

Article

Analyzes In Silico Indicate the lncRNAs MIR31HG and LINC00939 as Possible Epigenetic Inhibitors of the Osteogenic Differentiation in PDLCs

Rogério S. Ferreira ¹, Rahyza I. F. Assis ² , Francesca Racca ³, Ana Carolina Bontempi ¹, Rodrigo A. da Silva ⁴ , Malgorzata Wiench ^{5,*} and Denise C. Andia ^{1,*} 

¹ School of Dentistry, Health Science Institute, Paulista University, São Paulo 04026-002, SP, Brazil; rogerio.ferreira25@aluno.unip.br (R.S.F.); anacarolina.bontempi@gmail.com (A.C.B.)

² Department of Clinical Dentistry, Federal University of Espírito Santo, Vitória 29043-910, ES, Brazil

³ Periodontology Department, The Ohio State University College of Dentistry, Columbus, OH 43210-1267, USA; francesca_racca@hotmail.com

⁴ Program in Environmental and Experimental Pathology, Paulista University, São Paulo 04026-002, SP, Brazil; dasilva.rodrigo.a@gmail.com

⁵ School of Dentistry, Institute of Clinical Sciences, Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham B5 7EG, UK

* Correspondence: m.d.wiench@bham.ac.uk (M.W.); denise.andia@docente.unip.br (D.C.A.)

Abstract: Chromatin conformation, DNA methylation pattern, transcriptional profile, and non-coding RNAs (ncRNAs) interactions constitute an epigenetic pattern that influences the cellular phenotypic commitment and impacts the clinical outcomes in regenerative therapies. Here, we investigated the epigenetic landscape of the SP7 transcription factor (*SP7*) and Distal-Less Homeobox 4 (*DLX4*) osteoblastic transcription factors (TFs), in human periodontal ligament mesenchymal cells (PDLCs) with low (l-PDLCs) and high (h-PDLCs) osteogenic potential. Chromatin accessibility (ATAC-seq), genome DNA methylation (Methylome), and RNA sequencing (RNA-seq) assays were performed in l- and h-PDLCs, cultured at 10 days in non-induced (DMEM) and osteogenic (OM) medium in vitro. Data were processed in *HOMER*, *Genome Studio*, and *edgeR* programs, and metadata was analyzed by online bioinformatics tools and in *R* and *Python* environments. ATAC-seq analyses showed the TFs genomic regions are more accessible in l-PDLCs than in h-PDLCs. In Methylome analyses, the TFs presented similar average methylation intensities (AMIs), without differently methylated probes (DMPs) between l- and h-PDLCs; in addition, there were no differences in the expression profiles of TFs signaling pathways. Interestingly, we identified the long non-coding RNAs (lncRNAs), *MIR31HG* and *LINC00939*, as upregulated in l-PDLCs, in both DMEM and OM. In the following analysis, the web-based prediction tool *LncRRISearch* predicted RNA:RNA base-pairing interactions between *SP7*, *DLX4*, *MIR31HG*, and *LINC00939* transcripts. The machine learning program *TriplexFPP* predicted DNA:RNA triplex-forming potential for the *SP7* DNA site and for one of the *LINC00939* transcripts (*ENST0000502479*). PCR data confirmed the upregulation of *MIR31HG* and *LINC00939* transcripts in l-PDLCs (× h-PDLCs) in both DMEM and OM ($p < 0.05$); conversely, *SP7* and *DLX4* were downregulated, confirming those results observed in the RNA-Seq analysis. Together, these results indicate the lncRNAs *MIR31HG* and *LINC00939* as possible epigenetic inhibitors of the osteogenic differentiation in PDLCs by (post)transcriptional and translational repression of the *SP7* and *DLX4* TFs.

Keywords: PDLc; epigenetic; osteogenic; *SP7*; *DLX4*; *MIR31HG*; *LINC00939*



Citation: Ferreira, R.S.; Assis, R.I.F.; Racca, F.; Bontempi, A.C.; da Silva, R.A.; Wiench, M.; Andia, D.C. Analyzes In Silico Indicate the lncRNAs MIR31HG and LINC00939 as Possible Epigenetic Inhibitors of the Osteogenic Differentiation in PDLCs. *Genes* **2023**, *14*, 1649. <https://doi.org/10.3390/genes14081649>

Academic Editors: Naila Francis Paulo De Oliveira and Maria Cristina Leme Godoy dos Santos

Received: 16 July 2023

Revised: 10 August 2023

Accepted: 14 August 2023

Published: 18 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Over the past decades, stem cell-based treatment associated with regenerative therapies has become increasingly promising for the treatment of several diseases such as diabetes, cardiac ischemia, and osteoarthritis [1]. Mesenchymal stem cells (MSCs) can

be obtained from several sources, such as bone marrow (BMSCs), adipose tissue (ASCs), peripheral blood (HSCs) [2], and teeth (periodontal ligament cells—PDLCs, dental pulp cells—DPCs, and stem cell from exfoliated human dentition—SHED) [3]. MSCs are characterized by self-renewing, differentiation into cell multilineages capacities, and specific surface markers [4]. Moreover, MSCs show inherent tropism toward damaged tissues and the ability to regenerate them [5,6].

PDLCs can differentiate into osteoblastic, adipocyte, neuronal, and chondrogenic-like cells [7], although they might present distinct capacities to produce mineral nodules in vitro [8–12], which could impact clinical applications. This heterogeneity may be related to cell fate commitment, which is “the commitment of cells to specific cell fates and their capacity to differentiate into particular kinds of cells” (Gene_Ontology_Term_Definition_GO:0045165).

The cell fate commitment is determined by lineage-specific transcription factors (TFs), which are proteins that bind in DNA sites and drive the cellular phenotype acquisition [13]. These TFs are expressed through signaling pathways, which are cascades of extra and intracellular molecular events that culminate in the TFs gene expression. This entire molecular process is regulated by epigenetic mechanisms such as chromatin conformation, DNA methylation marks, and non-coding RNA (ncRNAs) interferences. Chromatin can remodel into a more condensed (heterochromatin) or less condensed (euchromatin) structure, determining the degree of RNA polymerase accessibility, responsible for gene transcription, at the DNA strand [14]. This conformation is modulated by epigenetic modifications in histones and DNA methylation patterns. The DNA methylation patterns, in turn, are determined by the intensity of the methyl group aggregation at carbon 5 of cytosine (5 mC) [15], favoring or not binding TFs in gene promoter regions [16]. ncRNAs are RNA molecules that are not translated into protein and can be distinguished according to their size: microRNAs (miRNAs), about 22 nucleotides long, and long non-coding RNAs (lncRNAs), over 200 nucleotides long. Mostly, the miRNAs act through RNA-induced silencing complex (RISC) as post-transcriptional silencers, promoting mRNA degradation by the cleavage mechanism or inhibiting its translation by the base pairing mechanism [17]. The lncRNAs can act as transcriptional regulators by the DNA:RNA triple-helix (triplex) forming, through Hoogsteen or reverse Hoogsteen base pairing, i.e., when a polypurine (A–G) or pyrimidine (C–U) motif of an lncRNA interacts with the major groove of a Watson–Crick double-stranded DNA (dsDNA), forming triplets of canonical bases, such as C•G–C and U•A–T (where ‘•’ and ‘–’ represent Hoogsteen and Watson–Crick interactions, respectively) [18,19]. In addition, they can also act as post-transcriptional and translational regulators by the RNA:RNA duplex intermolecular hybridization through nucleotide base pairing interactions between a lncRNA and an mRNA [20], among other mechanisms. However, such mechanisms remain poorly understood.

Previous studies from our group have demonstrated the individual epigenetic profile influences the capacity of extracellular matrix deposition, and, consequently, the osteogenic phenotype acquisition [9,11]; in addition, we also highlighted the *Sp7 Transcription Factor (SP7)* and *Distal-Less Homeobox 4 (DLX4)* genes, key TFs involved in osteoblastic differentiation, as downregulated in PDLCs with a low capacity of mineral matrix deposition [12]. Here, our aim is to investigate, in silico, the epigenetic landscape of both TFs, *SP7*, and *DLX4*, in PDLCs showing the distinct capacity of mineral matrix deposition in vitro.

2. Material and Methods

2.1. Cell Acquisition and Culture

After signing an informed consent approved by the Ethics Committee of Piracicaba Dental School, University of Campinas (CAAE55588816.4.0000.5418), PDLCs were collected, isolated, and cultured as described by Silverio et al., 2010 [21]. Then, PDLCs were characterized into low (l-PDLCs) and high osteogenic potential (h-PDLCs), according to their capacity of mineral deposition, in vitro, and according to our previous publications [9–12]. Briefly, PDLCs were characterized according to Dominici et al. [4] to confirm the ability

to differentiate into osteogenic and adipogenic cell lineages and the expression/lack of expression of specific cell surface markers, such as CD166, CD34, and CD45 [22]. The levels of CD34 and CD45 were very similar between h- and l-PDLCs, showing less than 1% of the expression of positive cells. Regarding multipotency marker CD166, more than 95% of cells in both populations showed positive expressions [9–12]. Alizarin red staining was performed to assess the amount of mineral matrix produced in vitro by each cell population [10]. Consequently, PDLCs were classified either as high osteogenic potential PDLCs (h-PDLCs), which was the cell population with the capacity to produce higher amounts of the mineral matrix or low osteogenic potential PDLCs (l-PDLCs with a lower capacity to produce mineral matrix). Based on our previous studies [10,12], we chose day 10 of the osteogenic media (OM) induction as the time point to analyze epigenomic and transcriptomic changes. l- and h-PDLCs were cultured in high glucose Dulbecco's Modified Eagle Medium (DMEM), 10% Fetal Bovine Serum (FBS), 100 U/mL of penicillin, and 100 mg/mL of streptomycin and maintained at 37 °C in a humidified atmosphere containing 5% CO₂. Three independent experiments were performed for each PDLC, with three technical replicates for each one, with cells in the passages P5-P8, except when stated otherwise.

2.2. Osteogenic Stimulation

Both l- and h-PDLCs were seeded into 6-well plates (2.5×10^5 cells/well) in DMEM, 10% FBS, and antibiotics. After 24 h of incubation, for cell adhesion, the culture medium was removed, and the cells were cultivated in non-induced medium (DMEM) supplemented with 10% FBS and 1% antibiotics (penicillin 100 U/mL and streptomycin 100 mg/mL) or in induced osteogenic medium (OM) with DMEM, 10% FBS, 1% antibiotics, and supplemented with ascorbic acid (50 µg/mL), β-glycerophosphate (10 mM), and dexamethasone (10 nM). Cells were incubated and collected after 10 days, with media change every three days. PDLCs were divided according to the following groups:

- (i) l-DMEM: l-PDLCs cultivated in DMEM, standard medium.
- (ii) l-OM: l-PDLCs cultivated in OM, osteogenic medium.
- (iii) h-DMEM: h-PDLCs cultivated in DMEM, standard medium.
- (iv) h-OM: h-PDLCs cultivated in OM, osteogenic medium.

For all comparisons, the h-DMEM was set as the control group x l-DMEM, and the h-OM was set as the control group x l-OM.

2.3. Assay for Transposase-Accessible Chromatin Using Sequencing (ATAC-Seq)

A total of 5×10^4 cells were harvested from each group and were incubated in a transposition reaction, as preconized by Buenrostro et al. [23]. The Tn5 transposase enzyme was used to insert an adapter sequence into the accessible chromatin regions, combined with single-step library digestion and preparation. Digitonin was included to reduce contamination with mitochondrial DNA [24]. Then, the transposase-containing DNA fragments were amplified by PCR and purified to select the appropriate size of the fragments. Sample quality was evaluated by TapeStation, quantified by PCR with the Kapa Sybr Fast LightCycler 480 kit, and pooled for subsequent sequencing at Illumina NextSeq 500 platform (Illumina Inc., Foster City, CA, USA) in the Genomics Birmingham Facility (Birmingham, UK). Two independent experiments were performed.

