6 Effect of siRNAs directed to *ankyrin repeat domain 34B* on specific gene expression markers in embryonic stem cells (ESCs). (a) Transfection of ESCs cultured in culture Petri-dishes (approximately 50% confluence) with the appropriate siRNA (100 nM) was carried out for 24 h. Upon isolation of total RNA, RT-PCR was carried out using the respective primers shown in additional data file 9 (see Experimental procedures). As an internal loading control, PCR reactions were carried out using the GAPDH primers (additional data file 9). Diagram summarizes densitometric analysis of three independent experiments (mean ± SD, $n = 3$, $\star P < 0.05$ for the siRNA-treated vs. vehicle-treated cells). (b, c, d) Effect of siRNA on gene expression NFH expression PPARγ and PECAM1, respectively, in ESCs and embryoid bodies generated by the hanging drop protocol. After RNA isolation the expression level was determined by RT-PCR (for more details see Materials). Experiments were carried out in triplicate (mean ± SD, $n = 3$, $\star P < 0.05$ for the siRNA-treated vs. vehicle-treated cells).

processes regulating specific organ development such as cardiac and notochord development in vertebrates (Martinsen 2005; Malikova *et al.* 2007). Interestingly, transcripts associated with mitogen-activated protein kinase (MAPK) SIGNALING Kyoto Encyclopedia of Genes and Genomes (KEGG) PATHWAY, p38 MAPK BIOCARTA pathway and the MAPKKK cascade GO biological process annotation (GO: 0000196) were found to be enriched in brachyury$^+$ cells. The MAPK pathway is speculated to promote inhibition of the self-renewal state of ESCs and is involved in the development of the early mesoderm differentiation. p38MAPK displays a large spectrum of action from neurons to the mesodermal-derived adipocytes [for review see (Binetruy *et al.* 2007)]. Transcripts associated with transforming growth factor beta (TGFβ) SIGNALING PATHWAY were also found to be enriched in brachyury$^+$ cells. TGFβ superfamily plays a key role for the induction of mesoderm and endoderm (Shen 2007). From these results, we might conclude that similarly to the *in vivo* situation similar signaling pathways are identified to be enriched in mesodermal brachyury$^+$ cells. Interestingly, Tbx6, Gm397, Lif, Follistatin and Wnt9a are all up-regulated in brachyury$^+$ cells in comparison with both ESCs and control EBs (Table S1). In this context, recently, it has been shown that Wnt signaling and Tbx6, a T-box transcription factor, regulate the presomitic mesoderm formation (Wittler *et al.* 2007; Dunty *et al.* 2008). Follistatin is a BMP-2 antagonist that is required for formation of the neural plate and dorsal mesoderm (Khokha *et al.* 2005). More recently, Gm397 has been identified to be a predominant transcript in ESCs and renamed as Zscan4c (Falco *et al.* 2007).

Promoter analysis of the transcripts upregulated in brachyury$^+$ cells

Transfac motifs M00242 peroxisome proliferator-activated receptor α (PPARα) and M00258 ISRE showed several fold enrichment in the brachyury$^+$-specific list. Of these, ISRE is a strong inducer of apoptosis acting through expression of caspase 8 (Casciano *et al.* 2004). This finding is in accordance with our observation that the GOs containing genes regulating apoptosis and cell death are also highly enriched (see Table 1). PPARα belongs to the steroid hormone receptor superfamily and is involved in transcriptional regulation of genes regulating lipid metabolism, cell proliferation and cell differentiation. Recently, it has been shown that PPARα is involved also in cardiomyocytes differentiation of murine ESCs (Ding *et al.* 2007; Sharifpanah *et al.* 2008).

Functional evaluation of *Larp2* and *Ankrd34b*, two TUFs specifically up-regulated in brachyury$^+$ cells

As *Larp2* is up-regulated only in the mesodermal brachyury and BMP-2 cell lineages (see Fig. 4a) and also temporally and spatially regulated during mouse embryonic development (see Fig. 4b), we assumed that the rest of the protein motif(s) apart from the La domain might play a role in the early developmental stage. Together, our findings from siRNA-mediated knockdown of *Larp2* suggest that the Larp2 is positively involved in the regulation of BMP-2 expression. No differences in the morphology of the EBs were observed (data not shown).

