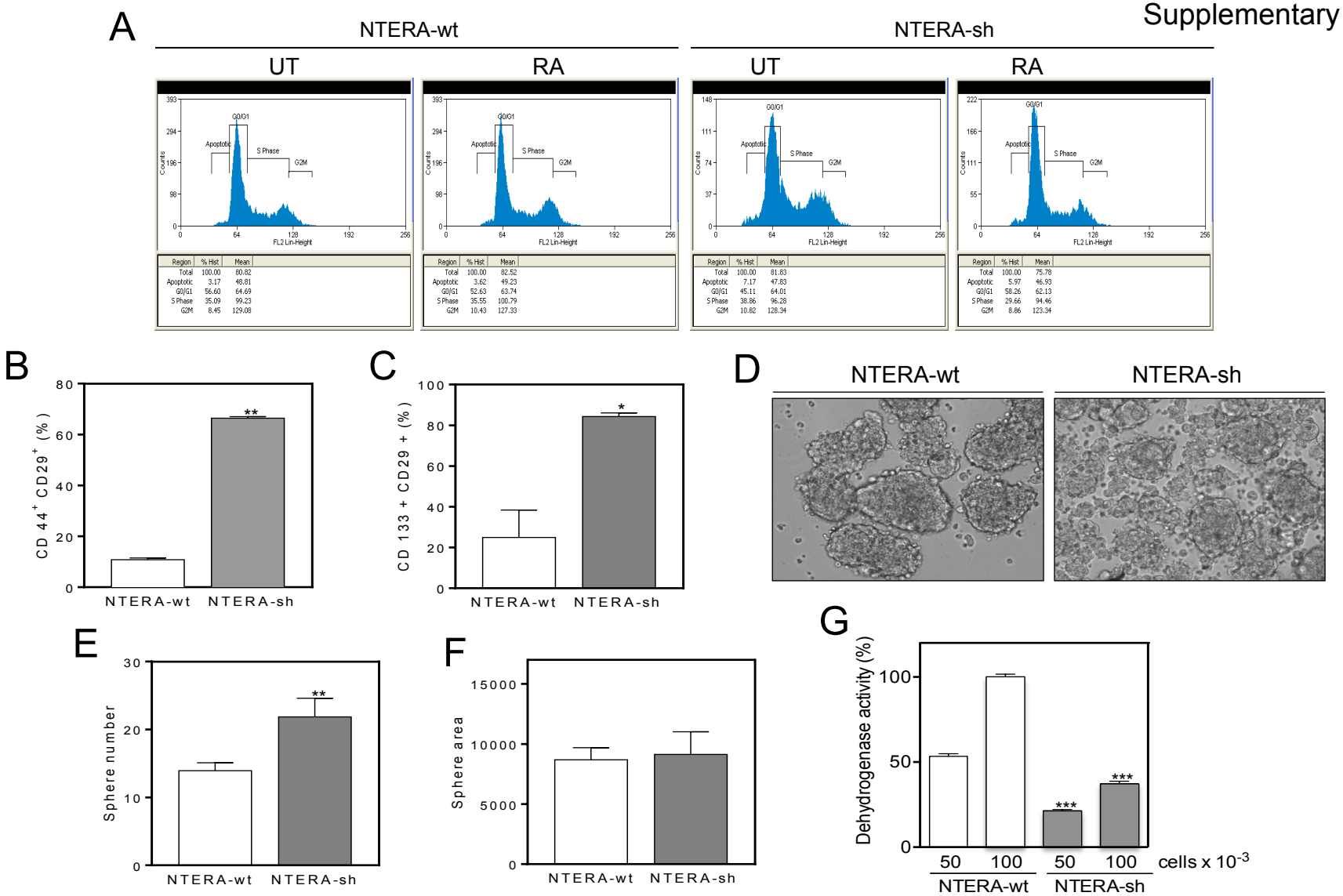
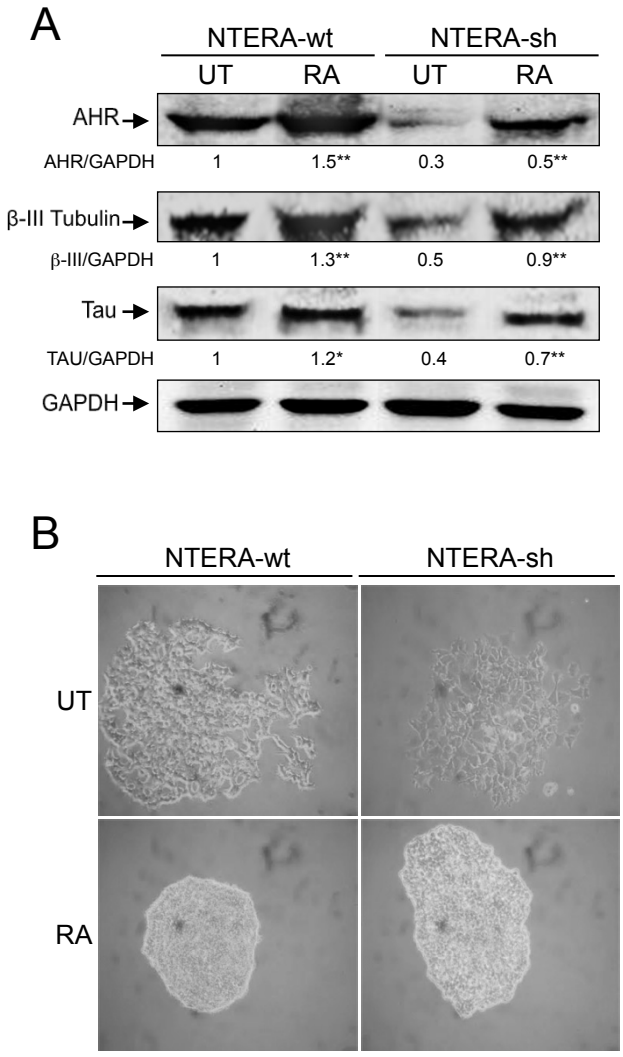
Supplementary Figure 1

