Differential expression profiles of IncRNAs and mRNAs during male germ-like cell differentiation in human umbilical mesenchymal stem cells by RNA-seq

Lichun Xie  
Shenzhen Children's hospital

Guichi Zhou  
Shenzhen Children's Hospital

Lian Ma  
Shenzhen Children's Hospital

Feiqiu Wen (fen62@163.com)  
Shenzhen Children's hospital  https://orcid.org/0000-0001-5559-633X

Research

Keywords: Human umbilical mesenchymal stem cell, long-noncoding RNA, male germ-like cell, cell differentiation, RNA sequencing

DOI: https://doi.org/10.21203/rs.3.rs-281787/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background: Long non-coding RNAs (lncRNAs) are key regulators of various biological processes and crucial for cell development and differentiation. However, their roles in the differentiation of human umbilical mesenchymal stem cells (HUMSCs) into male germ-like cells remain largely unknown.

Method: Here, the expression of lncRNAs and mRNAs in undifferentiated HUMSCs and HUMSCs undergoing differentiation into male germ-like cells was analyzed. RNA-sequencing was performed to profile the expression of non-coding RNAs. We analyzed the total expression of lncRNAs/mRNAs at three time points during HUMSC differentiation [day (D)7, D14, and D21].

Result: Expression profiling revealed 110 lncRNAs, 584 mRNAs, and 21 miRNAs common to the three experimental groups during HUMSC male germ-like cell differentiation. The maximum and minimum total overall lncRNA expression occurred on D14 (638) and D21 (283), respectively. The maximum and minimum numbers of up-regulated mRNAs were observed on D21 (2,398) and D7 (2,106), respectively. The maximum and minimum numbers of down-regulated mRNAs were observed on D14 (3,357) and D21 (202), respectively. The expression level of mRNA ENST00000486554 was up-regulated on D7, D14, and D21 after induction. Pathway analysis identified meiotic signaling pathways and nitrogen metabolism as being associated with the differentiation potential of HUMSC male germ-like cells. Non-coding RNA expression profiles significantly differed in HUMSC male germ-like cell differentiation. One mRNA, ENST00000486554, was crucial for differentiation.

Conclusions: Our results provide a systematic perspective on the potential functions of non-coding RNAs and novel insights into the complicated regulatory mechanisms underlying the differentiation of HUMSCs into male germ-like cells.

1 Introduction

Human umbilical mesenchymal stem cells (HUMSCs) are considered a promising cell type in regenerative therapy for male infertility. A previous study showed that HUMSCs can be induced to differentiate into male germ-like cells under certain conditions (1–3). In addition, several studies have shown that male germ cells express a large number of different non-coding RNAs, such as long non-coding RNAs (lncRNAs) and small RNAs, which play key roles in male germ cell-specific processes (4).

To date, the roles of lncRNAs and mRNAs during male germ-like cell differentiation of HUMSCs remain largely unknown. In this study, we constructed a co-expression network to identify potential target relationships between the lncRNAs and mRNAs based on microarray analysis and bioinformatic predictions.

2 Materials And Methods
2.1 Animals, subjects, and ethics statement

Five- to seven-day-old male Kunming mice were obtained from the Laboratory Animal Centre of Shantou University Medical College. The animals were maintained in a temperature- and humidity-controlled room and given free access to water and food. The Five- to seven-day-old mice were sacrificed by CO$_2$ asphyxiation, and the testes were collected for subsequent experiments.

All animal experiments were conducted in accordance with protocols approved by the Animal Care and Use Committee of Shantou University Medical College (SUMC-45-2014).

Human umbilical cords were obtained from women delivering full-term male infants by caesarean section.

Signed informed consent was obtained from each mother, and the study was approved by the Human Ethics Committee of the Shenzhen Children's Hospital (approval No. SUMC-37-2014 and performed in accordance with the ethical standards on human experimentation.

2.2 Cell culture and identification

2.2.1 Preparation of HUMSC cultures

HUMSCs were isolated, cultured, and identified as previously described (1, 3). Briefly, Wharton's jelly tissue from individual umbilical cords was cut into 1–2-mm$^2$ pieces, seeded into 24-well plates (Corning, Inc., Corning, NY, USA), and cultured in Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM-F12) medium supplemented with 2% penicillin/streptomycin, 10% fetal bovine serum (FBS, Gibco, Grand Island, NY, USA), 5 ng/ml epidermal growth factor (R&D Systems, Minneapolis, MN, USA), and 5 ng/ml basic fibroblast growth factor (Sigma-Aldrich, St. Louis, MO, USA). HUMSCs were cultured in a humidified 5% CO$_2$ atmosphere at 37°C without disturbance for 5–7 days, and the medium was changed every 2 days. The cells were preserved via serial passage when they reached 80–90% confluence.

