THE INTERNATIONAL JOURNAL OF DEVELOPMENTAL BIOLOGY
www.intjdevbiol.com

doi: 10.1387/ijdb.113342cl

SUPPLEMENTARY MATERIAL

corresponding to:

A possible role of *Reproductive homeobox 6* in primordial germ cell differentiation

CHANG LIU, PAICHI TSAI, ANA-MARIE GARCÍA, BRANDON LOGEMAN and TETSUYA S. TANAKA

Vector construction

Unless otherwise noted, the restriction enzymes and DNA modifying enzymes used in this study were obtained from New England Biolabs (NEB; Ipswich, MA), and the oligonucleotides used were synthesized by Invitrogen (Carlsbad, CA). PCR was carried out with Phusion DNA polymerase (NEB) to introduce new restriction enzyme sites at the ends of PCR products. The DNA constructs used as templates for PCR were kindly provided by Dr. William L. Stanford, University of Toronto, Ontario, Canada.

The CAG promoter (Niwa *et al.*, 1991) was obtained from pCX-eCFP (Hadjantonakis *et al.*, 2002) by double-digestion with SalI and EcoRI and cloned into pBluescript II KS(-), which was designated pBS_CAG. Then, a bovine growth hormone polyadenylation signal was obtained from ROSAβ-gal (Friedrich and Soriano, 1991) by XbaI digestion and cloned into pBS_CAG. The resulting plasmid was named pCAG_BGHpA, which has EcoRI, PstI, SmaI and BamHI sites uniquely available. The puromycin-resistant gene (*Puro*^{r}) was PCR-amplified with EcoRI-Puro-F (5'-ACG AAT TCC ATG ACC GAG TAC AAG CCC ACG G-3') and XbaI-Puro-R (5'-AAT CTA GAT CAG GCA CCG GGC TTG CG-3') as primers using pEGFP-IRES-Puro (Clontech, Mountain View, CA) as a template. This product was cloned into pIRES-Neo (Clontech), which had been digested with EcoRI and XbaI. This plasmid was designated pPuro. Then, using pCXeCFP as a template, ECFP was PCR-amplified and cloned into pPuro digested with EcoRV and EcoRI to generate ECFP fused with the puromycin-resistant gene. The primers used were as follows: EcoRV-ECFP-F, 5'-ATG ATA TCG CCA CCA TGG TGA GCA-3'; EcoRI-ECFP-Fus-R, 5'-TCG AAT TCT TGT ACA GCT CGT CC-3'. ECFP expression and puromycin resistance were functionally validated by transfection of the construct into Phoenix cells (Pear *et al.*, 1993) by the CaPO4 method. Because XbaI-Puro-R introduced a sequence motif for Dam methylation, the ECFP-Puro cassette was further PCR-amplified with EcoRV-ECFP-F and XbaI-Puro-R primers. After digestion with EcoRV and XbaI, followed by performing a fill-in reaction using the Klenow fragment (Invitrogen), the ECFP-Puro cassette was blunt-end ligated with SmaI-digested pCAG_BGHpA. The resulting vector was designated pCAG_CFPuro.

To clone the self-cleaving peptide T2A (Szymczak *et al.*, 2004), sense and antisense oligonucleotides encoding the T2A peptide were designed, as follows: EcoRI-2A-S1, 5'-AAT TCA AGA GCT AAA AGA GAG GGC AGA GGA AGT CTG CTA-3'; 2A-EcoRI*-S2, 5'-ACA TGC GGT GAC GTC GAG GAG AAT CCT GGC CCA GGA TCA GGA-3'; 2A-EcoRI*-AS1, 5'-AAT TTC CTG ATC CTG GGC CAG GAT TCT CCT CGA CGT CAC C-3'; EcoRI-2A-AS2, 5'-GCA TGT TAG CAG ACT TCC TCT GCC CTC TCT TTT AGC TCT TG-3' [Integrated DNA Technologies (IDT) Coralville, IA]. These oligonucleotides were self-assembled in 0.3 x SSC (50 mM Na^+) at 65 ºC, overnight, and cloned into EcoRI-digested