2.4. Global DNA Methylation Analysis (Methylome)

DNA Isolation and Oxidative Bisulfite Conversion

The groups were cultured (8.7×10^5 cells/100 mm dishes) as described above. After 10 days, the culture medium was removed, and the cells were washed two times with PBS and scrapped off in extraction buffer with proteinase K. Total DNA was purified by extraction with phenol/chloroform/isoamyl alcohol and stored at −20 °C. DNA's concentrations and quality were assessed using Qubit (Thermo Fisher Scientific Inc., Rockford, IL, USA) and spectrophotometer (Nanodrop 1000; Nanodrop Technologies LLC, Wilmington, NC,

USA). The oxidative bisulfite conversion reaction was performed according to the protocol described by Assis et al. [12].

2.5. RNA Sequencing (RNA-SEQ)

2.5.1. I. RNA Extraction

Cells from all groups were cultivated at 1.5×10^5 per well in 6-well plate, as previously described. After 10 days, the culture medium of each well from each group was removed, the cells were washed with PBS and scrapped off with TRIzol reagent (Invitrogen, Cat #15596-018, Waltham, MA, USA), according to the manufacturer's recommendation. Total RNA extraction was performed, and RNA samples were treated with DNA-free Turbo solution to remove genomic DNA (Ambion, Cat #1907, Austin, TX, USA). Then, the samples were submitted to integrity and concentration analysis by the Agilent 2100 bioanalyzer, with an RNA Integrity Number (RIN) value greater than 8.

2.5.2. II. RNA Sequencing

The samples were pooled in equal concentration, prepared, and sequenced with Illumina TruSeq Stranded mRNA Sample Prep Kit, according to the manufacturer's instructions, in Illumina NextSeq 500 platform (Illumina Inc., Foster City, CA, USA). Briefly, 1 µg of DNA-free total RNA samples were processed. The mRNA was fragmented and copied into the first strand cDNA, followed by second strand cDNA synthesis. cDNA fragments were submitted to the final repair process, addition of single adenosine base, and adapter ligation. Finally, the processed cDNA was amplified by 15 cycles of PCR to create the cDNA library, which was read on the HiSeq 2500 (v3) sequencer (Illumina, San Diego, CA, USA).

For more details about these genome wide methodologies, please go to Assis et al., 2022.

2.6. Bioinformatics Analysis

2.6.1. III. ATAC-Seq

The reads were aligned to the human genome (h19) through the *Bowtie2 tool* [25] and the duplicates were removed. The ENCODE consortium identified the blacklist reads, defined as anomalous. Non-exclusively mapped reads have been filtered and peaks have been called using the "factor mode" in HOMER [26], based on the default settings. A Bigwig file was also generated and uploaded in the UCSC Genome Navigator [22] to allow visualization of the accumulated reads. Analyses using HOMER and subsequent analyses were conducted with the help of Dr. Samuel Clokie (West Midlands Regional Genetics Laboratory, Birmingham Women's Hospital, Birmingham, UK).

2.6.2. IV. DNA Methylation

Epic BeadChips data was processed through Illumina Genome Studio program [27] and minfi [28], implemented in R, associated with dplyr [29] and tidyr [30] packages. Data normalization was performed using quartiles methods. Probes were considered differentially methylated (DMP) when presented values of delta $\beta > 0.2$ (hypermethylated) or < -0.2 (hypomethylated) and p-value < 0.01 . The average methylation intensity (AMI) of the TFs was calculated by the sum of the average β (AVG- β) values of the detected probes (sd), divided by the number of probes:

$$\overline{AMI} = \frac{\sum AVG-\beta(sd)}{n(sd)}. \quad (1)$$

The scatter plots representing the annotation, AMI, and AVG- β values and AMI of the TFs probes were generated in Microsoft Excel 365.

2.6.3. RNA-Seq

The reads were aligned to the hg19 reference genome and counted using the R *Rsubread* package [31]. Quantification was performed according to the last recommended pipeline, as defined in the *edgeR* software [32]. Genes were considered differentially expressed (DEGs) when presented values of log₂FoldChange (logFC) ≥ 1.5 (upregulated) or ≤ -1.5 (downregulated), and *p*-value ≤ 0.05 . “Heatmaps” representing the logFC intensity were generated in R environment with the ComplexHeatmap package [33], and the “volcano plots” representing both logFC amplitude and statistical significance ($-\log_{10}p$ value), with the EnhancedVolcano package [34].

2.6.4. Selection of lncRNAs and Prediction Analyses

The lncRNAs found DEGs (DElncRNAs) were upregulated in common between the RNA-seq datasets of the l-DMEM (\times h-DMEM) and l-OM (\times h-OM) groups and were selected and submitted to prediction analysis for RNA:RNA base-pairing interactions with the TFs, using the Rblast prediction program, based on the LncRRlsearch web server [35]. Since the TruSeq Stranded mRNA Sample Prep Kit, used in RNA Sequencing, is not quite suitable for non-coding RNA without polyA tails, we check if the selected DElncRNAs have polyadenylation sites from 3' end sequencing, in the PolyASite database [36]. Next, they verified the potentials of DNA:RNA triplex formation of these predicted DElncRNAs and of the TFs, using the TriplexFPP machine learning program, in the Python environment, which (1) predicted the probabilities of triplex-forming oligonucleotide (TFO), in practice, for these DElncRNAs and (2) the potentials triplex target sites (TTSs) on the TFs DNA sequences, based on experimentally verified data, considering positive triplex-forming for score > 0.5 [37]. The FASTA sequences input in the program were obtained in the Ensembl genome browser, under the “Human (GRCh37.p13)” parameter [38]. The pie charts and Venn diagrams representing the distribution of the DElncRNAs among the groups was generated in Microsoft Excel 365 and with the online tool Interactivenn, respectively [39].

2.7. PCR Analysis

cDNA synthesis was performed using 1 mg RNA as described previously [40]. Quantitative polymerase chain reaction (PCR) was carried out for each one of the three independent experiments, using LightCycler 96 Real-Time PCR System (Roche Diagnostics GmbH, Mannheim, Germany) and FastStart Essential DNA Green Master (Roche Diagnostic Co., Indianapolis, IN, USA), according to the manufacturer’s instructions and in technical triplicates. The primers’ sequences and reaction details are shown in Supplementary Table S1. The results of *MIR31HG*, *LINC00939*, *SP7* and *DLX4* were obtained from three biological replicates, in technical triplicates, analyzed by $\Delta\Delta C_t$ method [41] and are presented as relative amounts of the target gene using *18S* as inner reference gene.

2.8. Statistical Analysis of the PCR Data

Data were initially examined for normality by Shapiro–Wilk test and expressed as mean \pm standard deviation (SD). After normal data distribution was confirmed, one-way analysis of variance (ANOVA $\alpha \leq 0.05$) followed by pairwise multiple-comparison test (Tukey) were used to identify the difference amongst groups (GraphPad Prism 7—GraphPad Software Inc., San Diego, CA, USA).

3. Results

3.1. l-PDLCs Show More Accessible Chromatin Regions in *SP7* and *DLX4* Genes than h-PDLCs

The chromatin conformation coordinates the DNA accessibility of the transcriptional machinery, composed of the RNA polymerase II, TFs, among other molecular elements, responsible for gene expression. The chromatin accessibility analysis on *SP7* and *DLX4* genomic regions in both h- and l-PDLCs, induced (h-OM and l-OM) and non-induced (h-DMEM and l-DMEM), to osteogenesis in vitro at 10 days showed more accessible chromatin peaks in l-PDLCs compared to h-PDLCs, in both culture conditions, i.e., induced, and non-

induced (Figure 1). In summary, l-PDLCs exhibit chromatin conformation more favorable to *SP7* and *DLX4* genes transcription than h-PDLCs.

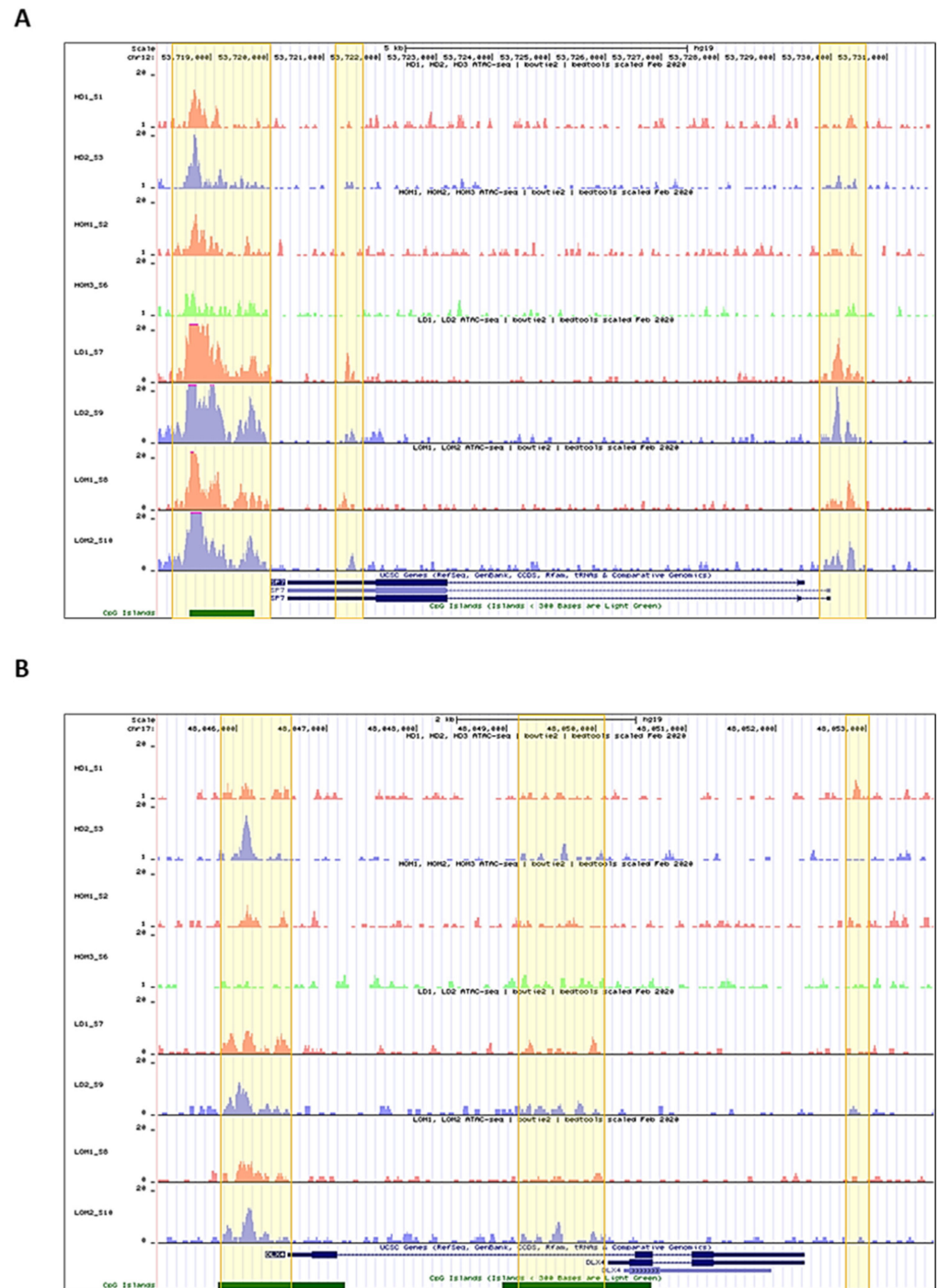


Figure 1. Overview of the chromatin accessibility on the *SP7* and *DLX4* gene regions in each h-PDLCs and l-PDLCs sample: The *bigwig* annotation tracks represent the accessible peaks on the *SP7* (A) and *DLX4* (B) gene regions in each h-PDLCs (DMEM and OM) and l-PDLCs (DMEM and OM) sample. Highlighted in yellow are the genomic regions with higher peak scores. *HD1_S1* and *HD2_S3* = h-DMEM samples; *HOM1_S2* and *HOM3_S6* = h-OM samples; *LD1_S7* and *LD2_S9* = l-DMEM samples; *LOM1_S8* and *LOM2_S20* = l-OM samples.