2.2.2 Preparation of testicular cell-conditioned medium

Testicular cell-conditioned medium was prepared as described previously (1). Briefly, the testes were isolated from 5–7-day-old male Kunming mice. The testes were decapsulated and disrupted physically into tissue fractions. Sertoli cells were obtained by repeated washing and centrifugation steps to remove the supernatant containing tissue debris. The cells were cultured in DMEM-F12 growth medium supplemented as described above in a humidified 5% CO$_2$ atmosphere at 37°C. This conditioning medium was collected on day 10 of primary culture, and every 3 days thereafter until day 16. The conditioning medium was centrifuged, filtered, and stored at −20°C or used immediately.

2.2.3 Male germ-like cell differentiation of HUMSCs

The HUMSCs used in each experiment were from passages 3–5. The male germ-like cell differentiation medium consisted of HDMEM-F12 (Gibco) supplemented with 5% FBS, 50% filtered testicular cell-
conditioned medium, 2 mM all-trans retinoic acid (RA, Sigma-Aldrich) and 1 mM testosterone (Solarbio, Beijing, China). HUMSCs were grown in differentiation medium for 7, 14, or 21 days; half of the differentiation induction medium was replaced every 3–4 days. HUMSCs of the control group were cultured in normal medium (DMEM-F12 containing 5% FBS and 2% penicillin/streptomycin).

2.3 RNA extraction

Total RNA was extracted using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer’s instructions. RNA purity and concentration were evaluated with a Nano Drop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). RNA quality was verified using a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) and examined via RNase-free agarose gel electrophoresis. In addition, total RNA concentrations were measured using a 2100 Bioanalyzer at 260 and 280 nm. The results indicated that the extracted RNA was suitable for downstream processing.

2.4 Raw read processing

The raw reads were generated by removing adaptor sequences (overlap = 5) using CutAdapt, poor-quality sequences (< 10% ‘N’ bases and > 85% QA > 20 bases) using the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/index.html), and ribosomal sequences using Tophat (4).

2.5 RNA-sequencing and data analysis

The extracted RNA samples were used for cDNA synthesis. Total mRNA was isolated by removing rRNA using a Ribo-Zero™ Magnetic Gold Kit (Epicentre, Madison, WI, USA), and the resultant total mRNA was broken into short fragments (~ 300 nt) by adding fragmentation buffer. First-strand cDNA was generated by random hexamer-primed reverse transcription. Next, second-strand cDNA was synthesized using RNase H and DNA polymerase I. The dsDNA fragments were washed with buffer prior to end-repair and poly (A) addition and ligated to strand-marker adapters. Following polymerase chain reaction and sequencing adapter ligation, the final cDNA library was generated (IlluminaRibo-Zero Gold [MRZG12324, Illumina, San Diego, CA, USA]). The cDNA library was sequenced on an Illumina sequencing platform (IlluminaHiSeq™ 2500, Illumina) using paired-end technology in a single run. Processing of the original image for sequence identification, base calling, and quality value calculations was performed with the IlluminaHiSeq 2500, and 125-bp paired-end reads were obtained. Three undifferentiated cell samples at day 0 (CK1-1) were used as the control group, and three samples of differentiated male germ-like cells at days 7 (T1-1), 14 (T2-1), and 21 (T3-1) were used as the experimental groups.

2.6 Profiles of lncRNA and mRNA expression

Reads that could be mapped to a gene were used to calculate expression levels. Gene expression levels were measured according to the number of mapped fragments (paired reads) per kilobase of exon region per million mapped reads (RPKM) using the formula shown below:

\[
\text{RPKM} = \frac{10^6 \text{C}}{\text{NL}/10^3} \quad \text{RPM} = \frac{10^6 \text{C}}{\text{NL}/10^3}
\]
After the expression level of each gene was determined, we compared the same genes between each combination of two samples to analyze the expression variance of lncRNAs and mRNAs. The false discovery rate (FDR) was calculated to determine the p value threshold in multiple tests; for this analysis, a threshold of $\text{FDR} \leq 0.001$ and absolute value of log$_2$ fold-change $\geq 1$ were used to assess the significance of the transcript expression differences. To compare more than two samples, the union set (or intersection set) of all differentially expressed genes was used for further analysis.

