Analysis of Differential Translation Profiles of Human Bone Marrow Mesenchymal Stem Cells during Osteogenic Differentiation

Hua LIU[1], Zhi Peng FAN[2,3,4], Qiu Bo YANG[1], Hui Na LIU[2,5]

Objective: *To explore the differential translation profiles and coding products of human jaw bone marrow mesenchymal stem cells (h-JBMMSCs) during osteogenic differentiation.*

Methods: *Ribo-seq was used to examine the differential translated genes (DEGs), open reading frames (ORFs) and genes associated with the osteogenic differentiation phase of h-JBMMSCs. Western blotting (WB) was performed to detect the expression of osteocalcin (OCN) and bone sialoprotein (BSP). Alkaline phosphatase (ALP) activity and alizarin red staining (ARS) were used to detect osteogenic differentiation. A lentivirus containing 5'UTR-ORF-GFPmut was designed to transfect h-JBMMSCs, and fluorescence and green fluorescent protein (GFP) expression were analysed. The SNHG1 peptide was synthesised for osteogenic induction and to detect osteogenic markers.*

Results: *A total of 53,432 ORFs were detected and 199 candidate translation sORFs, including lncRNA SNHG1, were identified after removing the annotated protein-coding genes. In addition, the 5'UTR-ORF-GFPmut showed green fluorescence and expressed GFP. Knockdown of the lncRNA SNHG1 increased the ALP activity of h-JBMMSCs, promoted the expression of OCN and BSP, and enhanced the intensity of ARS and calcium ion content. However, overexpression of lncRNA SNHG1 and the SNHG1 polypeptide inhibited the osteogenic differentiation of h-JBMMSCs.*

Conclusion: *LncRNA SNHG1 inhibited the osteogenic differentiation of h-JBMMSCs. LncRNA SNHG1 can encode a peptide of 19-amino acid and inhibit the osteogenic differentiation of h-JBMMSCs.*

Keywords: *h-JBMMSCs, lncRNA SNHG1, osteogenic differentiation, peptide, ribosome profiling*

Chin J Dent Res 2025;28(1):31–43; doi: 10.3290/j.cjdr.b6097608

1 Beijing Stomatological Hospital, School of Stomatology, Capital Medical University, Beijing, P.R. China.
2 Beijing Key Laboratory of Tooth Regeneration and Function Reconstruction, Beijing Stomatological Hospital, School of Stomatology, Capital Medical University, Beijing, P.R. China.
3 Beijing Laboratory of Oral Health, Capital Medical University, Beijing, P.R. China.
4 Research Unit of Tooth Development and Regeneration, Chinese Academy of Medical Sciences, Beijing, P.R. China.
5 Department of General Dentistry and Integrated Emergency Dental Care, Capital Medical University School of Stomatology, Beijing, P.R. China.

Corresponding authors: Dr Qiu Bo YANG, Beijing Stomatological Hospital, School of Stomatology, Capital Medical University, Beijing 100050, P.R. China. Tel: 86-10-57099114. Email: qiuboyang2003@163.com

Dr Hui Na LIU, Beijing Key Laboratory of Tooth Regeneration and Function Reconstruction, Beijing Stomatological Hospital, School of Stomatology, Capital Medical University, Beijing 100050, P.R. China. Tel: 86-10-57099114. Email: 948581537@qq.com

Bone marrow mesenchymal stem cells (BMMSCs) are generally derived from the jaws, long bones (the femur and tibia) and ilium, and BMMSCs from different bone sites have different characteristics.[1,2] Their multipotency allows them to differentiate into osteoblasts, adipocytes and chondrocytes, which can be used to treat bone defects or cartilage injuries.[3,4] Jaw BMMSCs have stronger odontogenic and osteogenic abilities, and are considered more suitable for maxillofacial stem cell therapy.[5,6] Therefore, JBMMSCs are more suitable for jaw defect reconstruction, periodontal tissue regeneration and dentine regeneration.[7] However, osteopet-

This study was supported by grants from the National Natural Science Foundation of China (82130028), National Key Research and Development Program (2022YFA1104401, 2022YFC2504201) and the National Natural Science Foundation of China (82100970).

rosis is a bone developmental disorder characterised by extremely dense bones and easy fracture. It often causes delayed or failed tooth eruption, root deformity and mandibular osteomyelitis in the maxillofacial region. Thus, the osteogenic differentiation of JBMMSCs has been explored to provide ideas for treating osteopetrosis.

At present, many studies have shown that the multidirectional differentiation of BMMSCs, especially osteogenic differentiation, is regulated by noncoding RNAs (ncRNAs).[8-10] Long ncRNAs (lncRNAs) are more than 200 nucleotides (nt) in length and can regulate osteogenic differentiation by regulating transcription and translation.[8,9] For example, heterogeneous ribonucleoprotein K regulates lncRNA-OG to promote the osteogenic differentiation of BMMSCs.[10] lncRNA-H19 can enhance the osteogenic differentiation of postmenopausal osteoporosis patients by regulating forkhead box C2 (Foxc2).[11] The lncRNA HOTAIR inhibits the osteogenic differentiation of BMMSCs by negatively regulating miR-378g.[12] Apart from being therapeutic targets, a few studies have indicated that lncRNAs have the potential to serve as biomarkers for bone-related diseases. For example, the lncRNA DANCR can be a diagnostic biomarker for postmenopausal osteoporosis, and the lncRNA HOTAIR may be used for the clinical diagnosis of rheumatoid arthritis.[13-15] However, clinical translation based on RNA therapy is impeded by delivery, tolerance and specificity issues.[16] Polypeptide drugs can effectively circumvent these problems, possess strong penetration and low immunogenicity, and are easier to chemically synthesise and modify.

As understanding of lncRNAs has increased, it has been discovered that they can also encode peptides and play a biological role.[17,18] Ribosome profiling sequencing (Ribo-seq) has been used to detect RNA translation levels.[19] Ribo-seq can shield mRNA fragments of 20 to 30 nucleotides in length from nuclease digestion by translating ribosomes and can be used to identify small open reading frames (sORFs) with translation potential.[20] An ORF is defined as a potential translation sequence, starting from the start codon to the end of the stop codon.[21] The length of an sORF differs from that of an ORF, theoretically ranging from 2 to 100 codons.[22] Initially considered noncoding due to its extremely short length, it has been found that the molecular sequence of ncRNA contains an sORF capable of encoding micropeptides with fewer than 100-amino acid (aa).[23-27] Prediction analysis of sORFs includes the prediction of the start codon, the sORF position and the translation potential of ncRNAs. In addition to three-nucleotide (3-nt) periodicity and RPF distribution, the ribosome P site is also used as a discriminant condition.[28,29] During the initial stage of peptide chain synthesis, the ribosome subunit, mRNA and initial aminoacyl-tRNA with a priming function are assembled into the initial complex. Once the initial complex is formed, the P position is occupied by the initial tRNA, which binds to the initial codon. Subsequently, based on the codon sequence on mRNA, various aminoacyl-tRNAs bind to ribosomes in sequence, leading to the gradual extension of the peptide chain is gradually extended from the N-terminus to the C-terminus. Some studies have indicated that these micropeptides can interact with other proteins or RNAs after processing and modification, playing a role in various physiological and pathophysiological processes.[30,31] For example, the 90-amino-acid (aa) polypeptide SPAR, encoded by the lncRNA LINC00961, inhibits mTORC1 activity via lysosomes. This action modulates the regenerative response of skeletal muscle.[31] Additionally, these micropeptides regulate mitochondrial metabolism and embryonic development and influence the onset and progression of cancer. However, research on lncRNAs related to the osteogenic differentiation of JBMMSCs remains limited.

The lncRNA small nucleolar RNA host gene 1 (SNHG1), situated on chromosome 11, consists of 11 exons.[32,33] Some studies have confirmed that SNHG1 significantly influences osteogenic differentiation, osteoclast differentiation and angiogenesis. It regulates osteogenic differentiation by regulating the p38 MAPK signalling pathway,[34] a key trigger of BMSCs.[35] In periodontal ligament stem cells, SNHG1 inhibits osteogenic differentiation through EZH2-mediated methylation of the KLF2 promoter H3K27me3.[36] SNHG1 also promotes osteoclast differentiation of BMMSCs[37] and angiogenesis of endothelial cells.[38,39] However, the regulatory effects of SNHG1 on human JBMMSCs (h-JBMMSCs) require further investigation. The presence of a sORF within SNHG1 and its potential to encode an endogenous peptide are still unclear.