3.2. DNA Methylation Patterns amongst l- and h-PDLCS Are Similar for SP7 and DLX4

In turn, DNA methylation is one of the main epigenetic mechanisms that regulate chromatin conformation and gene transcription. We investigated the DNA methylation patterns of the *SP7* and *DLX4* genes in the Methylome metadata. Despite the probes presenting different AVG- β values in *SP7* and *DLX4*, the average methylation intensities (AMI) were similar, exhibiting a less methylated pattern, except for the 3'UTR region in *DLX4*, whose AMI pattern was more methylated (Figure 2). No differentially methylated probes (DMPs) were found between l- and h-PDLCS, in both culture conditions. These results reveal similar DNA methylation patterns among *SP7* and *DLX4* (except in the 3'UTR region), without significant differences between l- and h-PDLCS. There were no other found osteogenic TFs or signaling pathways that were differently expressed among l- and h-PDLCS.

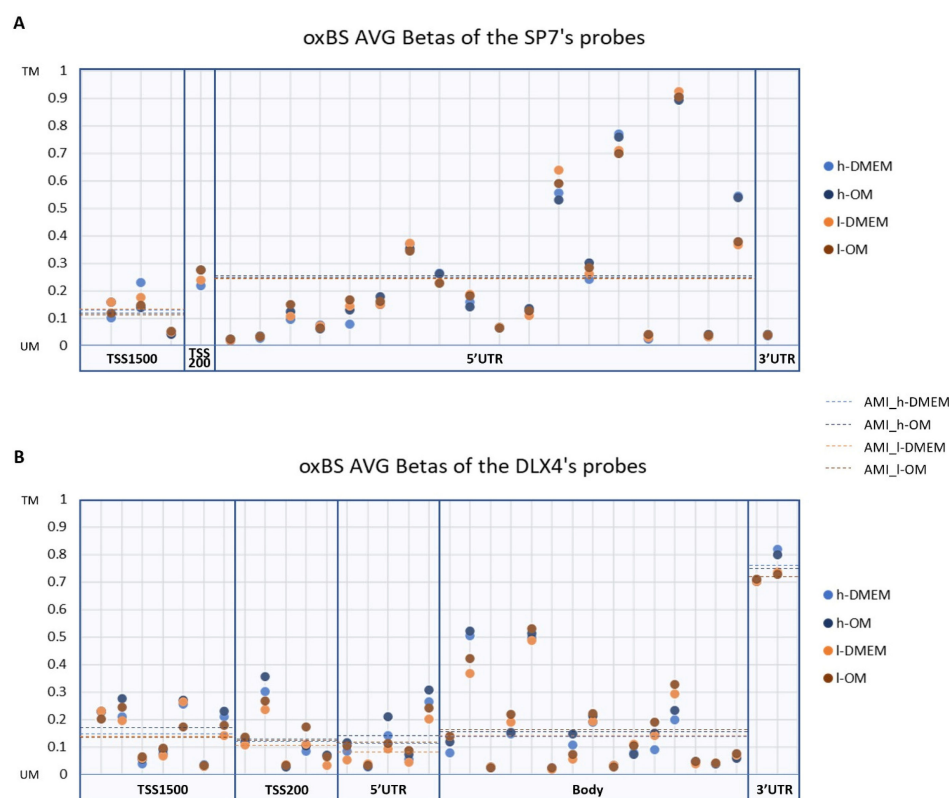


Figure 2. DNA methylation patterns in *SP7* and *DLX4* gene regions: The scatter plots represent the average β values (AVG- β) of each probe detected for *SP7* (A) and *DLX4* (B) gene regions in the oxidative bisulfite sequencing (oxBS) dataset. The dotted lines indicate the average methylation intensity (AMI) of the probes according to *UCSC_RefGene_Group*. TM: Totally Methylated; UM: Unmethylated.

The gene transcription mechanism is the biological consequence of molecular events that occur in extra and intracellular cascades called signaling pathways. Due to the inconsistency observed between the transcriptional profile and the epigenetic background of the TFs *SP7* and *DLX4* in our previous research [12], we analyzed the transcriptional profiles of the main signaling pathways involved in osteogenesis, in the l-DMEM (x h-DMEM), and in the l-OM (x h-OM) RNA-seq dataset. In addition, we analyzed the transcriptional profiles of other osteogenic TFs, like *RUNX2*, *SATB2*, *ATF4*, *MSX2*, *CTNNB1*, *DLX5*, and *TWIST1/2*. No differentially expressed genes (DEGs) were found for these TFs (Figure 3A) or signaling pathways, except for *Sclerostin* (*SOST*), a negative regulator of the WNT/ β -catenin pathway, founded DEG (downregulated) in both l-PDLCS groups (Figure 3B). This result points to no significant differences in the transcriptional profiles of other osteogenic TFs or signaling pathways between l- and h-PDLCS, in both culture conditions.

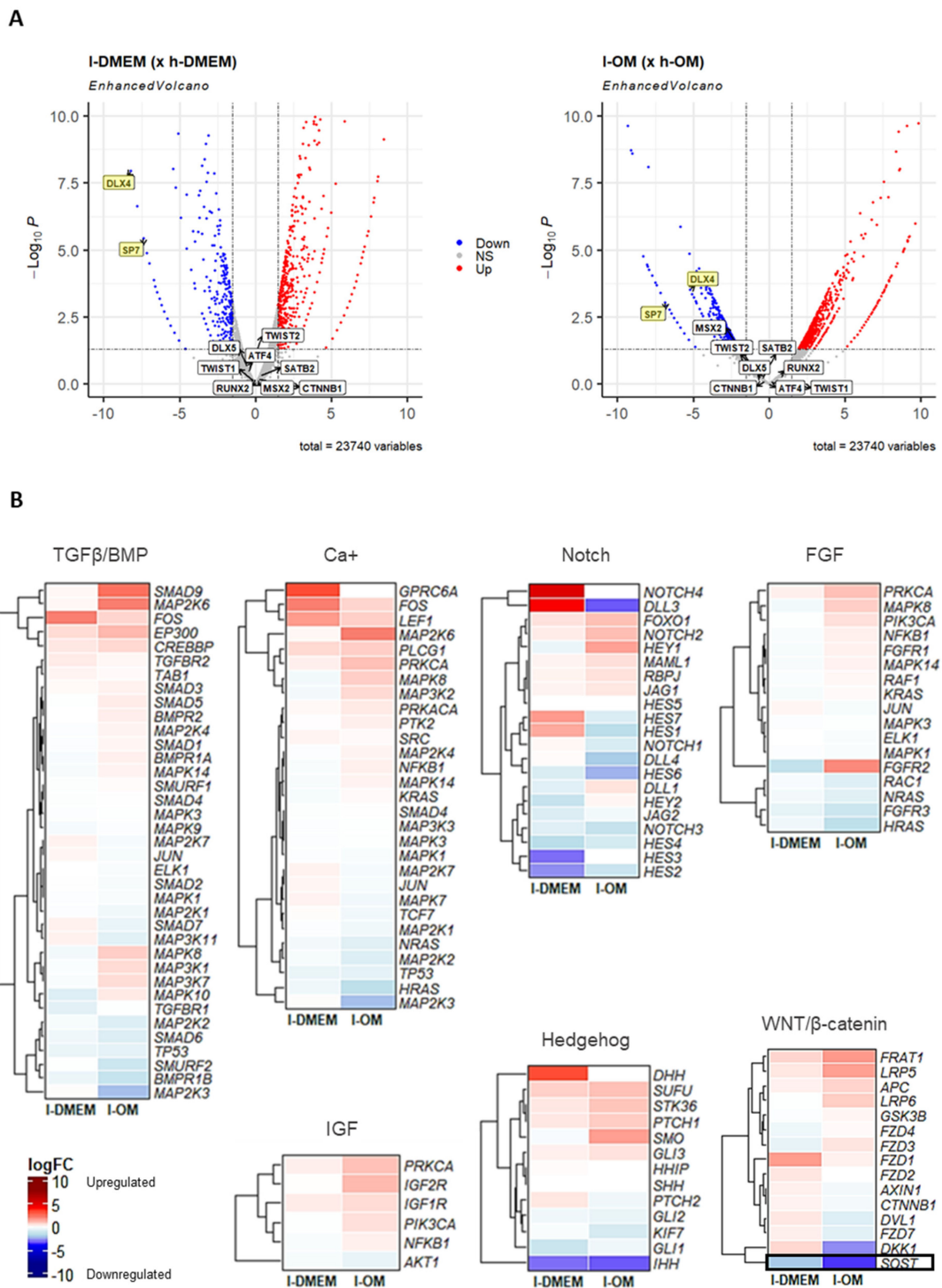


Figure 3. Analysis of other osteogenic TFs and signaling pathways in the RNA-seq dataset: The Volcano plots (A) show the magnitude of change in gene expression (logFC) and statistical significance (p -value) of the major osteogenic TFs, in I-DMEM (\times h-DMEM) (left) and I-OM (\times h-OM) (right). Highlighted in yellow are the *SP7* and *DLX4*, unique differentially expressed genes (logFC < -1.5 or > 1.5 and p -value < 0.05) in both groups. Down: downregulated; Up: upregulated; NS: not significant. The Heatmap graphs (B) represent the logFC of the genes involved in the major osteogenic signaling pathways, in I-DMEM (\times h-DMEM) and I-OM (\times h-OM). It highlighted the *Sclerostin* (*SOST*) gene, a negative regulator of the WNT/ β -catenin pathway, and was found to be differentially expressed (DEG) in both groups in I-PDLCs (\times h-PDLCs). Reddish colors = upregulated; bluish colors = downregulated.

3.3. The lncRNAs LINC00939 and MIR31HG Are Upregulated in l-PDLCs

The lncRNAs are important agents of the epigenetic machinery that act on the transcriptional, post-transcriptional, and translational regulations. Therefore, we also investigated the lncRNAs transcriptional profiles. Of all 1643 lncRNAs identified, 73 ($\approx 4.5\%$) were DEGs (DElncRNAs) in the l-DMEM group and 82 ($\approx 5\%$) in the l-OM group. There were 18 DElncRNAs in common among the groups, of which four long intergenic non-protein coding RNAs (LINC00939, LINC01260, LINC01347 and LINC02210), four uncharacterized (LOCs 286178, 408186, 100506476 and 100507547), and one host-gene (MIR31HG) were upregulated in both groups (Figure 4).