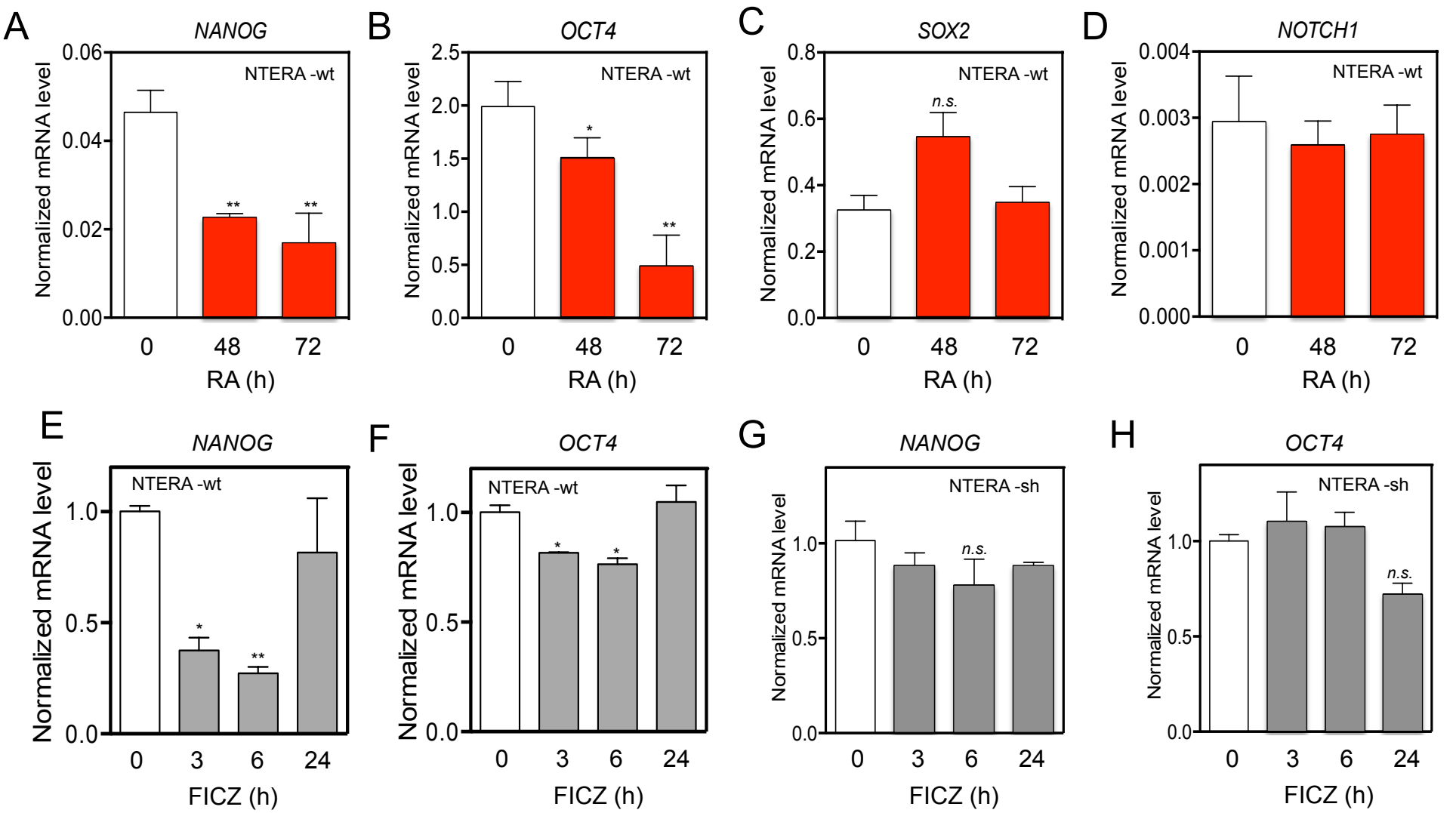


Supplementary Fig. 1. AHR knockdown increases undifferentiation markers in NTERA cells. (A) Cell cycle distribution and apoptosis levels were measured in NTERA-wt and NTERA-sh cultures using a FC500 flow cytometer and propidium iodide staining. **(B,C)** The percentages of $CD44^{+}/CD29^{+}$ and $CD133^{+}/CD29^{+}$ cells were determined by flow cytometry in NTERA-wt and NTERA-sh cells. **(D)** Both cell lines were grown in ultra-low adherence culture plates and their ability to form spheres analyzed after 48 h. **(E,F)** The number **(E)** and the area **(F)** of the spheres formed by each cell line were determined by conventional microscopy and the ImageJ software. **(G)** Dehydrogenase activity was measured in NTERA-wt and NTERA-sh cells using cell densities of $5x10^4$ and 10^5. Panels A-E: $n = 6$, two technical replicates in three biological replicates; panel F: $n = 6$, three technical replicates in two biological replicates. *p<0.05, **p<0.01 and p***<0.001. Data are shown as mean ± SD.



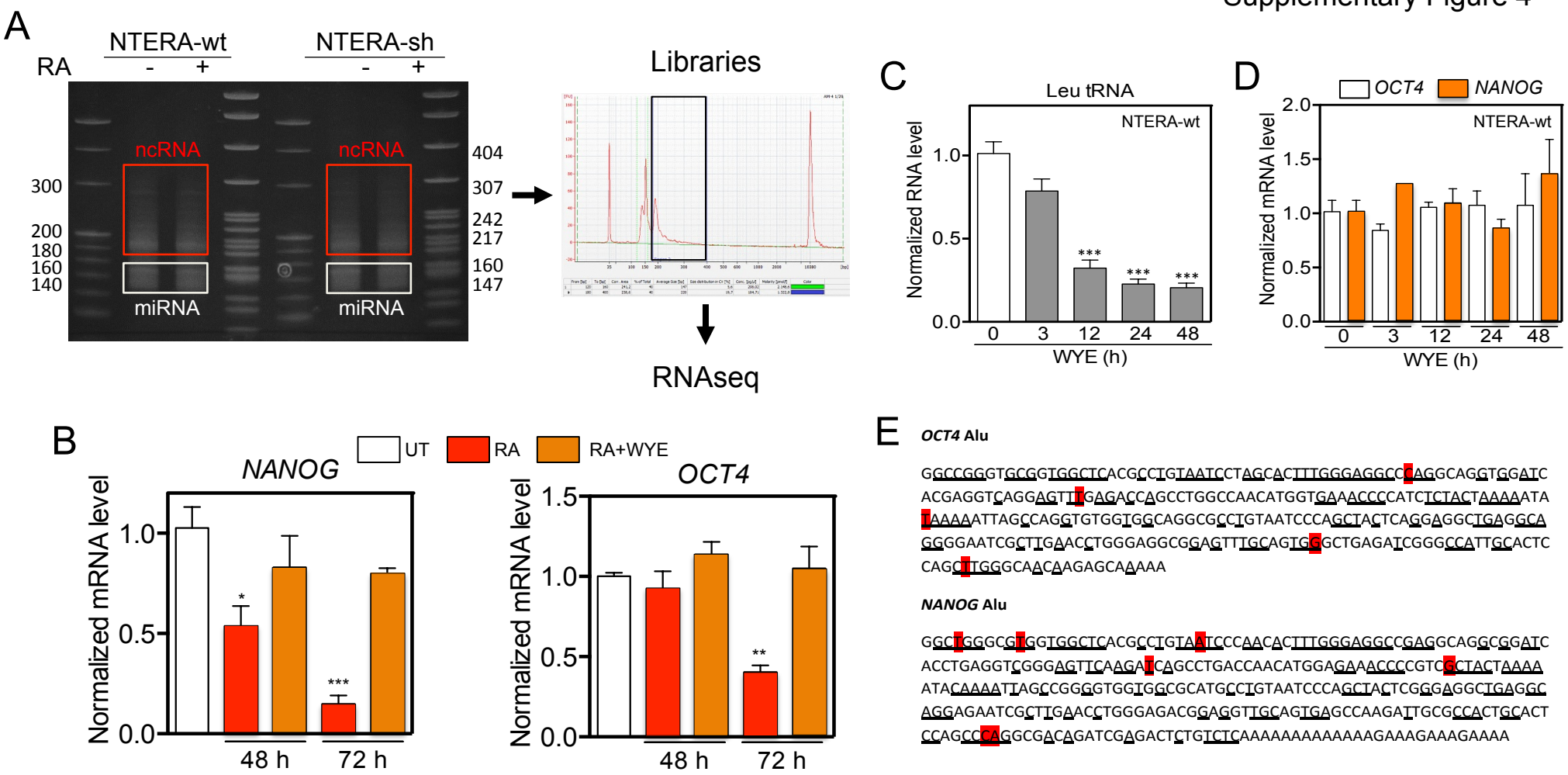
Supplementary Fig. 2. Long-term exposure to RA rescues differentiation in NTERA-sh cells. (A) NTERA-wt and NTERA-sh cells were plated at low cell density and the clones formed cultured for 4 weeks without (UT) or with 1 μM RA. Protein expression of AHR, Tau and βIII-tubulin were determined by immunoblotting. GAPDH was used as normalization control. **(B)** Clone morphology and spreading were also analyzed by conventional light transmitted microscopy. NTERA-wt cells were analyzed under the same experimental conditions. Panels A,B: $n = 4$ biological replicates. *$p<0.05$ and **$p<0.01$. Data are shown as mean ± SD.