To investigate the translation of h-JBMMSCs during osteogenic differentiation, Ribo-seq of h-JBMMSCs was conducted at 0 and 7 days after osteogenic differentiation. Through the screening of sORFs and differentially translational efficiency genes (DTEGs), the SNHG1 gene was chosen for further study. To explore the role of SNHG1 in the osteogenic differentiation of h-JBMMSCs, the present authors preliminarily verified its coding potential.

Materials and methods

Cells and cell culture

All the experimental procedures were approved by the Ethics Committee of Beijing Stomatological Hospital of Capital Medical University (license no: CMUSH-IRB-KJ-PJ-2023-50). Bone samples extracted from patients after orthognathic surgery were washed with phosphate-buffered saline (PBS) (HyClone, Logan, UT, USA) and cut into small fragments (< 1 mm^3). The fragments were cultured using mesenchymal stem cell medium (MSCM; ScienCell, Carlsbad, CA, USA). Osteogenic differentiation induction by h-JBMMSCs for 0 and 7 days was assessed via Ribo-seq.

Library preparation and sequencing

In sample selection, we established the criteria to include healthy individuals aged 18 to 35 years, excluding those with oral infections (acute or chronic), a family history of genetic disorders, diabetes, hypertension, heart disease or jaw bone diseases. One type of h-JBMMSCs was selected for osteogenic differentiation at 0 and 7 days, with three replicates in each group (n = 3). Translation inhibitors were applied to halt translation. The ribosome-peptide complex underwent treatment with a low concentration of RNase, allowing degradation of RNA fragments not shielded by ribosomes. Then, ribosomes were removed, and the RNA fragments protected by the ribosome, termed ribosome protect frame (RPF), were identified. The raw sequencing data acquired were compressed and stored in fastq format. To avoid affecting the subsequent data analysis, the original data were partitioned and quality controlled to produce clean data for subsequent data analysis. FastQC analysed the quality of the sequencing data. To confirm biological reproducibility among the groups, Pearson correlation analysis was executed on FPKM values.

Triplet periodicity analysis

Three-nucleotide (3-nt) periodicity reflects the distribution pattern of ribosome reading density and serves as an intrinsic characteristic of translation, proceeding three nucleotides at a time. Ribosomes display a pause after every three bases (one codon) during transcript translation. By assigning the ribosome P site to each base position of each codon, distribution counts were determined and the RPF distribution graph was created. Given that ribosomes reside longest at the first base of the codon, the ratio of RPFs at this position typically peaks.[40]

Differential gene expression analysis

Differentially translated genes (DTGs) were evaluated by using DESeq2 R, and the genes with different translations were screened using $P < 0.05$ and $Log_2FC > 2$. Pathway annotation of the identified differentially translated genes was conducted using the KEGG database. Significant pathway terms were determined for differentially translated genes through *P* value calculation via a Fisher test.

Estimation of translational efficiency

Translational efficiency (TE) refers to the ratio of the total RNA molecules of each gene in the sample to the ribosome for translation. The calculation formula was TE = (FPKM in Ribo-seq) / (FPKM in RNA-seq). TE represents RNA utilisation rather than a direct measure of protein output.[29] DESeq2 R was used to analyse the differential TEs, and the DTEGs were screened for $P < 0.05$ and $log_2FC > 2$. The DESeq2 R package was used to calculate the log_2FoldChange, *P* value and padj values of the translation efficiency difference multiple for subsequent analysis, and KEGG analysis was subsequently performed on the DEGs.

Detection of actively translated open reading frames (ORFs)

Open reading frame (ORF) analysis can be used to examine small peptides encoded by ncRNAs. Translated ORFs are identified based on the 3-nt periodicity, ribosome P site and RPF distribution.[28,29] To more accurately identify ORFs with coding potential, RPF readings need to be analysed using various indicators and methods. Annotated protein-coding genes were excluded from the ORF prediction table to isolate the ORF of ncRNA and extract the amino acid sequence. Logistic regression analysis determines the likelihood of finding the start codon.

Lentiviral transduction

After excluding annotated coding genes, we screened the remaining ORFs, prioritising genes showing significant expression differences between h-JBMMSCs on day 0 and day 7 of osteogenic differentiation. Additionally, the ORF's start codon frequency is relatively high, and the predicted peptide chain length exceeds two amino acids. Finally, drawing upon previous re-

search, we selected lncRNA SNHG1 for further study. The cDNA of the human lncRNA SNHG1 and short hairpin RNAs (shRNAs) were inserted into the LV6 lentiviral vector (Genepharma, Suzhou, China) and GV112 lentiviral vector (Genechem, Shanghai, China) respectively. The culture medium was changed 12 hours after lentivirus infection, and puromycin (2 μg/ml) was used for screening 48 hours after infection. The target sequences were as follows: lncRNA SNHG1 shRNA (sh-SNHG1), 5′-GGTTTCAAGGCCATAGCTTTA-3′; and control shRNA (sh-Control), 5′-TTCTCCGAACGTGTCACGT-3′. To generate GFP fusion protein constructs with the lncRNA SNHG1 5' untranslated region (UTR)-ORF (5'UTR-ORF-GFPmut), the 5'UTR-ORF sequence was cloned and inserted into the GV348 vector, in which the GFP initiation codon (ATG) was mutated to ATT. H-JBMMSCs were transfected with 5'UTR-ORF-GFPmut for 24 hours and observed under a microscope.

Peptide synthesis

The amino acid sequence that could be translated was obtained from the gene sequence predicted by the ORF, and the peptide was synthesised (Jietai, Nanjing, China). During the osteogenic induction process, peptides were added simultaneously at a concentration of 10 μg/ml when the solution was changed. Osteogenic indexes were assessed after 5 days, 7 days and 2 weeks of induction.

Osteogenic differentiation

H-JBMMSCs are seeded in 6-well plates and cultured with osteogenic-induced conditioned medium (mesenchymal stem cell medium supplemented with 10 nM dexamethasone, 1.8 mM KH_2PO_4, 2 mM β-glycerophosphate and 100 μM/ml ascorbic acid). Cells were treated with 10 μg/ml control peptide or SNHG1 peptide once every 3 days during the culture process. After 5 days, an alkaline phosphatase (ALP) activity kit (Sigma-Aldrich, St Louis, MO, USA) was used to detect ALP activity. The protein concentration was used to standardise the OD value. After 7 days, western blot (WB) analysis was performed to measure the expression of target proteins. After 14 days, the cells were fixed with 70% alcohol, washed in PBS twice, and treated with Alizarin Red (Sigma-Aldrich). After drying the plate, 10% cetylpyridinium chloride was added for 30 minutes to detect the calcium ion content. The absorbance at 562 nm was measured and compared with the standard calcium curve to determine the concentration.

RNA isolation and qPCR

Total RNA was extracted from the cells using the TRIzol reagent (Invitrogen). One-microgram aliquots of RNA were reverse transcribed into cDNA (Invitrogen, Waltham, MA, USA). A QuantiTect SYBR Green PCR kit (Qiagen, Hilden, Germany) and an Archimed-X6 real-time PCR detection system (Rocgene, Beijing, China) were used for relative quantitative PCR. The relative expression levels of SNHG1 were normalised to those of GAPDH. The primers used in this study were as follows: human GAPDH, (forward) 5′- CGGACCAATACGACCAAATCCG -3′ and (reverse) 5′- AGCCACATCGCTCAGACACC -3′; and human SNHG1, (forward) 5′- ACAGCAGTTGAGGGTTTGCT -3′ and (reverse) 5′- GGGCCTGGATCATGTAAGAA -3′.