SUPPLEMENTARY TABLE 1

PRIMERS USED FOR RT-PCR

Transcript (sq or q)*		Sequence	Product size (bp)	Reference
c-Kit (sq)	Forward Reverse	CCCTCAGCCTCAGCACATAGCC CTGTAGTGTTCTTCACAGGG	224	Elliott *et al.*, 2007
Dazl (sq & q)	Forward Reverse	GCCAGCACTCAGTCTTCATC GTTGGAGGCTGCATGTAAGT	419	Geijsen *et al.*, 2004
Ef1α (sq)	Forward Reverse	ATATTACCCCTAACACCTGG CTGTGACAGATTTTTGGTCAAG	247	Tanaka *et al.*, 2002
Ef1α (q)	Forward Reverse	CAATGGAAGCAGCTGGCTTCACTGC CCTCATGTCACGAACAGCAAAGCGAC	290	This study
Esg1 (sq)	Forward Reverse	ATAAGCTTGATCTCGTCTTCC CTTGCTAGGATGTAACAAAGC	501	Tanaka *et al.*, 2002
Esg1 (q)	Forward Reverse	GCCGTGCGTGGTGGATAAGC GCCAAACAGATATTTCAGCACCAGC	175	Tanaka *et al.*, 2002
Fgls (sq)	Forward Reverse	TTGCTCCGCACCATGAACCA TGAAGCACTTCAGGACCGGA	461	Geijsen *et al.*, 2004
Mvh (sq & q)	Forward Reverse	GCTCAAACAGGGTCTGGGAAG GGTTGATCAGTTCTCGAG	145	Toyooka *et al.*, 2003
Nanog (sq)	Forward Reverse	AGGGTCTGCTACTGAGATGCTCTG CAACCACTGGTTTTTCTGCCACCG	364	Mitsui *et al.*, 2003
Nanos3 (sq)	Forward Reverse	TCCCGTGCCATCTATCAG GGATGTTGAGGCAACACC	494	Tsuda *et al.*, 2003
Oct3/4 (sq & q)	Forward Reverse	GGCGTTCTCTTTGGAAAGGTGTTC CTCGAACCACATCCTTCTCT	313	Nichols *et al.*, 1998
Piwil2 (sq & q)	Forward Reverse	CCGTCATGAAGGAGAGCTCG GGAACGACTCTGTGCTGGAT	348	Geijsen *et al.*, 2004
Rex1 (sq)	Forward Reverse	CGAGTGGCAGTTTCTTCTTGG CTTCTTGAACAATGCCTATGACTCACTTCC	302	Toyooka *et al.*, 2008
Rhox6 (sq)	Forward Reverse	TGTTCTGAATAGGCTGGCTCAACTGCGGTACAG CATCCTCATCTGGCTCCATGACAGGGCTG	310	This study
Rhox6 (sq & q)	Forward Reverse	TGTTCTGAATAGGCTGGCTCAACTGCGGTACAG GGAGAGTCGCTCTGGGGAAGAGGC (This pair detects both endogenous transcripts and the transgene)	252	This study
Rhox6 transgene (sq & q)	Forward Reverse	TGTTCTGAATAGGCTGGCTCAACTGCGGTACAG GCATGTTAGCAGACTTCCTCTGCCCTCTCTTTTAGCTCTTG (see "EcoRI-2A-AS2" above)	298	This study
Rhox9 (sq)	Forward Reverse	GGCTGGGAACTATCTGGCTCACCAGCGGACCC CTTTATTGTTGAAATAATTATAGAGAAGTA	349	This study
Rhox9 (q)	Forward Reverse	GGCTGGGAACTATCTGGCTCACCAGCGGACCC GGGAGAGTTGTTCTCTGTAATCGGTG	253	This study

*: "sq" and "q" indicate that primer pairs were used for sq- and q-RT-PCR, respectively.

pCAG_CFPuro. This T2A adaptor contains compatible ends with EcoRI, although the one at the 3' end (*) was mutated, such that the EcoRI site at the 5' end remained unique in the final construct, pCAG_T2APuro. *Rhox6* was first PCR amplified from cDNA and cloned into pBluescript II KS (-) to assure that the cDNA encoded *Rhox6* specifically and not *Rhox9*. The cDNA used was prepared using total RNA extracted from the extraembryonic tissues of embryos at E9.5-10.5 from matings of CD1 mice. The primers used were as follows: EcoRV-Rhox6-F, 5'-ATG ATA TCG GAA GCC TCT TCG GGA GCA GCG TC-3', and EcoRI-Rhox6-R, 5'-GCG AAT TCA GAA TGC TCA TCT TTA TTG CTG AAA TAA TTG TAG AG-3'. After the sequence was validated, the protein-coding region of *Rhox6* was PCR-amplified, then cloned into pCAG_T2APuro digested with EcoRI. The primers used for this step were as follows: EcoRI-Rhox6-F, 5'-GGA ATT CGC CAC CAT GGA AAC TCC TCA AGA CAG-3', and EcoRI-Rhox6-R, 5'-AAG AAT TCG GAG AGT CGC TCT GGG GAA GAG G-3' (IDT). The forward primer introduced Kozak's consensus sequence for strong translational initiation (GCC ACC ATG G) (Kozak, 1987).