Interestingly, knockdown of *Ankr34b* led to significant changes in the expression patterns of NF-H, PPARγ and PECAM1 on differentiation. NF-H is a pan-ectodermal marker (Petzold *et al.* 2007). PPARγ is an adipogenic marker (Szabo *et al.* 2008) and PECAM1, a haematopoietic and endothelial progenitor cell-specific marker (Ross *et al.* 2008; Mariappan *et al.* 2009).

To date, the role of Ankyrins in mesoderm-derived adipocytes and endothelial progenitor cells has not yet been reported. These results demonstrate a negative regulatory effect of the *Ankrd34b* gene on the expression of PPARγ and PECAM1 and a positive regulatory role on the expression of NF-H. These results suggest that *Ankrd34b* might be a positive regulator of the ectoderm-dependent neurogenesis and negative regulator of the mesoderm-dependent adipogenesis and hematopoiesis. The present findings support also the observation that Ankyrins are involved in neurogenesis as observed as previously described (Lahaye *et al.* 2002; Diks *et al.* 2006; Hortsch *et al.* 2009). In conclusion, substantial attempts were made to identify the complete transcriptome of the brachyury$^+$ mesodermal cells and to address the role of two TUFs in terms of their possible role in regulation of key genes involved in germ layer and somatic cell development. The transcriptomic analysis of the brachyury$^+$ cells reported here defines the cellular and genetic identities of the mesodermal cells for the first time. Moreover, this will in turn significantly contribute to building lineage-specific transcriptomic atlas, which will be of an immense application to the developmental biologists and also for promoting the protocols for the derivation of selective lineages for the future cell replacement therapy. Finally, the

identification of the TUFs by this approach in combination with 'loss of function' tools offers a new approach in elucidating the role of the TUFs during the differentiation processes.

Experimental procedures

Materials

All cell culture media, supplements, DNAse I and Superscript™ II Reverse Transcriptase were purchased from Invitrogen (Karlsruhe, Germany). CGR8 ESCs were obtained from European Collection of Cell Cultures (ECACC), catalogue 95011018 (Genoa, Italy). The Brachyury BAC clone (RP23-376B1) and the ThermoScript™ Reverse Transcriptase were obtained from Invitrogen. The BAC subcloning kit was obtained from Gene Bridges, Dresden, Germany. Oligos for subcloning of the brachyury promoter were purchased from Metabion (Munich, Germany). Primary antibodies anti-*brachyury* (catalogue ab20680) were from Abcam (Cambridge, UK). Anti-mouse IgG1-AlexaFluor488 and anti-rabbit-Ig-AlexaFluor647 were purchased from Molecular Probes (Invitrogen GmbH, Karlsruhe, Germany). The TaqMan Gene Expression Assays were purchased from Applied Biosystems (Foster City, CA, USA). The QIAGEN's RNeasy Total RNA Isolation kit was obtained from Qiagen GmbH (Hilden, Germany). The REDTaq ReadyMix was purchased by Sigma (Munich, Germany).

ESC Culture and differentiation of EBs

CGR8 ESCs were cultured without feeder cells in Glasgow Minimum Essential Medium supplemented with 10% FBS, 2 mM L-glutamine, 100 units/mL leukemia inhibitory factor and 50 μM β-mercaptoethanol (β-ME) in 0.2% gelatine-coated flasks as previously described (Gissel *et al.* 2005). To induce differentiation, the hanging drop protocol was used as described earlier (Gissel *et al.* 2005) using the differentiation medium containing Iscove's Modified Dulbecco's Medium (IMDM), supplemented with 20% FCS, 1% nonessential amino acids (v/v), 2 mM L-glutamine and 100 μM β-ME). On day 2, the formed multicellular aggregates (EBs) were transferred to suspension. On day 3, the EBs were either treated with puromycin (5 μg/ml) for the first 48 h and then with 2 μg/ml for the subsequent 24 h or left without treatment for 3 days. On day 6, the EGFP-fluorescing EBs were trypsinized and used for either FACS analysis or RNA extraction or immunohistochemistry.