Western blot analysis

RIPA buffer containing protease and phosphatase inhibitors was used to extract proteins. Protein samples were blocked for 1 hour after gel electrophoresis and membrane transfer. The first antibody was incubated overnight in the refrigerator at 4°, and the second was incubated at room temperature for 1 hour. The samples were observed with western ECL substrates (Bio-Rad, Hercules, CA, USA). The primary antibodies used were anti-OCN (cat no. A20800, ABclonal), anti-BSP (cat no. bs-23258, Bioss), anti-GFP (cat no. 66002-1-lg, Proteintech), and anti-glyceraldehyde 3-phosphate dehydrogenase (GAPDH; cat no. G8795, Sigma-Aldrich).

Statistical analysis

Statistically significant differences ($P < 0.05$) were determined using a Student *t* test. GraphPad Prism 8 software (GraphPad, La Jolla, CA, USA) was utilised for all statistical analyses.

Results

Differential gene analysis

Ribo-seq was performed on h-JBMMSCs induced by osteogenic differentiation for 0 and 7 days, with 3 replicates in each group. The Pearson correlation coefficients (r) between the two replicate samples of each group were greater than 0.99, indicating high reproducibility of our experiments (Fig 1a and b). For any codon, frame 0/1/2 represents three nucleotides from 5' to 3', respectively. We found that the length of RPFs was approximately 29

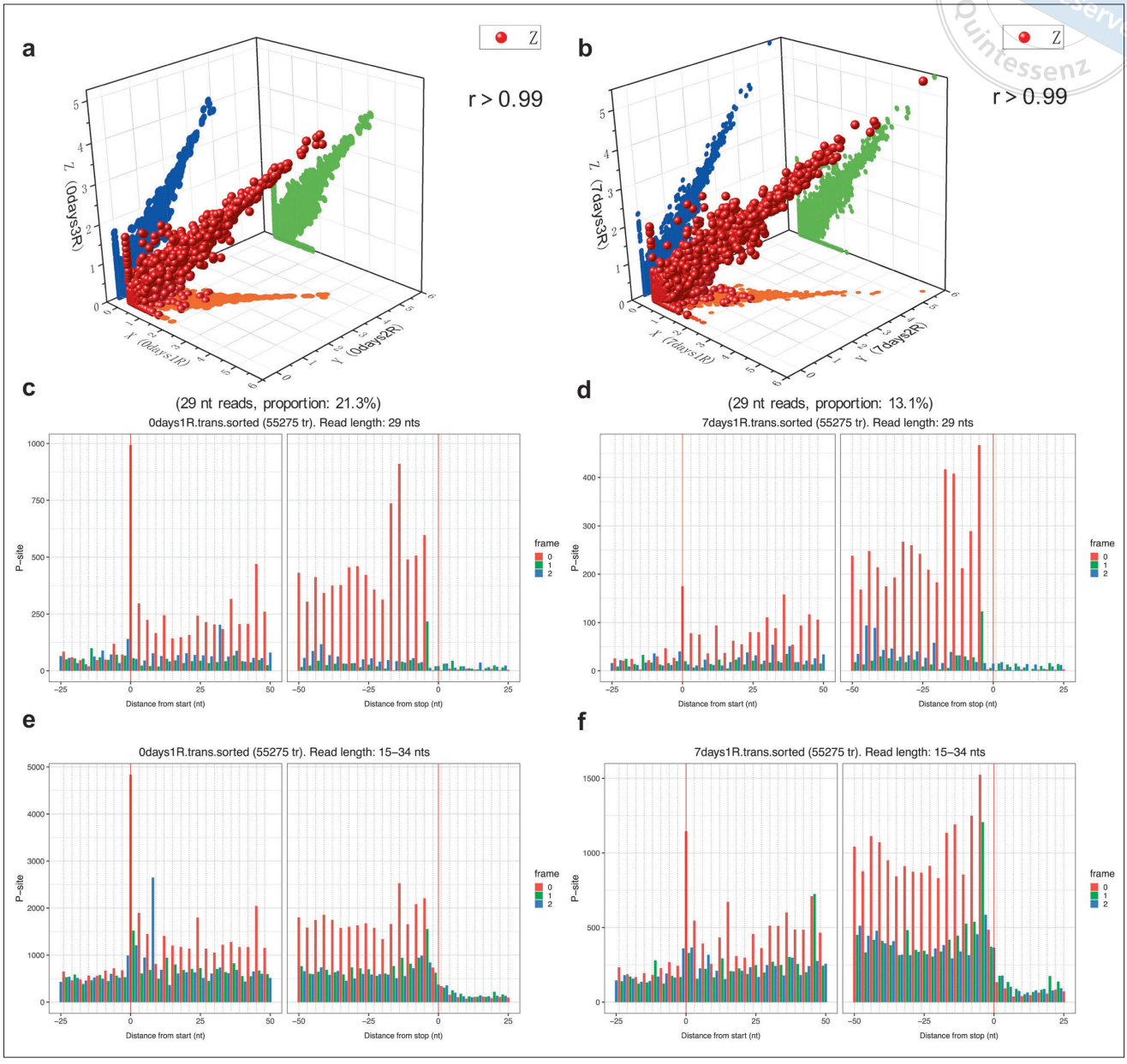

Fig 1a to f Ribo-seq data characteristics. 3D scatter plot of biological repeatability in the 0- and 7-day groups **(a and b)**. The Pearson correlation coefficients **(r)** between the two replicate samples of each group were greater than 0.99. The 29-nt **(c and d)** and 15-34-nt RPF distributions around the ribosomal P site **(e and f)**. Frame 0/1/2 represents three nucleotides of 5' to 3', respectively.

to 30 nt. When we used 29 nt readings to scan the start codon and stop codon, we observed a clear 3 nt periodicity from the distribution of the first base at the 5' end of the RPF (Fig 1c and d and Fig S1a [provided on request]). Trinucleotide periodicity was also observed when we scanned the start codon and stop codon with readings of 15-34 nt (Fig 1e and f and Fig S1e to h [provided on request]). These results demonstrate a characteristic feature of reliable translation-specific data.

Analysis of ribosome sequencing data revealed 1,548 differentially expressed translation genes compared with those at 0 days of osteogenic treatment, including 919 upregulated genes and 629 downregulated genes (Fig S2a, provided on request). KEGG analysis of the differential translation genes (DTGs) showed that they were enriched in p53, ECM-receptor interaction, cell cycle and other pathways (Fig 2a). TE can reflect the RNA utilisation rate, and gene TE was further analysed

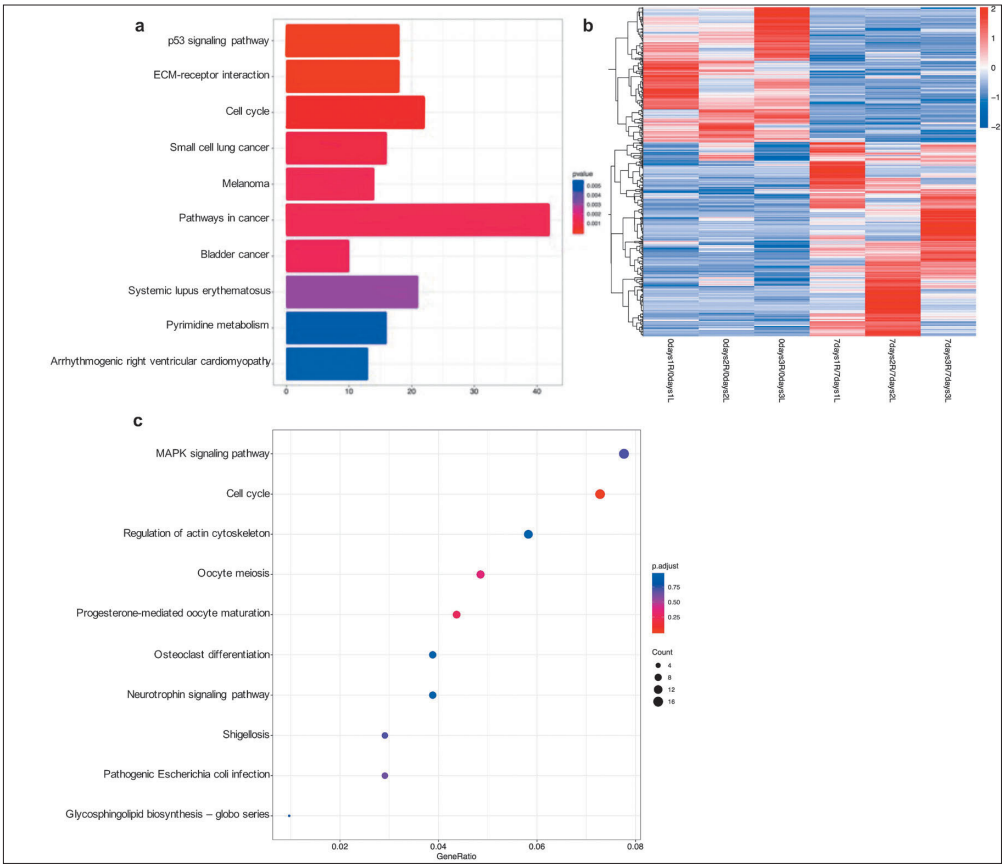

Fig 2a to c Differential translation gene analysis and differential TE gene analysis. KEGG analysis of differentially translated genes **(a)**. Cluster thermogram of genes with different translation efficiency **(b)**. KEGG analysis of differentially TE genes **(c)**.