DsRedT3 was PCR-amplified using MST-B (Bevis and Glick, 2002; Vintersten *et al*., 2004) as a template and cloned into pBS_SAGtxVenus, which was used to build pGTIV (Tanaka *et al*., 2008). The primers used were as follows: EcoRV-Red-F, 5'-ATG ATA TCG CCA CCA TGG CCT CCT C-3'; EcoRI-Red-Fus-R, 5'-TTG AAT TCA GGA ACA GGT GGT GGC GG-3'. The resulting plasmid was designated pBS_SAGtxDsRedT3. Then, the T2A_Puro^r cassette from pCAG_T2APuro was cloned between EcoRI and XbaI sites in pBS_SAGtxDsRedT3. Finally, the Gtx_DsRedT3_T2A_Puro^r cassette was digested with NsiI and BamHI and then cloned into pCAG_T2APuro digested with PstI and BamHI. This vector was designated pCAG_DsRedT2AP (Chowdhury *et al*., 2010).

To build a construct that expressed short-hairpin (sh) RNA (Brummelkamp *et al*., 2002), the human H1 (hH1) promoter in pcDNA_RasGap (Kunath *et al*., 2003) was first PCR-amplified with the following primers: EcoRV-XhoI-hH1-F, 5'-GCG ATA TCT CGA GCA ATA TTT GCA TGT CGC TAT GTG-3'; SalI-HK-hH1-R, 5'-ATG TCG ACA AGC TTA AGG TAC CGA GTG GTC TCA TAC AG-3'. Next, pCAG_CFPuro was digested with KpnI, which was blunt-ended with mung bean nuclease and further digested with XhoI. This was used to clone PCR-amplified hH1 digested with EcoRV and SalI. The resulting vector was designated phH1CCP. Oligonucleotides for shRNA expression were cloned between KpnI and HindIII sites downstream of hH1 (underlined in SalI-HK-hH1-R).

phH1CCP was further modified as follows. First, the blasticidin-resistant gene (*Bsd^r*) was obtained from pGTLox4 (http://www.cmhd.ca/genetrap/vectors.html) by digestion with EcoRI and ScaI and cloned into pBluescript II KS(-) digested with EcoRI and SmaI. Because the translational initiation site of *Bsd^r* includes an NcoI

SUPPLEMENTARY TABLE 2

SUMMARY OF SHRNA EXPERIMENTS

	Vector	Amount (Delivery method)	cDNA/ shRNA	# of clones (EBs)	# of EBs with EGFP +	# of EBs with EGFP ±	# of EBs with EGFP -	Sum of EGFP expression level**
Series I	pCAG_2AP	10 µg						
		Electroporation	Rhox6	12 (39)	17	11	11	7.9
			DsRedT3	12 (61)	28	29	4	10.5
		10 µg						
		Electroporation	Rhox6	4 (26)	19	5	2	3.5
			DsRedT3	5 (14)	13	1	0	5
		Total	Rhox6	16 (65)	36	16	13	11.4
			DsRedT3	17 (75)	41	30	4	15.5
Series II	phH1CCP	10 µg						
		Electroporation	Scrambled	9 (17)	8	n/a	9	2.7
			shRhox6	4 (6)	0	n/a	6	0
			shRhox6&9	8 (13)	2	n/a	11	1
		10 µg						
		Electroporation	Scrambled	15 (18)	7	n/a	11	5.5
			shRhox6	14 (23)	0	n/a	23	0
			shRhox6&9	18 (31)	0	n/a	31	0
		10 µg						
		Electroporation	Scrambled	7 (13)	5	n/a	8	2
			shRhox6	9 (13)	1	n/a	12	1
			shRhox6&9	8 (16)	3	n/a	13	1.3
		Total	Scrambled	31 (48)	20	-	28	10.2
			shRhox6	27 (42)	1	-	41	1
			shRhox6&9	34 (60)	5	-	55	2.3
Series III	phH1CCP	1 µg	Scrambled	13 (22)	20	2	0	13
		FuGene HD	shRhox6	13 (33)	23	7	3	11
			shRhox9	12 (24)	21	2	1	11.5
			shRhox6&9	13 (34)	17	13	4	11.7
	phH1CRB	7 µg	Scrambled	11 (37)	26	10	1	10.9
		Electroporation	shRhox6	13 (43)	16	15	12	9.9
			shRhox9	7 (21)	11	3	7	4.9
			shRhox6&9	11 (31)	15	10	6	8.8
		Total	Scrambled	24 (59)	46	12	1	23.9
			shRhox6	26 (76)	39	22	15	20.9
			shRhox9	19 (45)	32	5	8	16.4
			shRhox6&9	24 (65)	32	23	10	20.5