Subcloning of brachyury promoter in the ESC-targeting vector by BAC recombineering

Five-kilobase promoter region upstream of the translation initiation site of the brachyury gene was subcloned from a BAC into the ESC targeting construct pPuroIRES2 EGFP by a recombineering approach as described in Fig. S1. Oligos with the following sequences GCT GAG TCT TCT CTG TCA TTT AAA TTA GGC TCT GTC TCA GTT TGC CAT TCA GAT CTT CAC AGC TTG TCT GTA AGC GGA TG and CCT CCC GCC ACC CTC TCC ACC TTC CAG GAG TCT TGA CTC CCT ACC CAA CAA GAT CTG CTC TCC TGA GTA GGA CAA ATC were used to amplify the linear minimal vector containing ColE1 origin and Ampicillin resistance cassette. The linear amplified product was electroporated into the *Escherichia coli* harboring the BAC and induced to express Red ET proteins. The minimal vector containing the brachyury promoter was positively selected with ampicillin.

Vector constructs and cell line generation

The basic ESC targeting construct pPuro IRES2 EGFP was constructed as described previously (Doss *et al.* 2007a). The minimal vector with brachyury promoter was digested with *Bgl II* to release the brachyury promoter, which is then ligated in *Bgl II*-digested 5′ dephosphorylated pPuroIRES2 EGFP to get pT-bra^{p} PuroIRES2 EGFP. This was then linearized with *Afe*I and electroporated into CGR8 ESC line with 500 μF and 240 V in a Bio-Rad (Munich, Germany) Gene Pulser™. The transfected clones were selected by 400 μg/mL neomycin, and after selection, the clones were maintained with 200 μg/mL neomycin.

Confocal fluorescence microscopy

Brachyury-GFP EBs from day 3 to day 5 were washed twice with prewarmed phosphate buffered saline (PBS) just before analysis using fluorescence microscopy. EBs suspended in PBS were analyzed using a confocal microscope (LSM 510 META) on an inverted stand (Axiovert 200M, Carl-Zeiss, Jena, Germany) using the objective Plan-NEOFLUAR 10× and the appropriate filter set for GFP detection. Images were taken using Zeiss software LSM 5 version 3.5.

Immunohistochemistry

Puromycin-resistant and EGFP-fluorescing brachyury^{+} cells obtained, as outlined in Fig. 1d, were plated on the fibronectin coated cover slips. After 24-h post-plating, fixation was carried out with 99% methanol. Blocking was carried out in 5% bovine serum albumin. Incubations with the primary antibodies anti-*brachyury* (Abcam, catalogue ab20680, polyclonal, 5 μg/ml) and anti-GFP (Millipore, Billerica, MA, USA, monoclonal, 1 : 500) were carried out overnight at 4 °C in 1% bovine serum albumin. After extensive washing, secondary detection was carried out for 2 h at room temperature. Hoechst dye (Hoechst 33324, Molecular Probes, Carlsbad, CA) was used to stain nuclei. Samples were mounted using ProLong Gold mounting medium (Molecular Probes) and observed using a Zeiss Axiovert 200 fluorescence microscope.

Semiquantitative RT-PCR analysis

Total RNA was extracted using RNeasy Mini kit (Qiagen) with on-column DNase I (Qiagen) digestion according to the manufacturer's instructions. Five micrograms of total RNA was reverse transcribed using SuperScript II Reverse transcriptase (Invitrogen) with random primers according to the manufacturer's recommended protocol. PCR amplification was carried out using REDTaq ReadyMix (Sigma) with 0.4 μM each primer. GAPDH was used as an internal control. The following conditions were used: an initial denaturation at 95 °C for 2 min, followed by 22–35 cycles of 30-s denaturation at 95 °C, 30 s annealing at 60 °C and 60 s of elongation at 72 °C. A final extension at 72 °C for 5 min was included. Electrophoretic separation of PCR products was carried out on 2% agarose gels with 0.001% ethidium bromide. The primer sequences are listed in the Table S5.

Quantitative real-time PCR

Validation of the Affymetrix data was carried out by qPCR analysis using ABI Prism 7900HT Sequence Detection System (Applied Biosystems). An amount of 100 ng RNA each from every biological triplicate of brachyury ESCs, control EBs and brachyury$^+$ cells was reverse transcribed using ThermoScript™ Reverse Transcriptase (Invitrogen). Then real-time PCR was carried out in technical triplicates for every sample from three biologically independent experiments using TaqMan Gene Expression Assays from Applied Biosystems. The gene expression assays included for the validation were *Sox17* (Mm00488363_m1), *Bmp2* (Mm01340178_m1), *NF-H* (Mm01191456_m1), *GAPDH* (Mm99999915_g1) and *Nanog* (Mm02019550_s1). The threshold line is set in the exponential phase of the amplification, and the cycle at which the sample reaches this level is called the cycle threshold (C_t). Therefore, the lower the C_t value is the more the gene is expressed. Averaged C_t values of each qRT-PCR from the target gene were normalized with the average C_t values of the housekeeping gene, GAPDH, that ran in the same reaction plate to get ΔC_t value. An increase in ΔC_t value by 1 unit corresponds to a twofold lowering of the expression level. The fold change was calculated by using the formula, fold change = $2^{-(Ct\ gene1\ -\ Ct\ gene2)}$. The resulting fold change is expressed as percentage of the maximum.

