The IncRNAs RP1-261G23.7, RP11-69E11.4 and SATB2-AS1 are a novel clinical signature for predicting recurrent osteosarcoma

Tang Ying¹,#, Jin-ling Dong², CenYuan¹, Peng Li²,* and Qingshan Guo¹,*

¹. Trauma Center, State Key Laboratory of Trauma, Burns and Combined Injury, Institute of Surgery Research, Daping Hospital, Army Medical University, No. 10 ChangjiangZhi Road, Yuzhong district, ChongQing 400042, China;

². Department of Pharmacognosy and Traditional Chinese Medicine, Department of Pharmacy and Laboratory Medicine, Army Medical University, No. 30 Gaotanyan Centre Street, Shapingba District, ChongQing 400038, China

# The first author

*Corresponding author: Qingshan Guo and Peng Li

Email: guoqingshanroom@163.com; Tel: 0236-8757997;

Abstract

Background: Osteosarcoma is the most common primary bone malignancy in children and adolescents. In order to find factors related to its recurrence, and thus improve recovery prospects, a powerful clinical signature is needed. IncRNAs are essential in osteosarcoma processes and development, and here we report on significant IncRNAs to aid in earlier diagnosis of osteosarcoma.

Methods: A univariate Cox proportional hazards regression analysis and a multivariate Cox regression analysis were used to analyze osteosarcoma patients’ IncRNA expression data from the Therapeutically Applicable Research To Generate Effective Treatments (TARGET), a public database.

Results: A IncRNA signature consisting of three IncRNAs (RP1-261G23.7, RP11-69E11.4 and SATB2-AS1) was selected. The signature was used to sort patients into high-risk and low-risk groups with meaningful recurrence rates (median recurrence time 16.80 vs. >128.22 months, log-rank test, P<0.001) in the training group, and
predictive ability was validated in a test dataset (median 16.32 vs. >143.80 months, log-rank test, P=0.006). A multivariate Cox regression analysis showed that the significant lncRNA was an independent prognostic factor for osteosarcoma patients. Functional analysis suggests that these lncRNAs were related to the PI3K-Akt signaling pathway, the Wnt signaling pathway, and the G-protein coupled receptor signaling pathway, all of which have various, important roles in osteosarcoma development. The significant 3-lncRNA set could be a novel prediction biomarker that could aid in treatment and also predict the likelihood of recurrence of osteosarcoma in patients.

Key words: osteosarcoma; recurrence rate; biomarker; lncRNA

Introduction

Osteosarcoma is a common malignant bone tumor, which is predominant in childhood and adolescence[1]. Multidrug chemotherapy and surgical removal of primary tumors have shown ameliorating effects[2]. Although improvements in therapeutic strategies have been achieved, the prognosis remains poor for most patients with recurrent osteosarcoma. It is therefore imperative to identify effective and novel recurrent biomarkers and therapeutic targets for osteosarcoma.

Long noncoding RNAs (lncRNAs) are a novel class of RNA molecules, defined as transcripts >200 nucleotides that lack protein coding potential. There have been many studies on the mechanistic role of lncRNA in osteosarcoma. Dysregulation of lncRNAs is implicated in different physiological processes in osteosarcoma, such as cell growth and metastasis [3-5]. For example, nine lncRNAs (91H, FGFR3-AS1, BCAR4, TUG1, UCA1, HIF2PUT, HOTTIP, HULC, and MALAT1) are upregulated in osteosarcoma, which is considered oncogenic for osteosarcoma[6]. The lncRNAs MALAT1 and HNF1A-AS1 promote cell proliferation and metastasis in osteosarcoma by activating the PI3K/Akt [7] and Wntβ-catenin [8] signaling pathways, respectively. Moreover, lncRNA TUG1 downregulation inhibits proliferation of osteosarcoma cells and promotes apoptosis [9].
There have also been many lncRNA studies focusing on clinical prognosis in osteosarcoma. Dysregulation of lncRNA MEG3 could act as a potential predictor of the progression and poor prognosis of osteosarcoma [10], and overexpression of lncRNAs UCA1 and TUG1 are essential in the poor prognosis of osteosarcoma [11, 12]. In addition, SNHG4 promotes tumor growth and is a predictor of poor survival and recurrence in human osteosarcoma [13]. HNF1A-AS1 is a negative prognostic biomarker that promotes tumorigenesis and progression in osteosarcoma [14]. However, prognostic signatures of recurrence are still rare, meaning that the identification of such signatures is necessary to improve clinical treatment.

In this study, we identified a lncRNA signature in osteosarcoma from the Therapeutically Applicable Research To Generate Effective Treatments (TARGET) database, and the signature was an independent prognostic factor for osteosarcoma. The presence of the signature allowed patients to be separated into high-risk and low-risk groups, and this signature may enable the development of novel recurrence biomarkers and therapeutic targets for osteosarcoma.

Materials and methods

Clinical Cohorts and RNA-Seq data

Clinical information (age, gender) of 92 patients with osteosarcoma was obtained from the TARGET data portal (https://ocg.cancer.gov/), including 47 recurrent samples, 32 non-recurrent samples, and 13 censored samples. The RNA expression data (level 3) of 92 osteosarcoma samples were downloaded from the freely available TARGET data portal. The RNA sequence data from the 92 samples was derived from the IlluminaHiSeq_RNASeq sequencing platforms, and the differentially expressed mRNAs and lncRNAs were obtained by using the edgeR software [15] to further analyze the recurrence vs. non-recurrence data. Fold changes (log₂ absolute) ≥2 and P < 0.05 indicated a statistically significant difference.

Construction of co-expression network
We examined the correlation between the expression levels of the 42 differentially expressed lncRNAs and each differentially expressed mRNA using two-sided Pearson correlation coefficients and the z-test. The mRNAs positively or negatively correlated with the 42 lncRNAs were considered to be lncRNA-related mRNAs (|Pearson correlation coefficient| > 0.4 [16-20] and P-value < 0.05). The lncRNA-mRNA co-expression network was then constructed using Cytoscape V.3.8.5(https://cytoscape.org/).

**Construction of a lncRNA signature in the training dataset**

We performed a univariate Cox proportional hazards regression analysis[21] to assess the relationship between patient recurrence and differentially expressed lncRNAs in the training group. In order to screen the most powerful diagnostic and prognostic lncRNAs, we next used multivariate Cox regression analysis and built a model to assess the prognosis risk according to the following:

\[
\text{Risk Score (RS)} = \sum_{i=1}^{N} \text{Exp}_i \times \text{Coef}_i
\]

where \(N\) is the representative number of prognostic lncRNAs, \(\text{Exp}_i\) is the expression value of the lncRNAs, and \(\text{Coef}_i\) is a single factor Cox regression coefficient. Risk Score (RS) is the multi-node weighted sum of risk scores.

**Statistical analysis**

We adopted the median risk score as a cutoff value in the training dataset[22], using it to divide the osteosarcoma patients into low- and high-risk groups. We used Kaplan-Meier survival analysis to estimate recurrence time and compare recurrence curves between the two groups, and a two-sided log-rank test to assess the statistical significance. We then validated the prognostic ability of the lncRNA signature in the test dataset using the ROC values and Kaplan-Meier survival analysis. We carried out data stratification analysis and multivariate Cox regression analysis in order to identify whether the lncRNA were an independent factor, and to assess the signature within the
available data. We considered a value of P<0.05 as significant. All analyses were performed with the R statistical program (version 3.5.1), using the survival and pROC packages downloaded from Bioconductor (http://www.bioconductor.org/).

Construction of the IncRNA-function network

The Kyoto encyclopedia of genes and genomes pathway (KEGG) and gene ontology functional enrichment analyses (GO) were carried out to better illustrate the potential functional mechanism of recurrence-related IncRNA modules of osteosarcoma. The DAVID Bioinformatics Tool (https://david.ncifcrf.gov/), which consists of an integrated biological information base and analytical tools aimed at systematically extracting biological meaning from gene/protein lists was used. We first screened the 3 IncRNAs correlated with differentially expressed mRNAs (|Pearson correlation coefficient| > 0.4 and P-value < 0.05), and then screened the mRNAs involved in the functional analysis. These database inputs and screening parameters allowed for construction of the IncRNA-function network using Cytoscape, version 3.8.5 (https://cytoscape.org/).

Result

Patient characteristics

All expression data used in this study were from patients clinically and pathologically diagnosed with osteosarcoma. The median age was 15 years (4-40 years) for the enrolled patients. Other clinical information about the patients is in Table 1. Subsequently, we randomly divided the 92 patients into two groups (training group, n=61; test group, n=31) to seek and validate the prognostic IncRNA found in the osteosarcoma patients. The selection scheme of the recurrent IncRNA signature can be found in Figure 1.

Identification of significantly differentially expressed mRNAs and IncRNAs

A total of 19,495 mRNAs and 14,589 IncRNAs were identified from the TARGET database. Using |Fold changes| ≥2 and P < 0.05 as cutoffs, we identified 184
differentially expressed mRNAs (71 downregulated and 113 upregulated) and 42
differentially expressed lncRNAs (22 downregulated and 20 upregulated) (Table S1).

Construction of the IncRNA-mRNA co-expression network in osteosarcoma

In order to better study the role of differentially expressed lncRNAs in osteosarcoma
and to further elucidate the interaction between these differentially expressed mRNAs,
we constructed an osteosarcoma lncRNA-mRNA related regulatory network. We used
the 36 differentially expressed lncRNAs and 101 differentially expressed mRNAs to
construct this network (Figure 2, Table S2), and screening criteria were the |Pearson
correlation coefficient| > 0.4 and P-value < 0.05. Using these criteria, we found that six
upregulated lncRNAs (RP1-261G23.7, RP11-81A22.5, RP11-69E11.4,
RP11-817J15.3, SATB2-AS1 and CTB-4E7.1) functioned in regulating many mRNAs,
suggesting an important regulatory role for these lncRNAs in osteosarcoma.

Construction of the recurrent IncRNA signature with the training dataset

The training group (n = 61), which included complete clinical information, was used to
explore the association of recurrence with 42 differentially expressed lncRNAs. We
initially performed a univariate Cox proportional hazards regression analysis of the
expression profiling data of differentially expressed lncRNAs, with recurrence time
and recurrence status as the dependent variables. We identified four lncRNAs that were
significantly correlated with recurrence of patients (P value <0.05, Table S3). In order
to screen the most powerful diagnostic and prognostic lncRNAs, we next used a
multivariate Cox regression analysis (Table S4) and built a 3-lncRNA (RP1-261G23.7,
RP11-69E11.4, and SATB2-AS1) model to assess the recurrence risk. This set of
3-lncRNAs also functions in regulating the lncRNA-mRNA co-expression network.
The risk score (Table S5) of the combination of RP1-261G23.7, RP11-69E11.4, and
SATB2-AS1 was determined as follows:

\[
RS = (0.01 \times ev_{RP1-261G23.7}) + (0.01 \times ev_{RP11-69E11.4}) + (0.02 \times ev_{SATB2-AS1})
\]
where $RS$ is the risk score and $ev$ is the expression value.

**Determining the survival power of the lncRNA signature in the training and test datasets**

The selected lncRNA signature returned a risk score for each patient. We used the median risk score to divide the training group patients into either a low-risk group ($n = 31$) or a high-risk group ($n = 30$). Using the Kaplan-Meier survival analysis, we found that the low-risk group patients had a significantly lower recurrence rate than those in the high-risk group (median recurrence time: $>128.22$ months vs. $16.80$ months, log-rank test $P<0.001$; Figure 3A). The 5-year recurrence rate of the patients in the high-risk group was less than $20\%$, while that of the patients in the low-risk group was greater than $60\%$.