Supplementary Figure 3

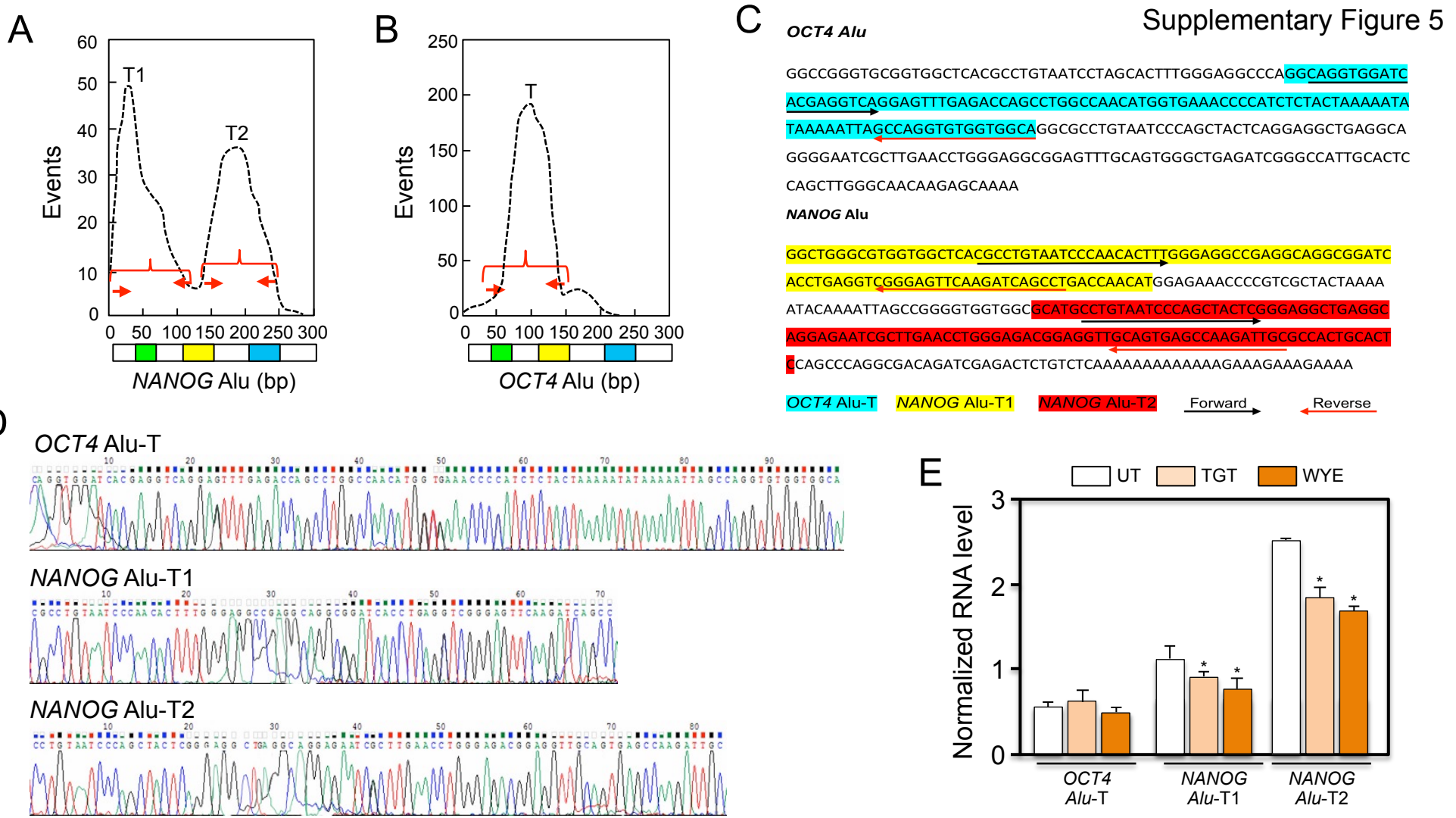


Supplementary Fig. 3. RA treatment reduces the mRNA levels of *OCT4* and *NANOG* in AHR-expressing cells. (A-D) NTERA-wt cells were left untreated (UT) or treated with 1 μM RA for 48 or 72 h and the mRNA levels of *NANOG* **(A)**, *OCT4* **(B)**, *SOX2* **(C)** and *NOTCH1* **(D)** were determined by RT-qPCR using the oligonucleotides indicated in Supplementary Table 1. **(E-H)** The mRNA levels of *OCT4* and *NANOG* were determined in NTERA-wt **(E,F)** and NTERA-sh **(G-H)** cells treated for the indicated times with 10 nM of the AHR ligand 6-formylindolo[3,2-b]carbazole (FICZ). *GAPDH* mRNA was used to normalize gene expression (ΔCt) and $2^{-\Delta\Delta Ct}$ to calculate variations with respect to control or untreated conditions. Panels A-F: $n = 6$, two technical replicates in three biological replicates*p<0.05, **p<0.01 and ***p<0.001. Data are shown as mean ± SD.