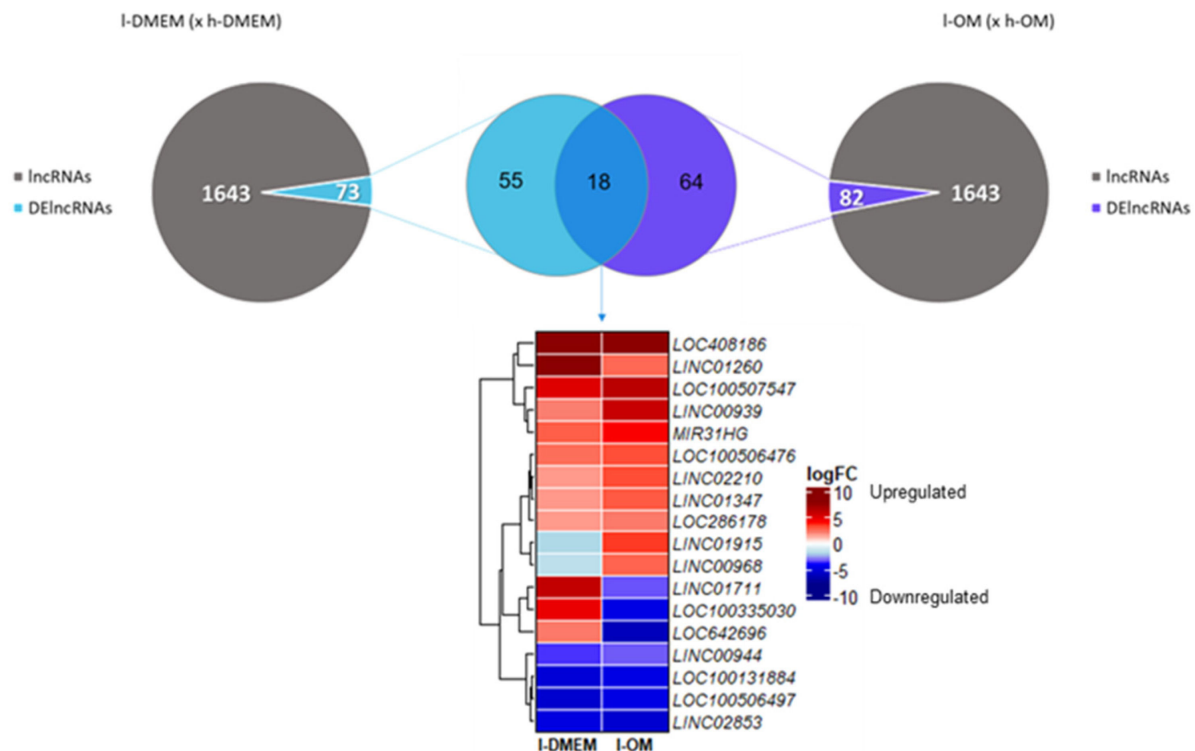


Figure 4. TFs Transcriptional Profiles: The Volcano plots show the magnitude of change in gene expression (logFC) and statistical significance (p -value) of the TFs, in l-DMEM (\times h-DMEM) (A) and l-OM (\times h-OM) (B). As highlighted, the SP7 and FABP4 genes are differentially expressed ($\log_{2}FC < -1.5$ or >1.5 and p -value < 0.05). Down: downregulated; Up: upregulated; NS: not significant.

LncRNAs, MIR31HG and LINC00939 Are Predicted to Interact with SP7 and DLX4 by RNA:RNA Base-Pairing

The RNA:RNA base-pairing interaction is an epigenetic mechanism played by lncRNAs to regulate the genic post-transcription, in order to repress or promote the transcripts translation process. Aiming to verify the possibility of RNA:RNA base-pairing interaction between the TFs *SP7* and *DLX4*, and the DElncRNAs that are upregulated in common among l-DMEM and l-OM groups used the prediction program RIBlast, based on the LncRRsearch webserver. There were predicted interactions between SP7:MIR31HG, SP7:LINC00939 (Figure 5), DLX4:MIR31HG, and DLX4:LINC00939 (Figure 6). All interactions presented negative minimum energy (MinEnergy) <12 kcal/mol, which characterizes a high affinity of binding between the transcripts. Both MIR31HG and LINC00939 were identified as 3' polyadenylated (Figure S1). The results corroborate the probability of post-transcriptional and/or translational repressive regulation on *SP7* and *DLX4* by MIR31HG and LINC00939 base-pairing interactions and suggest that these interactions occur passively, without the need for enzymes or catalytic molecular agents, due to the negative binding energy values presented.

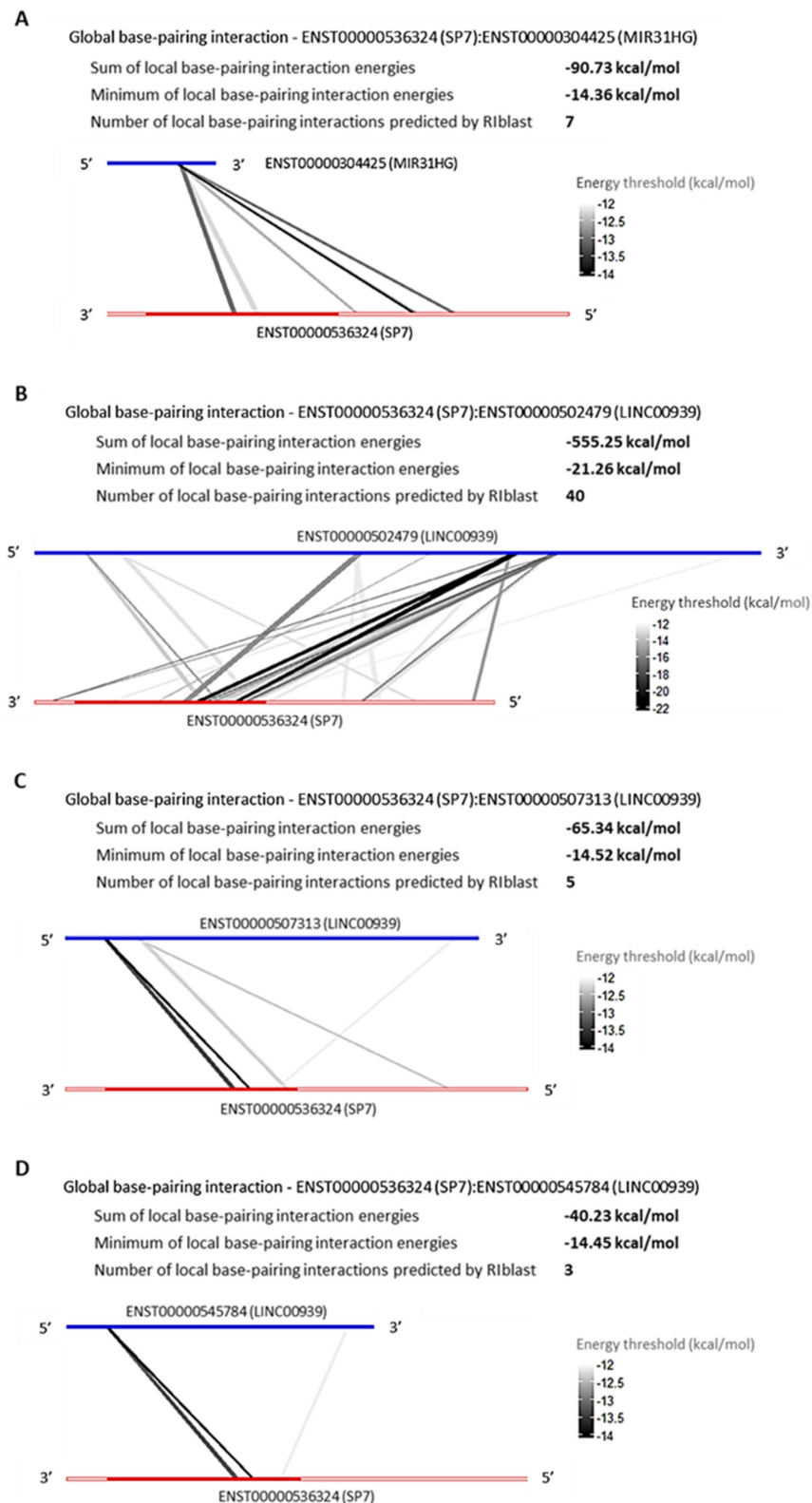


Figure 5. Global base-pairing interactions between SP7 and predicted DElncRNAs transcripts; Global base-pairing interactions between ENST00000536324 (SP7) and ENTS00000304425 (MIR31HG) (A), ENST00000502479 (LINC00939) (B), ENST00000507313 (LINC00939) (C), and ENST00000545784 (LINC00939) (D) transcripts. The blue bars represent the lincRNA transcripts and the red bars represent the SP7 transcript. The grayish lines represent the predicted interactions according to the energy threshold and SP7's annotation (3'UTR, CDS, and 5'UTR regions).

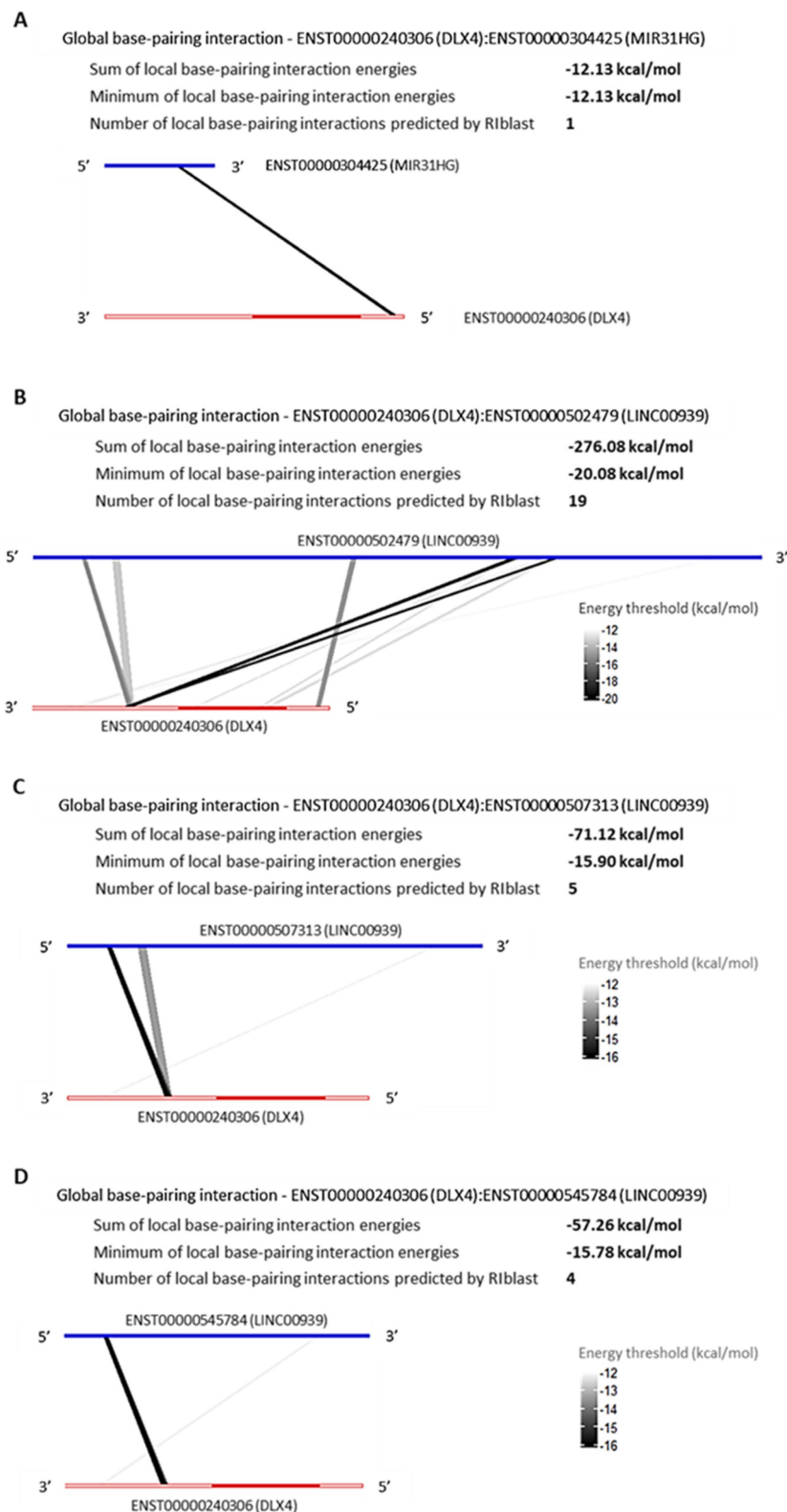


Figure 6. Global base-pairing interactions between DLX4 and predicted DElncRNAs transcripts; Global base-pairing interactions between ENST00000240306 (DLX4) and ENTS00000304425 (MIR31HG) (A), ENST00000502479 (LINC00939) (B), ENST00000507313 (LINC00939) (C), and ENST00000545784 (LINC00939) (D) transcripts. The blue bars represent the lincRNA transcripts and the red bars represent the DLX4 transcript. The grayish lines represent the predicted interactions according to the energy threshold and DLX4's annotation (3'UTR, CDS, and 5'UTR regions).