The recurrent risk score model used to calculate the lncRNA signature-based risk scores of the 31 patients in the test dataset was also used to validate the prediction power of the lncRNA signature. Patients were divided into low-risk and high-risk groups using the median risk score as a demarcation point, in the same fashion as the training dataset. Kaplan-Meier curves were used to show the low-risk and high-risk groups in the test dataset (Figure 3B), and the results were similar to those from the training dataset. In the test dataset, the low-risk group had a significantly higher recurrence rate than the high-risk group (median recurrence time: $>143.80$ months vs. $16.32$, log-rank test $P=0.006$). The recurrence rate of patients in the high-risk group was less than $20\%$ at 5 years, in contrast to more than $60\%$ in the low-risk group.

**The survival prediction power of the lncRNA signature in the training and test groups**

We performed a separate ROC analysis to test the prediction power of the lncRNA signature, which considered the larger area under the ROC curve as a better model for predicting recurrence of the osteosarcoma. In the training dataset, the predictive ability of the 3 lncRNA signature was high ($AUC_{\text{Signature}}=0.68$, Figure 3C), which further demonstrated that this signature was a novel and highly accurate biomarker of
recurrence. A similar result was seen in the test group (AUC Signature = 0.75, Figure 3D).

**The selected 3-lncRNA signature is an independent prognostic factor**

A multivariate Cox regression analysis using the signature-based risk score and other clinical characteristics (e.g., age and gender) demonstrated that the prognostic power of the lncRNA signature risk score for the prediction of recurrence was independent of clinical characteristics for the entire dataset (High-risk group vs. Low-risk group, HR = 2.06, 95% CI: 1.53–2.77, P<0.001, n=92, Table 2).

**Functional prediction of differentially expressed mRNAs related to 3 lncRNAs**

KEGG and GO analyses were used to investigate the potential involvement of the 3 lncRNAs in biological processes associated with osteosarcoma development. We performed functional analysis with co-expressed mRNAs of 3 lncRNAs in each lncRNA module (|Pearson correlation coefficient| > 0.4 and P-value < 0.05, Table S6). The 3 lncRNAs-functional network (Figure 4, Table S6) was created using Cytoscape. The results indicated that the 3 lncRNA signature was related to a collagen trimer, endoplasmic reticulum lumen, and protein digestion and absorption. We also found that only RP1-261G23.7 was involved in the wnt signal pathway, which is in a variety of biological processes, including cell cycle regulation and cancer. RP11-69E11.4 was involved in the G-protein coupled receptor signaling pathway, which is linked to the activation of adenylyl cyclase activity and subsequent cAMP/PKA/CREB signaling[23] as potential drivers of tumorigenesis. Finally, SATB2-AS1 is involved in the PI3K-Akt signaling pathway, which is essential in cell proliferation and the biological characteristics of malignant cells [24].

**Discussion**

Individuals with recurrent osteosarcoma have less than a 20% chance of long-term survival, despite aggressive therapies [25]. Identification of a clinically relevant signature for prognosis is of critical need for patients. Recent studies suggested that lncRNAs could be good candidates for tumor signatures as they possess high sensitivity and high specificity [26-30]. Increasing data has demonstrated that lncRNAs are
associated with osteosarcoma prognosis. For example, up-regulated FOXC2-AS1 was correlated with longer survival time [31], while HULC was associated with distant metastasis and clinical stage, and with shorter overall survival time [32, 33]. Moreover, BANCR was associated with large tumor size, distant metastasis, and advanced clinical stage, and was an independent predictor of poor survival rate[34-36]. However, lncRNA signatures and their relation to recurrence of osteosarcoma has rarely been investigated.

In our study, different statistical methods were used to identify a 3-lncRNA signature associated with the recurrence of osteosarcoma in patients. We found that the lncRNA signature was an independent predictor in patient recurrence. A multivariate Cox regression analysis was used to assess the independence of the selected signature in predicting recurrence of osteosarcoma. With age and gender as covariables, the risk scores of patients based on the signature maintained an independent association with recurrence. Taken together, these results suggested that the prognostic power of the lncRNA signature for predicting recurrence of osteosarcoma in patients was independent of other clinical characteristics.

Moreover, the expression levels of RP1-261G23.7, RP11-69E11.4, and SATB2-AS1 were increased in recurrent osteosarcoma samples vs. non-recurrent samples. These lncRNAs also had important regulatory roles in osteosarcoma based on the lncRNA-mRNA co-expression network findings. Finally, we analyzed the functions of these lncRNAs, and the results showed that RP1-261G23.7 was involved in growth factor activity, single organismal cell-cell adhesion, the G-protein coupled receptor signaling pathway, and cell adhesion.

Vascular endothelial growth factor A (VEGF-A) is known to function directly in blood vessel formation and maintenance, and is a promising therapeutic target for activation in cancer. RP1-261G23.7 is antisense VEGF-A lncRNA, and a recent report showed that RP1-261G23.7 functions as a VEGF-A promoter enhancer-like element, possibly by acting as a local scaffold for proteins[37]. RP11-69E11.4 was involved in the Wnt
signaling pathway, which is sensitive to alterations at various nodes of the pathway, and these alterations were identified in multiple tumor types after hyperactivation. These alterations converge into increased tumorigenicity, sustained proliferation and enhanced metastatic potential[24]. In addition, SATB2-AS1 is specifically involved in the PI3K-Akt signaling pathway, which can affect the epithelial-mesenchymal transition in numerous ways to influence tumor aggressiveness[38]. SATB2-AS1 was reported as overexpressed in osteosarcoma, and was associated with increased cell proliferation and growth[39]. These results indicated that recurrent IncRNA signatures executed a variety of biological functions in osteosarcoma tumorigenesis and development. There has been relatively little research on the three signature IncRNAs, and our results provide a foundation for further investigation into the function of the three IncRNAs.

There are a few limitations of this research regarding osteosarcoma. Most importantly, the specific prediction mechanism of the 3 IncRNAs in osteosarcoma requires further study. Moreover, the IncRNA signature has not yet been tested prospectively in multiple clinical trials. Despite these drawbacks, the significant and consistent correlation of our IncRNA signature with recurrence in two independent datasets indicated that it was a potentially powerful prognostic signature for osteosarcoma.

In conclusion, our data clearly indicate that the IncRNA signature can predict the recurrence of osteosarcoma in patients. The signature provides accurate prediction, which is clinically very significant.

Author contributions

Qingshan Guo contributed to the conception and design of the study. Tang Ying contributed to the collection of the clinical and TARGET data of the patients. Tang Ying, Jinling Dong and Cen Yuan contributed to the interpretation of the data (statistical and computational analysis) and wrote the manuscript. Peng Li, Qingshan Guo, and Tang Ying contributed to the review and revision of the manuscript. All authors read and approved the final manuscript.
Funding

The present study was funded by the National Natural Science Foundation of China (81472596).

Competing interests

The authors declare no conflicts of interest.

REFERENCES

1. Tian W, Li Y, Zhang J, Li J, Gao J: Combined analysis of DNA methylation and gene expression profiles of osteosarcoma identified several prognosis signatures. Gene 2018, 650:7-14.

2. Yu D, Kahen E, Cubitt CL, McGuire J, Kreahling J, Lee J, Altiok S, Lynch CC, Sullivan DM, Reed DR: Identification of Synergistic, Clinically Achievable, Combination Therapies for Osteosarcoma. Sci Rep 2015, 5:16991.

3. Ba Z, Gu L, Hao S, Wang X, Cheng Z, Nie G: Downregulation of lncRNA CASC2 facilitates osteosarcoma growth and invasion through miR-181a. Cell Prolif 2018, 51(1).

4. Ye H, Lin J, Yao X, Li Y, Lin X, Lu H: Overexpression of Long Non-Coding RNA NNT-AS1 Correlates with Tumor Progression and Poor Prognosis in Osteosarcoma. Cell Physiol Biochem 2018, 45(5):1904-1914.

5. Yan L, Wu X, Yin X, Du F, Liu Y, Ding X: LncRNA CCAT2 promoted osteosarcoma cell proliferation and invasion. J Cell Mol Med 2018, 22(5):2592-2599.

6. Li Z, Dou P, Liu T, He S: Application of Long Noncoding RNAs in Osteosarcoma: Biomarkers and Therapeutic Targets. Cell Physiol Biochem 2017, 42(4):1407-1419.

7. Dong Y, Liang G, Yuan B, Yang C, Gao R, Zhou X: MALAT1 promotes the proliferation and metastasis of osteosarcoma cells by activating the PI3K/Akt pathway. Tumour Biol 2015, 36(3):1477-1486.

8. Zhou S, Yu L, Xiong M, Dai G: LncRNA SNHG12 promotes tumorigenesis and metastasis in osteosarcoma by upregulating Notch2 by sponging miR-195-5p. Biochem Biophys Res Commun 2018, 495(2):1822-1832.

9. Zhang Q, Geng PL, Yin P, Wang XL, Jia JP, Yao J: Down-regulation of long non-coding RNA TUG1 inhibits osteosarcoma cell proliferation and promotes apoptosis. Asian Pac J Cancer Prev 2013, 14(4):2311-2315.

10. Tian ZZ, Guo XJ, Zhao YM, Fang Y: Decreased expression of long non-coding RNA MEG3 acts
as a potential predictor biomarker in progression and poor prognosis of osteosarcoma. *Int J Clin Exp Pathol* 2015, 8(11):15138-15142.

11. Zhou M, Wang X, Shi H, Cheng L, Wang Z, Zhao H, Yang L, Sun J: Characterization of long non-coding RNA-associated ceRNA network to reveal potential prognostic lncRNA biomarkers in human ovarian cancer. *Oncotarget* 2016, 7(11):12598-12611.

12. Ma B, Li M, Zhang L, Huang M, Lei JB, Fu GH, Liu CX, Lai QW, Chen QQ, Wang YL: Upregulation of long non-coding RNA TUG1 correlates with poor prognosis and disease status in osteosarcoma. *Tumour Biol* 2016, 37(4):4445-4455.

13. Xu R, Feng F, Yu X, Liu Z, Lao L: LncRNA SNHG4 promotes tumour growth by sponging miR-224-3p and predicts poor survival and recurrence in human osteosarcoma. *Cell Prolif* 2018, 51(6):e12515.

14. Cai L, Lv J, Zhang Y, Li J, Wang Y, Yang H: The lncRNA HNF1A-AS1 is a negative prognostic factor and promotes tumorigenesis in osteosarcoma. *J Cell Mol Med* 2017, 21(11):2654-2662.

15. Robinson MD, McCarthy DJ, Smyth GK: *edgeR*: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010, 26(1):139-140.

16. Xiong HG, Li H, Xiao Y, Yang QC, Yang LL, Chen L, Bu LL, Zhang WF, Zhang JL, Sun ZJ: Long noncoding RNA MYOSLID promotes invasion and metastasis by modulating the partial epithelial-mesenchymal transition program in head and neck squamous cell carcinoma. *J Exp Clin Cancer Res* 2019, 38(1):278.

17. Larsen TV, Hussmann D, Nielsen AL: PD-L1 and PD-L2 expression correlated genes in non-small-cell lung cancer. *Cancer Commun (Lond)* 2019, 39(1):30.