2.7 lncRNA-gene correlation analysis

We analyzed the correlations between lncRNAs that were consistently up- or down-regulated in more than 50% of samples with all mRNAs detected at three differentiation time points. The expression levels (FPKM values) of the lncRNAs and genes in each group were used to calculate correlation coefficients (R values). First, the fold-change method was used to calculate fold-changes and a negative binomial distribution was utilized to evaluate the FDR for each lncRNA and gene. Next, an FDR $\leq 0.05$ and absolute value of log$_2$ fold-change $\geq 2$ were used as thresholds to evaluate the significance of the expression differences and identify significant correlations of the lncRNAs and genes between different samples. The absolute value of the Pearson correlation coefficient was used to calculate the distance between lncRNAs, and the pairwise average-linkage method was used for hierarchical clustering. Common up-regulated or down-regulated transcripts at the three time points during differentiation were determined by Venn analysis.

2.8 Gene ontology (GO) and pathway analysis

GO analysis was conducted based on Gene Ontology (www.geneontology.org) which provides three different structured networks, namely biological process, cellular component, and molecular function, to analyze the probable functions of the differentially expressed lncRNAs from the functions of their target genes. In addition, pathway analysis of the differentially expressed mRNAs was performed using the latest Kyoto Encyclopedia of Genes and Genomes (KEGG) database. GO terms and KEGG pathways with a p value $< 0.05$ were considered as significantly enriched.

2.9 Construction of coding–non-coding gene co-expression network

We constructed a coding–non-coding network based on the correlation analysis of normalized signal intensity between the lncRNAs and mRNAs that are often differentially expressed during differentiation of HUMSCs male germ-like cells, according to the clustering coefficient and degree. lncRNAs and mRNAs with the most significant Pearson correlation coefficients (at least 0.99) were used to create the network using the open-source bioinformatics software Cytoscape (Institute of Systems Biology, Seattle, WA, USA).

3 Results
3.1 Identification of differentiated HUMSC male germ-like cells

Differentiated HUMSC male germ-like cells were identified as described previously (1). On day 14, the HUMSCs formed “tadpole-like” cells in differentiation medium and showed both mRNA and protein expression of the male germ cell-specific markers Oct4 (POUF5), Ckit, CD49f (a6), Stella (DDPA3), and Vasa (DDX4) (1–3).

3.2 Differential expression of lncRNAs, mRNAs, and miRNAs induced in HUMSCs

We identified the expression profiles of lncRNAs and mRNAs in male germ-like cell-differentiated HUMSCs by comparing them with the expression profile of paired undifferentiated HUMSCs. Cluster analysis showed that the expression levels of lncRNAs and mRNAs significantly differed between the induced and control groups (Fig. 1). Following male germ-like cell induction, 1456 differentially expressed lncRNAs were detected; the entire expression profile is presented in Fig. 1a. The lncRNAs were predominantly down-regulated. The maximum lncRNA expression occurred on D14 (638), whereas minimum expression occurred on D21 (283; Fig. 2a).

In addition, 12,712 mRNAs were identified; the whole expression pattern is shown in Fig. 1b. Among them, 2106 and 2286 mRNAs were up-regulated and down-regulated, respectively, in T1-1 samples compared to those in CK-1 samples; 2363 and 3357 mRNAs were up-regulated and down-regulated, respectively, in T2-1 samples compared to those in CK-1 samples; and 2398 and 202 mRNAs were up-regulated and down-regulated, respectively, in T3-1 samples compared to those in CK-1 samples (Fig. 2b); the mRNAs were predominantly up-regulated. The number of up-regulated mRNAs reached a maximum on D21 (2,398) and minimum on D7 (2,106).

A total of 610 differentially expressed miRNAs were detected (Fig. 2c), and they were predominantly down-regulated. For down-regulated miRNAs, the maximum and minimum numbers were observed on D14 (3357) and D21 (202), respectively.

Scatter plots showing differentially expressed lncRNAs between male germ-like cells at different stages and undifferentiated HUMSCs are shown in Fig. 3. Additionally, a Venn diagram showing differentially expressed lncRNAs, mRNAs, and miRNAs in three comparisons was also constructed (Fig. 4; CK-1 vs. T1-1, CK-1 vs. T2-1, and CK-1 vs. T3-1).

3.3 GO analysis

The common up-regulated GO functions at all stages related to male germ-like cell differentiation, including the regulation of immunoglobulin production, immune responses, and metabolic processes, were investigated. Pathway analysis indicated that the nitrogen metabolism, ubiquitin-mediated proteolysis, and cytosolic DNA-sensing pathways, which are closely related to cellular activities, were...
significantly enriched during HUMSC male germ-like cell differentiation. In contrast, the common down-regulated GO functions were related to male meiosis, protein modification, and protein ubiquitination. These pathways were enriched with down-regulated genes involved in the MAPK signaling pathway, homologous recombination, and collecting duct acid secretion.