3.4. Machine Learning Program Predicted Potential for DNA:RNA Triplex-Forming for the SP7 DNA Site and for the lncRNA LINC00939

Another epigenetic mechanism mediated by lncRNAs is the DNA:RNA triplex-forming, which consists of interactions between lncRNAs and DNA-specific sequences, via triple-helix (triplex) formation, to inhibit or induce gene transcriptions. We performed prediction analysis of DNA:RNA triplex-forming potential for the SP7 and DLX4 DNA sites and for the MIR31HG and LINC00939 transcripts. Triplex-forming potential was predicted for the SP7 DNA sites (score = 0.68) (Table 1) and for the first exon (ENSE00002048191) of one LINC00939 transcript (ENST00000502479.1) (score = 0.75) (Table 2). Conversely, there was no prediction of triplex-forming potential for the DLX4 DNA site, nor for the lncRNA MIR31HG. These results point to a potential transcriptional regulation on the SP7 DNA site (inhibition) by triplex formation with the lncRNA LINC00939.

Table 1. Triplex prediction for the SP7 and DLX4 DNA sites.

TRIPLEX DNA SITE PREDICTION					
Gene Symbol	Assembly	Chromosome	Map Location	Score	Triplex-Forming
SP7	GRCh37	12	53720362:53739099	0.68235606	Triplex
DLX4	GRCh37	17	48046334:48052321	0.13426343	Nontriplex

3.5. MIR31HG and LINC00939 Transcripts Are Upregulated while SP7 and DLX4 Are Downregulated in l-PDLCs

In agreement with RNA-seq data, the expression of the lncRNAs, MIR31HG and LINC00939, were upregulated in l-PDLCs (\times h-PDLCs) in both conditions, non-induced (DMEM) and induced (OM), with statistical significance for MIR31HG (l-DMEM \times l-OM, $p = 0.001$; l-OM \times h-OM, $p = 0.001$) and for LINC00939 (l-DMEM \times l-OM, $p = 0.004$; l-DMEM \times h-DMEM, $p = 0.015$; l-OM \times h-OM, $p < 0.0001$). In addition, the transcript levels of the osteogenic TFs SP7 and DLX4 in l-PDLCs were lower, with statistical significance for SP7 (l-DMEM \times l-OM, $p = 0.002$; l-DMEM \times h-DMEM, $p < 0.0001$; h-DMEM \times h-OM, $p = 0.004$; l-OM \times h-OM, $p < 0.0001$), and for DLX4 (h-DMEM \times h-OM, $p = 0.005$; l-OM \times h-OM, $p < 0.0002$) (Figure 7). These results show a negative correlation between MIR31HG/LINC00939 and SP7/DLX4 expression profiles, reinforcing the hypothesis of repressive action of these lncRNAs on the (post)transcriptional regulation of SP7 and DLX4.

Table 2. Triplex-forming potential of the lncRNAs MIR31HG and LINC00939.

LNCRNAS TRIPLEX-FORMING POTENTIAL					
Gene Symbol	Transcript ID	Annotation	Exon ID	Score	Triplex-Forming
MIR31HG	ENST00000304425.3	exon	ENSE00001540342	0.0	Nontriplex
		exon	ENSE00001540341	0.0	Nontriplex
		exon	ENSE00001540339	0.0	Nontriplex
		exon	ENSE00001729409	0.0	Nontriplex
		intron 1		0.0	Nontriplex
		intron 2		0.0	Nontriplex
		intron 3		0.0	Nontriplex

Table 2. Cont.

LNCRNAs TRIPLEX-FORMING POTENTIAL					
Gene Symbol	Transcript ID	Annotation	Exon ID	Score	Triplex-Forming
LINC00939	ENST00000502479.1	exon	ENSE00002048191	0.7498492	Triplex
		exon	ENSE00002044944	0.0	Nontriplex
		exon	ENSE00001441774	0.0	Nontriplex
		exon	ENSE00002049964	0.033063784	Nontriplex
		exon	ENSE00002021220	0.30899337	Nontriplex
		exon	ENSE00002063907	0.0	Nontriplex
		intron 1		0.0	Nontriplex
		intron 2		0.0	Nontriplex
		intron 3		0.0	Nontriplex
		intron 4		0.0	Nontriplex
		intron 5		0.0	Nontriplex
LINC00939	ENST00000507313.1	exon	ENSE00002060110	0.0	Nontriplex
		exon	ENSE00002044944	0.0	Nontriplex
		exon	ENSE00001441774	0.0	Nontriplex
		exon	ENSE00002033127	0.0	Nontriplex
		exon	ENSE00002081221	0.0	Nontriplex
		exon	ENSE00002046518	0.0	Nontriplex
		intron 1		0.0	Nontriplex
		intron 2		0.0	Nontriplex
		intron 3		0.0	Nontriplex
		intron 4		0.124810636	Nontriplex
		intron 5		0.0	Nontriplex
LINC00939	ENST00000545784.1	exon	ENSE00002206498	0.0	Nontriplex
		exon	ENSE00002044944	0.0	Nontriplex
		exon	ENSE00002228181	0.0	Nontriplex
		exon	ENSE00002271284	0.0	Nontriplex
		exon	ENSE00002021220	0.30899337	Nontriplex
		exon	ENSE00002216415	0.0	Nontriplex
		intron 1		0.0	Nontriplex
		intron 2		0.0	Nontriplex
		intron 3		0.124810636	Nontriplex
		intron 4		0.0	Nontriplex
		intron 5		0.0	Nontriplex

Altogether, the results of chromatin accessibility, DNA methylation levels, RNA-seq, PCR, and predicted base-pairing interactions suggest *MIR31HG* and *LINC00939* might be good predictors of osteogenic commitment.

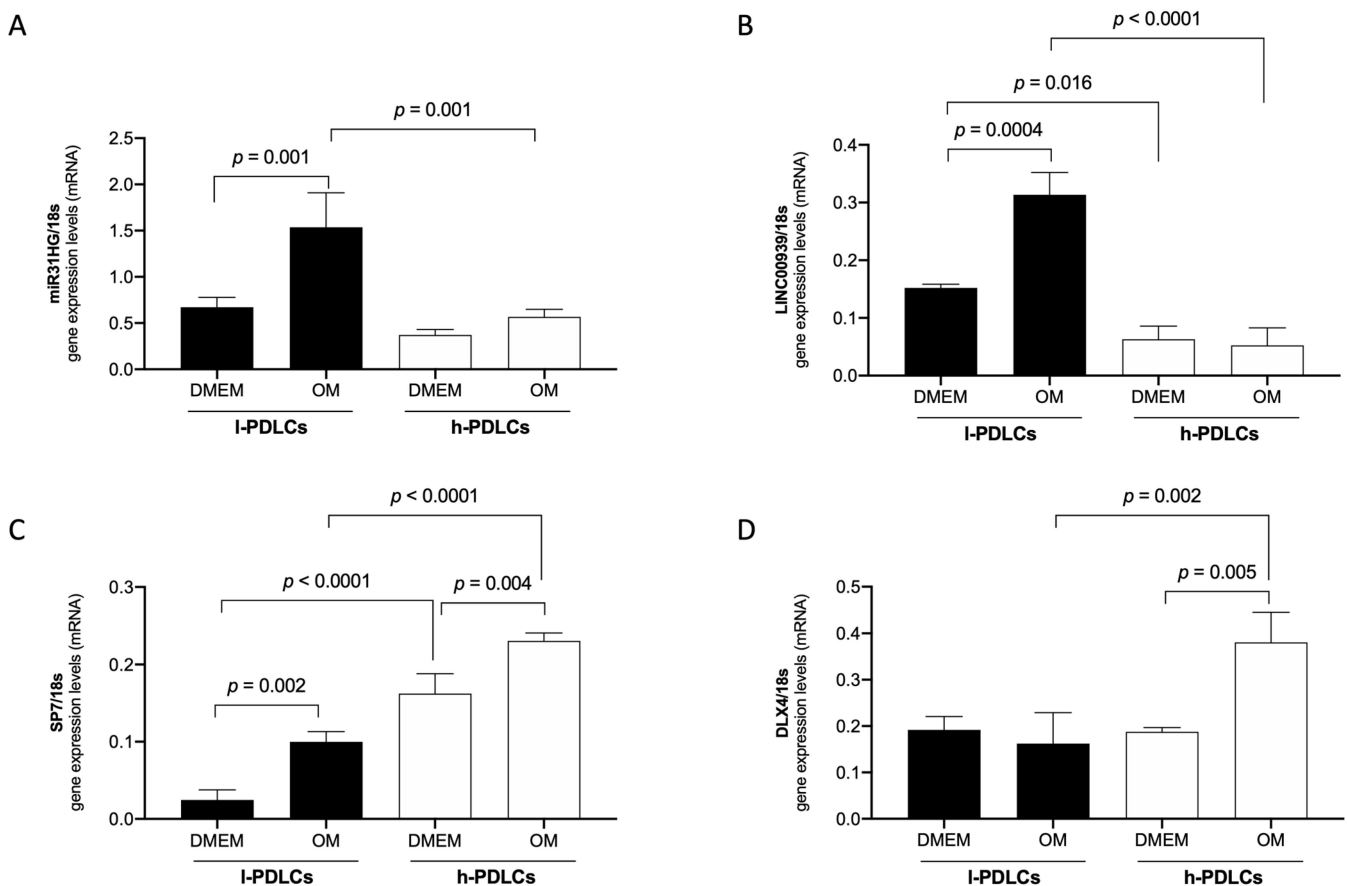


Figure 7. Gene expression levels of MIR31HG, LINC00939, SP7, and DLX4 in l- and h-PDLCs, cultured at 10 days in non-induced (DMEM) and osteogenic (OM) medium: qPCR analysis of MIR31HG (A), LINC00939 (B), SP7 (C), and DLX4 (D) in l- and h-PDLCs, cultured at 10 days in a non-induced (DMEM) and osteogenic (OM) medium. The results represent a mean \pm standard deviation of three biological replicates, considering differential expression for $p < 0.05$.

4. Discussion

The advancement of genomic sequencing techniques and bioinformatic tools has helped us to better understand the epigenetic machinery involved in the phenotype acquisition of MSCs. Chromatin conformation, DNA methylation pattern, transcriptional profile, and ncRNAs interactions are part of this machinery and impact cell differentiation potential. Here, we aimed to investigate the epigenetic machinery of the transcription factors (TFs), *SP7* and *DLX4*, involved in the osteoblastic differentiation, in PDLCs, using omic techniques and bioinformatic tools. We provided data that indicate *SP7* and *DLX4* might be regulated at transcriptional, post-transcriptional, and translational steps by the epigenetic mediators MIR31HG and LINC00939 and this could impact their osteogenic phenotype acquisition. So far, our results indicate *MIR31HG* and *LINC00939* might be good predictors of osteogenic commitment.