18. Wang B, Ran Z, Liu M, Ou Y: Prognostic Significance of Potential Immune Checkpoint Member HHLA2 in Human Tumors: A Comprehensive Analysis. *Front Immunol* 2019, 10:1573.

19. Kawaguchi T, Azuma K, Sano M, Kim S, Kawahara Y, Sano Y, Shimodaira T, Ishibashi K, Miyaji T, Basch E et al: The Japanese version of the National Cancer Institute’s patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE): psychometric validation and discordance between clinician and patient assessments of adverse events. *J Patient Rep Outcomes* 2017, 2(1):2.

20. Bie LY, Li D, Mu Y, Wang S, Chen BB, Lyu HF, Han LL, Nie CY, Yang CC, Wang L et al: Analysis of cyclin E co-expression genes reveals nuclear transcription factor Y subunit alpha is an oncogene in gastric cancer. *Chronic Dis Transl Med* 2019, 5(1):44-52.
21. Jeong HH, Kim S, Wee K, Sohn KA: Investigating the utility of clinical outcome-guided mutual information network in network-based Cox regression. *BMC Syst Biol* 2015, 9 Suppl 1:S8.

22. Zhou M, Guo M, He D, Wang X, Cui Y, Yang H, Hao D, Sun J: A potential signature of eight long non-coding RNAs predicts survival in patients with non-small cell lung cancer. *J Transl Med* 2015, 13:231.

23. Alakus H, Babicky ML, Ghosh P, Yost S, Jepsen K, Dai Y, Arias A, Samuels ML, Mose ES, Schwab RB et al: Genome-wide mutational landscape of mucinous carcinomatosis peritonei of appendiceal origin. *Genome Med* 2014, 6(5):43.

24. Tai D, Wells K, Arcaroli J, Vanderbilt C, Aisner DL, Messersmith WA, Lieu CH: Targeting the WNT Signaling Pathway in Cancer Therapeutics. *Oncologist* 2015, 20(10):1189-1198.

25. Ferrari S, Smeland S, Mercuri M, Bertoni F, Longhi A, Ruggieri P, Alvegard TA, Picci P, Capanna R, Bernini G et al: Neoadjuvant chemotherapy with high-dose Ifosfamide, high-dose methotrexate, cisplatin, and doxorubicin for patients with localized osteosarcoma of the extremity: a joint study by the Italian and Scandinavian Sarcoma Groups. *J Clin Oncol* 2005, 23(34):8845-8852.

26. Shi T, Gao G, Cao Y: Long Noncoding RNAs as Novel Biomarkers Have a Promising Future in Cancer Diagnostics. *Dis Markers* 2016, 2016:9085195.

27. Yarmishyn AA, Kurochkin IV: Long noncoding RNAs: a potential novel class of cancer biomarkers. *Front Genet* 2015, 6:145.

28. Hu HB, Jie HY, Zheng XX: Three Circulating LncRNA Predict Early Progress of Esophageal Squamous Cell Carcinoma. *Cell Physiol Biochem* 2016, 40(1-2):117-125.

29. Qiu ZL, Shen CT, Sun ZK, Wei JW, Zhang XY, Song HJ, Luo QY: Circulating Long Non-Coding RNAs Act as Biomarkers for Predicting 131I Uptake and Mortality in Papillary Thyroid Cancer Patients with Lung Metastases. *Cell Physiol Biochem* 2016, 40(6):1377-1390.

30. Tang Q, Ni Z, Cheng Z, Xu J, Yu H, Yin P: Three circulating long non-coding RNAs act as biomarkers for predicting NSCLC. *Cell Physiol Biochem* 2015, 37(3):1002-1009.

31. Zhu KP, Zhang CL, Shen GQ, Zhu ZS: Long noncoding RNA expression profiles of the doxorubicin-resistant human osteosarcoma cell line MG63/DXR and its parental cell line MG63 as ascertained by microarray analysis. *Int J Clin Exp Pathol* 2015, 8(8):8754-8773.

32. Sun XH, Yang LB, Geng XL, Wang R, Zhang ZC: Increased expression of lncRNA HULC indicates a poor prognosis and promotes cell metastasis in osteosarcoma. *Int J Clin Exp Pathol* 2015, 8(3):2994-3000.
33. Uzan VR, Lengert A, Boldrini E, Penna V, Scapulatempo-Neto C, Scrideli CA, Filho AP, Cavalcante CE, de Oliveira CZ, Lopes LF et al: High Expression of HULC Is Associated with Poor Prognosis in Osteosarcoma Patients. PLoS One 2016, 11(6):e0156774.

34. Peng ZQ, Lu RB, Xiao DM, Xiao ZM: Increased expression of the IncRNA BANCR and its prognostic significance in human osteosarcoma. Genet Mol Res 2016, 15(1).

35. Flockhart RJ, Webster DE, Qu K, Mascarenhas N, Kovalski J, Kretz M, Khavari PA: BRAFV600E remodels the melanocyte transcriptome and induces BANCR to regulate melanoma cell migration. Genome Res 2012, 22(6):1006-1014.

36. Fang D, Yang H, Lin J, Teng Y, Jiang Y, Chen J, Li Y: 17beta-estradiol regulates cell proliferation, colony formation, migration, invasion and promotes apoptosis by upregulating miR-9 and thus degrades MALAT-1 in osteosarcoma cell MG-63 in an estrogen receptor-independent manner. Biochem Biophys Res Commun 2015, 457(4):500-506.

37. Nieminen T, Scott TA, Lin FM, Chen Z, Yla-Herttuala S, Morris KV: Long Non-Coding RNA Modulation of VEGF-A during Hypoxia. Noncoding RNA 2018, 4(4).

38. Lunardi A, Webster KA, Papa A, Padmani B, Clohessy JG, Bronson RT, Pandolfi PP: Role of aberrant PI3K pathway activation in gallbladder tumorigenesis. Oncotarget 2014, 5(4):894-900.

39. Liu SH, Zhu JW, Xu HH, Zhang GQ, Wang Y, Liu YM, Liang JB, Wang YX, Wu Y, Guo QF: A novel antisense long non-coding RNA SATB2-AS1 overexpresses in osteosarcoma and increases cell proliferation and growth. Mol Cell Biochem 2017, 430(1-2):47-56.

Figure and Table legends

Figure 1  Flowchart of the study

The order of analyses to develop the risk score model and validate the IncRNA signature to predict recurrence outcomes.

Figure 2  The IncRNA-mRNA co-expression network

The IncRNA-mRNA co-expression network. Red balls represent upregulated mRNAs and red cones represent upregulated IncRNAs. Blue balls represent downregulated
mRNAs and blue cones represent downregulated lncRNAs.

**Figure 3**  **IncRNA signature predicts recurrence of patients with osteosarcoma.** (A-B) Kaplan-Meier survival curves classified osteosarcoma patients into high-risk and low-risk groups using the lncRNAs signature in the training and test datasets. P-values were calculated by log-rank test. (C-D) Results of receiver operating characteristic (ROC) analysis.

**Figure 4**  **The IncRNA-function network**

The IncRNA-function network. Red cones represent lncRNAs and green diamonds represent functions.

**Table 1**  **Summary of patient demographics and clinical characteristics.**

**Table 2**  **Univariate and multivariate Cox regression analysis of the 3 IncRNA signature and recurrence of osteosarcoma patients in the entire group.**

**Table S1**  **Differentially expressed mRNAs and IncRNAs**

**Table S2**  **lncRNAs-mRNAs co-expression analysis (|Pearson correlation coefficient| > 0.4 and P-value < 0.05).**

**Table S3**  **Univariate Cox proportional hazards regression analysis (P <0.05) of the differentially expressed IncRNAs profiling data in the training dataset.**

**Table S4**  **Multivariate Cox regression analysis (P <0.05) of the 4 lncRNAs profiling data in the training dataset.**

**Table S5**  **The signature riskScore composed of 3 IncRNA combinations in the training and test dataset.**

**Table S6**  **The 3 IncRNAs-mRNAs co-expression and the 3 IncRNAs-function network**
TARGET database

92 samples from the mRNA expression database

47 samples from recurrent and 32 samples from non-recurrent

Different screen (P<0.05, Fold change >2)

184 differentially expressed mRNAs profile

Pearson correlation coefficient

IncRNA-mRNA co-expression network

3 IncRNAs - function network

92 samples from the IncRNA expression database

42 samples from recurrent and 32 samples from non-recurrent

Different screen (P<0.05, Fold change >2)

42 differentially expressed IncRNAs profile

95 samples divided into two groups

samples in training group and in test group

training group (n=62)

Univariable Cox (P<0.05)

4 differentially expressed IncRNAs

Multivariable Cox (P<0.05)

3 remained to construct predictive model

Validation in the test group (n=31)

KM analysis

ROC analysis
A  Training group, n=61

- Low-risk, n=31
- High-risk, n=30

Log Rank P= 5e-04

B  Test group, n=31

- Low-risk, n=20
- High-risk, n=11

Log Rank P= 0.006

D  ROC Curves

- Sensitivity
- 1 - Specificity

- Signature=0.68
- Signature=0.75
| Characteristic | Training set | Testing set | Total |
|---------------|--------------|-------------|-------|
| Gender        |              |             |       |
| Male          | 37           | 15          | 160   |
| Female        | 24           | 16          | 27    |
| Age           |              |             |       |
| Median        | 15           | 15          | 15    |
| Range         | 4~40         | 6~21        | 4~40  |
| Vital status  |              |             |       |
| Non-recurrence| 28           | 16          | 44    |
| Recurrence    | 33           | 15          | 48    |
Table 2: Univariable and multivariable Cox regression analysis of the association between the 3-lncRNA signature and the recurrence of osteosarcoma patients (n=91)