As lncRNAs can regulate genes in cis and trans, GO analysis of the common regulated cis-acting lncRNAs showed that mitotic G1 phase, meiosis, and hippo signaling were involved; the results of pathway analysis showed that cell cycle, meiosis, MAPK signaling pathway, and hippo signaling pathway were involved. GO analysis of the common regulated trans-acting lncRNAs showed that RNA transcription and pyruvate transport were involved; the results of pathway analysis showed that adherence junction, RNA polymerase, and Fc gamma R-mediated phagocytosis were involved (Fig. 5).

3.4 Pathway analysis

Enriched KEGG pathways (p < 0.05) were selected and ranked by gene counts. Twenty significant signaling pathways for the up-regulated and down-regulated lncRNAs are listed in Fig. 6.

3.5 Candidate mRNA of frequently regulated mRNAs

Based on our results, we found that the commonly regulated mRNA that was up-regulated in all comparison groups was mRNA ENST00000486554, which is encoded by the glucocorticoid-induced leucine zipper gene (Gilz/Tsc22d3-2). As reported previously (5), the Tsc22d3-2 gene plays a crucial role in male fertility.

3.6 Co-expression network of lncRNA-mRNA

A total of 79 lncRNAs and 67 mRNAs involved in 3180 pathways were used to generate a network map for identifying the interactions between lncRNAs and mRNAs that are differentially expressed for both and the interactions between the pathways. The network indicated the existence of a complex regulatory relationship between lncRNAs and mRNAs, where one lncRNA could regulate one or more core genes in different ways, and vice versa (Fig. 7).

4 Discussion

HUMSCs can be easily isolated and differentiated into male germ-like cells in vitro (1). Although our study showed that HUMSCs can be differentiated into male germ-like cells, these cells could not fully undergo germ cell differentiation, and hence, further study of the molecular mechanisms underlying differentiation is required (1–3, 6–8). The role of epigenetic regulation, including that mediated via various RNAs in the HUMSC-derived germ-like cell differentiation is not well-understood. The complex regulatory role of RNAs acts as a “bridge” between DNA and proteins (4). Among all non-coding RNAs, miRNAs (20–24 nt) have been shown to regulate cellular functions (9, 16). miRNAs bind to the 3′-untranslated regions of target mRNAs called miRNA-binding elements and inhibit protein translation or modulate mRNA stability at the
post-transcriptional level. IncRNAs are known to function via their interaction with DNA, RNA, or proteins by a variety of mechanisms, such as by altering the expression of coding genes via modulation of transcription (10, 11) and by serving as direct enhancers or as RNAs with enhancer-like functions (12, 13). Although studies have shown that IncRNAs are important regulators of human tissue regeneration, such as regulating stem cell differentiation (14–16), the functions of IncRNAs in the male germ-like cell differentiation of HUMSCs are largely unknown. In this study, we performed RNA-seq to investigate the transcriptome data set of HUMSC-derived male germ-like cell differentiation. The differential expression profiles of IncRNAs and mRNAs were significantly different (p < 0.05) between the induction and control groups.

The commonly expressed 110 IncRNAs, 584 mRNAs, and 21 miRNAs were altered at all time points during HUMSC-derived male germ-like cell differentiation. The significant differences (p < 0.05) between the experimental and control groups indicated that IncRNAs play a crucial role in the differentiation of HUMSCs. Therefore, the results of our study improve the understanding of the roles played by IncRNAs and mRNAs in HUMSCs during male germ-like cell differentiation.

We observed that a significant proportion of identified GO terms was related to the regulation of cell cycle and metabolism. As previous studies have shown that metabolism is important for cell development (17), this result indicates the important role of organelles in stem cell differentiation (i.e., coupling of organelle inheritance with mitosis to balance growth and differentiation). Furthermore, the results of pathway analysis showed that among the most significantly regulated pathways, metabolism and biosynthesis may play core roles in the differentiation of HUMSCs.

In our study, the expression of mRNA ENST00000486554 was up-regulated at days 7, 14, and 21 after induction compared to that in undifferentiated HUMSCs, indicating its importance in HUMSC differentiation into male germ-like cells. The mRNA, ENST00000486554 is encoded by TSC22D3-2. A previous study reported that loss of TSC22D3-2 leads to male sterility (5), and our data supports the suggested role of TSC22D3-2 in HUMSC differentiation into male germ-like cells. In addition, our pathway analysis suggests that meiosis signaling pathways and nitrogen metabolism are involved in the differentiation of HUMSCs. Thus, our results are consistent with those of previous reports and show that TSC22D3-2 plays a crucial role in testis development and male fertility. The underlying mechanisms of the identified IncRNAs should be further examined.