*: These EBs expressed EGFP in small parts, but considered as EGFP+ to calculate the "EGFP level". n/a, not applicable.

**: The number of EBs with EGFP expression was divided by the total number of EBs counted for each clone, which indicates the "level" of EGFP expression per clone. Values of the "EGFP level" were added within each condition. Thus, the closer the sum is to the number of clones per condition, the higher the "EGFP level" is.

site (CCATGG), a modified T2A cassette was inserted between the EcoRI and NcoI sites. Then, the T2A_Bsd^r cassette digested with EcoRI and BamHI was cloned into phH1CCP. This plasmid was designated phH1C2AB. Finally, the Gtx_DsRedT3 cassette was obtained from pCAG_DsRedT2AP by EcoRI-digestion and cloned into phH1C2AP. The resulting vector was designated phH1CRB.

To select target regions for shRNA expression, lists of candidate sites were obtained using the following web-based tools: siDESIGN Center (http://www.dharmacon.com/designcenter/designcenterpage.aspx); siRNA Target Designer (http://www.promega.com/siRNADesigner/program/); siRNA Target Finder (http://www.ambion.com/techlib/misc/siRNA_finder.html); BLOCK-iT™ RNAi Designer (https://rnaidesigner.invitrogen.com/rnaiexpress/). The specificity of the target sequence was validated using the Ensembl Genome Browser. A scrambled oligonucleotide sequence was designed as a negative control and showed matches with three intronic regions. According to their specificity, three target sites were selected for each *Rhox* gene, and corresponding oligonucleotides were designed (IDT). These oligonucleotides were self-assembled in 0.3 x SSC (50 mM Na^+) at 65 ºC overnight and ligated with phH1CCP and phH1CRB, which had been double-digested with KpnI and HindIII. Their nucleotide sequence was validated by sequencing.

To evaluate the levels of knockdown obtained, mouse ESCs were transiently transfected with each shRNA expression vector (phH1CCP) by FuGene HD (Roche Applied Science, Indianapolis, IN). Two days after selection with 2 μg/ml puromycin (InvivoGen, San Diego, CA), total RNA was extracted (Chomczynski and Sacchi, 1987) and used to synthesize cDNA. The oligonucleotide sequences that induced specific knockdown of their targets are listed below. Only sequences for sense strands are shown. For oligonucleotides against *Rhox6* or *Rhox9*, the position of the corresponding nucleotide sequence is shown by counting the "A" in the translational initiation codon (ATG) as 1: for scrambled1-S, 5'-CGG TTG TGG TGC GGA ACA CA-3', there are three hits with intronic regions with e-values of 0.32 or greater; for shRhox6-627S (627-647), 5'- CCA GGA GAG TGC TGA TGT TCT G-3', the e-value for the match with *Rhox6* is 0.0013, and there are seven additional matches with either non-coding genomic or intronic regions with e-values of 0.32 or greater; for shRhox9-652S (652-670), 5'-CCT GCC ACC GAT TAC AGA GA-3', the e-value for the match with *Rhox9* is 0.021, and there is another match with an intronic region with an e-value of 4.8; for shRhox6&9-542S (542-560 in Rhox6), 5'-CGG AGG GAT CTT GCA CGA TG-3', the e-value for the match with *Rhox6* and *Rhox9* is 0.021, and there is another match with a non-coding genomic region with an e-value of 1.3. The closer the E-value is to zero, the more significant the similarity is.

References

BEVIS, B. J. and GLICK, B. S. (2002). Rapidly maturing variants of the Discosoma red fluorescent protein (DsRed), *Nat Biotechnol* 20: 83-87.

BRUMMELKAMP, T. R., BERNARDS, R. and AGAMI, R. (2002). A system for stable expression of short interfering RNAs in mammalian cells, *Science* 296: 550-553.

CHOMCZYNSKI, P. and SACCHI, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction, *Anal Biochem* 162: 156-159.