by calculating the log2TE cumulative scores at 0 days and 7 days. The overall expression at 7 days was found to be greater than that at 0 days (Fig S2b, provided on request). Cluster analysis of the thermogram also revealed that the overall consistency of the upregulated genes in the group was significantly greater than that of the downregulated genes (Fig 2b). A total of 830 DTEGs were identified, with 488 genes being upregulated and 342 genes downregulated (Fig S2c, provided on request). KEGG analysis revealed enrichment of DTEGs in the MAPK pathway (Fig 2c).

In addition, most of the DTEGs (83.7%) were located on annotated protein-coding genes, but 16.4% of them were located in antisense RNA, long interspersed ncRNA (lincRNA), processing transcription and other noncoding genes (Fig 3a). Genes can be classified according to their biotype, an indicator of biological significance. In addition to protein-coding genes, they also include processed transcripts and pseudogenes. Processed transcripts encompass lncRNAs, ncRNAs and unclassified processed transcripts. Antisense RNAs and lincRNAs are types of lncRNAs. These results suggest that lncRNAs may have coding potential as ncRNAs.

Prediction and screening of ORFs

In addition to the 3-nt periodicity and RPF distribution, ribosome P sites are also used as discriminant conditions in the prediction and analysis of ORFs (Fig 3b). The frequency analysis of different codons at the ribosome P site showed that AGC, CUA, CUC, CUG and other codons tended to increase (Fig 3c). Moreover, a total of 53,432 ORFs were obtained, and 199 candidate translation ORFs were detected by removing the annotated protein-coding genes. Additionally, the cumulative analysis of the number of peptides of different lengths and the frequency of initial codon usage revealed 162 peptides of 0 to 10 aa, 27 peptides of 11 to 20 aa, 8 peptides of 21 to 30 aa and 2 peptides of 31 to 40 aa. The frequency of TTG and CTG usage was significantly greater than that of other codons (Fig 3d). The predicted lncRNA SNHG1 has three ORFs, two of which encode a 2-aa peptide and another initiation codon, CTG, and a 19-aa peptide (Fig S2d, provided on request). We consider that the lncRNA SNHG1 has potential translational value in h-JBMMSCs.

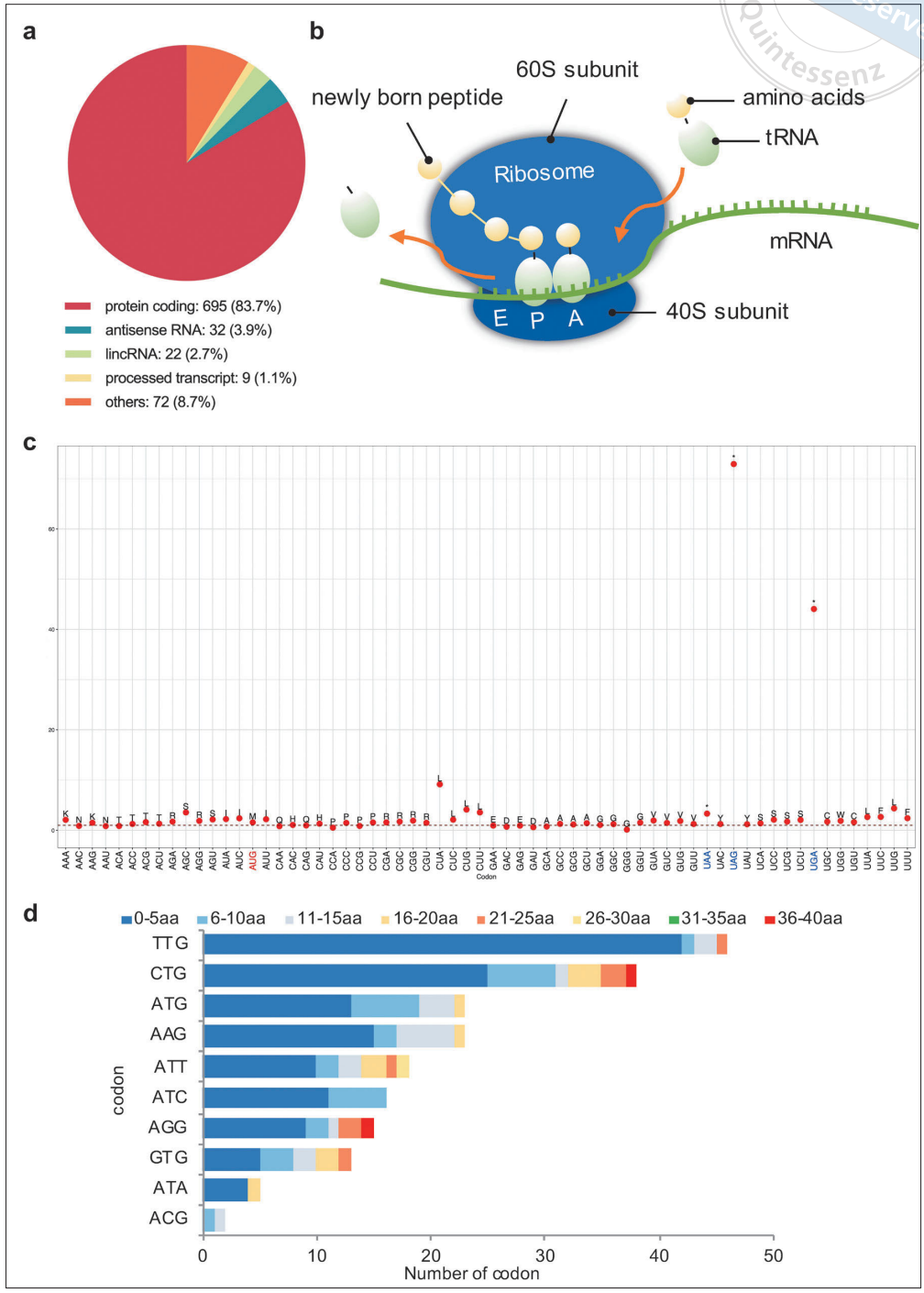

Fig 3a to d Analysis of the coding potential of noncoding RNAs. Proportion of genes with different TEs **(a)**. Schematic diagram of ribosome translation **(b)**. Frequency of differential codon usage at the ribosome P site. The abscissa represents 64 codons, the ordinate represents codon frequency, the red dots represent amino acids, and the uppercase letter corresponds to amino acid abbreviations. A frequency greater than 1 indicates upregulated amino acids, and a frequency less than 1 indicates downregulated amino acids. In the horizontal differential codon, red represents the classical start codon, blue represents the stop codon and black represents other codons **(c)**. The ORF predicts the initial codon usage frequency of noncoding RNA ORFs and the cumulative distribution of peptides of different lengths after removing the protein-coding gene **(d)**.