In our previous investigation, we showed an epigenetic landscape more accessible in l-PDLCs, when compared to h-PDLCs [12]; however, several TFs such as *DLX4* and *SP7* were pointed out to be downregulated in l-PDLCs. Here, in the ATAC-seq analyses, we found more accessible chromatin peaks for both, *SP7* and *DLX4* genomic regions in l-PDLCs, when compared to h-PDLCs, except to the transcription end site (TES) region of the *DLX4*. One of the hypotheses for this discrepancy would be the differential presence, concentration, and action of some epigenetic chromatin regulators, such as DNA methyltransferases (DNMTs) and histone deacetylases (HDACs), among the PDLCs populations. Nevertheless, to verify this hypothesis, it would be necessary to integrate other analyses that were not performed

in this study, for example, chromatin immunoprecipitation (ChIP). The difference in peak score amplitude between l- and h-PDLCs was greater in *SP7* than in *DLX4*, and the peaks annotation also was distinct among the TFs genomic regions, with peaks concentration in exonic, TSS, and TES regions for *SP7*, and peaks concentration in intronic and TSS regions for *DLX4*. A study performed by Tai et al., in 2017, discovered a similar profile of chromatin accessibility on the *SP7* genomic region during the MC3T3 pre-osteoblast differentiation [42], with high chromatin accessibility centered within ~1 kb downstream of TES. Furthermore, this region coincided with CpG islands enrichments in *SP7*, indicating its importance in gene regulation. In the same, abovementioned study, accessible peaks were localized in the intronic region and in all osteoblastogenesis stages (pre-osteoblast, matrix deposition, nodule formation, and mineralization). These findings show chromatin conformation more favorable to *SP7* and *DLX4* transcription in l-PDLCs than h-PDLCs and suggest an important role of TES and intronic regions, i.e., non-promoter regions, on their gene regulations.

The TFs also presented distinct DNA methylation patterns to each other, without differentially methylated probes (DMPs) between l- and h-PDLCs. The average methylation intensities (AMIs) remained below 0.5 in all groups, indicating an overall DNA methylation pattern tending to unmethylated, except in the 3'UTR region of the *DLX4*, which presented AMIs above 0.5, ranged from 0.7 to 0.8, indicating DNA methylation pattern tending to totally methylated in this region. Conversely, very low methylation intensities were found in the 3'UTR region of the *SP7*, ranging around 0.03, which matches the high chromatin accessibility presented in ATAC-seq results for this TES region. In 2019, Lhoumaud et al. examined the interplay between chromatin accessibility and DNA methylation in mouse embryonic stem cells, using the *EpiMethylTag* method, that combined ATAC-seq with bisulfite conversion and showed that DNA methylation rarely coexists with chromatin accessibility [43].

The signaling pathways are complex cascades that are dependent on a “perfect syntony” among many molecular events to promote gene transcriptions. In our RNA-seq analyses, we did not find differentially expressed genes (DEGs) amongst l- and h-PDLCs in the main signaling pathways involved in the TFs gene transcription, except for the *Sclerostin* (*SOST*), which was found to be downregulated in l-PDLCs. Recent studies show a correlation of expression between *SOST* and several osteoblastic TFs, such as *RUNX2* and *SP7* [44,45]. This protein plays an important role in the osteoblast development, acting as an inhibitor of the Wnt/ β catenin pathway by binding to the cell transmembrane receptors low-density lipoprotein receptor protein 5 and 6 (LRP5/6) and Frizzled. However, the other genes of this pathway were not found in DEGs. Several studies show the repressor effect of Sirtuin 1 (*SIRT1*) on the *SOST* expression of osteocytes [46,47]. In our RNA-seq dataset, the *SIRT1* was found to be slightly upregulated in l-DMEM (logFC = 0.17) and l-OM (logFC = 0.49), without significant difference. So far, the results suggest a transcriptional status of the signaling pathways more favorable to TFs' expression in l-PDLCs compared to h-PDLCs.

Regarding the RNA-seq results for these TFs, our previous study [12] showed *SP7* and *DLX4* as DEGs and downregulated in both l-DMEM (x h-DMEM) and l-OM (x h-OM). However, these transcriptional profiles did not match the chromatin conformation or neither the DNA methylation pattern found here for these TFs. Both presented more accessible peaks, with higher peak scores in l-PDLCs than h-PDLCs, with DNA methylation patterns tending to be unmethylated in both populations. We expected to find opposite chromatin accessibility and DNA methylation patterns in l-PDLCs. Other studies also found discrepancies between ATAC-seq and RNA-seq results, and a positive correlation between DNA methylation and gene expression [48–50]. This paradox suggested the hypothesis of a fine-tuning gene regulation on *SP7* and *DLX4* in l-PDLCs, mediated by lncRNA interactions.

After we identified 18 lncRNAs DEGs in common between l-DMEM (x h-DMEM) and l-OM (x h-OM), only nine were upregulated in both groups, of which two (MIR31HG and LINC00939) were predicted by the web-based prediction tool to interact with *SP7* and *DLX4* by RNA:RNA base-pairing. This program is based on up-to-date benchmark data to compute the base-pairing probabilities inter and intramolecular between RNA sequences [51].

The MIR31HG is a host gene of the MIR31, a miRNA validated as an *SP7* repressor and osteogenesis inhibitor [52,53]. Nevertheless, to date, there are no studies correlating the *MIR31HG* and *SP7* expression profiles, nor predicting interactions between them. In 2018, Huang et al. cultivated BMSCs on titanium surfaces biofunctionalized with small interfering RNA (siRNA)-targeting MIR31HG and reported an increase in the relative expression of osteogenic genes such as *ALPL*, *RUNX2*, and *BGLAP*, without analyzing the *SP7* [54]. Since the MIR31HG hosts the MIR31, one possibility would be the MIR31HG produces precursors of MIR31 through intracellular shearing, and this represses the *SP7* [55]. However, the MIR31 was not found in DEG in our RNA-seq dataset, requiring a small RNA sequencing (miRNA-seq) to confirm this transcriptional profile. The LINC00939 is an Intergenic lncRNA, located on chromosome 12 (12q24.32), measuring 24,687 pairs of bases; although, we did not find previous publications about this lncRNA. Our hypothesis is both lncRNAs can regulate the post-transcriptional and/or translational processes of the *SP7* and *DLX4* by RNA:RNA base-pairing interactions. To the *SP7*, the interactions with MIR31HG were predicted in 5'UTR and coding sequence (CDS) regions, and LINC00939 was predicted with the majority being in the CDS region. In this case, both lncRNAs could repress the *SP7* post-transcription by "A to I" RNA edition, promoting hydrolytic deamination of adenosine to inosine [56], and/or the *SP7* translation by the impediment of ribosome subunit binding. To the *DLX4*, the interaction with MIR31HG was predicted in the 5'UTR region, and LINC00939 predicted the majority in the 3'UTR region. In this case, the MIR31HG could repress the *DLX4* translation by the impediment of ribosome subunit binding, and the LINC00939 could repress the *DLX4* post-transcription via Staufen-mediated decay, forming an intermolecular duplex with Alu element in the 3'UTR region [57,58]. Nonetheless, we should interpret these results with caution since the prediction tools "don't consider complex structure folding as RNA tertiary structures, non-canonical base-pairing and co-transcriptional folding process and may contain false-positive predictions" [35]. Thus, experimental validations involving silencing and/or overexpression of these lncRNAs followed by *SP7* and *DLX4* protein analyses are necessary to authenticate the causality of these predicted interactions.

Since RNA:RNA base-pairing is a post-transcriptional and translational regulatory mechanism, we also investigated the probability of transcriptional regulation by DNA:RNA triplex forming. For this purpose, we use the TriplexFPP, a machine learning program "based on the experimentally verified data, where the high-level features are learned by the convolutional neural networks" [37], which increases the potential of a triplex formation in practice. In contrast, the scarce number of validation assays limits the program training, which greatly restricts the scope of predictions. In our analysis, it was predicted a triplex target site (TTS) for the *SP7* gene and triplex-forming oligonucleotide (TFO) for the exon *ENSE00002048191* of the LINC00939 transcript *ENST00000502479.1*. We hypothesize that the *SP7*:LINC00939 triplex-forming represses the *SP7* transcription by steric hindrance of the RNA polymerase II at the promoter region [20,59]. However, the program does not predict matching between TTS–TFO, requiring experimental assays, such as chromatin oligo-affinity precipitation (ChOP), to confirm this interaction [60,61].

Finally, in an attempt to identify markers to assist with the selection of PDLCs with distinct osteogenic potential, we selected a panel of four genes for further confirmation by qPCR. The panel included both lncRNAs, *MIR31HG* and *LINC00939* and both osteogenic TFs, *SP7* and *DLX4*. The confirmation and combination of higher levels of *MIR31HG* and *LINC00939* transcripts alongside lower levels of *SP7* and *DLX4* gene expression in l-PDLCs after in vitro osteogenic induction indeed show potential in predicting distinct osteogenic

potential amongst PDLCs. Surprisingly, there was no correlation between qPCR and RNA-seq results for the *DLX4* gene expression among the l- and h-DMEM. Everaert et al. also reported discrepancies among qPCR and RNA-seq gene expression measurements, typically in smaller genes [62]. Coincidentally, *DLX4* was the smallest gene analyzed here (sizes: *DLX4* \approx 6 kb; *SP7* \approx 18 kb; *LINC00939* \approx 24 kb; *MIR31HG* \approx 104 kb).

Further studies with functional approach by silencing and overexpressing the lncRNAs, *MIR31HG* and *LINC00939*, and the impact on *SP7* and *DLX4* gene expression will be necessary to confirm the hypothesis here identified.

5. Conclusions

The set of results obtained here indicates the lncRNAs, *MIR31HG* and *LINC00939*, as possible epigenetic mediators on PDLCs osteogenic phenotype commitment through repressive regulation of the osteogenic TFs, *SP7* and *DLX4* (Figure 8). In this regard, *MIR31HG* and *LINC00939* might be good predictors of osteogenic commitment in PDLCs.