Abbreviations

IncRNAs, long non-coding RNAs; HUMSCs, human umbilical mesenchymal stem cells; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes

Declarations

Ethics approval and consent to participate
All animal experiments were conducted in accordance with protocols approved by the Institutional Animal Care and Use Committee of Shantou University Medical College. The study was approved by the Human Ethics Committee of the Shenzhen Children's Hospital (approval No. SUMC-37-2014) and performed in accordance with the ethical standards on human experimentation. Signed informed consent was obtained from each participant.

**Consent for publication**

Not applicable.

**Availability of data and materials:**

The datasets generated and/or analyzed during the current study are available in the NCBI Sequence Read Archive [https://www.ncbi.nlm.nih.gov/sra/PRJNA641247].

**Funding**

This work was supported by The National Natural Science Foundation of China [grant no. 81671525] and Sanming Project of Medicine in Shenzhen [grant no. SZSM201512033].

**Authors’ contributions**

XLC designed the data analysis and wrote the manuscript; LQL and ZGC provided imaging assistance; ML and WFQ contributed to critical revision. All authors have read and approved the final manuscript.

**Acknowledgements**

Not applicable

**References**

1. Xie L, Lin L, Tang Q, Li W, Huang T, Huo X, Liu X, Jiang J, He G, Ma L. Sertoli cell-mediated differentiation of male germ cell-like cells from human umbilical cord Wharton's jelly-derived mesenchymal stem cells in an in vitro co-culture system. Eur J Med Res. 2015;20:9.

2. Chen H, Tang QL, Wu XY, Xie LC, Lin LM, Ho GY, Ma L. Differentiation of human umbilical cord mesenchymal stem cells into germ-like cells in mouse seminiferous tubules. Mol Med Rep. 2015;12:819–28.

3. Huang P, Lin LM, Wu XY, Tang QL, Feng XY, Lin GY, LinX, Wang HW, Huang TH, Ma L. Differentiation of human umbilical cord Wharton's jelly-derived mesenchymal stem cells into germ-like cells in vitro. J Cell Biochem. 2010;109:747–54.

4. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013;14:R36.
5. Suarez PE, Rodriguez EG, Soundararajan R, Merillat AM, Stehle JC, Rotman S, Roger T, Voirol MJ, Wang J, Gross O, et al. The glucocorticoid-induced leucine zipper (gilz/Tsc22d3-2) gene locus plays a crucial role in male fertility. Mol Endocrinol. 2012;26:1000–13.

6. Wang HS, Hung SC, Peng ST, Huang CC, Wei HM, Guo YJ, Fu YS, Lai MC, Chen CC. Mesenchymal stem cells in the Wharton's jelly of the human umbilical cord. Stem Cells. 2004;22:1330–7.

7. Hsieh JY, Fu YS, Chang SJ, Tsuang YH, Wang HW. Functional module analysis reveals differential osteogenic and stemness potentials in human mesenchymal stem cells from bone marrow and Wharton's jelly of umbilical cord. Stem Cells Dev. 2010;19:1895–910.

8. Mitchell KE, Weiss ML, Mitchell BM, Martin P, Davis D, Morales L, Helwig B, Beerenstrauch M, Abou-Easa K, Hildreth T, et al. Matrix cells from Wharton's jelly form neurons and glia. Stem Cells. 2003;21:50–60.

9. Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J. Natural RNA circles function as efficient microRNA sponges. Nature. 2013;495:384–8.

10. Knowling S, Morris KV. Non-coding RNA and antisense RNA. Nature's trash or treasure? Biochimie. 2011;93:1922–7.

11. Costa FF. Non-coding RNAs: meet thy masters. Bioessays. 2010;32:599–608.

12. Mattick JS, Gagen MJ. The evolution of controlled multitasked gene networks: the role of introns and other noncoding RNAs in the development of complex organisms. Mol Biol Evol. 2001;18:1611–30.

13. Mattick JS. Linc-ing long noncoding RNAs and enhancer function. Dev Cell. 2010;19:485–6.

14. Ørom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, Lai F, Zytnicki M, Notredame C, Huang Q, et al. Long noncoding RNAs with enhancer-like function in human cells. Cell. 2010;143:46–58.

15. Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe. Nat Rev Genet. 2009;10:94–108.

16. Guttman M, Rinn JL. Modular regulatory principles of large non-coding RNAs. Nature. 2012;482:339–46.

17. Shyh-Chang N, Ng HH. The metabolic programming of stem cells. Genes Dev. 2017;31:336–46.