LncRNA SNHG1 inhibited the osteogenic differentiation of h-JBMMSCs

The mRNA expression level of SNHG1 was downregulated in h-JBMMSCs 7 days after osteogenic differentiation (Fig 4a). Then, we knocked down SNHG1 expression by infecting h-JBMMSCs with retrovirus (Fig 4b). After osteogenic induction for 5 days, compared with that in the sh-control group, ALP activity in the sh-SNHG1 group increased significantly (Fig 4c). In addition, after 7 days of induction, the expression levels of the OCN and BSP proteins in the sh-SNHG1 group were upregulated (Fig 4f and g). Two weeks later, the mediation of ARS in the sh-SNHG1 group was enhanced, and

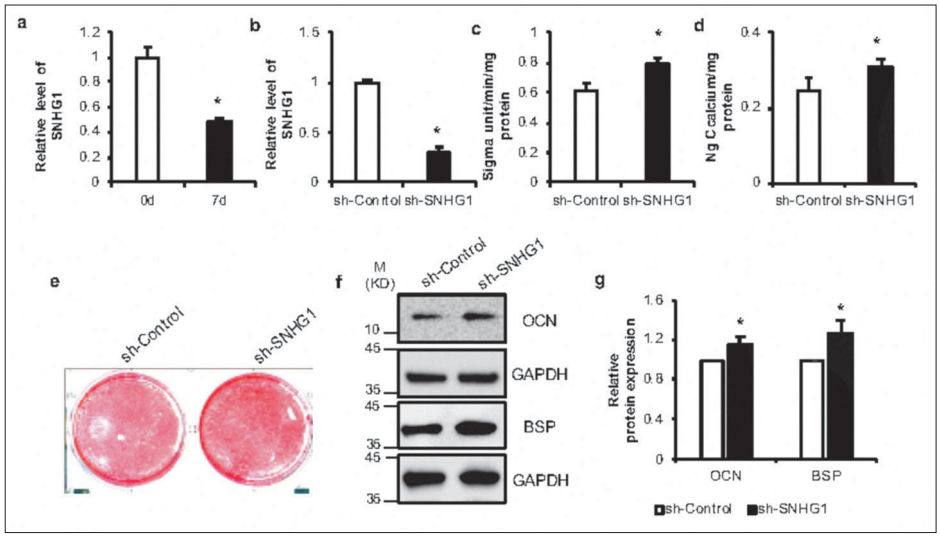

Fig 4a to g LncRNA SNHG1 knockdown enhanced osteogenic differentiation in h-JBMMSCs. The mRNA expression level of SNHG1 at 7 days after osteogenic differentiation, with GAPDH serving as a control **(a)**. SNHG1 was knocked down in h-JBMMSCs, and cells transfected with the empty vector served as controls **(b)**. ALP activity on day 5 **(c)**. Alizarin red staining **(d)** and calcium ion quantification on day 14 **(e)**. The protein levels of OCN, BSP and GAPDH were measured via Western blot assay **(f and g)**. Statistical analysis was performed using a Student *t* test. In the bar graphs, data are presented as mean ± SD. **P* < 0.05.

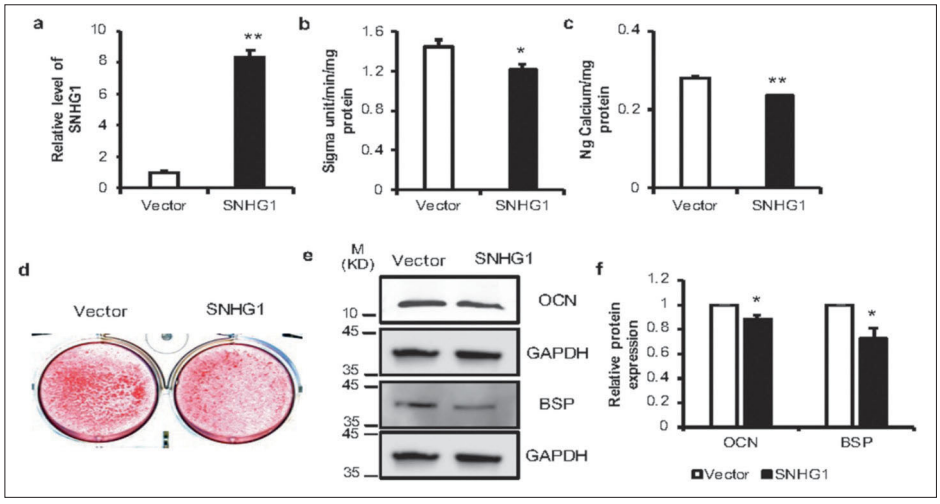

Fig 5a to f LncRNA SNHG1 overexpression inhibited osteogenic differentiation in h-JBMMSCs. SNHG1 was overexpressed in h-JBMMSCs, and cells transfected with the empty vector served as controls **(a)**. ALP activity on day 5 **(b)**. Calcium ion quantitative analysis **(c)** and Alizarin red staining on day 14 **(d)**. The protein levels of OCN, BSP and GAPDH were measured via Western blot assay **(e and f)**. Statistical analysis was performed using a Student *t* test. In the bar graphs, data are presented as mean ± SD.**P* < 0.05. ***P* < 0.01.

the calcium ion concentration was increased (Fig 4d and e). These results indicate that knocking down SNHG1 in h-JBMMSCs promotes osteogenic differentiation.

Moreover, SNHG1 was overexpressed, and its expression of SNHG1 was increased (Fig 5a). After 5 days of osteogenic induction, ALP activity in the SNHG1 group decreased significantly compared to the Vector group (Fig 5b). After 7 days, the WB results showed that the expression of OCN and BSP was downregulated (Fig 5e and f). After 2 weeks, in the SNHG1 group, there was a decrease in the number of ARS and the calcium ion concentration (Fig 5c and d). These results indicate that the overexpression of SNHG1 in h-JBMMSCs inhibits osteogenic differentiation.

lncRNA SNHG1 encoded a small peptide and SNHG1-peptide inhibits the osteogenic differentiation of h-JBMMSCs

To investigate the activity of the start codon of the SNHG1 ORF, we created a 5'UTR-ORF-GFPmut construct (Fig 6a). Green fluorescence was observed under a microscope following the transfection of h-JBMMSCs (Fig 6b). WB results indicated the presence of the GFP fusion protein post-transfection of 5'UTR-ORF-GFPmut (Fig 6c). Moreover, h-JBMMSCs transfected with 5'UTR-ORF-GFPmut showed inhibited osteogenic differentiation (Fig 6d to h). Therefore, 57-nt sORF of the lncRNA SNHG1 might encode a 19-aa peptide.

To verify the function of the SNHG1 peptide, we synthesised a control peptide and an SNHG1-peptide.