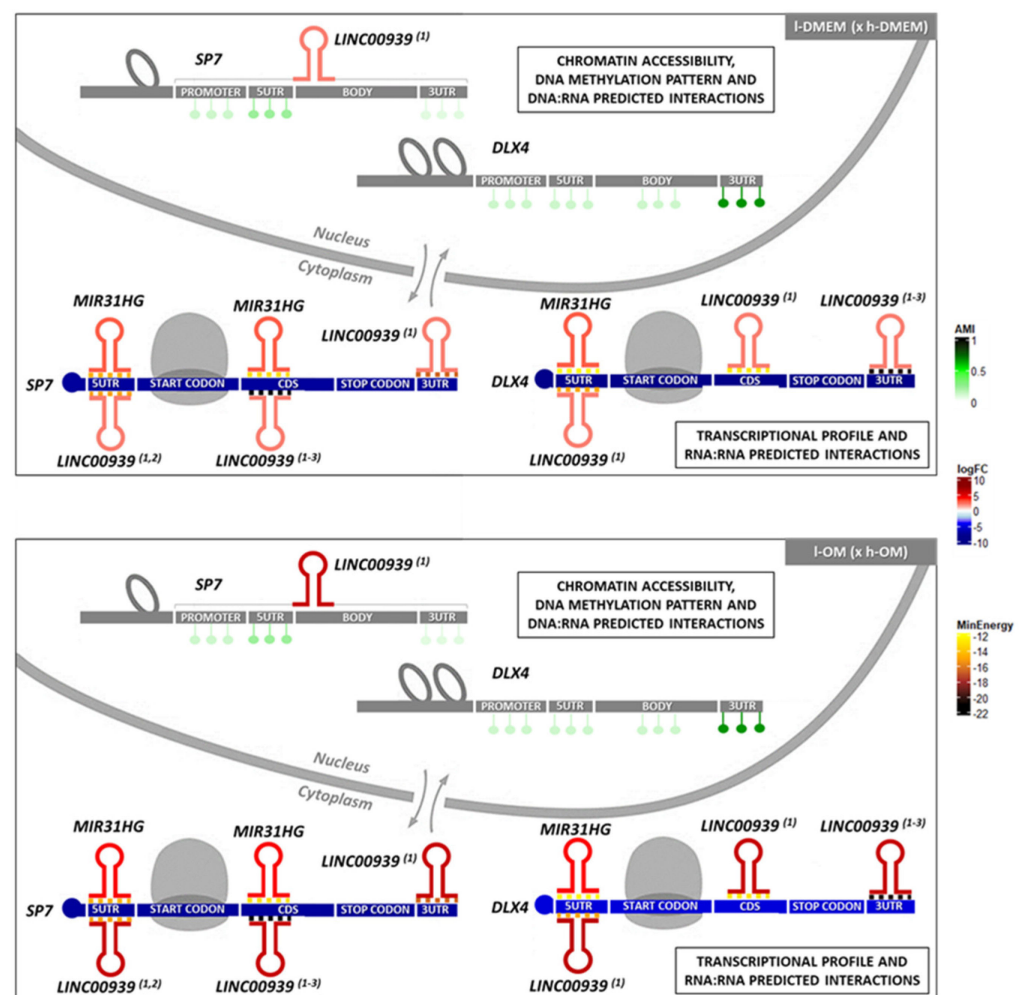


Figure 8. Schematic representation of the *SP7* and *DLX4* epigenetic status and their predicted interactions with the lncRNAs, *MIR31HG* and *LINC00939*, in I-PDLCs (\times h-PDLCs) at 10 days, cultured in DMEM and OM medium: The results of the ATAC-seq dataset show the chromatin is more accessible for the *SP7* and *DLX4* genes in I-PDLCs (DMEM and OM) compared with h-PDLCs (DMEM and OM). The results of the Methylome dataset show a less methylated pattern for the *SP7* and *DLX4* genes, in all groups, except to the 3'UTR region of the *DLX4*, which exhibit a more methylated pattern. The lncRNAs analyses in the RNA-seq dataset show the DElncRNAs *MIR31HG* and *LINC00939* (predicted to interact with *SP7* and *DLX4* by RNA:RNA base-pairing) upregulated in

both l-DMEM (\times h-DMEM) and l-OM (\times h-OM) groups. The DNA:RNA interactions analyses predicted potential triplex-forming for one LINC00939 transcript (ENST00000502479) and for the SP7 DNA site, in practice. These data together allow us to hypothesize the existence of a possible transcriptional repressor regulation on the SP7 by triplex-forming with LINC00939, and a post-transcriptional repressor regulation on the SP7 and DLX4 transcripts by MIR31HG and LINC00939 base-pairing interactions, in l-PDLCs at 10 days, cultured in DMEM and OM medium. (1) = ENST00000502479; (2) = ENST00000507313; (3) = ENST00000545784; CDS = coding DNA sequence; reddish colors = upregulated; bluish colors = downregulated; greenish colors = average methylation intensity (AMI); yellowish and brownish colors = minimum energy (kcal/mol) of the predicted RNA:RNA interactions. Adapted from Gomes et al., 2020 [63].

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes14081649/s1>, Table S1. Gene expression primers sequences and polymerase chain reaction cycle conditions. Figure S1. Search results for poly(A) sites on the genomic regions of the MIR31HG and LINC00939 lncRNAs.

Author Contributions: Conceptualization, R.S.F., R.A.d.S., M.W. and D.C.A.; methodology, R.S.F. and R.I.F.A.; validation, R.S.F., R.I.F.A. and F.R., formal analysis, R.S.F., F.R. and A.C.B.; investigation, R.S.F., R.I.F.A., F.R. and A.C.B.; resources, R.A.d.S., M.W. and D.C.A.; data curation, R.S.F.; writing—original draft preparation, R.S.F.; writing—review and editing, R.S.F., R.I.F.A., F.R., A.C.B., R.A.d.S., M.W. and D.C.A.; visualization, R.S.F. and D.C.A.; supervision, D.C.A.; project administration, D.C.A.; funding acquisition, R.S.F., M.W. and D.C.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by São Paulo Research Foundation-FAPESP/The University of Birmingham, UK Collaborative Research Program (grant number 2017/07944-5) and The APC was funded by Paulista University.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Piracicaba Dental School, University of Campinas, São Paulo, Brazil (CAAE55588816.4.0000.5418).

Informed Consent Statement: Written informed consent has been obtained from the patient(s) to publish this paper.

Data Availability Statement: The data presented in this study are available on request from the corresponding authors. The data are not publicly available due to potentially identifiable genomic information.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hmadcha, A.; Martin-Montalvo, A.; Gauthier, B.R.; Soria, B.; Capilla-Gonzalez, V. Therapeutic potential of mesenchymal stem cells for cancer therapy. *Front. Bioeng. Biotechnol.* **2020**, *8*, 43. [[CrossRef](#)] [[PubMed](#)]
2. Meirelles, L.D.S.; Chagastelles, P.C.; Nardi, N.B. Mesenchymal stem cells reside in virtually all post-natal organs and tissues. *J. Cell. Sci.* **2006**, *119*, 2204–2213. [[CrossRef](#)] [[PubMed](#)]
3. Hargreaves, K.M.; Diogenes, A.; Teixeira, F.B. Treatment options: Biological basis of regenerative endodontic procedures. *Pediatr. Dent.* **2013**, *35*, 129–140. [[CrossRef](#)]
4. Dominici, M.L.B.K.; Le Blanc, K.; Mueller, I.; Slaper-Cortenbach, I.; Marini, F.C.; Krause, D.S.; Deans, R.J.; Keating, A.; Prockop, D.J.; Horwitz, E.M. Minimal criteria for defining multipotent mesenchymal stromal cells: The International Society for Cellular Therapy position statement. *Cytotherapy* **2006**, *8*, 315–317. [[CrossRef](#)] [[PubMed](#)]
5. Salgado, A.J.; Reis, R.L.; Sousa, N.; Gimble, J.M. Adipose tissue derived stem cells secretome: Soluble factors and their roles in regenerative medicine. *Curr. Stem Cell Res. Ther.* **2010**, *5*, 103–110. [[CrossRef](#)]
6. Jiang, Y.; Jahagirdar, B.N.; Reinhardt, R.L.; Schwartz, R.E.; Keene, C.D.; Ortiz-Gonzalez, X.R.; Reyes, M.; Lenvik, T.; Lund, T.; Blackstad, M.; et al. Pluripotency of mesenchymal stem cells derived from adult marrow. *Nature* **2002**, *418*, 41–49. [[CrossRef](#)]
7. Sevari, S.P.; Ansari, S.; Moshaverinia, A. A narrative overview of utilizing biomaterials to recapitulate the salient regenerative features of dental-derived mesenchymal stem cells. *Int. J. Oral Sci.* **2021**, *13*, 22. [[CrossRef](#)]
8. Saito, M.T.; Salmon, C.R.; Amorim, B.R.; Ambrosano, G.M.; Casati, M.Z.; Sallum, E.A.; Nociti, F.H., Jr.; Silvério, K.G. Characterization of highly osteoblast/cementoblast cell clones from a CD105-enriched periodontal ligament progenitor cell population. *J. Periodontol.* **2014**, *85*, e205–e211. [[CrossRef](#)]

9. Assis, R.I.; Feltran, G.D.S.; Silva, M.E.S.; do Rosário Palma, I.C.; Rovai, E.S.; de Miranda, T.B.; Ferreira, M.R.; Zambuzzi, W.F.; Birbrair, A.; Andia, D.C.; et al. Non-coding RNAs repressive role in post-transcriptional processing of RUNX2 during the acquisition of the osteogenic phenotype of periodontal ligament mesenchymal stem cells. *Dev. Biol.* **2021**, *470*, 37–48. [[CrossRef](#)]
10. Assis, R.I.; Schmidt, A.G.; Racca, F.; da Silva, R.A.; Zambuzzi, W.F.; Silvério, K.G.; Nociti, F.H., Jr.; Pecorari, V.G.; Wiench, M.; Andia, D.C. DNMT1 inhibitor restores RUNX2 expression and mineralization in periodontal ligament cells. *DNA Cell Biol.* **2021**, *40*, 662–674. [[CrossRef](#)]
11. Ferreira, R.S.; Assis, R.I.; Feltran, G.D.S.; do Rosário Palma, I.C.; Françoso, B.G.; Zambuzzi, W.F.; Andia, D.C.; da Silva, R.A. Genome-wide DNA (hydroxy) methylation reveals the individual epigenetic landscape importance on osteogenic phenotype acquisition in periodontal ligament cells. *J. Periodontol.* **2022**, *93*, 435–448. [[CrossRef](#)] [[PubMed](#)]
12. Assis, R.I.; Racca, F.; Ferreira, R.S.; Ruiz, K.G.; da Silva, R.A.; Clokie, S.J.; Wiench, M.; Andia, D.C. Osteogenic commitment of human periodontal ligament cells is predetermined by methylation, chromatin accessibility and expression of key transcription factors. *Cells* **2022**, *11*, 1126. [[CrossRef](#)] [[PubMed](#)]
13. Stein, G.S.; Zaidi, S.K.; Stein, J.L.; Lian, J.B.; Van Wijnen, A.J.; Montecino, M.; Young, D.W.; Javed, A.; Pratap, J.; Choi, J.; et al. Transcription-factor-mediated epigenetic control of cell fate and lineage commitment. *Biochem. Cell Biol.* **2009**, *87*, 1–6. [[CrossRef](#)] [[PubMed](#)]
14. Morrison, O.; Thakur, J. Molecular complexes at euchromatin, heterochromatin and centromeric chromatin. *Int. J. Mol. Sci.* **2021**, *22*, 6922. [[CrossRef](#)]
15. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **2002**, *16*, 6–21. [[CrossRef](#)]
16. Lambert, S.A.; Jolma, A.; Campitelli, L.F.; Das, P.K.; Yin, Y.; Albu, M.; Chen, X.; Taipale, J.; Hughes, T.R.; Weirauch, M.T. The human transcription factors. *Cell* **2018**, *172*, 650–665. [[CrossRef](#)]
17. Pratt, A.J.; MacRae, I.J. The RNA-induced silencing complex: A versatile gene-silencing machine. *J. Biol. Chem.* **2009**, *284*, 17897–17901. [[CrossRef](#)]
18. Kunkler, C.N.; Hulewicz, J.P.; Hickman, S.C.; Wang, M.C.; McCown, P.J.; Brown, J.A. Stability of an RNA•DNA–DNA triple helix depends on base triplet composition and length of the RNA third strand. *Nucleic Acids Res.* **2019**, *47*, 7213–7222. [[CrossRef](#)]
19. Li, Y.; Syed, J.; Sugiyama, H. RNA-DNA triplex formation by long noncoding RNAs. *Cell Chem. Biol.* **2016**, *23*, 1325–1333. [[CrossRef](#)]
20. Statello, L.; Guo, C.J.; Chen, L.L.; Huarte, M. Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* **2021**, *22*, 96–118. [[CrossRef](#)]
21. Silvério, K.G.; Rodrigues, T.L.; Coletta, R.D.; Benevides, L.; Da Silva, J.S.; Casati, M.Z.; Sallum, E.A.; Nociti, F.H., Jr. Mesenchymal stem cell properties of periodontal ligament cells from deciduous and permanent teeth. *J. Periodontol.* **2010**, *81*, 1207–1215. [[CrossRef](#)] [[PubMed](#)]
22. UCSC: Genome Browser. University of California Santa Cruz. Available online: <https://genome.ucsc.edu/> (accessed on 28 May 2023).
23. Buenrostro, J.D.; Wu, B.; Chang, H.Y.; Greenleaf, W.J. ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* **2015**, *109*, 21–29. [[CrossRef](#)]
24. Corces, M.R.; Buenrostro, J.D.; Wu, B.; Greenside, P.G.; Chan, S.M.; Koenig, J.L.; Snyder, M.P.; Pritchard, J.K.; Kundaje, A.; Greenleaf, W.J.; et al. Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nat. Genet.* **2016**, *48*, 1193–1203. [[CrossRef](#)] [[PubMed](#)]
25. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)] [[PubMed](#)]
26. Heinz, S.; Benner, C.; Spann, N.; Bertolino, E.; Lin, Y.C.; Laslo, P.; Cheng, J.X.; Murre, C.; Singh, H.; Glass, C.K. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **2010**, *38*, 576–589. [[CrossRef](#)]
27. Illumina: GenomeStudio Software. Available online: <https://www.illumina.com/techniques/microarrays/array-data-analysis-experimental-design/genomestudio.html> (accessed on 28 May 2023).
28. Aryee, M.J.; Jaffe, A.E.; Corrada-Bravo, H.; Ladd-Acosta, C.; Feinberg, A.P.; Hansen, K.D.; Irizarry, R.A. Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **2014**, *30*, 1363–1369. [[CrossRef](#)]
29. Wickham, H.; François, R.; Henry, L.; Müller, K. dplyr: A Grammar of Data Manipulation. Available online: <https://github.com/tidyverse/dplyr> (accessed on 28 May 2023).
30. Wickham, H.; Averick, M.; Bryan, J.; Chang, W.; McGowan, L.D.A.; François, R.; Grolemond, G.; Hayes, A.; Henry, L.; Hester, J.; et al. Welcome to the Tidyverse. *J. Open Source Softw.* **2019**, *4*, 1686. [[CrossRef](#)]
31. Liao, Y.; Smyth, G.K.; Shi, W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* **2019**, *47*, e47. [[CrossRef](#)]
32. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140. [[CrossRef](#)]
33. Gu, Z.; Eils, R.; Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **2016**, *32*, 2847–2849. [[CrossRef](#)]
34. Blighe, K.; Rana, S.; Lewis, M. EnhancedVolcano: Publication-ready volcano plots with enhanced coloring and labeling. In *R Package Version 1.8.0*; R Foundation for Statistical Computing: Vienna, Austria, 2020.