Supplementary Figure 4

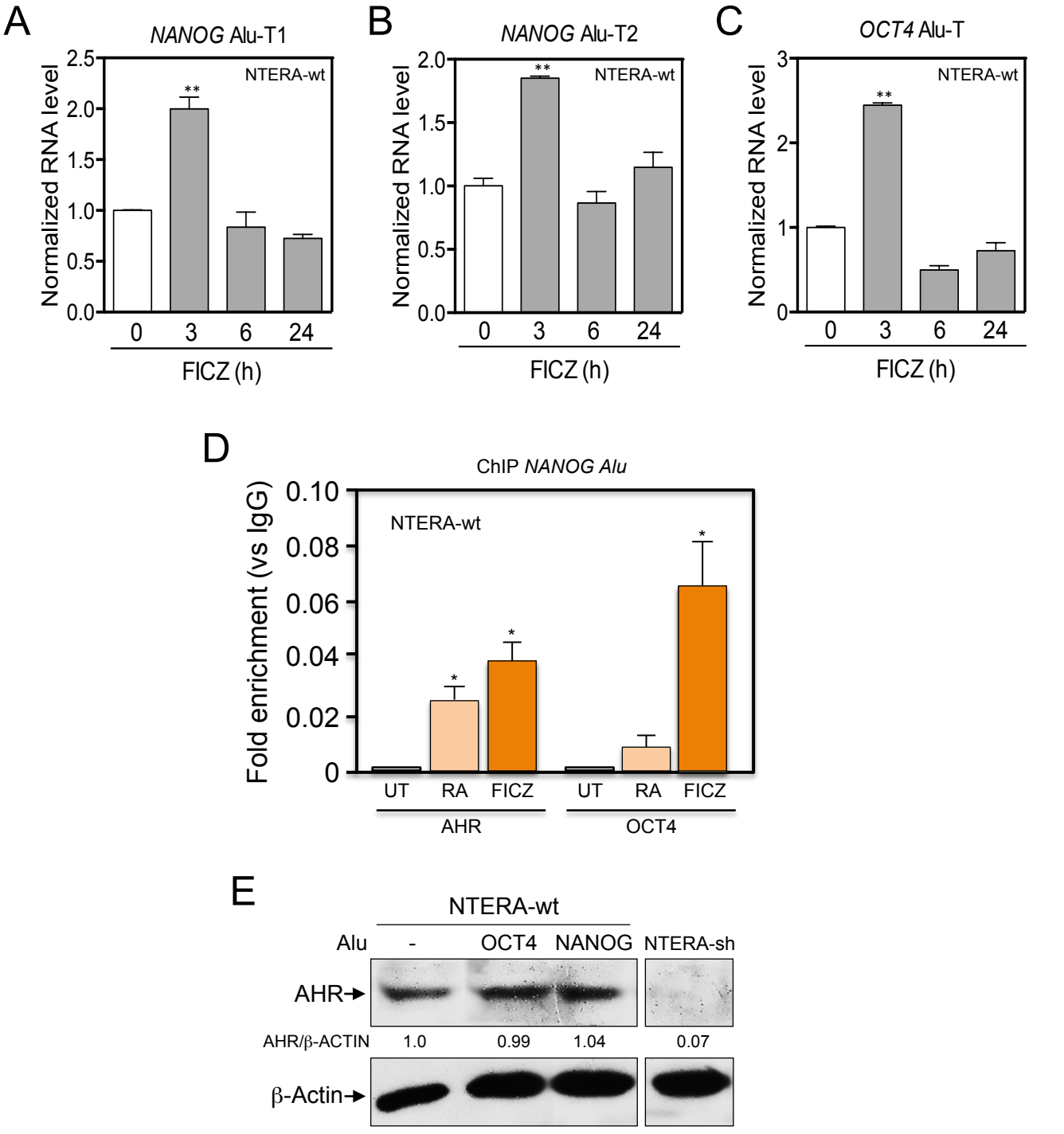


Supplementary Fig. 4. WYE132 and tagetitoxin are efficient inhibitors of RNA polIII in NTERA cells. (A) NTERA-wt and NTERA-sh cells were left untreated or treated with 1 μM RA for 48 h. Non-coding RNA (nc-RNA) libraries were made using molecules ranging from 75 to 250 bp in length. Libraries were sequenced as indicated in the Materials and Methods. **(B)** NTERA-wt cells were left untreated (UT), treated with 1 mM RA or with 1 mM RA+1 mM WYE132 for 48 or 72 h and the expression of *NANOG* and *OCT4* quantified by RT-qPCR using the oligonucleotides indicated in Additional file 9. **(C)** NTERA-wt cells were treated with 1 μM WYE132 for 3, 12, 24 and 48 h and the levels of *Leu* tRNA determined by real-time RT-qPCR using the oligonucleotides listed in Additional file 9. **(D)** *NANOG* and *OCT4* mRNA levels were determined by RT-qPCR in NTERA-wt cells treated with 1 μM WYE132 for the indicated times in the absence of RA. **(E)** Sequence analysis of *OCT4 Alu*-Sq and *NANOG Alu*-Sx for the presence of the core nucleotides reported to confer transposable activity to *Alu*S transposons (Bennett et al. 2008). Matching nucleotides are underlined and those differing from the consensus sequence are indicated by a red background. *GAPDH* mRNA was used to normalize gene expression (ΔCt) and $2^{-\Delta\Delta Ct}$ to calculate variations with respect to control or untreated conditions. Panels A,B: n = 4, two technical replicates in two biological replicates; panels C,D: n = 6, three technical replicates in two biological replicates; panel E: n = 4, two technical replicates in two biological replicates, *p<0.05 and ***p<0.001. Data are shown as mean ± SD.