| Variables            | HR    | 95% CI of HR | P      |
|----------------------|-------|--------------|--------|
|                      |       | lower        | upper  |
| Univariable analysis |       |              |        |
| Gender               | male vs. female | 1.24 | 0.69 | 2.21 | 0.47 |
| Age                  | >15 vs. ≤15 | 0.96 | 0.90 | 1.02 | 0.17 |
| LncRNA signature     | High risk vs. low risk | 2.10 | 1.57 | 2.81 | 0.00 |
| Multivariable analysis |     |              |        |
| Gender               | male vs. female | 1.17 | 0.63 | 2.16 | 0.62 |
| Age                  | >15 vs. ≤15 | 0.96 | 0.89 | 1.02 | 0.18 |
| LncRNA signature     | High risk vs. low risk | 2.06 | 1.53 | 2.77 | 0.00 |
| mRNA    | logFC  | logCPM  | PValue  | IncRNA   | logFC  |
|---------|--------|---------|---------|----------|--------|
| ASPD     | -4.273 | 3.337   | 3.75E-13| RP1-169D4| -5.168 |
| SCGB3A1  | 6.053  | 6.440   | 8.00E-12| FLJ26245 | -5.558 |
| SFTPB    | 6.715  | 5.107   | 9.15E-11| AP003025.2| -5.038 |
| PRB2     | 8.150  | 6.067   | 9.03E-10| RP1-383J24| -3.121 |
| FMO1     | -3.875 | 3.473   | 1.06E-09| EAF1-AS1 | -2.781 |
| MEPE     | 3.899  | 5.653   | 9.15E-11| AP003025.2| -5.038 |
| PAGE2    | -3.093 | 4.805   | 3.42E-09| RP1-345P4.| -2.441 |
| ZNF362   | -1.936 | 6.089   | 8.86E-09| AC132217.4| -3.002 |
| HCRT     | 5.979  | 4.050   | 2.58E-08| RP1-162J8.| -1.304 |
| SCGB1A1  | 8.646  | 5.961   | 4.42E-08| RP11-345P4.| -2.441 |
| ARID3B   | -2.150 | 3.522   | 5.99E-08| CTC-425F1.| 1.889 |
| NAPSA    | 4.003  | 3.267   | 1.10E-07| RP11-169D4| -5.168 |
| PRB1     | 7.114  | 4.250   | 1.83E-07| RP11-299J3.| -1.304 |
| SFTPC    | 9.068  | 6.702   | 1.96E-07| IQCH-AS1 | 2.596 |
| SFTPA2   | 8.052  | 4.758   | 4.28E-07| RP11-345P4.| -2.441 |
| SFTPA1   | 7.576  | 4.302   | 5.99E-08| CTC-425F1.| 1.889 |
| SERTAD3  | -1.189 | 4.876   | 4.02E-07| RP1-261G23.| 2.493 |
| SCGB3A2  | 2.900  | 4.692   | 5.60E-07| RP11-81A22.| 1.820 |
| GLTSCR1  | 3.899  | 5.653   | 5.99E-08| AC132217.4| -3.002 |
| USP28    | -1.509 | 3.029   | 7.00E-07| CTB-4E7.1| 1.889 |
| RGS4     | -2.078 | 3.181   | 1.00E-07| RP11-358L22| 1.238 |
| TCF7L2   | 1.869  | 4.661   | 1.00E-07| SATB2-AS1| 1.174 |
| IGF2     | 1.869  | 4.661   | 1.24E-07| RP1-162J8.3| -1.511 |
| PSMC4    | -1.005 | 6.758   | 1.76E-06| CTB-4E7.1| 1.889 |
| CHRLD2   | 3.515  | 4.544   | 1.99E-06| RP1-3P17.5| 2.595 |
| MUC1     | 1.640  | 4.661   | 2.03E-06| RP1-358L22| 1.238 |
| SAP25    | -2.259 | 3.071   | 2.16E-06| RP1-426C22| -1.232 |
| AHCYL2   | -1.486 | 4.482   | 2.40E-06| CECR7    | 1.110 |
| NETO2    | -1.624 | 5.153   | 2.43E-06| SATB2-AS1| 1.174 |
| CWF19L2  | -1.403 | 4.900   | 2.51E-06| SATB2-AS1| 1.174 |
| AGER     | 1.808  | 3.664   | 2.89E-06| SATB2-AS1| 1.174 |
| PLEKHA2  | -1.452 | 5.437   | 2.91E-06| SATB2-AS1| 1.174 |
| MYH3     | -2.105 | 3.162   | 3.39E-06| SATB2-AS1| 1.174 |
| NUCKS1   | -1.071 | 7.578   | 4.71E-06| SATB2-AS1| 1.174 |
| DHX57    | -1.148 | 4.540   | 4.81E-06| RP1-426C22| -1.232 |
| CNBD2    | 1.538  | 3.329   | 4.85E-06| RP1-162J8.3| -1.511 |
| TMCC2    | 1.378  | 3.493   | 5.49E-06| RP1-162J8.3| -1.511 |
| PDLIM3   | -2.312 | 5.301   | 5.64E-06| RP1-162J8.3| -1.511 |
| STAC3    | -2.015 | 4.535   | 5.70E-06| RP1-162J8.3| -1.511 |
| WRAP53   | -1.860 | 5.179   | 6.11E-06| RP1-162J8.3| -1.511 |
| C10orf10 | 1.822  | 5.038   | 6.56E-06| RP1-162J8.3| -1.511 |
| BMP8A    | 1.515  | 3.357   | 8.21E-06| RP1-162J8.3| -1.511 |
| APOD     | 2.222  | 4.047   | 9.75E-06| RP1-162J8.3| -1.511 |
| DEFA1B   | 6.446  | 3.292   | 1.00E-05| RP1-162J8.3| -1.511 |
| TAC4     | 1.935  | 3.945   | 1.24E-05| RP1-162J8.3| -1.511 |
| AURKB    | -1.229 | 6.238   | 1.48E-05| RP1-162J8.3| -1.511 |
| RASD1    | 1.610  | 5.178   | 1.53E-05| RP1-162J8.3| -1.511 |
| TBC1D19  | -1.108 | 3.609   | 1.67E-05| RP1-162J8.3| -1.511 |
| ITPKC    | -1.093 | 3.672   | 1.87E-05| RP1-162J8.3| -1.511 |
| Gene   | Value1  | Value2  | Value3     |
|--------|---------|---------|------------|
| SNX14  | -1.039460238 | 6.058735277 | 1.89E-05   |
| CHRNA1 | -2.421522693  | 4.268443643 | 1.90E-05   |
| WiFi1  | 2.647174161   | 5.05289459   | 2.09E-05   |
| LGR6   | 1.739164028   | 3.257566152  | 2.60E-05   |
| DENND3 | -1.312302412  | 3.705449673  | 2.67E-05   |
| EPCAM  | 1.765967746   | 3.841916784  | 2.78E-05   |
| ZNF710 | -1.132463     | 4.091575447  | 3.04E-05   |
| GBP1   | -1.22557533   | 4.896390902  | 3.09E-05   |
| BAHCC1 | -1.445015263  | 4.214001669  | 3.22E-05   |
| SYT12  | 1.454443278   | 3.807653337  | 3.74E-05   |
| RASGRP2| 1.530044715   | 4.248493461  | 4.21E-05   |
| HGF    | -1.549831738  | 3.608516734  | 4.49E-05   |
| KHL41  | -2.648228645  | 4.142729977  | 5.55E-05   |
| ZNF461 | -1.145597683  | 3.567924535  | 6.24E-05   |
| VEGFA  | 1.362891711   | 7.790020997  | 7.97E-05   |
| GRAMD1B| 1.084746284   | 4.077250007  | 8.49E-05   |
| AOC3   | 1.259533406   | 4.716003574  | 8.50E-05   |
| FST    | -1.186380123  | 4.059175087  | 9.58E-05   |
| SSX2B  | -2.221271497  | 3.880073978  | 9.89E-05   |
| STPG1  | -1.247583693  | 3.596473834  | 0.00010633 |
| RANGRF | -1.128271333  | 5.764356439  | 0.00011908 |
| VCAM1  | -1.334008912  | 4.790016347  | 0.00013417 |
| GPRC5A | 1.55489394    | 3.255798716  | 0.00013896 |
| NOX4   | -1.430023856  | 3.171048662  | 0.00015257 |
| MYL1   | -2.923136868  | 5.255334233  | 0.00015546 |
| BMP8B  | 1.263049337   | 4.998038797  | 0.0001591  |
| CORT   | 1.41004454    | 3.892331172  | 0.00017539 |
| SUGCT  | -1.077803554  | 4.3605691    | 0.00018793 |
| GNG4   | 1.564791632   | 4.337042003  | 0.0001903  |
| GOLGA8A| 1.108619825   | 3.631739412  | 0.00020135 |
| CHMP4C | 1.552517239   | 3.832391403  | 0.00022177 |
| ITGA10 | 1.187089974   | 5.645040177  | 0.00024619 |
| TPDS2  | 1.223482995   | 4.5360376    | 0.00025295 |
| SLC8A3 | 1.525656835   | 4.866163038  | 0.0002571  |
| C1QTNF3| -1.892631207  | 6.207310325  | 0.00026478 |
| EFHD1  | 1.235076965   | 5.147629317  | 0.00029826 |
| RPL26  | -1.0125847    | 11.99168194  | 0.00029853 |
| NDN    | 1.44427702    | 4.208415795  | 0.00036067 |
| TNNC2  | -2.397409413  | 5.040637607  | 0.00038851 |
| ASPHD1 | -1.178765028  | 3.13139949   | 0.0003901  |
| MEST   | -1.516330388  | 5.541945702  | 0.00039173 |
| AMELX  | 5.053963801   | 4.963222201  | 0.00041301 |
| CRIP1  | 1.142745417   | 8.736147293  | 0.00041878 |
| DLL1   | 1.31638978    | 3.159522163  | 0.00045101 |
| ISM1   | 1.454570644   | 3.175211782  | 0.00045703 |
| PPP1R14A| 1.559239755   | 3.727535817  | 0.00047987 |
| TDRP   | 1.097972819   | 3.385736177  | 0.00049244 |
| RASL11B| -1.217763801  | 3.311301469  | 0.00050157 |
| JCHAIN | 2.642871345   | 5.530548797  | 0.00050183 |
| LYNX1  | 1.139080832   | 3.205399373  | 0.00051824 |
| SYT15  | 1.087090861   | 3.758298027  | 0.00057117 |
| Gene     | Log2 Fold Change | p-value  |
|----------|-----------------|----------|
| CPE      | 1.042028007     | 9.58813628 | 0.00057487 |
| TSPAN11  | 1.097661318     | 5.052414732 | 0.00058602 |
| COL13A1  | 1.082410054     | 5.378097787 | 0.00059587 |
| CYFIP2   | 1.326096142     | 5.19711638 | 0.00065792 |
| TLL7     | 1.673647002     | 4.15675553 | 0.00067156 |
| CRYBA2   | 2.325760595     | 3.160781935 | 0.00070322 |
| KRT18    | 1.117290092     | 3.621054513 | 0.00072489 |
| GBP5     | -1.27695774     | 3.074606371 | 0.00078567 |
| GAL      | 1.815241213     | 4.488189646 | 0.00078753 |
| SAA1     | -2.133726748    | 3.132263656 | 0.00083226 |
| COL8A2   | -1.446250232    | 5.380727241 | 0.00084346 |
| CPNE5    | 1.058740431     | 4.152398159 | 0.0008724 |
| RHBDL2   | 1.045312759     | 6.132268873 | 0.00097703 |
| NPW      | 1.661795479     | 5.308660554 | 0.00107409 |
| COL22A1  | 1.637922537     | 5.933291028 | 0.00111255 |
| MLC1     | 1.431505847     | 4.318626477 | 0.00123669 |
| TNSF5F11B| 1.11796063      | 3.858025708 | 0.00130304 |
| CHST13   | 1.20578232      | 3.89732194  | 0.00130933 |
| NTNG1    | -1.223801089    | 3.39713111  | 0.00139898 |
| IGF1R    | 1.092230502     | 4.604757689 | 0.00145261 |
| TPSAB1   | 1.569588104     | 3.89503023  | 0.00158598 |
| IQCF3    | 1.322088303     | 3.378362315 | 0.00166872 |
| SEMA6D   | 1.047662865     | 3.879087643 | 0.00168935 |
| DMP1     | 1.706225181     | 5.373542261 | 0.00170426 |
| INADL    | -1.07809899     | 4.534311602 | 0.00188551 |
| HBB      | 1.829961239     | 9.85497514  | 0.00206485 |
| ATF3     | 1.042517788     | 3.94218687  | 0.00214931 |
| CCKT2    | 1.450120173     | 3.643376295 | 0.00226876 |
| OR4F17   | 1.129626909     | 3.86343791  | 0.002604 |
| FOS      | 1.156332999     | 7.10701183  | 0.00263045 |
| CGREF1   | 1.090372093     | 5.826510092 | 0.00263515 |
| C11orf70 | 1.007164914     | 3.279804146 | 0.00266626 |
| TPS2B    | 1.613775485     | 3.650070644 | 0.00268917 |
| DSCR8    | -1.732497079    | 3.686536986 | 0.00274629 |
| CDYL2    | 1.220460554     | 3.190196988 | 0.00305359 |
| PTN      | -1.15535371     | 8.192544841 | 0.0032828 |
| CXCL8    | 1.62274194      | 3.587795701 | 0.00350428 |
| MSC      | -1.057807226    | 4.262290802 | 0.00381951 |
| GLI1     | 1.206975812     | 3.39206332  | 0.00401056 |
| WFDC2    | 1.616522908     | 3.503997676 | 0.00411246 |
| ADGRD1   | 1.027781742     | 4.463971697 | 0.00438649 |
| PYGM     | 1.178648751     | 3.913496689 | 0.00497737 |
| COL11A2  | 1.309188742     | 7.692556486 | 0.00520437 |
| FOSB     | 1.37955124      | 4.232493194 | 0.00563983 |
| CHI3L1   | 1.673107347     | 3.451849033 | 0.00574495 |
| HRC      | 1.518364483     | 3.047277945 | 0.00597747 |
| FRZB     | 1.206054426     | 3.510020165 | 0.00663679 |
| TMEM145  | 1.233112901     | 3.165275784 | 0.00669925 |
| MMP1     | -1.247190585    | 3.471305185 | 0.00672908 |
| ALDH1A1  | 1.344676387     | 4.426009831 | 0.00695665 |
| PTRZ1    | 1.277704864     | 5.25252576  | 0.00744095 |
| Gene     | Log2 Fold Change | Log2 Fold Change | p-value   |
|----------|-----------------|-----------------|-----------|
| ANGPTL4  | 1.054255939     | 5.685659489     | 0.00818924|
| CAMK2B  | 1.172853278     | 3.532462371     | 0.00837671|
| PEG3    | -1.599265392    | 3.109561147     | 0.00907183|
| SCARA5  | 1.229166487     | 3.629477163     | 0.00909507|
| DAZ1    | 2.326886263     | 3.614268323     | 0.00911625|
| SCRGl   | 1.40433905      | 3.094886127     | 0.00968083|
| IL17B   | 2.024031261     | 3.70835196      | 0.00984527|
| MFI2    | -1.069851581    | 6.840966655     | 0.01024922|
| PI15    | 1.204395938     | 3.775356414     | 0.01049743|
| SLC38A5 | 1.013845055     | 6.038054857     | 0.01065674|
| HBA2    | 1.435376231     | 9.802864684     | 0.01126334|
| NKG7    | -1.017564634    | 3.506814288     | 0.01181224|
| CILP2   | -1.142370676    | 3.62863524      | 0.01209858|
| PHEX    | 1.03414063      | 3.30066864      | 0.01245601|
| C2orf82 | 1.327191312     | 4.710250734     | 0.01289324|
| NPB     | 1.083177926     | 5.019660588     | 0.01483066|
| SCUBE3  | 1.127329898     | 3.378299645     | 0.01570375|
| SLC13A5 | 1.141026344     | 3.76158152      | 0.01572656|
| MT1A    | 1.017102762     | 4.080451982     | 0.01844636|
| ASPN    | -1.137290124    | 6.3561574       | 0.019844  |
| MFAP4   | -1.487709024    | 6.372355643     | 0.01989345|
| FAM133A | -1.13267864     | 3.22665608      | 0.02020882|
| ACTC1   | -1.635864662    | 4.197965647     | 0.02044519|
| CA3     | 1.176995431     | 7.642981555     | 0.02181363|
| OGN     | -1.178699331    | 6.258590051     | 0.02319615|
| CHAD    | 1.104942336     | 5.667159825     | 0.0279692 |
| BGLAP   | 1.247073395     | 7.27597136      | 0.02968101|
| FMOB    | 1.056185618     | 5.803326811     | 0.03001712|
| MYLFP   | -1.12875465     | 5.196491591     | 0.03037025|
| TP53    | -1.274560019    | 9.140405276     | 0.03325026|
| NEB     | -1.081916047    | 4.135265126     | 0.03377208|
| KCNK17  | 1.074969972     | 3.525778708     | 0.04229337|
| SLN     | -1.061866423    | 4.573555423     | 0.04388763|
| logCPM     | PValue    |
|-----------|-----------|
| 9.06632665 | 3.26E-14  |
| 9.436033721| 1.17E-11  |
| 8.246819423| 8.76E-11  |
| 8.354029117| 1.35E-10  |
| 8.026641272| 6.53E-09  |
| 9.16481264  | 1.30E-07  |
| 7.712199903 | 7.67E-07  |
| 8.060141953 | 1.37E-06  |
| 9.412127367 | 2.06E-06  |
| 12.8618976  | 1.22E-05  |
| 8.066502309 | 1.24E-05  |
| 9.464426655 | 1.55E-05  |
| 7.817671974 | 1.85E-05  |
| 8.324041388 | 1.85E-05  |
| 7.827670877 | 4.08E-05  |
| 9.650842714 | 6.20E-05  |
| 9.69484159  | 6.24E-05  |
| 10.06145228 | 0.000231336|
| 8.173276608 | 0.000332981|
| 8.066567676 | 0.000390659|
| 9.749938607 | 0.000467686|
| 9.981538449 | 0.000709075|
| 11.60922847 | 0.000779237|
| 9.604746052 | 0.00081731 |
| 8.337389281 | 0.000819217|
| 8.020275542 | 0.000850903|
| 7.712639548 | 0.001681984|
| 8.034895756 | 0.001692665|
| 8.622211845 | 0.001749888|
| 8.180298617 | 0.002432551|
| 7.752098765 | 0.003843043|
| 8.099978216 | 0.003920206|
| 7.882605218 | 0.004907711|
| 9.271372712 | 0.005573257|
| 8.305501295 | 0.007391792|
| 8.16306522  | 0.008049707|
| 8.288825635 | 0.008466702|
| 7.905389313 | 0.022636647|
| 8.428488585 | 0.026818372|
| 9.47477093  | 0.030924362|
| 7.878108638 | 0.031380042|
| 8.359503387 | 0.034377325|
Table S2 lncRNAs-mRNAs co-expression analysis ([Pearson correlation coefficient] > 0.4 and P-value <