35. Fukunaga, T.; Iwakiri, J.; Ono, Y.; Hamada, M. LncRRsearch: A web server for lncRNA-RNA interaction prediction integrated with tissue-specific expression and subcellular localization data. *Front. Genet.* **2019**, *10*, 462. [[CrossRef](#)] [[PubMed](#)]
36. Herrmann, C.J.; Schmidt, R.; Kanitz, A.; Artimo, P.; Gruber, A.J.; Zavolan, M. PolyASite 2.0: A consolidated atlas of polyadenylation sites from 3' end sequencing. *Nucleic Acids Res.* **2020**, *48*, D174–D179. [[CrossRef](#)] [[PubMed](#)]
37. Zhang, Y.; Long, Y.; Kwok, C.K. Deep learning based DNA: RNA triplex forming potential prediction. *BMC Bioinform.* **2020**, *21*, 522. [[CrossRef](#)]
38. Cunningham, F.; Allen, J.E.; Allen, J.; Alvarez-Jarreta, J.; Amode, M.R.; Armean, I.M.; Austine-Orimoloye, O.; Azov, A.G.; Barnes, I.; Bennett, R.; et al. Ensembl 2022. *Nucleic Acids Res.* **2022**, *50*, D988–D995. [[CrossRef](#)] [[PubMed](#)]
39. Heberle, H.; Meirelles, G.V.; da Silva, F.R.; Telles, G.P.; Minghim, R. InteractiVenn: A web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinform.* **2015**, *16*, 169. [[CrossRef](#)]
40. Assis, R.I.; Wiench, M.; Silverio, K.G.; da Silva, R.A.; Feltran, G.D.S.; Sallum, E.A.; Casati, M.Z.; Nociti, F.H., Jr.; Andia, D.C. RG108 increases NANOG and OCT4 in bone marrow-derived mesenchymal cells through global changes in DNA modifications and epigenetic activation. *PLoS ONE* **2018**, *13*, e0207873. [[CrossRef](#)] [[PubMed](#)]
41. Schmittgen, T.D.; Livak, K.J. Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.* **2008**, *3*, 1101–1108. [[CrossRef](#)]
42. Tai, P.W.; Wu, H.; Van Wijnen, A.J.; Stein, G.S.; Stein, J.L.; Lian, J.B. Genome-wide DNase hypersensitivity, and occupancy of RUNX2 and CTCF reveal a highly dynamic gene regulome during MC3T3 pre-osteoblast differentiation. *PLoS ONE* **2017**, *12*, e0188056. [[CrossRef](#)]
43. Lhoumaud, P.; Sethia, G.; Izzo, F.; Sakellaropoulos, T.; Snetkova, V.; Vidal, S.; Badri, S.; Cornwell, M.; Giammartino, D.C.D.G.; Kim, K.T.; et al. EpiMethylTag: Simultaneous detection of ATAC-seq or ChIP-seq signals with DNA methylation. *Genome Biol.* **2019**, *20*, 248. [[CrossRef](#)]
44. Pérez-Campo, F.M.; Santurtún, A.; García-Ibarbia, C.; Pascual, M.A.; Valero, C.; Garcés, C.; Sañudo, C.; Zarrabeitia, M.T.; Riancho, J.A. Osterix and RUNX2 are transcriptional regulators of sclerostin in human bone. *Calcif. Tissue Int.* **2016**, *99*, 302–309. [[CrossRef](#)]
45. Ohyama, Y.; Nifuji, A.; Maeda, Y.; Amagasa, T.; Noda, M. Spatiotemporal association and bone morphogenetic protein regulation of sclerostin and osterix expression during embryonic osteogenesis. *Endocrinology* **2004**, *145*, 4685–4692. [[CrossRef](#)] [[PubMed](#)]
46. Cohen-Kfir, E.; Artsi, H.; Levin, A.; Abramowitz, E.; Bajayo, A.; Gurt, I.; Zhong, L.; D'Urso, A.; Toiber, D.; Mostoslavsky, R.; et al. Sirt1 is a regulator of bone mass and a repressor of Sost encoding for sclerostin, a bone formation inhibitor. *Endocrinology* **2011**, *152*, 4514–4524. [[CrossRef](#)] [[PubMed](#)]
47. Kim, J.M.; Yang, Y.S.; Xie, J.; Lee, O.; Kim, J.; Hong, J.; Boldyreff, B.; Filhol, O.; Chun, H.; Greenblatt, M.B.; et al. Regulation of sclerostin by the SIRT1 stabilization pathway in osteocytes. *Cell Death Differ.* **2022**, *29*, 1625–1638. [[CrossRef](#)] [[PubMed](#)]
48. Starks, R.R.; Biswas, A.; Jain, A.; Tuteja, G. Combined analysis of dissimilar promoter accessibility and gene expression profiles identifies tissue-specific genes and actively repressed networks. *Epigenetics Chromatin* **2019**, *12*, 16. [[CrossRef](#)]
49. Anastasiadi, D.; Esteve-Codina, A.; Piferrer, F. Consistent inverse correlation between DNA methylation of the first intron and gene expression across tissues and species. *Epigenetics Chromatin* **2018**, *11*, 37. [[CrossRef](#)]
50. Wang, K.; Dai, R.; Xia, Y.; Tian, J.; Jiao, C.; Mikhailova, T.; Zhang, C.; Chen, C.; Liu, C. Spatiotemporal specificity of correlated DNA methylation and gene expression pairs across different human tissues and stages of brain development. *Epigenetics* **2022**, *17*, 1110–1127. [[CrossRef](#)]
51. Fukunaga, T.; Hamada, M. Riblast: An ultrafast RNA–RNA interaction prediction system based on a seed-and-extension approach. *Bioinformatics* **2017**, *33*, 2666–2674. [[CrossRef](#)]
52. Baglio, S.R.; Devescovi, V.; Granchi, D.; Baldini, N. MicroRNA expression profiling of human bone marrow mesenchymal stem cells during osteogenic differentiation reveals Osterix regulation by miR-31. *Gene* **2013**, *527*, 321–331. [[CrossRef](#)]
53. McCully, M.; Conde, J.; Baptista, P.V.; Mullin, M.; Dalby, M.J.; Berry, C.C. Nanoparticle-antagomiR based targeting of miR-31 to induce osterix and osteocalcin expression in mesenchymal stem cells. *PLoS ONE* **2018**, *13*, e0192562. [[CrossRef](#)]
54. Huang, Y.; Zheng, Y.; Xu, Y.; Li, X.; Zheng, Y.; Jia, L.; Li, W. Titanium surfaces functionalized with siMIR31HG promote osteogenic differentiation of bone marrow mesenchymal stem cells. *ACS Biomater. Sci. Eng.* **2018**, *4*, 2986–2993. [[CrossRef](#)]
55. Zhang, X.; Wang, W.; Zhu, W.; Dong, J.; Cheng, Y.; Yin, Z.; Shen, F. Mechanisms and functions of long non-coding RNAs at multiple regulatory levels. *Int. J. Mol. Sci.* **2019**, *20*, 5573. [[CrossRef](#)] [[PubMed](#)]
56. Slotkin, W.; Nishikura, K. Adenosine-to-inosine RNA editing and human disease. *Genome Med.* **2013**, *5*, 105. [[CrossRef](#)] [[PubMed](#)]
57. Gong, C.; Maquat, L.E. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* **2011**, *470*, 284–288. [[CrossRef](#)]
58. Park, E.; Maquat, L.E. Staufen-mediated mRNA decay. *Wiley Interdiscip. Rev. RNA* **2013**, *4*, 423–435. [[CrossRef](#)] [[PubMed](#)]
59. Browning, D.F.; Busby, S.J. The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.* **2004**, *2*, 57–65. [[CrossRef](#)]
60. Mondal, T.; Subhash, S.; Vaid, R.; Enroth, S.; Uday, S.; Reinius, B.; Mitra, S.; Mohammed, A.; James, A.R.; Hoberg, E.; et al. MEG3 long noncoding RNA regulates the TGF- β pathway genes through formation of RNA–DNA triplex structures. *Nat. Commun.* **2015**, *6*, 7743. [[CrossRef](#)] [[PubMed](#)]
61. Rossi, M.N.; Maione, R. Identification of chromatin binding sites for long noncoding RNAs by chromatin oligo-affinity precipitation (ChOP). In *RNA–Chromatin Interactions: Methods and Protocols*; Humana: New York, NY, USA, 2020; pp. 17–28. ISBN 978-1-0716-0680-3.

62. Everaert, C.; Luypaert, M.; Maag, J.L.; Cheng, Q.X.; Dinger, M.E.; Hellemans, J.; Mestdagh, P. Benchmarking of RNA-sequencing analysis workflows using whole-transcriptome RT-qPCR expression data. *Sci. Rep.* **2017**, *7*, 1559. [[CrossRef](#)] [[PubMed](#)]
63. Gomes, C.P.D.C.; Schroen, B.; Kuster, G.M.; Robinson, E.L.; Ford, K.; Squire, I.B.; Heymans, S.; Martelli, F.; Emanuelli, C.; Devaux, Y.; et al. Regulatory RNAs in heart failure. *Circulation* **2020**, *141*, 313–328. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.