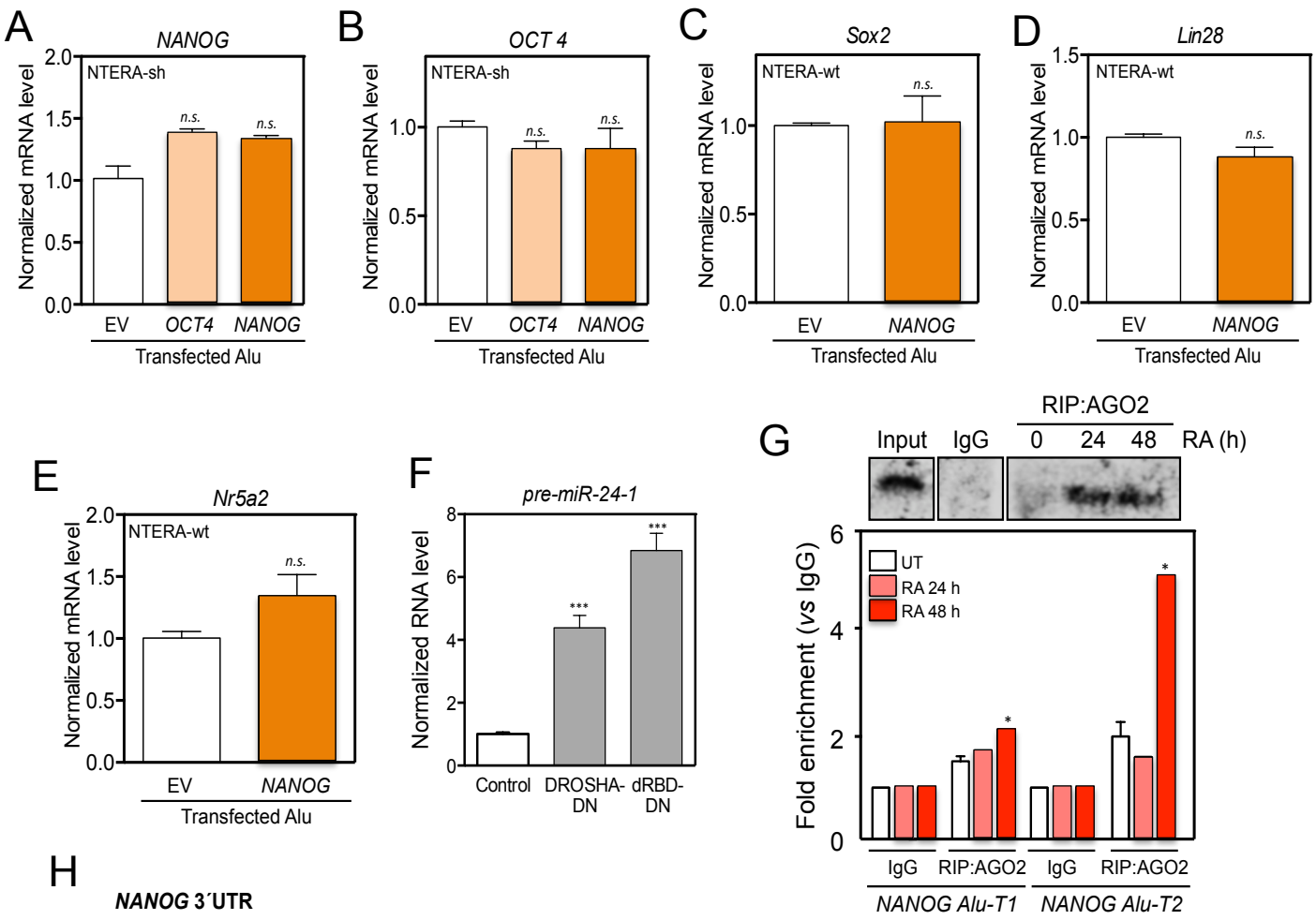
Supplementary Fig. 5. RNA sequencing and transcript-specific qPCR to quantify transcription of *NANOG* and *OCT4 Alus*. (A,B) *NANOG Alu*-T1 and *Alu*-T2 and *OCT4 Alu*-T profiles were used to select transcript specific oligonucleotides (red arrows and Additional file 9) to quantify *Alu*´s transcription by RT-qPCR. **(C)** Location and sequence of the transcripts *NANOG Alu*-T1 (yellow), *Alu*-T2 (red) and *OCT4 Alu*-T (cyan) produced from the *NANOG* and *OCT4* elements. The oligonucleotide sequences used for RT-qPCR are indicated by black (forward) and red (reverse) arrows. **(D)** Amplicons produced by RT-qPCR from the *NANOG* and *OCT4 Alus* were sequenced in both strands by fluorescent capillary dideoxi-sequencing. **(E)** Basal NTERA-wt cells were treated with 1 μM WYE132 or 10 mM TGT for 48 h and the mRNA levels of *OCT4 Alu*-T, *NANOG Alu*-T1 and *Alu*-T2 transcripts were determined by RT-qPCR. *GAPDH* mRNA was used to normalize gene expression (ΔCt) and $2^{-\Delta\Delta Ct}$ to calculate variations with respect to control or untreated conditions. Panel E: $n = 4$, two technical replicates in two biological replicates. *p<0.05 and **p<0.01. Data are shown as mean ± SD.