Flow cytometry

Single cell suspension was prepared by trypsinization. Cell clumps were removed by passing through cell strainer cap of a round bottom tube from Falcon® (BD, Heidelberg, Germany). Propidium iodide (PI) staining (Sigma) was included to exclude dead cells. Acquisition of 10 000 live (PI negative) cells was made with FACscan (BD Biosciences), and the data analysis was carried out with CellQuest software (BD). The respective wild type EBs on the same day as the sample EBs were used as the control.

Affymetrix analysis

RNA was extracted from ESCs or EBs using the QIAGEN's RNeasy Total RNA Isolation kit. For the synthesis of double-stranded cDNA from 15 μg total RNA the cDNA synthesis system kit (Roche, Mannheim, Germany) has been used. Biotinylated cRNA was synthesized with Perkin–Elmer's nucleotide analogues using the Ambion's MEGAScript T7 kit. According to the standard Affymetrix protocol, 15 μg fragmented cRNA was hybridized for 16 h at 45 °C. RNA preparations from three biologically independent experiments with three different conditions (day 0 ES, day 6 control EBs and day 6 puromycin treated EBs) were processed and hybridized on the Mouse Genome 430 Version 2 Array (Affymetrix) which includes 45 101 probe sets. After hybridization, arrays were washed and stained with streptavidin–phycoerythrin in the Affymetrix GeneChip Fluidics Station 450 and further scanned using the Affymetrix GeneChip Scanner 3000 7G. The image data were analyzed with GCOS 1.4 using Affymetrix default analysis settings and global scaling as normalization method. All chips passed our quality criteria. Comparability of the experiment conditions has been checked over all 45 101 probe sets after RMA normalization (Irizarry *et al.* 2003) by a principal component analysis using Partek Genomic Suite (6.3 beta) correlation as a dispersion matrix and normalized Eigenvector scaling.

Promoter analysis

The transcription factor binding site analysis was carried out on the promoters of the genes up-regulated in the brachyury$^+$ cells with the genes up-regulated in BMP-2$^+$ cells and α-MHC$^+$ cells as reference sets (from our previous work) (Doss *et al.* 2007a,b). First, the transcripts up-regulated in the brachyury$^+$ cells compared to both ESCs and corresponding EBs were obtained using t-test with significance threshold 0.01 for false discovery rate. The transcripts which were differentially expressed between the ESCs and EBs were discarded. The same procedure was repeated for BMP-2$^+$ and α-MHC$^+$ cells to obtain the BMP-2$^+$ and α-MHC$^+$-specific transcript lists. After mapping transcripts to genes and discarding 200 genes contained in more than one list, we had 105 genes in brachyury$^+$-specific list, 266 genes in BMP-2$^+$-specific list and 584 genes in α-MHC$^+$-specific list. The promoters of these genes were compared against each other with respect to transcription factor binding sites. For the source of binding motifs, we used the position weight matrices (PWMs) from Transfac (version 11.4) (Wingender *et al.* 1996) and Jaspar (Sandelin *et al.* 2004) databases. Promoters of size 2-kb upstream of the transcription start site were obtained from the UCSC Genome Browser database (Karolchik *et al.* 2008). In case two promoters of the same gene had overlapped more than 1 kb, we discarded one of the promoters. Each PWM was scored in every promoter by taking the average of the three highest scoring hits within the promoter. A PWM was considered interesting if the proportion of promoters scoring over some threshold was significantly higher in the brachyury$^+$ list than in the other

two lists combined. The threshold was optimized for the fold change in the proportion, while requiring the Bonferroni corrected hypergeometric *P*-value to be below 0.05.