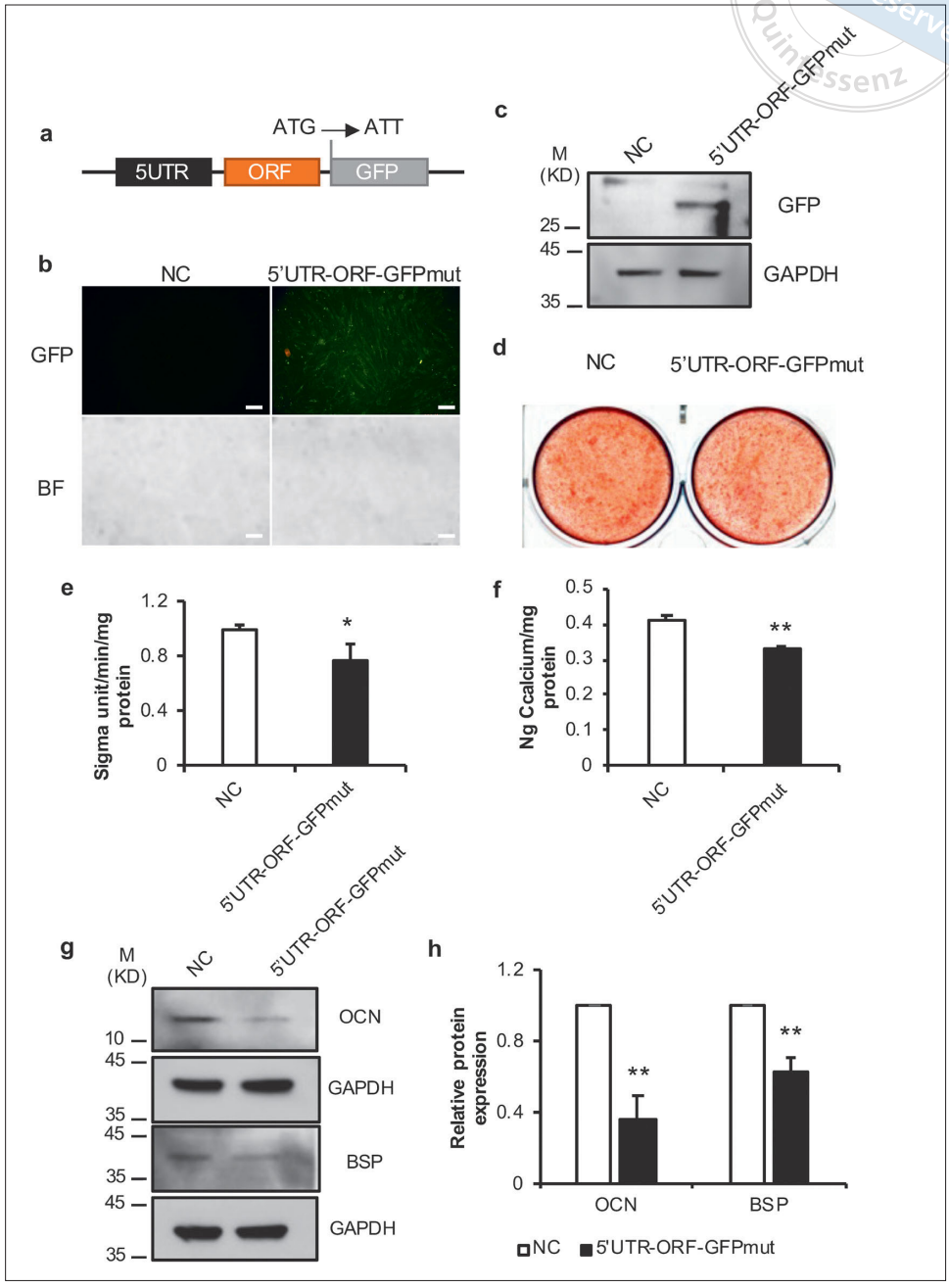

Fig 6a to h Transfection of the lncRNA SNHG1-encoded peptide and h-JBMMSCs with 5'UTR-ORF-GFPmut inhibited osteogenic differentiation. The 5'UTR-ORF-GFPmut virus was designed, and the sequence of the GFP-ORF start codon (ATG) was mutated to ATT **(a)**. GFP was expressed by h-JBMMSCs transfected with 5'UTR-ORF-GFPmut under a fluorescence microscope **(b)**. Western blot analysis showed that GFP was expressed in h-JBMMSCs transfected with 5'UTR-ORF-GFPmut **(c)**. ALP activity on day 5 **(d)**. Alizarin red staining **(e)** and calcium ion quantitative analysis on day 14 **(f)**. Protein levels of OCN, BSP and GAPDH were measured via Western blot assay **(g and h)**. Statistical analysis was performed using a Student *t* test. In the bar graphs, data are presented as mean ± SD. $^{*}P < 0.05$. $^{**}P < 0.01$.

Compared with the control peptide, SNHG1-peptide decreased ALP activity after 5 days of osteogenic induction (Fig 7a). At 7 days, SNHG1-peptide decreased the protein expression levels of OCN and BSP (Fig 7d and e). By 2 weeks, the ARS staining became lighter, and the level of calcium ions decreased (Fig 7b and c). These results indicate that the SNHG1-peptide inhibits osteogenic differentiation in h-JBMMSCs.

Discussion

Trinucleotide periodicity often indicates the coverage of ribosome fragments mapped to the subcodon position and tends to the first position.[40] It is commonly used to assess the quality of Ribo-seq data. Our results demonstrated good biological repeatability and a 3-nt periodicity in the group, which ensures the reliability of subse-

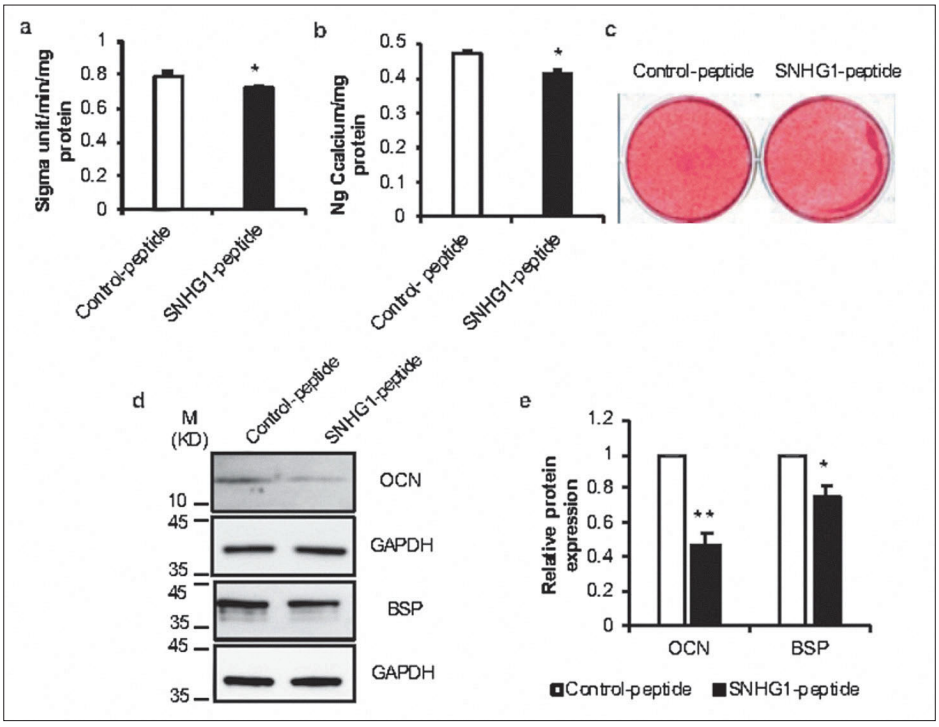

Fig 7a to e SNHG1-peptide inhibited the osteogenic differentiation of h-JBMMSCs. ALP activity on day 5 **(a)**. Calcium ion quantitative analysis **(b)** and Alizarin red staining on day 14 **(c)**. The protein levels of OCN, BSP and GAPDH were measured via Western blot assay **(d and e)**. Statistical analysis was performed using a Student *t* test. In the bar graphs, data are presented as mean ± SD. **P* < 0.05. ***P* < 0.01.

quent data analysis. The ribosome plays a crucial role in protein translation, thus impacting the growth, development and differentiation of cells. Knocking down RIOK3 significantly impairs ribosome function, influences the translation of intracellular proteins and diminishes the osteogenic differentiation of mouse BMMSCs (mBMMSCs).[41] In this study, compared to h-JBMMSCs at 0 days and 7 days post-osteogenic differentiation, there were more upregulated than downregulated genes in both the DTGs and DTEGs. The increase in the ribosome binding rate during osteogenic differentiation suggests that gene translation may have a regulatory function. Further bioinformatics analysis revealed that KEGG pathway analysis identified several pathways, such as the MAPK, p53 and ECM-receptor interaction pathways, were associated with osteogenic differentiation. The MAPK pathway is closely linked to osteogenic differentiation,[42,43] and lncRNA differentiation antagonising nonprotein coding RNA (DANCR) inhibits osteogenic differentiation of BMSCs.[44] The lncRNA SNHG1 negatively regulates the p38 MAPK signalling pathway, thus inhibiting the osteogenic differentiation of BMSCs.[34] The genes enriched in the MAPK signalling pathway include RAC1, MAPK8IP2, MAPK13, RPS6KA1, NFKB2 and FGF7, suggesting that these genes may act as regulatory factors in osteogenic differentiation. For instance, RAC1 can contribute to atherosclerosis,[45] but in MC3T3-E1 cells, OCN and runt-related transcription factor 2 (RUNX2) are upregulated by enhancing the expression of RAC1.[46]

We detected 53,432 ORFs, 199 of which were unannotated protein-coding genes according to ORF prediction. These 199 ORFs were predicted to encode micropeptides less than 40-aa in length. However, TE cannot distinguish between translating ribosomes and non-translating ribosomes based on the number of RPFs associated with transcripts.[29] To further highlight the active translation characteristics of RPFs, a predicted ribosome P site was proposed.[47] Through the prediction of ORFs and the analysis of ribosome P sites, we found that UUG and CUG were used more frequently as initial codons than was AUG. Among them, lncRNA SNHG1 was 57 nt long and its initial codon was CUG. Eukaryotic translation starts from the first AUG encountered by ribosomes on mRNA. However, the unexpected ribosome occupancy observed in the analysis of the initiation site indicates a significant number of non-AUG initiation events.[48] While the usage of start codons depends on genes and research, it has been demonstrated that CUG and UUG are utilised at a frequency of approximately 1% to 10% compared to AUG under optimal conditions, while AAG and AGG are largely unrecognised.[49,50] In mammals, when the POLG gene is encoded under optimal conditions, the initiation efficiency of the CUG codon is comparable to that of the AUG codon (~60% to 70%).[51] By integrating ORF

prediction, KEGG analysis and initial codon analysis, we selected SNHG1 for further investigation.