Supplementary Figure 6



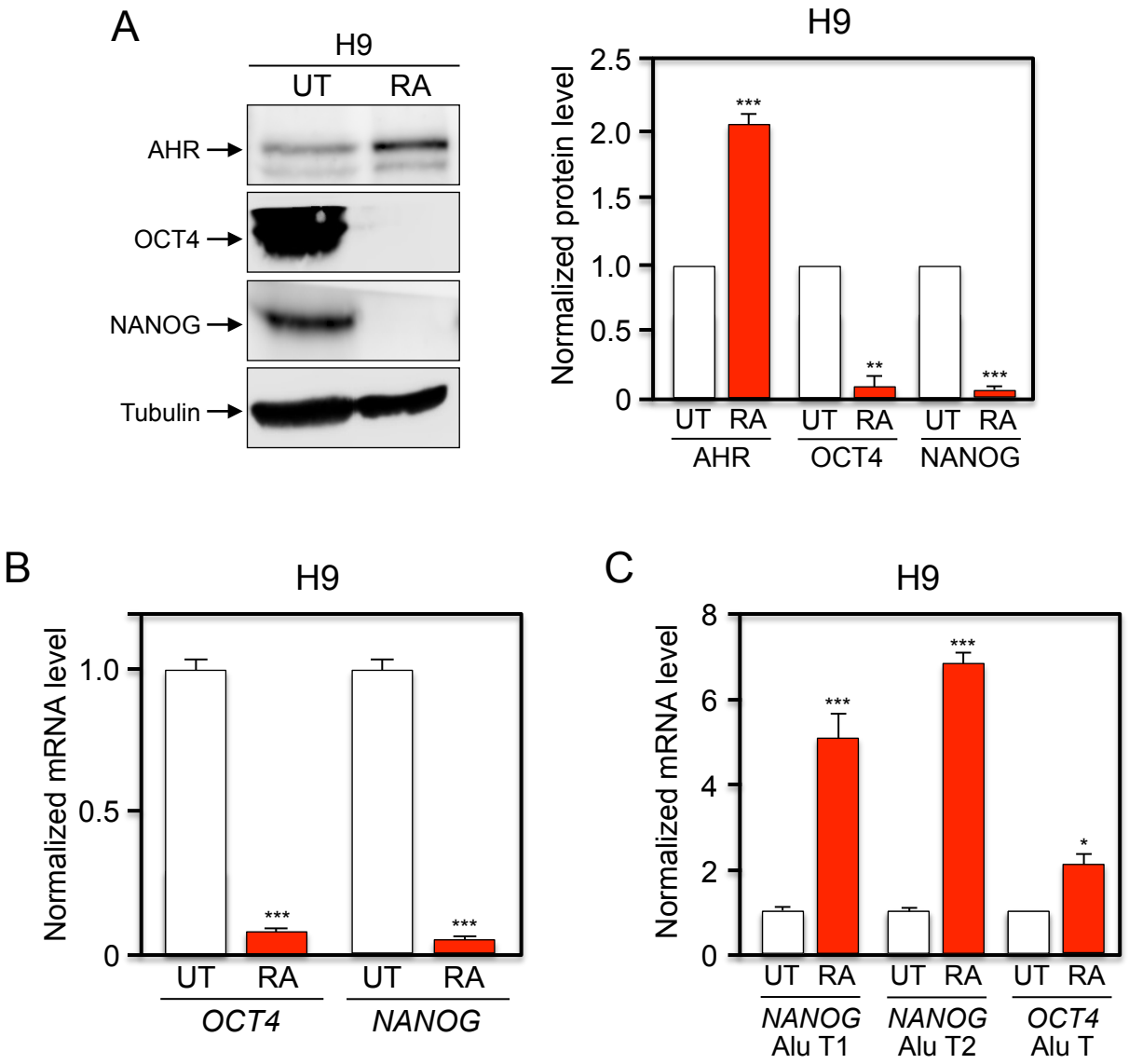
Supplementary Fig. 6. AHR activation by FICZ increases the expression of *Alu* derived ncRNAs. *Alus* can be specifically expressed in NTERA cells. (A-C) NTERA-wt cells were treated with 10 nM FICZ for up to 24 h and the expression of *NANOG Alu*-T1 **(A)**, *NANOG Alu*-T2 **(B)** and *OCT4 Alu*-T **(C)** was analyzed by RT-qPCR. **(D)** Binding of AHR and OCT4 to the *NANOG Alu* was analyzed by ChIP in NTERA-wt cells treated with 10 nM FICZ and quantified by qPCR using the oligonucleotides indicated in Supplementary Table 1. **(E)** NTERA-wt cells were transfected with *NANOG* and *OCT4 Alu*s and their effects on AHR protein levels analyzed by immunoblotting. NTERA-sh cells were used as negative control and β-Actin to normalize AHR expression. The oligonucleotides used are indicated in Additional file 9. *GAPDH* mRNA was used to normalize gene expression (ΔCt) and $2^{-\Delta\Delta Ct}$ to calculate variations with respect to control or untreated conditions. Panels A-C: n = 4, two technical replicates in two biological replicates; Panel D: n = 3, three biological replicates. *p<0.05 and **p<0.01. Data are shown as mean ± SD.

Supplementary Figure 7



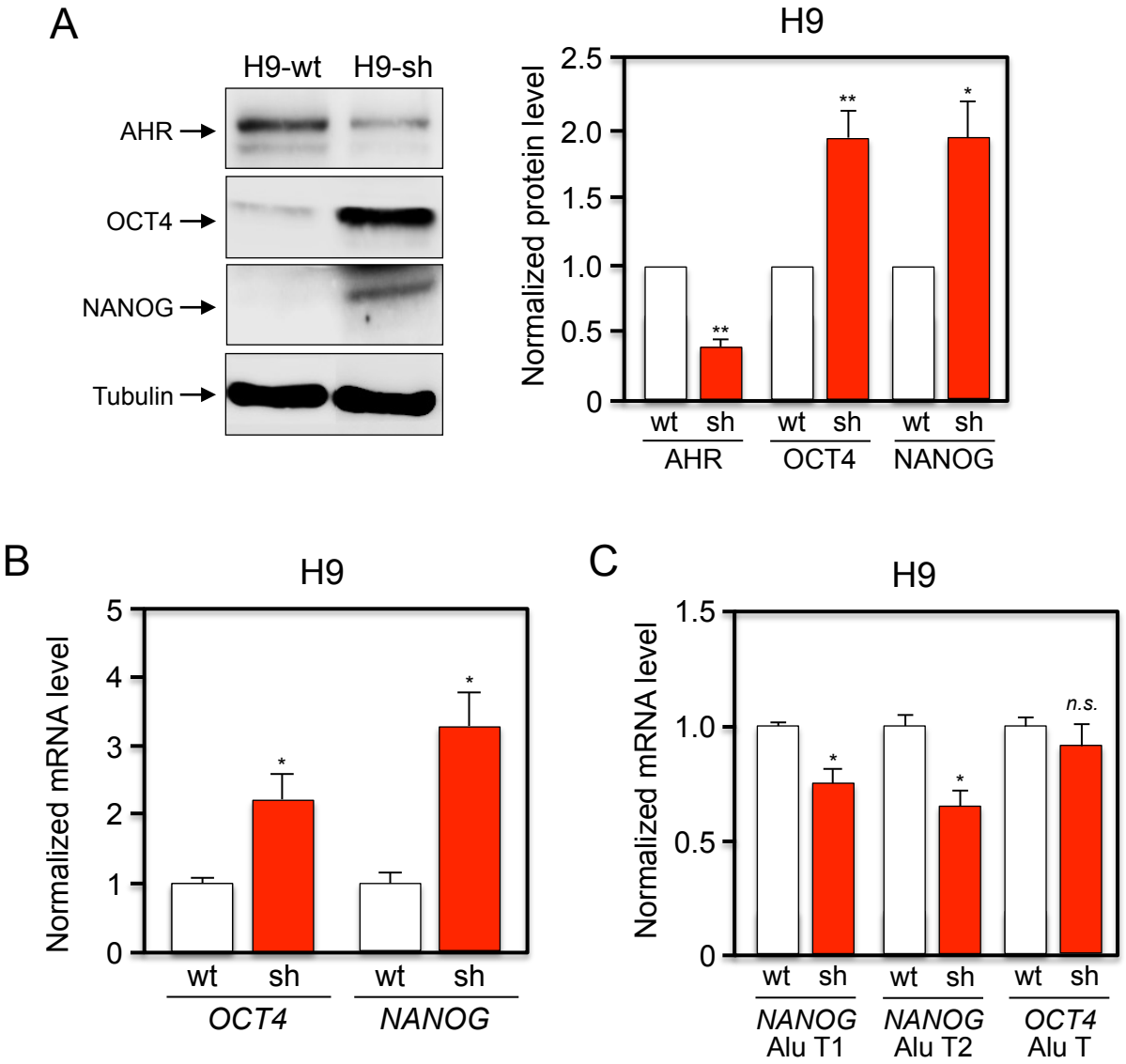
Supplementary Fig. 7. Gene repression by *NANOG Alu*-derived transcripts. Loading of *NANOG Alu*-T1 and *Alu*-T2 transcripts into AGO2 is time-dependent. (A,B) *NANOG* and *OCT4 Alu*s were individually transfected in NTERA-sh cells and the mRNA levels of *NANOG* and *OCT4* determined by RT-qPCR. **(C-E)** NTERA-wt cells were transfected with *NANOG Alu* and the mRNA expression of *SOX2*, *LIN28* and *NR5A2* determined by RT-qPCR. Control transfections were done with the empty pUC57 plasmid (EV). **(F)** NTERA-wt cells were transfected with dominant-negative (DN) expression vectors for DROSHA and DGCR8 and the levels of *pre-miR-24-1* determined by RT-qPCR using the oligonucleotides indicated in Supplementary Table 1. **(G)** RNA immunoprecipitation (RIP) was done in NTERA-wt cells left untreated or treated with 1 μM RA for 24 or 48 h using anti-AGO2 antibody. Loading of *NANOG Alu*-T1 and *Alu*-T2 ncRNA transcripts was analyzed by qPCR using the oligonucleotides indicated in Supplementary Table 1. **(H)** The *NANOG* 3´UTR was amplified from total RNA by RT-qPCR and cloned into the pGL2-Basic-Luc plasmid. This construct was used to produce 3´UTR deletion mutants lacking the homology region for the *NANOG Alu*-T1 (yellow), *Alu*-T2 (cyan) or both. *GAPDH* mRNA expression was used to normalize individual levels of gene expression (ΔCt) and $2^{-\Delta\Delta Ct}$ to calculate gene expression with respect to the control or untreated conditions. Panels A-E: n = 6, two technical replicates in three biological replicates; panel F: n = 4, two technical replicates in two biological replicates; panel H: n = 6, two technical replicates in three biological replicates. *p<0.05 and ***p<0.001. Data are shown as mean ± SD.