Functional annotation

Differentially expressed genes were analyzed according to predefined pathways and functional categories annotated by KEGG (Kanehisa & Goto 2000), BioCarta (BioCarta Pathways 2009) and GO using the DAVID bioinformatic resource (Dennis *et al.* 2003). For an overrepresented GO or KEGG pathway, a cutoff *P*-value of 0.01 has been chosen. In general, it should be noted that one gene can participate in more than one KEGG or BioCarta pathway and GO category.

siRNA knockdown

ON-TARGET plus Smart pool siRNA against Larp2 was purchased from Thermo Scientific Dharmacon, Lafayette, CO, USA. siRNA targeted against Anrd34b was purchased from Qiagen Inc. (Cat. no. SI00864199, against the target sequence ACGGAACAAATTAAGCAATTA). Brachyury ESCs or wild type cells (each 1×10^5 per well of 12-well plate) were cultured for 24 h before incubating with 100 nm siRNA and 1.0 μL Lipofectamine 2000 (Invitrogen) for optimal transfection Nontarget control siRNA was used as negative control for each assay. Hanging drop EBs were made 48 h post-siRNA transfection. EBs were harvested at regular daily intervals from day 0 to day 10, and the transcripts of interest were analyzed at the mRNA level by qRT-PCR.

Acknowledgements

This work was supported by a grant from the European Community (6th Framework Programme, Thematic Priority: Life sciences, genomics and biotechnology for health, Contract No: FunGenES LSHG-CT-2003-503494). We thank Johannes Winkler and Shuhua Chen for the critical reviewing of the manuscript.

References

Baron, M.H. (2005) Early patterning of the mouse embryo: implications for hematopoietic commitment and differentiation. *Exp. Hematol.* **33**, 1015–1020.

Binetruy, B., Heasley, L., Bost, F., Caron, L. & Aouadi, M. (2007) Concise review: regulation of embryonic stem cell lineage commitment by mitogen-activated protein kinases. *Stem Cells* **25**, 1090–1095.

BioCarta Pathways (2009) http://www.biocarta.com/genes/index.asp **23**, 25.

Casciano, I., De Ambrosis, A., Croce, M., Pagnan, G., Di Vinci, A., Allemanni, G., Banelli, B., Ponzoni, M., Romani, M. & Ferrini, S. (2004) Expression of the caspase-8 gene in neuroblastoma cells is regulated through an essential interferon-sensitive response element (ISRE). *Cell Death Differ.* **11**, 131–134.

Dennis, G. Jr, Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C. & Lempicki, R.A. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* **4**, 3.

Diks, S.H., Bink, R.J., van de, W.S., Joore, J., van Rooijen, C., Verbeek, F.J., den Hertog, J., Peppelenbosch, M.P. & Zivkovic, D. (2006) The novel gene asb11: a regulator of the size of the neural progenitor compartment. *J. Cell Biol.* **174**, 581–592.

Ding, L., Liang, X., Zhu, D. & Lou, Y. (2007) Peroxisome proliferator-activated receptor alpha is involved in cardiomyocyte differentiation of murine embryonic stem cells in vitro. *Cell Biol. Int.* **31**, 1002–1009.

Doss, M.X., Koehler, C.I., Gissel, C., Hescheler, J. & Sachinidis, A. (2004) Embryonic stem cells: a promising tool for cell replacement therapy. *J. Cell Mol. Med.* **8**, 465–473.

Doss, M.X., Chen, S., Winkler, J., Hippler-Altenburg, R., Odenthal, M., Wickenhauser, C., Balaraman, S., Schulz, H., Hummel, O., Hubner, N., Ghosh-Choudhury, N., Sotiriadou, I., Hescheler, J. & Sachinidis, A. (2007a) Transcriptomic and phenotypic analysis of murine embryonic stem cell derived BMP2+ lineage cells: an insight into mesodermal patterning. *Genome Biol.* **8**, R184.

Doss, M.X., Winkler, J., Chen, S., Hippler-Altenburg, R., Sotiriadou, I., Halbach, M., Pfannkuche, K., Liang, H., Schulz, H., Hummel, O., Hubner, N., Rottscheidt, R., Hescheler, J. & Sachinidis, A. (2007b) Global transcriptome analysis of murine embryonic stem cell-derived cardiomyocytes. *Genome Biol.* **8**, R56.

Dunty, W.C. Jr, Biris, K.K., Chalamalasetty, R.B., Taketo, M.M., Lewandoski, M. & Yamaguchi, T.P. (2008) Wnt3a/-catenin signaling controls posterior body development by coordinating mesoderm formation and segmentation. *Development* **135**, 85–94.