The expression of the lncRNA SNHG1 decreased after 7 days of osteogenic differentiation induced by h-JBMMSCs. Then, we knocked down and overexpressed SNHG1 in h-JBMMSCs for osteogenic induction. The activity of ALP in vitro is commonly used to detect the initial differentiation of BMSCs into osteoblasts, and is also a crucial factor in BMSC calcification.[52] The results of ALP activity, WB, ARS and calcium ion quantification all indicated that the lncRNA SNHG1 inhibited osteogenic differentiation. This finding is similar to previous research results showing that the lncRNA SNHG1 inhibits osteogenic differentiation via the miR-101/DKK1 axis.[53] To verify whether the lncRNA SNHG1 is translated, we designed lentivirus to add a GFP tag after the ORF of SNHG1 and mutated the start codon of GFP at the same time. However, our results showed that green fluorescence was observed under the microscope, and the WB results indicated that GFP was expressed in the cells. These results indicate that the lncRNA SNHG1 encodes a small peptide. Using the ORF sequence, we derived the corresponding amino acid sequence to synthesise small peptides in vitro and utilised them in the process of osteogenic differentiation induction. Compared to the control peptide, the SNHG1-peptide significantly reduced ALP activity and the expression of OCN and BSP. At the same time, the amount of ARS decreased, and the content of calcium ions decreased. These outcomes imply that the encoded products of lncRNAs may have similar function as lncRNAs.

Combining the results of the biological analysis and the SNHG1 experiment, we found that the lncRNA SNHG1 is a gene with upregulated TEs, but the SNHG1 mRNA level decreased after 7 days of osteogenic differentiation of h-JBMMSCs. Some studies have shown that genes are not completely consistent at the transcriptional and the translational levels, and the lower the transcription level, the greater the TE.[54,55] On the other hand, when h-JBMMSCs were transfected with 5'UTR-ORF-GFPmut or SNHG1-peptide acted on h-JBMMSCs, the osteogenic differentiation was inhibited. Combined with the upregulation of the TE of SNHG1, we speculated that SNHG1 expression would increase at the translational level but decrease at the protein level on the seventh day of osteogenic differentiation. For example, during the activation of Drosophila eggs, there is a poor correlation betwee changes in protein levels, and only 25% of the proteins encoded by translation upregulated genes are upregulated, and most of them are unchanged or downregulated.[56] It is speculated that the acceleration of translation is balanced by the degradation of proteins or that newly synthesised proteins account for only a small portion of the accumulated proteins, so the upregulation of translation expression of these proteins can hardly be detected.

These results also raise several questions. For example, 16.4% of genes with differential translation efficiency are noncoding genes, but only 199 noncoding genes contain sORFs. Thus, one question is: what is the role of these genes that may not be translated in binding to ribosomes? At the same time, there are still many limitations in this study. The existence of the endogenous SNHG1 peptide in h-JBMMSCs and its mechanism of inhibiting osteogenic differentiation need further study.

Conclusion

By analysing Ribo-seq data, it was found that the osteogenic differentiation of h-JBMMSCs may be regulated by translation. One hundred and ninety-nine candidate sORFs of ncRNAs were identified, and UUG and CUG were the main initiation codons, which provided evidence for non-AUG initiation. The lncRNA SNHG1, which encodes a small peptide of 19-aa, was shown to inhibit the osteogenic differentiation of h-JBMMSCs. This peptide inhibits the osteogenic differentiation of h-JBMMSCs and may have a potential role in the treatment of osteopetrosis.

Conflicts of interest

The authors declare no conflicts of interest related to this study.

Author contribution

Dr Hua LIU performed the experiments and was a major contributor to the writing of manuscript; Drs Zhi Peng FAN, Hui Na LIU and Qiu Bo YANG conceived and designed the study. All the authors have read and approved the final manuscript.

(Received Apr 09, 2024; accepted Aug 13, 2024)