| lncRNA     | mRNA     | cor  | pvalue     |
|------------|----------|------|------------|
| AC132217.4 | IGF2     | 0.614| 7.65E-11   |
| AP003025.2 | AMELX    | 0.433| 1.62E-05   |
| AP003025.2 | IQCF3    | 0.575| 2.03E-09   |
| AP003025.2 | OR4F17   | 0.524| 8.17E-08   |
| AP003025.2 | TMEM145  | 0.488| 7.85E-07   |
| CECR7      | TCF7L2   | 0.47 | 2.31E-06   |
| CECR7      | TTLL7    | 0.433| 1.65E-05   |
| CTB-4E7.1  | ADGRD1   | 0.517| 1.32E-07   |
| CTB-4E7.1  | BMP8A    | 0.622| 3.74E-11   |
| CTB-4E7.1  | BMP8B    | 0.677| 1.23E-13   |
| CTB-4E7.1  | C1QTNF3  | -0.419| 3.27E-05  |
| CTB-4E7.1  | CA3      | 0.727| 2.42E-16   |
| CTB-4E7.1  | CGREF1   | 0.649| 2.68E-12   |
| CTB-4E7.1  | CHMP4C   | 0.746| 1.53E-17   |
| CTB-4E7.1  | CHST13   | 0.594| 4.37E-10   |
| CTB-4E7.1  | COL11A2  | 0.504| 3.10E-07   |
| CTB-4E7.1  | COL13A1  | 0.444| 9.14E-06   |
| CTB-4E7.1  | COL22A1  | 0.716| 9.93E-16   |
| CTB-4E7.1  | COL8A2   | -0.606| 1.50E-10  |
| CTB-4E7.1  | CORT     | 0.558| 7.52E-09   |
| CTB-4E7.1  | CPE      | 0.605| 1.76E-10   |
| CTB-4E7.1  | CPNE5    | 0.471| 2.10E-06   |
| CTB-4E7.1  | CYFIP2   | 0.78 | 5.18E-20   |
| CTB-4E7.1  | DMP1     | 0.618| 5.11E-11   |
| CTB-4E7.1  | EFDH1    | 0.51 | 2.09E-07   |
| CTB-4E7.1  | EPCAM    | 0.511| 1.95E-07   |
| CTB-4E7.1  | GNG4     | 0.611| 9.68E-11   |
| CTB-4E7.1  | GRAMD1B  | 0.757| 2.51E-18   |
| CTB-4E7.1  | HCRT     | 0.413| 4.32E-05   |
| CTB-4E7.1  | ISM1     | 0.441| 1.09E-05   |
| CTB-4E7.1  | LGR6     | 0.594| 4.47E-10   |
| CTB-4E7.1  | MEPE     | 0.509| 2.28E-07   |
| CTB-4E7.1  | MFAP4    | -0.578| 1.60E-09  |
| CTB-4E7.1  | MLC1     | 0.439| 1.22E-05   |
| CTB-4E7.1  | MSC      | -0.554| 9.79E-09  |
| CTB-4E7.1  | NPB      | 0.426| 2.25E-05   |
| CTB-4E7.1  | NTNG1    | -0.428| 2.05E-05  |
| CTB-4E7.1  | OGN      | -0.48 | 1.31E-06   |
| CTB-4E7.1  | PHEX     | 0.58 | 1.36E-09   |
| CTB-4E7.1  | PTN      | -0.458| 4.54E-06  |
| CTB-4E7.1  | PTPRZ1   | 0.748| 1.10E-17   |
| CTB-4E7.1  | PYGM     | 0.412| 4.60E-05   |
| CTB-4E7.1  | RASGRP2  | 0.451| 6.50E-06   |
| CTB-4E7.1  | RHBDL2   | 0.648| 2.95E-12   |
| CTB-4E7.1  | SLC8A3   | 0.644| 4.46E-12   |
| CTB-4E7.1  | SYT12    | 0.604| 1.82E-10   |
| CTB-4E7.1  | TAC4     | 0.453| 5.84E-06   |
| CTB-4E7.1  | TCF7L2   | 0.462| 3.57E-06   |
| CTB-4E7.1  | TMCC2    | 0.735| 7.50E-17   |
| Gene       | Symbol  | Value  | p-value  |
|-----------|---------|--------|----------|
| CTB-4E7.1 | TPD52   | 0.474  | 1.83E-06 |
| CTB-4E7.1 | VCAM1   | -0.458 | 4.40E-06 |
| CTB-4E7.1 | WIFI    | 0.425  | 2.42E-05 |
| CTD-2003C8.2 | ASPN | 0.509  | 2.17E-07 |
| CTD-2003C8.2 | C1QTNF3 | 0.621  | 3.95E-11 |
| CTD-2003C8.2 | COL8A2 | 0.471  | 2.13E-06 |
| CTD-2003C8.2 | FMO1  | 0.423  | 2.63E-05 |
| CTD-2003C8.2 | MFAP4 | 0.474  | 1.86E-06 |
| CTD-2003C8.2 | MSC   | 0.471  | 2.15E-06 |
| CTD-2003C8.2 | OGN   | 0.555  | 9.73E-09 |
| CTD-2003C8.2 | PDLIM3 | 0.496  | 5.02E-07 |
| CTD-2003C8.2 | PTN   | 0.444  | 9.08E-06 |
| CTD-2003C8.2 | RGS4  | 0.492  | 6.50E-07 |
| CTD-2026K11.BMP8A | 0.449  | 7.32E-06 |
| CTD-2358C21.CDYL2 | 0.409 | 5.15E-05 |
| CTD-2358C21.TTLL7 | 0.505  | 2.84E-07 |
| EAF1-AS1 | AMELX   | 0.528  | 6.27E-08 |
| EAF1-AS1 | ARID3B  | 0.605  | 1.72E-10 |
| EAF1-AS1 | ASPDH   | 0.419  | 3.27E-05 |
| EAF1-AS1 | IQCF3   | 0.73   | 1.52E-16 |
| EAF1-AS1 | MYH3    | 0.432  | 1.73E-05 |
| EAF1-AS1 | OR4F17  | 0.739  | 3.94E-17 |
| EAF1-AS1 | SCGB3A2 | 0.512  | 1.80E-07 |
| EAF1-AS1 | SEMA6D  | 0.429  | 1.94E-05 |
| EAF1-AS1 | TMEM145  | 0.747  | 1.15E-17 |
| EAF1-AS1 | ZNF461  | 0.632  | 1.37E-11 |
| FLJ26245 | CWF19L2 | 0.443  | 9.86E-06 |
| FLJ26245 | PTPRZ1  | 0.432  | 1.73E-05 |
| FLJ26245 | SAP25   | 0.477  | 1.54E-06 |
| GS1-600G8.5 | FMO1  | 0.47   | 2.32E-06 |
| GS1-600G8.5 | MMP1   | 0.408  | 5.47E-05 |
| GS1-600G8.5 | SUGCT  | 0.444  | 9.48E-06 |
| HAS2-AS1 | ADGRD1  | -0.449 | 7.06E-06 |
| HAS2-AS1 | BMP8A   | -0.465 | 2.93E-06 |
| HAS2-AS1 | BMP8B   | -0.516 | 1.41E-07 |
| HAS2-AS1 | CA3     | -0.449 | 7.03E-06 |
| HAS2-AS1 | CHMP4C  | -0.499 | 4.24E-07 |
| HAS2-AS1 | CHST13  | -0.458 | 4.37E-06 |
| HAS2-AS1 | COL22A1 | -0.513 | 1.65E-07 |
| HAS2-AS1 | COL8A2  | 0.432  | 1.73E-05 |
| HAS2-AS1 | CORT    | -0.522 | 9.22E-08 |
| HAS2-AS1 | CPE     | -0.428 | 2.07E-05 |
| HAS2-AS1 | CYFIP2  | -0.553 | 1.05E-08 |
| HAS2-AS1 | GNG4    | -0.447 | 7.80E-06 |
| HAS2-AS1 | GRAMD1B | -0.547 | 1.64E-08 |
| HAS2-AS1 | LGR6    | -0.42  | 3.13E-05 |
| HAS2-AS1 | MFAP4   | 0.402  | 6.97E-05 |
| HAS2-AS1 | MSC     | 0.524  | 8.27E-08 |
| HAS2-AS1 | NPB     | -0.462 | 3.59E-06 |
| HAS2-AS1 | NTNG1   | 0.422  | 2.84E-05 |
| HAS2-AS1 | OGN     | 0.555  | 9.13E-09 |
| Gene Name   | Gene Symbol | Correlation | p-value  |
|------------|-------------|-------------|----------|
| HAS2-AS1   | PHEX        | -0.414      | 4.18E-05 |
| HAS2-AS1   | PI15        | 0.462       | 3.64E-06 |
| HAS2-AS1   | PTN         | 0.495       | 5.25E-07 |
| HAS2-AS1   | PTPRZ1      | -0.501      | 3.73E-07 |
| HAS2-AS1   | SLC8A3      | -0.469      | 2.45E-06 |
| HAS2-AS1   | SYT12       | -0.484      | 1.03E-06 |
| HAS2-AS1   | TMCC2       | -0.544      | 2.16E-08 |
| HAS2-AS1   | TSPAN11     | -0.451      | 6.60E-06 |
| IQCH-AS1   | IQCF3       | 0.537       | 3.34E-08 |
| IQCH-AS1   | OR4F17      | 0.534       | 4.08E-08 |
| IQCH-AS1   | TCF7L2      | 0.476       | 1.66E-06 |
| IQCH-AS1   | TMEM145     | 0.43        | 1.85E-05 |
| IQCH-AS1   | TTLL7       | 0.41        | 4.83E-05 |
| LINC01287  | CDYL2       | 0.455       | 5.16E-06 |
| LINC01549  | BMP8A       | 0.451       | 6.40E-06 |
| LINC01549  | BMP8B       | 0.546       | 1.80E-08 |
| LINC01549  | CA3         | 0.5         | 3.98E-07 |
| LINC01549  | CHMP4C      | 0.638       | 7.66E-12 |
| LINC01549  | CHST13      | 0.405       | 6.30E-05 |
| LINC01549  | COL22A1     | 0.487       | 8.51E-07 |
| LINC01549  | COL8A2      | -0.513      | 1.76E-07 |
| LINC01549  | CORT        | 0.452       | 6.14E-06 |
| LINC01549  | CYFIP2      | 0.589       | 6.72E-10 |
| LINC01549  | DMP1        | 0.466       | 2.82E-06 |
| LINC01549  | GNG4        | 0.532       | 4.77E-08 |
| LINC01549  | GRAMD1B     | 0.563       | 5.09E-09 |
| LINC01549  | LGR6        | 0.484       | 1.01E-06 |
| LINC01549  | MFAP4       | -0.428      | 2.13E-05 |
| LINC01549  | OGN         | -0.432      | 1.72E-05 |
| LINC01549  | PHEX        | 0.42        | 3.01E-05 |
| LINC01549  | PTN         | -0.425      | 2.39E-05 |
| LINC01549  | PTPRZ1      | 0.615       | 7.24E-11 |
| LINC01549  | RHBDL2      | 0.429       | 2.03E-05 |
| LINC01549  | SLC8A3      | 0.405       | 6.26E-05 |
| LINC01549  | SYT12       | 0.449       | 7.06E-06 |
| LINC01549  | TCF7L2      | 0.409       | 5.17E-05 |
| LINC01549  | TMCC2       | 0.591       | 5.52E-10 |