Falco, G., Lee, S.L., Stanghellini, I., Bassey, U.C., Hamatani, T. & Ko, M.S. (2007) Zscan4: a novel gene expressed exclusively in late 2-cell embryos and embryonic stem cells. *Dev. Biol.* **307**, 539–550.

Gadue, P., Huber, T.L., Nostro, M.C., Kattman, S. & Keller, G.M. (2005) Germ layer induction from embryonic stem cells. *Exp. Hematol.* **33**, 955–964.

Gissel, C., Voolstra, C., Doss, M.X., Koehler, C.I., Winkler, J., Hescheler, J. & Sachinidis, A. (2005) An optimized embryonic stem cell model for consistent gene expression and developmental studies: a fundamental study. *Thromb. Haemost.* **94**, 719–727.

Herrmann, B.G. (1991) Expression pattern of the Brachyury gene in whole-mount TWis/TWis mutant embryos. *Development* **113**, 913–917.

Hochberg, Y. & Benjamini, Y. (1990) More powerful procedures for multiple significance testing. *Stat. Med.* **9**, 811–818.

Hortsch, M., Nagaraj, K. & Godenschwege, T.A. (2009) The interaction between L1-type proteins and ankyrins–a master

switch for L1-type CAM function. *Cell. Mol. Biol. Lett.* **14**, 57–69.

Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B. & Speed, T.P. (2003) Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res.* **31**, e15.

Kanehisa, M. & Goto, S. (2000) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30.

Karolchik, D., Kuhn, R.M., Baertsch, R., *et al.* (2008) The UCSC Genome Browser Database: 2008 update. *Nucleic Acids Res.* **36**, D773–D779.

Khokha, M.K., Yeh, J., Grammer, T.C. & Harland, R.M. (2005) Depletion of three bmp antagonists from Spemann's organizer leads to a catastrophic loss of dorsal structures. *Dev. Cell* **8**, 401–411.

Kispert, A. & Hermann, B.G. (1993) The Brachyury gene encodes a novel DNA binding protein. *EMBO J.* **12**, 4898–4899.

Kispert, A. & Herrmann, B.G. (1994) Immunohistochemical analysis of the Brachyury protein in wild-type and mutant mouse embryos. *Dev. Biol.* **161**, 179–193.

Lacaud, G., Keller, G. & Kouskoff, V. (2004) Tracking mesoderm formation and specification to the hemangioblast in vitro. *Trends Cardiovasc. Med.* **14**, 314–317.

Lahaye, K., Kricha, S. & Bellefroid, E.J. (2002) XNAP, a conserved ankyrin repeat-containing protein with a role in the Notch pathway during Xenopus primary neurogenesis. *Mech. Dev.* **110**, 113–124.

Malikova, M.A., Van Stry, M. & Symes, K. (2007) Apoptosis regulates notochord development in Xenopus. *Dev. Biol.* **311**, 434–448.

Mariappan, D., Winkler, J., Chen, S., Schulz, H., Hescheler, J. & Sachinidis, A. (2009) Transcriptional profiling of CD31+ cells isolated from murine embryonic stem cells. *Genes Cells* **14**, 243–260.

Martinsen, B.J. (2005) Reference guide to the stages of chick heart embryology. *Dev. Dyn.* **233**, 1217–1237.

Petzold, A., Keir, G., Warren, J., Fox, N. & Rossor, M.N. (2007) A systematic review and meta-analysis of CSF neurofilament protein levels as biomarkers in dementia. *Neurodegener. Dis.* **4**, 185–194.

Potta, S.P., Liang, H., Pfannkuche, K., Winkler, J., Chen, S., Doss, M.X., Obernier, K., Kamisetti, N., Schulz, H., Hubner, N., Hescheler, J. & Sachinidis, A. (2009) Functional characterization and transcriptome analysis of embryonic stem cell-derived contractile smooth muscle cells. *Hypertension* **53**, 196–204.

Ross, E.A., Freeman, S., Zhao, Y., Dhanjal, T.S., Ross, E.J., Lax, S., Ahmed, Z., Hou, T.Z., Kalia, N., Egginton, S., Nash, G., Watson, S.P., Frampton, J. & Buckley, C.D. (2008) A novel role for PECAM-1 (CD31) in regulating haematopoietic progenitor cell compartmentalization between the peripheral blood and bone marrow. *PLoS ONE* **3**, e2338.