Supplementary Figure 8

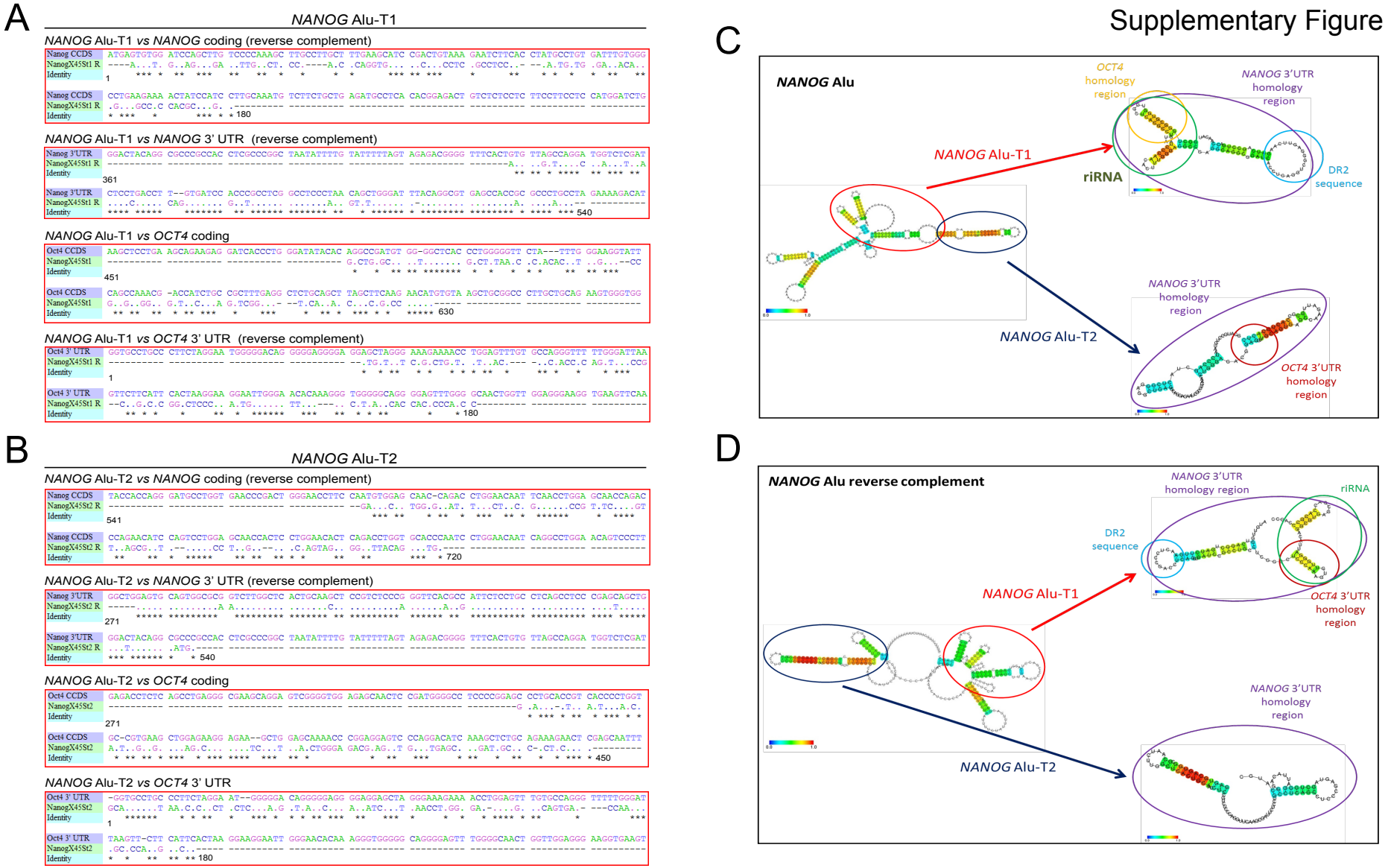


Supplementary Fig. 8. Retinoid acid treatment of H9 human embryonic stem cells represses *OCT4* and *NANOG* expression and increases the expression of *NANOG Alu*- and *OCT4 Alu*-derived transcripts. (A) H9 cells were left untreated (UT) or treated with 1 μM RA for 48 h (RA) and the protein expression of AHR, OCT4 and NANOG was analyzed by immunoblotting. Tubulin expression was used to normalize protein levels. Quantification of these results is shown in the right panel. **(B)** *OCT4* and *NANOG* mRNA levels were quantified by RT-qPCR in untreated or RA-treated H9 cells using the oligonucleotides indicated in Supplementary Table 1. **(C)** H9 cells left untreated of treated with 1 μM RA for 48 h were analyzed for the expression of the *NANOG Alu*-T1, *NANOG Alu*-T2 and *OCT4 Alu*-T by RT-qPCR using the oligonucleotides indicated in Supplementary Table 1. *GAPDH* mRNA was used to normalize gene expression (ΔCt) and $2^{-\Delta\Delta Ct}$ to calculate variations with respect to control or untreated conditions. Two technical replicates were performed in two biological replicates; *p<0.05, **p<0.01 and ***p<0.001 . Data are shown as mean ± SD.