| LINC01549  | TP5D2       | 0.459       | 4.15E-06 |
| LINC01551  | AMELX       | 0.43        | 1.88E-05 |
| LINC01551  | ARID3B      | 0.448       | 7.62E-06 |
| LINC01551  | ASPDH       | 0.436       | 1.37E-05 |
| LINC01551  | DHX57       | 0.468       | 2.60E-06 |
| LINC01551  | GLTSCR1     | 0.411       | 4.65E-05 |
| LINC01551  | IQCF3       | 0.762       | 1.09E-18 |
| LINC01551  | OR4F17      | 0.65        | 2.42E-12 |
| LINC01551  | SCGB3A2     | 0.447       | 7.99E-06 |
| LINC01551  | STPG1       | 0.444       | 9.42E-06 |
| LINC01551  | TMEM145     | 0.676       | 1.45E-13 |
| LINC01551  | TP53        | 0.422       | 2.73E-05 |
| LINC01551  | ZNF461      | 0.502       | 3.49E-07 |
| PVRL3-AS1  | DSCR8       | 0.421       | 2.94E-05 |
| Gene       | Symbol | Expression | p-value |
|-----------|--------|------------|---------|
| PVRL3-AS1 | HGF    | 0.41       | 4.83E-05|
| RP11-162J8.3 | ADGRD1 | -0.505     | 2.76E-07|
| RP11-162J8.3 | BMP8A | -0.407     | 5.61E-05|
| RP11-162J8.3 | CGREF1 | -0.417     | 3.49E-05|
| RP11-162J8.3 | EFHD1 | -0.409     | 5.23E-05|
| RP11-162J8.3 | RGS4  | 0.404      | 6.61E-05|
| RP11-162J8.3 | SUGCT | 0.401      | 7.45E-05|
| RP11-1L9.1   | ADGRD1 | -0.45      | 6.72E-06|
| RP11-1L9.1   | CHRNA1 | 0.509      | 2.24E-07|
| RP11-1L9.1   | COL1A2 | -0.45      | 6.85E-06|
| RP11-1L9.1   | COL13A1| -0.407     | 5.63E-05|
| RP11-1L9.1   | CPE    | -0.494     | 5.70E-07|
| RP11-1L9.1   | FMO1   | 0.475      | 1.74E-06|
| RP11-1L9.1   | HGF    | 0.436      | 1.42E-05|
| RP11-1L9.1   | PDLIM3 | 0.449      | 7.07E-06|
| RP11-1L9.1   | RGS4   | 0.418      | 3.46E-05|
| RP11-1L9.1   | RHBDL2 | -0.529     | 6.08E-08|
| RP11-299J3.8 | BAHCC1 | 0.407      | 5.66E-05|
| RP11-299J3.8 | CWF19L2| 0.407      | 5.68E-05|
| RP11-299J3.8 | DHX57  | 0.669      | 3.23E-13|
| RP11-299J3.8 | GLTSCR1| 0.588      | 7.09E-10|
| RP11-299J3.8 | NUCKS1 | 0.439      | 1.19E-05|
| RP11-299J3.8 | PLEKHA2| 0.534      | 4.21E-08|
| RP11-299J3.8 | STPG1  | 0.521      | 9.94E-08|
| RP11-299J3.8 | USP28  | 0.574      | 2.29E-09|
| RP11-299J3.8 | ZNF362 | 0.539      | 3.07E-08|
| RP11-345P4.1C | STPG1 | 0.439      | 1.21E-05|
| RP11-358L22.:CAMK2B | 0.839 | 1.65E-25  |
| RP11-358L22.:SFTPB | 0.449 | 7.30E-06  |
| RP11-383J24.1 | BAHCC1 | 0.575      | 2.00E-09|
| RP11-383J24.1 | DHX57 | 0.498      | 4.31E-07|
| RP11-383J24.1 | GLTSCR1| 0.544      | 2.06E-08|
| RP11-383J24.1 | PLEKHA2| 0.543      | 2.20E-08|
| RP11-383J24.1 | STPG1  | 0.404      | 6.56E-05|
| RP11-383J24.1 | USP28  | 0.509      | 2.23E-07|
| RP11-383J24.1 | ZNF362 | 0.443      | 9.53E-06|
| RP11-417E7.2 | CHRNA1 | 0.434      | 1.55E-05|
| RP11-417E7.2 | FMO1   | 0.413      | 4.25E-05|
| RP11-417E7.2 | RGS4   | 0.481      | 1.20E-06|
| RP11-417E7.2 | SUGCT  | 0.409      | 5.26E-05|
| RP11-417E7.2 | VCAM1  | 0.423      | 2.60E-05|
| RP11-426C22.:ADGRD1 | -0.436 | 1.41E-05  |
| RP11-426C22.:C1QTNF3 | 0.42 | 3.03E-05  |
| RP11-426C22.:CA3 | -0.417 | 3.50E-05  |
| RP11-426C22.:CHMP4C | -0.423 | 2.70E-05  |
| RP11-426C22.:COL1A2 | -0.41 | 4.83E-05  |
| RP11-426C22.:COL8A2 | 0.408 | 5.35E-05  |
| RP11-426C22.:CPE | -0.461 | 3.67E-06  |
| RP11-426C22.:CYFIP2 | -0.431 | 1.77E-05  |
| RP11-426C22.:EFHD1 | -0.443 | 9.84E-06  |
| RP11-426C22.:GBP1 | 0.464 | 3.14E-06  |
| Gene         | Value | p-value |
|-------------|-------|---------|
| GBP5        | 0.431 | 1.82E-05|
| HGF         | 0.476 | 1.60E-06|
| MFAP4       | 0.475 | 1.76E-06|
| MSC         | 0.452 | 5.99E-06|
| NPB         | -0.424| 2.55E-05|
| RGS4        | 0.526 | 7.38E-08|
| RHBDL2      | -0.521| 9.91E-08|
| SLC8A3      | -0.425| 2.41E-05|
| SUGCT       | 0.446 | 8.18E-06|
| AGER        | 0.409 | 5.09E-05|
| ARID3B      | 0.435 | 1.47E-05|
| IQCF3       | 0.65  | 2.27E-12|
| NPW         | 0.432 | 1.68E-05|
| OR4F17      | 0.547 | 1.65E-08|
| SYT15       | 0.537 | 3.33E-08|
| TAC4        | 0.415 | 3.87E-05|
| TMEM145     | 0.572 | 2.60E-09|
| DSCR8       | 0.426 | 2.33E-05|
| HGF         | 0.451 | 6.39E-06|
| SSX2B       | 0.419 | 3.28E-05|
| ADGRD1      | 0.454 | 5.52E-06|
| BMP8A       | 0.882 | 3.87E-31|
| BMP8B       | 0.786 | 1.85E-20|
| CA3         | 0.458 | 4.31E-06|
| CGREF1      | 0.498 | 4.27E-07|
| CHMP4C      | 0.607 | 1.44E-10|
| CHST13      | 0.539 | 3.05E-08|
| COL11A2     | 0.424 | 2.55E-05|
| COL13A1     | 0.458 | 4.35E-06|
| COL22A1     | 0.613 | 8.45E-11|
| CORT        | 0.503 | 3.22E-07|
| CPE         | 0.448 | 7.44E-06|
| CYFIP2      | 0.57  | 3.11E-09|
| DMP1        | 0.696 | 1.33E-14|
| EFHD1       | 0.52  | 1.10E-07|
| GRAMD1B     | 0.516 | 1.37E-07|
| HCRT        | 0.401 | 7.43E-05|
| ISM1        | 0.435 | 1.47E-05|
| ITGA10      | 0.404 | 6.40E-05|
| LGR6        | 0.586 | 8.77E-10|
| MEPE        | 0.56  | 6.57E-09|
| PHEX        | 0.57  | 2.93E-09|
| PTPRZ1      | 0.514 | 1.61E-07|
| RASGRP2     | 0.558 | 7.60E-09|
| RHBDL2      | 0.488 | 8.07E-07|
| SLC8A3      | 0.514 | 1.64E-07|
| SYT12       | 0.62  | 4.30E-11|
| TAC4        | 0.434 | 1.57E-05|
| TMCC2       | 0.588 | 6.99E-10|
| TSPAN11     | 0.503 | 3.14E-07|
| WIFI        | 0.531 | 5.25E-08|
| Gene Symbol       | Score | Significance |
|------------------|-------|--------------|
| RP11-817J15.3 ADGRD1 | 0.473 | 1.91E-06     |
| RP11-817J15.3 ASPN   | -0.481| 1.20E-06     |
| RP11-817J15.3 BMP8A  | 0.523 | 8.97E-08     |
| RP11-817J15.3 BMP8B  | 0.642 | 5.45E-12     |
| RP11-817J15.3 C11orf70 | 0.524 | 8.07E-08     |
| RP11-817J15.3 C1QTNF3 | -0.528| 6.44E-08     |
| RP11-817J15.3 CA3    | 0.598 | 3.06E-10     |
| RP11-817J15.3 CGREF1 | 0.471 | 2.17E-06     |
| RP11-817J15.3 CHMP4C | 0.706 | 3.91E-15     |
| RP11-817J15.3 CHST13 | 0.489 | 7.44E-07     |
| RP11-817J15.3 CILP2  | -0.458| 4.44E-06     |
| RP11-817J15.3 COL11A2 | 0.402 | 7.22E-05     |
| RP11-817J15.3 COL22A1 | 0.61  | 1.13E-10     |
| RP11-817J15.3 COL8A2 | -0.646| 3.71E-12     |
| RP11-817J15.3 CORT   | 0.472 | 2.07E-06     |
| RP11-817J15.3 CPE    | 0.504 | 3.01E-07     |
| RP11-817J15.3 CPNE5  | 0.416 | 3.78E-05     |
| RP11-817J15.3 CYFIP2 | 0.558 | 7.47E-09     |
| RP11-817J15.3 DMP1   | 0.519 | 1.17E-07     |
| RP11-817J15.3 EFHD1  | 0.519 | 1.18E-07     |
| RP11-817J15.3 GNG4   | 0.446 | 8.30E-06     |
| RP11-817J15.3 GRAMD1B | 0.608| 1.29E-10     |
| RP11-817J15.3 ITGA10 | 0.416| 3.72E-05     |
| RP11-817J15.3 KCNK17 | 0.558| 7.55E-09     |
| RP11-817J15.3 LGR6   | 0.506 | 2.67E-07     |
| RP11-817J15.3 MFAP4  | -0.583| 1.12E-09     |
| RP11-817J15.3 MLC1   | 0.523 | 8.79E-08     |
| RP11-817J15.3 MSC    | -0.428| 2.09E-05     |
| RP11-817J15.3 NPB    | 0.492 | 6.30E-07     |
| RP11-817J15.3 OGN    | -0.533| 4.55E-08     |
| RP11-817J15.3 PHEX   | 0.536 | 3.60E-08     |
| RP11-817J15.3 PTN    | -0.426| 2.31E-05     |