Sandelin, A., Alkema, W., Engstrom, P., Wasserman, W.W. & Lenhard, B. (2004) JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res.* **32**, D91–D94.

Sharifpanah, F., Wartenberg, M., Hannig, M., Piper, H.M. & Sauer, H. (2008) Peroxisome proliferator-activated receptor alpha agonists enhance cardiomyogenesis of mouse ES cells by utilization of a reactive oxygen species-dependent mechanism. *Stem Cells* **26**, 64–71.

Shen, M.M. (2007) Nodal signaling: developmental roles and regulation. *Development* **134**, 1023–1034.

Szabo, E., Qiu, Y., Baksh, S., Michalak, M. & Opas, M. (2008) Calreticulin inhibits commitment to adipocyte differentiation. *J. Cell Biol.* **182**, 103–116.

Tam, P.P. & Behringer, R.R. (1997) Mouse gastrulation: the formation of a mammalian body plan. *Mech. Dev.* **68**, 3–25.

Willems, E. & Leyns, L. (2008) Patterning of mouse embryonic stem cell-derived pan-mesoderm by Activin A/Nodal and Bmp4 signaling requires fibroblast growth factor activity. *Differentiation* **76**, 745–759.

Wingender, E., Dietze, P., Karas, H. & Knuppel, R. (1996) TRANSFAC: a database on transcription factors and their DNA binding sites. *Nucleic Acids Res.* **24**, 238–241.

Winkler, J., Hescheler, J. & Sachinidis, A. (2005) Embryonic stem cells for basic research and potential clinical applications in cardiology. *Biochim. Biophys. Acta* **1740**, 240–248.

Wittler, L., Shin, E.H., Grote, P., Kispert, A., Beckers, A., Gossler, A., Werber, M. & Herrmann, B.G. (2007) Expression of Msgn1 in the presomitic mesoderm is controlled by synergism of WNT signalling and Tbx6. *EMBO Rep.* **8**, 784–789.

Received: 21 August 2009
Accepted: 16 November 2009

Supporting Information/Supplementary Material

The following Supporting Information can be found in the online version of the article:

Figure S1 Overview of the Bacterial Artificial Chromosomes (BAC) subcloning by recombineering approach (right scheme, slightly modified from the Gene Bridges BAC recommbineering manual) to subclone the brachyury promoter from the BAC and then to clone the promoter into Puro/EGFP IRES vector.

Figure S2 Validation of Affymetrix data by quantitative real-time polymerase chain reaction (PCR) analyses. The fold change was calculated by using the formula: fold-change = $2^{-(\Delta Ctgene1 - \Delta Ctgene2)}$. ΔCt of the gene in the sample in which it is expressed lowest is taken as ΔCt gene2 to calculate the fold change using the above formula. The resulting fold change is expressed as percentage of the maximum fold change (= 100%) for that particular gene in every assay. Values are expressed as mean ± standard deviation (n = 3; technical replicates).

Figure S3 Hierarchical clustering of the transcripts up- or downregulated in brachyury$^+$ cells in comparison to at least one of brachyury ESCs and the 6-days old control EBs.

Figure S4 RT-PCR analysis for various developmental gene markers in ESCs and EBs treated with either siRNA against *Ankrd34b* or with vehicle of the siRNA.

Table S1 Transcription factors (a) and development related (b) transcripts upregulated in brachyury$^+$ cells compared to both undifferentiated brachyury ESCs and control 6-days old brachyury EBs

Table S2 Development (a), transcription factors (b) and signal transducers (c) related transcripts upregulated in brachyury$^+$ cells compared to undifferentiated brachyury ESCs but downregulated in comparison to the 6-days old EBs

Table S3 Development (a), transcription factors (b) and signal transducers (c) related transcripts upregulated in brachyury$^+$ cells compared to control 6 day EBs but downregulated in comparison to the brachyury ESCs

Table S4 Development (a), transcription factors (b) and signal transducers (c) related transcripts upregulated in brachyury$^+$ cells compared to either control 6 day EBs or to the brachyury ESCls but which don't show any twofold downregulation than anyone of the control population

Table S5 Primers and conditions used in this study for the semiquantitative PCR analysis

Additional Supporting Information may be found in the online version of this article.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.