Supplementary Figure 9



Supplementary Fig. 9. Transient AHR knockdown in H9 human embryonic stem cells upregulates *OCT4* and *NANOG* expression and reduces the expression of *NANOG Alu*- and *OCT4 Alu*-derived transcripts. (A) H9 cells were transiently transfected with an sh-RNA against AHR (sh) (see the Methods) and its effects on the protein levels of AHR, OCT4 and NANOG determined by immunoblotting. Wild type H9 cells (wt) were also analyzed in parallel. Tubulin expression was used to normalize protein levels. Quantification of these results is shown in the right panel. **(B)** *OCT4* and *NANOG* mRNA levels were quantified by RT-qPCR in wild type and AHR-knockdown H9 cells using the oligonucleotides indicated in Supplementary Table 1. **(C)** Wild type and AHR-knockdown H9 cells were analyzed for the expression of the *NANOG Alu*-T1, *NANOG Alu*-T2 and *OCT4 Alu*-T by RT-qPCR using the oligonucleotides indicated in Supplementary Table 1. *GAPDH* mRNA was used to normalize gene expression (ΔCt) and $2^{-\Delta\Delta Ct}$ to calculate variations with respect to control or untreated conditions. Two technical replicates were performed in two biological replicates; *p<0.05, **p<0.01 and ***p<0.001 . Data are shown as mean ± SD.



Supplementary Fig. 10. Homology of *NANOG Alu*-T1 and *Alu*-T2 transcripts with the 3´UTR and coding sequences of *NANOG* and *OCT4*. Conditional training analysis for RNA secondary structure prediction (CONTRAfold) of the *NANOG Alu*. **(A)** The *NANOG Alu*-T1 transcript was aligned with the coding sequences and 3´UTR regions of *NANOG* and *OCT4*. **(B)** The same analysis was done for the *NANOG Alu*-T2 ncRNA transcript. Alignments were done with the forward and reverse complementary strands. Asterisks indicate the coincidence of the same nucleotide in both sequences. **(C)** The forward sequence of the *NANOG Alu* was analyzed by CONTRAfold to estimate the secondary structures of its *Alu*-T1 and *Alu*-T2 transcripts. Homology with respect to the *NANOG* 3´UTR, *OCT4* 3´UTR and DR2 region are indicated for each transcript. **(D)** The same analysis was performed for the *NANOG Alu* reverse complementary strand. The following algorithm has been used: http://contra.stanford.edu/contrafold.

Supplementary Table 1

Primer sequences for mRNA expression

Gene name	Primer sequence (5´- 3´)
GAPDH	*Forward:* GAAGGTGAAGGTCGGAGTC *Reverse:* GAAGATGGTGATGGGATTTC
SOX2	*Forward:* TGCGAGCGCTGCACAT *Reverse:* TCATGAGCGTCTTGGTTTTCC
OCT4	*Forward:* GGAGGAAGCTGACAACAATGAAA *Reverse:* GGCCTGCACGAGGGTTT
NANOG	*Forward:* GATTTGTGGGCCTGAAGAAA *Reverse:* AAGTGGGTTGTTTGCCTTTG
NOTCH1	*Forward:* TGCCAGACCAACATCAAC *Reverse:* CTCATAGTCCTCGGATTGC
CYP1A1	*Forward:* GCTGACTTCATCCCTATTCT *Reverse:* GCTCCAGGAGATAGCAGTTG

Primer sequences for pre-MIR expression

Gene name	Primer sequence (5´- 3´)
pre-MIR24-1	*Forward:* AGCTGAGGCGCTGCTTCT *Reverse:* CCTCGGGCACTTACAGACA
pre-MIRLET7	*Forward:* GGGATGAGGTAGTAGGTTG *Reverse:* GGAAAGACAGTAGATTGTATAG
pre-MIR21	*Forward:* GCTTATCAGACTGATGTTGACTG *Reverse:* CAGCCCATCGACTGGTG
pre-MIR141	*Forward:* GGCCCTGGGTCCATCTTCC *Reverse:* ACCCGGGAGCCATCTTTACC
pre-MIR371	*Forward:* GTGGCACTCAAACTGTGG *Reverse:* GTAACACTCAAAAGATGGCGG
pre-MIR424	*Forward:* GGGGATACAGCAGCAATTC *Reverse:* CCACCTTCTACCTTCCCC

Primer sequences for Alu transposons expression

Gene name	Primer sequence (5′- 3′)
OCT4 Alu-T	*Forward:* CAGGTGGATCACGAGGTCA *Reverse:* TGCCACCACACCTGGC
NANOG Alu-T1	*Forward:* CGCCTGTAATCCCAACACTT *Reverse:* GGCTGATCTTGAACTCCCG
NANOG Alu-T2	*Forward:* CCTGTAATCCCAGCTACTC *Reverse:* GCAATCTTGGCTCACTGC

Primer sequences for ChIP assays

Gene name	Primer sequence (5′- 3′)
OCT4x36s	*Forward:* CCCATTTCCTGGATTTGAGA *Reverse:* CTTAAGCTCCCAAGCCTTCC
NANOGx45s	*Forward:* GCCAAAACCCTGTTTCCTTA *Reverse:* CGGCCTCTGCTCATCTTTTA
NANOGx45s2	*Forward:* AGGAGAATCGCTTGAACCTG *Reverse:* CGGCCTCTGCTCATCTTTTA

Primer sequences for DNA sequencing

Primer name	Primer sequence (5′- 3′)
pUC/M13	*Forward:* CCCAGTCACGACGTTGTAAAACG *Reverse:* CAGGAAACAGCTATGAC

Primer sequences for NANOG 3′UTR cloning

Gene name	Primer sequence (5′- 3′)
Nanog 3'UTR:	*Forward:* GGTACCCGTGTGAAGATGAGTGAAAC *Reverse:* GAGCTCATGTTTAAGCTGTATATTTACTCATTG

Primer sequences for Leu tRNA

Gene name	Primer sequence (5′- 3′)
Leu tRNA:	*Forward:* ATGGCCGAGTGGTCTAAGG
	Reverse: ACCAGAAGACCCGAACACAG