References

1. Ullah I, Subbarao RB, Rho GJ. Human mesenchymal stem cells - Current trends and future prospective. Biosci Rep 2015;35:e00191.
2. Lloyd B, Tee BC, Headley C, Emam H, Mallery S, Sun Z. Similarities and differences between porcine mandibular and limb bone marrow mesenchymal stem cells. Arch Oral Biol 2017;77:1–11.
3. Zhang W, Dong ZW, Li DK, et al. Cathepsin K deficiency promotes alveolar bone regeneration by promoting jaw bone marrow mesenchymal stem cells proliferation and differentiation via glycolysis pathway. Cell Prolif 2021;54:e13058.
4. Fu X, Liu G, Halim A, Ju Y, Luo Q, Song AG. Mesenchymal stem cell migration and tissue repair. Cells 2019;8:784.
5. Zong C, Zhao L, Huang C, Chen Y, Tian L. Isolation and culture of bone marrow mesenchymal stem cells from the human mandible. J Vis Exp 2022;13:(182).
6. Li C, Wang FF, Zhang R, Qiao P, Liu H. Comparison of proliferation and osteogenic differentiation potential of rat mandibular and femoral bone marrow mesenchymal stem cells in vitro. Stem Cells Dev 2020;29:728–736.
7. Luo S, Pei F, Zhang W, et al. Bone marrow mesenchymal stem cells combine with treated dentin matrix to build biological root. Sci Rep 2017;12:44635.
8. Song W, Xie JH, Li JY, Bao C, Xiao Y. The emerging roles of long noncoding RNAs in bone homeostasis and their potential application in bone-related diseases. DNA Cell Biol 2020;39: 926–937.
9. Patil S, Dang K, Zhao X, Gao Y, Qian A. Role of LncRNAs and CircRNAs in bone metabolism and osteoporosis. Front Genet 2020;13:584118.
10. Tang S, Xie ZY, Wang P, et al. LncRNA-OG promotes the osteogenic differentiation of bone marrow-derived mesenchymal stem cells under the regulation of hnRNPK. Stem Cells 2019;37:270–283.
11. Zhou P, Li Y, Di RL, et al. H19 and Foxc2 synergistically promotes osteogenic differentiation of BMSCs via Wnt-β-catenin pathway. J Cell Physiol 2019;234:13799–13806.
12. Wang W, Li T, Feng S. Knockdown of long non-coding RNA HOTAIR promotes bone marrow mesenchymal stem cell differentiation by sponging microRNA miR-378g that inhibits nicotinamide N-methyltransferase. Bioengineered 2021;12:12482–12497.
13. Tong X, Gu PC, Xu SZ, Lin XJ. Long non-coding RNA-DANCR in human circulating monocytes: A potential biomarker associated with postmenopausal osteoporosis. Biosci Biotechnol Biochem 2015;79:732–737.
14. Gregson CL, Armstrong DJ, Bowden J, et al. UK clinical guideline for the prevention and treatment of osteoporosis. Arch Osteoporos 2022;17:58 [erratum 2022;17:80].
15. Song J, Kim D, Han J, Kim Y, Lee M, Jin EJ. PBMC and exosome-derived Hotair is a critical regulator and potent marker for rheumatoid arthritis. Clin Exp Med 2015;15:121–126.
16. Winkle M, EI-Daly SM, Fabbri M, Calin GA. Noncoding RNA therapeutics - Challenges and potential solutions. Nat Rev Drug Discov 2021;20:629–651.
17. Huang JZ, Chen M, Chen D, et al. A peptide encoded by a putative lncRNA HOXB-AS3 suppresses colon cancer growth. Mol Cell 2017;68:171–184.e6.
18. Zhu S, Wang JZ, Chen D, et al. An oncopeptide regulates m6A recognition by the m6A reader IGF2BP1 and tumorigenesis. Nat Commun 2020;11:1685.
19. Makarewich CA, Olson EN. Mining for micropeptides. Trends Cell Biol 2017;27:685–696.
20. Ingolia NT, Hussmann JA, Weissman JS. Ribosome profiling: Global views of translation. Cold Spring Harb Perspect Biol 2019;11:a032698.
21. Sieber P, Platzer M, Schuster S. The definition of open reading frame revisited. Trends Genet 2018;34:167–170.
22. Malekos E, Carpenter S. Short open reading frame genes in innate immunity: From discovery to characterization. Trends Immunol 2022;43:741–756.
23. Chen Y, Long WL, Yang LQ, et al. Functional peptides encoded by long non-coding RNAs in gastrointestinal cancer. Front Oncol 2021;23:777374.
24. Plaza S, Menschaert G, Payre F. In search of lost small peptides. Annu Rev Cell Dev Biol 2017;33:391–416.
25. Xing J, Liu H, Jiang W, Wang L. LncRNA-encoded peptide: Functions and predicting methods. Front Oncol 2020;10:622294.
26. Orr MW, Mao YH, Storz G, Qian SB. Alternative ORFs and small ORFs: Shedding light on the dark proteome. Nucleic Acids Res 2020;48:1029–1042.
27. Sandmann CL, Schulz J, Ruiz-Orera J, et al. Evolutionary origins and interactomes of human, young microproteins and small peptides translated from short open reading frames. Mol Cell 2023;83:994-1011.e18.
28. Calviello L, Mukherjee N, Wyler E, et al. Detecting actively translated open reading frames in ribosome profiling data. Nat Methods 2016;13:165–170.
29. Choi SW, Kim HW, Nam JW. The small peptide world in long noncoding RNAs. Brief Bioinform 2019;20:1853–1864.
30. Batista PJ, Chang HY. Long noncoding RNAs: Cellular address codes in development and disease. Cell 2013;152:1298–1307.
31. Matsumoto A, Pasut A, Mastumoto M, et al. mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. Nature 2017;541:228–232.
32. Li B, Li A, You Z, Xu J, Zhu S. Epigenetic silencing of CDKN1A and CDKN2B by SNHG1 promotes the cell cycle, migration and epithelial-mesenchymal transition progression of hepatocellular carcinoma. Cell Death Dis 2020;11:823.
33. Tan X, Chen WB, Lv DJ, et al. LncRNA SNHG1 and RNA binding protein hnRNPL form a complex and coregulate CDH1 to boost the growth and metastasis of prostate cancer. Cell Death Dis 2021;12:138 [erratum 2024;15:213].
34. Jiang YP, Wu W, Jiao G, Chen Y, Liu H. LncRNA SNHG1 modulates p38 MAPK pathway through Nedd4 and thus inhibits osteogenic differentiation of bone marrow mesenchymal stem cells. Life Sci 2019;228:208–214.
35. Ba P, Duan X, Fu G, Lv S, Yang P, Sun Q. Differential effects of p38 and Erk1/2 on the chondrogenic and osteogenic differentiation of dental pulp stem cells. Mol Med Rep 2017;16:63–68.
36. Li Z, Guo X, Wu S. Epigenetic silencing of KLF2 by long noncoding RNA SNHG1 inhibits periodontal ligament stem cell osteogenesis differentiation. Stem Cell Res Ther 2020;11:435.
37. Yu X, Rong PZ, Song MS, et al. lncRNA SNHG1 induced by SP1 regulates bone remodeling and angiogenesis via sponging miR-181c-5p and modulating SFRP1/Wnt signaling pathway. Mol Med 2021;27:141.
38. Mi S, Du JY, Liu J, et al. FtMt promotes glioma tumorigenesis and angiogenesis via lncRNA SNHG1/miR-9-5p axis. Cell Signal 2020;75:109749.
39. Wang Z, Wang RH, Wang K, Liu X. Upregulated long noncoding RNA Snhg1 promotes the angiogenesis of brain microvascular endothelial cells after oxygen-glucose deprivation treatment by targeting miR-199a. Can J Physiol Pharmacol 2018;96:909–915.
40. Guo H, Ingolia NT, Weissman JS, Bartel DP. Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature 2010;466:835–840.

41. Zong HX, Liu YQ, Wang XL, et al. RIOK3 potentially regulates osteogenesis-related pathways in ankylosing spondylitis and the differentiation of bone marrow mesenchymal stem cells. Genomics 2023;115:110730.
42. Zhang Y, Gu XF, Li D, Cai L, Xu Q. METTL3 regulates osteoblast differentiation and inflammatory response via Smad signaling and MAPK signaling. Int J Mol Sci 2019;21:199.
43. Li J, Du H, Ji X, et al. ETV2 promotes osteogenic differentiation of human dental pulp stem cells through the ERK/MAPK and PI3K-Akt signaling pathways. Stem Cell Res Ther 2022;13:495.
44. Zhang J, Tao Z, Wang Y. Long non-coding RNA DANCR regulates the proliferation and osteogenic differentiation of human bone-derived marrow mesenchymal stem cells via the p38 MAPK pathway. Int J Mol Med 2018;41:213–219.
45. Healy A, Berus JM, Christensen JL, et al. Statins disrupt macrophage Rac1 regulation leading to increased atherosclerotic plaque calcification. Arterioscler Thromb Vasc Biol 2020;40:714–732.
46. Li X, Ji JB, Wei W, Liu L. MiR-25 promotes proliferation, differentiation and migration of osteoblasts by up-regulating Rac1 expression. Biomed Pharmacother 2018;99:622–628.
47. Ingolia NT. Ribosome footprint profiling of translation throughout the genome. Cell 2016;165:22–33.
48. Fedorova AD, Kiniry SJ, Andreev DE, Mudge JM, Baranov PV. Thousands of human non-AUG extended proteoforms lack evidence of evolutionary selection among mammals. Nat Commun 2022;13:7910.
49. Peabody DS. Translation initiation at non-AUG triplets in mammalian cells. J Biol Chem 1989;264:5031–5035.
50. Kearse MG, Wilusz JE. Non-AUG translation: A new start for protein synthesis in eukaryotes. Genes Dev 2017;31:1717–1731.
51. Loughran G, Zhdanov AV, Mikhaylova MS, et al. Unusually efficient CUG initiation of an overlapping reading frame in POLG mRNA yields novel protein POLGARF. Proc Natl Acad Sci U S A 2020;117:24936–24946.
52. Chen X, Yang L, Ge D, et al. Long non-coding RNA XIST promotes osteoporosis through inhibiting bone marrow mesenchymal stem cell differentiation. Exp Ther Med 2019;17:803–811.
53. Xiang J, Fu HQ, Xu Z, Fan WJ, Liu F, Chen B. lncRNA SNHG1 attenuates osteogenic differentiation via the miR-101/DKK1 axis in bone marrow mesenchymal stem cells. Mol Med Rep 2020;22:3715–3722.
54. Sendoel A, Dunn JG, Rodriguez EH, et al. Translation from unconventional 5' start sites drives tumour initiation. Nature 2017;541:494–499.
55. Lukoszek R, Feist P, Ignatova Z. Insights into the adaptive response of Arabidopsis thaliana to prolonged thermal stress by ribosomal profiling and RNA-Seq. BMC Plant Biol 2016;16:221.
56. Kronja I, Yuan B, Eichhorn SW, et al. Widespread changes in the posttranscriptional landscape at the Drosophila oocyte-to-embryo transition. Cell Rep 2014;7:1495–1508.