CHOWDHURY, F., NA, S., LI, D., POH, Y. C., TANAKA, T. S., WANG, F. and WANG, N. (2010). Material properties of the cell dictate stress-induced spreading and differentiation in embryonic stem cells, *Nat Mater* 9: 82-88.

ELLIOTT, A. M., DE MIGUEL, M. P., REBEL, V. I. and DONOVAN, P. J. (2007). Identifying genes differentially expressed between PGCs and ES cells reveals a role for CREB-binding protein in germ cell survival, *Dev Biol* 311: 347-358.

FRIEDRICH, G. and SORIANO, P. (1991). Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice, *Genes Dev* 5: 1513-1523.

GEIJSEN, N., HOROSCHAK, M., KIM, K., GRIBNAU, J., EGGAN, K. and DALEY, G. Q. (2004). Derivation of embryonic germ cells and male gametes from embryonic stem cells, *Nature* 427: 148-154.

HADJANTONAKIS, A.-K., MACMASTER, S. and NAGY, A. (2002). Embryonic stem cells and mice expressing different GFP variants for multiple non-invasive reporter usage within a single animal, *BMC Biotechnol.* 2: 11.

KOZAK, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs, *Nucleic Acids Res* 15: 8125-8148.

KUNATH, T., GISH, G., LICKERT, H., JONES, N., PAWSON, T. and ROSSANT, J. (2003). Transgenic RNA interference in ES cell-derived embryos recapitulates a genetic null phenotype, *Nat Biotechnol* 21: 559-561.

MITSUI, K., TOKUZAWA, Y., ITOH, H., SEGAWA, K., MURAKAMI, M., TAKAHASHI, K., MARUYAMA, M., MAEDA, M. and YAMANAKA, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells, *Cell* 113: 631-642.

NICHOLS, J., ZEVNIK, B., ANASTASSIADIS, K., NIWA, H., KLEWE-NEBENIUS, D., CHAMBERS, I., SCHOLER, H. and SMITH, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4, *Cell* 95: 379-391.

NIWA, H., YAMAMURA, K. and MIYAZAKI, J. (1991). Efficient selection for high-expression transfectants with a novel eukaryotic vector, *Gene* 108: 193-199.

PEAR, W. S., NOLAN, G. P., SCOTT, M. L. and BALTIMORE, D. (1993). Production of high-titer helper-free retroviruses by transient transfection, *Proc Natl Acad Sci USA* 90: 8392-8396.

SZYMCZAK, A. L., WORKMAN, C. J., WANG, Y., VIGNALI, K. M., DILIOGLOU, S., VANIN, E. F. and VIGNALI, D. A. A. (2004). Correction of multi-gene deficiency *in vivo* using a single 'self-cleaving' 2A peptide-based retroviral vector, *Nat Biotech* 22: 589-594.

TANAKA, T. S., DAVEY, R. E., LAN, Q., ZANDSTRA, P. W. and STANFORD, W. L. (2008). Development of a gene trap vector with a highly-sensitive fluorescent protein reporter system aiming for the real-time single cell expression profiling, *Genesis* 46: 347-356.

TANAKA, T. S., KUNATH, T., KIMBER, W. L., JARADAT, S. A., STAGG, C. A., USUDA, M., YOKOTA, T., NIWA, H., ROSSANT, J. and KO, M. S. (2002). Gene expression profiling of embryo-derived stem cells reveals candidate genes associated with pluripotency and lineage specificity, *Genome Res* 12: 1921-1928.

TOYOOKA, Y., SHIMOSATO, D., MURAKAMI, K., TAKAHASHI, K. and NIWA, H. (2008). Identification and characterization of subpopulations in undifferentiated ES cell culture, *Development* 135: 909-918.

TOYOOKA, Y., TSUNEKAWA, N., AKASU, R. and NOCE, T. (2003). Embryonic stem cells can form germ cells in vitro, *Proc Natl Acad Sci USA* 100: 11457-11462.

TSUDA, M., SASAOKA, Y., KISO, M., ABE, K., HARAGUCHI, S., KOBAYASHI, S. and SAGA, Y. (2003). Conserved role of nanos proteins in germ cell development, *Science* 301: 1239-1241.

VINTERSTEN, K., MONETTI, C., GERTSENSTEIN, M., ZHANG, P., LASZLO, L., BIECHELE, S. and NAGY, A. (2004). Mouse in red: red fluorescent protein expression in mouse ES cells, embryos, and adult animals, *Genesis* 40: 241-246.