| RP11-817J15.3 PTPRZ1 | 0.628| 2.07E-11     |
| RP11-817J15.3 PYGM   | 0.508 | 2.32E-07     |
| RP11-817J15.3 RASGRP2 | 0.55 | 1.35E-08     |
| RP11-817J15.3 RHBDL2 | 0.542| 2.36E-08     |
| RP11-817J15.3 SLC8A3 | 0.557| 8.35E-09     |
| RP11-817J15.3 SYT12  | 0.6  | 2.66E-10     |
| RP11-817J15.3 TAC4   | 0.516| 1.37E-07     |
| RP11-817J15.3 TMCC2  | 0.581| 1.28E-09     |
| RP11-81A22.5 AMELX   | 0.429| 2.00E-05     |
| RP11-81A22.5 ASPN    | -0.491| 6.54E-07     |
| RP11-81A22.5 BMP8A   | 0.46 | 3.88E-06     |
| RP11-81A22.5 BMP8B   | 0.5 | 3.81E-07     |
| RP11-81A22.5 C1QTNF3 | -0.482| 1.15E-06     |
| RP11-81A22.5 CA3     | 0.558| 7.81E-09     |
| RP11-81A22.5 CGREF1  | 0.51 | 2.09E-07     |
| RP11-81A22.5 CHMP4C  | 0.729| 1.85E-16     |
| RP11-81A22.5 CHST13  | 0.4  | 7.85E-05     |
| RP11-81A22.5 COL11A2 | 0.464| 3.19E-06     |
| RP11-81A22.5 COL22A1 | 0.503| 3.23E-07     |
| Gene       | Value 1 | Value 2 |
|------------|---------|---------|
| COL8A2     | -0.598  | 3.19E-10|
| CORT       | 0.627   | 2.22E-11|
| CPE        | 0.418   | 3.41E-05|
| CPNE5      | 0.507   | 2.53E-07|
| CYFIP2     | 0.749   | 8.67E-18|
| DMP1       | 0.489   | 7.42E-07|
| EFHD1      | 0.487   | 8.66E-07|
| EPCAM      | 0.402   | 7.26E-05|
| GNG4       | 0.54    | 2.80E-08|
| GRAMD1B    | 0.592   | 5.17E-10|
| ISM1       | 0.409   | 5.27E-05|
| LGR6       | 0.513   | 1.72E-07|
| MEPE       | 0.426   | 2.29E-05|
| MFAP4      | -0.542  | 2.45E-08|
| MSC        | -0.536  | 3.59E-08|
| NPB        | 0.515   | 1.51E-07|
| OGN        | -0.519  | 1.15E-07|
| PTN        | -0.491  | 6.56E-07|
| PTGRZ1     | 0.7     | 8.17E-15|
| RASGRP2    | 0.458   | 4.47E-06|
| RHBDL2     | 0.459   | 4.16E-06|
| SLC8A3     | 0.469   | 2.39E-06|
| SYT12      | 0.528   | 6.17E-08|
| TMCC2      | 0.701   | 7.36E-15|
| TPD52      | 0.501   | 3.74E-07|
| VCAM1      | -0.426  | 2.29E-05|
| GPRC5A     | 0.44    | 1.16E-05|
| SNX14      | -0.488  | 8.24E-07|
| AMELX      | 0.429   | 2.00E-05|
| ASPN       | -0.491  | 6.54E-07|
| BMP8A      | 0.46    | 3.88E-06|
| BMP8B      | 0.5     | 3.81E-07|
| C1QTNF3    | -0.482  | 1.15E-06|
| CA3        | 0.558   | 7.81E-09|
| CGREF1     | 0.51    | 2.09E-07|
| CHMPC4C    | 0.729   | 1.85E-16|
| CHST13     | 0.4     | 7.85E-05|
| COL11A2    | 0.464   | 3.19E-06|
| COL22A1    | 0.503   | 3.23E-07|
| COL8A2     | -0.598  | 3.19E-10|
| CORT       | 0.627   | 2.22E-11|
| CPE        | 0.418   | 3.41E-05|
| CPNE5      | 0.507   | 2.53E-07|
| CYFIP2     | 0.749   | 8.67E-18|
| DMP1       | 0.489   | 7.42E-07|
| EFHD1      | 0.487   | 8.66E-07|
| EPCAM      | 0.402   | 7.26E-05|
| GNG4       | 0.54    | 2.80E-08|
| GRAMD1B    | 0.592   | 5.17E-10|
| ISM1       | 0.409   | 5.27E-05|
| LGR6       | 0.513   | 1.72E-07|
| Gene Name | Value 1  | Value 2 |
|-----------|---------|---------|
| RP1-261G23.7 MEPE | 0.426 | 2.29E-05 |
| RP1-261G23.7 MFAP4 | -0.542 | 2.45E-08 |
| RP1-261G23.7 MSC | -0.536 | 3.59E-08 |
| RP1-261G23.7 NPB | 0.515 | 1.51E-07 |
| RP1-261G23.7 OGN | -0.519 | 1.15E-07 |
| RP1-261G23.7 PTN | -0.491 | 6.56E-07 |
| RP1-261G23.7 PTPRZ1 | 0.7 | 8.17E-15 |
| RP1-261G23.7 RASGRP2 | 0.458 | 4.47E-06 |
| RP1-261G23.7 RHBDL2 | 0.459 | 4.16E-06 |
| RP1-261G23.7 SLC8A3 | 0.469 | 2.39E-06 |
| RP1-261G23.7 SYT12 | 0.528 | 6.17E-08 |
| RP1-261G23.7 TMCC2 | 0.701 | 7.36E-15 |
| RP1-261G23.7 TPD52 | 0.501 | 3.74E-07 |
| RP1-261G23.7 VCAM1 | -0.426 | 2.29E-05 |
| SATB2-AS1 ADGRD1 | 0.625 | 2.73E-11 |
| SATB2-AS1 BMP8A | 0.648 | 3.02E-12 |
| SATB2-AS1 BMP8B | 0.748 | 1.07E-17 |
| SATB2-AS1 C1QTNF3 | -0.437 | 1.32E-05 |
| SATB2-AS1 CA3 | 0.686 | 4.54E-14 |
| SATB2-AS1 CGREF1 | 0.581 | 1.24E-09 |
| SATB2-AS1 CHAD | 0.417 | 3.53E-05 |
| SATB2-AS1 CHMP4C | 0.679 | 1.04E-13 |
| SATB2-AS1 CHST13 | 0.747 | 1.32E-17 |
| SATB2-AS1 COL11A2 | 0.564 | 4.67E-09 |
| SATB2-AS1 COL13A1 | 0.526 | 7.18E-08 |
| SATB2-AS1 COL22A1 | 0.79 | 7.31E-21 |
| SATB2-AS1 COL8A2 | -0.43 | 1.92E-05 |
| SATB2-AS1 CORT | 0.626 | 2.43E-11 |
| SATB2-AS1 CPE | 0.585 | 8.80E-10 |
| SATB2-AS1 CPNE5 | 0.444 | 9.26E-06 |
| SATB2-AS1 CYFIP2 | 0.621 | 3.99E-11 |
| SATB2-AS1 DMP1 | 0.657 | 1.13E-12 |
| SATB2-AS1 EFHD1 | 0.585 | 9.18E-10 |
| SATB2-AS1 EPCAM | 0.4 | 7.74E-05 |
| SATB2-AS1 GNG4 | 0.644 | 4.23E-12 |
| SATB2-AS1 GRAMD1B | 0.654 | 1.64E-12 |
| SATB2-AS1 ISM1 | 0.459 | 4.21E-06 |
| SATB2-AS1 KCNK17 | 0.404 | 6.42E-05 |
| SATB2-AS1 LGR6 | 0.669 | 3.25E-13 |
| SATB2-AS1 LYNX1 | 0.42 | 3.09E-05 |
| SATB2-AS1 MEPE | 0.469 | 2.37E-06 |
| SATB2-AS1 MFAP4 | -0.524 | 8.21E-08 |
| SATB2-AS1 MLC1 | 0.484 | 1.05E-06 |
| SATB2-AS1 MSC | -0.416 | 3.81E-05 |
| SATB2-AS1 NPB | 0.582 | 1.21E-09 |
| SATB2-AS1 OGN | -0.459 | 4.08E-06 |
| SATB2-AS1 PHEX | 0.599 | 2.75E-10 |
| SATB2-AS1 PTPRZ1 | 0.646 | 3.72E-12 |
| SATB2-AS1 PYGM | 0.485 | 9.42E-07 |
| SATB2-AS1 RASGRP2 | 0.462 | 3.55E-06 |
| SATB2-AS1 RHBDL2 | 0.689 | 3.34E-14 |
| Gene 1     | Gene 2     | Score 1 | Score 2 |
|-----------|-----------|---------|---------|
| SATB2-AS1 | SEMA6D    | 0.458   | 4.37E-06|
| SATB2-AS1 | SLC13A5   | 0.445   | 8.94E-06|
| SATB2-AS1 | SLC8A3    | 0.711   | 2.15E-15|
| SATB2-AS1 | SYT12     | 0.57    | 3.13E-09|
| SATB2-AS1 | TAC4      | 0.455   | 5.08E-06|
| SATB2-AS1 | TMCC2     | 0.77    | 2.99E-19|
| SATB2-AS1 | TPDS2     | 0.516   | 1.39E-07|
| SATB2-AS1 | TSPAN11   | 0.552   | 1.14E-08|
| SATB2-AS1 | TF1       | 0.453   | 5.65E-06|
| SNHG23    | APOD      | 0.455   | 5.08E-06|
| SNHG23    | ARID3B    | 0.411   | 4.79E-05|
| SNHG23    | ASPDH     | 0.509   | 2.19E-07|
| SNHG23    | IL17B     | 0.412   | 4.50E-05|
| SNHG23    | IQCF3     | 0.519   | 1.12E-07|
| SNHG23    | MFAP4     | 0.489   | 7.53E-07|
| SNHG23    | MYH3      | 0.532   | 4.89E-08|
| SNHG23    | NOX4      | 0.422   | 2.74E-05|
| SNHG23    | NPB       | -0.446  | 8.47E-06|
| SNHG23    | OR4F17    | 0.595   | 3.93E-10|
| SNHG23    | PDLIM3    | 0.464   | 3.13E-06|
| SNHG23    | RGS4      | 0.496   | 4.96E-07|
| SNHG23    | RHBDL2    | -0.403  | 6.89E-05|
| SNHG23    | SCGB3A2   | 0.433   | 1.64E-05|
| SNHG23    | TMEM145   | 0.482   | 1.15E-06|
| SNHG23    | ZNF461    | 0.555   | 9.44E-09|
| TRMT2B-AS1| AMELX     | 0.515   | 1.49E-07|
| TRMT2B-AS1| ARID3B    | 0.566   | 4.00E-09|
| TRMT2B-AS1| CRYBA2    | 0.457   | 4.58E-06|
| TRMT2B-AS1| IQCF3     | 0.595   | 3.94E-10|
| TRMT2B-AS1| OR4F17    | 0.57    | 2.91E-09|
| TRMT2B-AS1| SEMA6D    | 0.486   | 9.11E-07|
| TRMT2B-AS1| SYT15     | 0.443   | 9.66E-06|
| TRMT2B-AS1| TMEM145   | 0.686   | 4.40E-14|
| TRMT2B-AS1| ZNF461    | 0.497   | 4.55E-07|
| XXbac-BPG25| ARID3B  | 0.449   | 6.99E-06|
| XXbac-BPG25| TCF7L2  | 0.52    | 1.08E-07|
| XXbac-BPG25| TMCC2   | 0.435   | 1.47E-05|
Table S3  Univariate Cox proportional hazards regression analysis (P <0.05) of the differentially expressed IncRNA:

| IncRNAID      | coef   | p       | HR      | confidence-lower | confidence-upper |
|---------------|--------|---------|---------|------------------|-----------------|
| SATB2.AS1     | 0.019733087 | 0.00822798 | 1.0199291 | 1.005110287 | 1.034966335 |
| RP1.261G23.7   | 0.010574055 | 0.00994847 | 1.0106302 | 1.002536981 | 1.018788668 |
| RP11.69E11.4   | 0.011166789 | 0.01452316 | 1.0112294 | 1.002214245 | 1.020325588 |
| RP11.817J15.3  | 0.016746135 | 0.02627327 | 1.0168871 | 1.001977778 | 1.032018347 |
expressed IncRNAs profiling data in the training dataset.
| IncRNA       | coef  | exp(coef) | se(coef) | p       |
|-------------|-------|-----------|----------|---------|
| RP1.261G23.7| 0.01488| 1.014991  | 0.004258 | 0.000475 *** |
| SATB2.AS1   | 0.02085| 1.021069  | 0.008271 | 0.011708 * |
| RP11.69E11.4| 0.010541 | 1.010597 | 0.005126 | 0.039751 * |
ing dataset.
Table S5 The signature riskScore composed of 3 lncRNA combinations in the training and test training

| id     | futex  | fustat | riskScore   | risk  | id          |
|--------|--------|--------|-------------|-------|-------------|
| TARGET.40.0A4HXS | 8.076712329 | 0 | 0.494084086 | low   | TARGET.40.PANSX |
| TARGET.40.PAMHY | 10.68493151 | 0 | 0.519979269 | low   | TARGET.40.PAKXL |
| TARGET.40.PAVALI | 1.509589041 | 0 | 0.429255376 | low   | TARGET.40.PAMHI |
| TARGET.40.0A4I8U | 2.693150685 | 0 | 0.253904261 | low   | TARGET.40.0A4I4A |
| TARGET.40.PAMRHI | 0.106849315 | 1 | 3.983957344 | high  | TARGET.40.PARGT |
| TARGET.40.PAUUM | 1.84109589 | 0 | 0.409227004 | low   | TARGET.40.PATMI |
| TARGET.40.PATJVI | 4.315068493 | 0 | 0.201279501 | low   | TARGET.40.0A4HM |
| TARGET.40.PAKFVX | 0.989041096 | 1 | 0.904008157 | high  | TARGET.40.PATKS |
| TARGET.40.PANMIC | 1.405479452 | 1 | 0.66447752 | high  | TARGET.40.PARBG |
| TARGET.40.PALFYN | 1.783561644 | 1 | 0.552796941 | low   | TARGET.40.0A4I6O |
| TARGET.40.PAMTC | 8.224657534 | 0 | 0.272585586 | low   | TARGET.40.PAME |
| TARGET.40.PARJXU | 2.221917808 | 1 | 1.911894204 | high  | TARGET.40.PALG |
| TARGET.40.PAPIJR | 4.709589041 | 0 | 0.451577357 | low   | TARGET.40.PASNZ |
| TARGET.40.PALWW | 5.926027397 | 0 | 0.617118688 | low   | TARGET.40.PAUTW |
| TARGET.40.PAMJXS | 8.723287671 | 1 | 1.183158824 | high  | TARGET.40.0A4HX |
| TARGET.40.0A4I65 | 5.460273973 | 1 | 0.916753452 | high  | TARGET.40.PANZ |
| TARGET.40.PAPXGT | 6.920547945 | 0 | 0.569954756 | low   | TARGET.40.PASEB |
| TARGET.40.PATMIX | 0.18630137 | 1 | 2.63461591 | low   | TARGET.40.PATUX |
| TARGET.40.PATMPI | 1.364383562 | 1 | 2.682254816 | high  | TARGET.40.PARKA |
| TARGET.40.PAPNVI | 1.62739726 | 1 | 2.330271 | low   | TARGET.40.PAUYT |
| TARGET.40.PATPBS | 4.021917808 | 0 | 0.928600529 | low   | TARGET.40.PASY |
| TARGET.40.0A4I0U | 6.071232877 | 0 | 0.742399221 | low   | TARGET.40.PALEC |
| TARGET.40.NAASJI | 1.063013699 | 1 | 0.488103249 | low   | TARGET.40.PASRN |
| TARGET.40.0A4I0Q | 1.421917808 | 1 | 1.18310682 | low   | TARGET.40.PANSE |
| TARGET.40.PAMYY | 0.350684932 | 1 | 0.599961478 | low   | TARGET.40.0A4I3S |
| TARGET.40.PALHR1 | 11.983561644 | 0 | 2.63461591 | low   | TARGET.40.PATUX |
| TARGET.40.PALZGU | 3.978082192 | 1 | 1.411669784 | high  | TARGET.40.PANVJ |
| TARGET.40.0A4I8S | 0.863013699 | 1 | 1.219711363 | high  | TARGET.40.0A4I48 |
| TARGET.40.0A4I8O | 1.095890411 | 1 | 0.35445718 | low   | TARGET.40.0A4HL |
| TARGET.40.PAVDTY | 1.010958904 | 0 | 0.768173721 | low   | TARGET.40.0A4I4V |
| TARGET.40.0A4HY5 | 0.597260274 | 1 | 0.612404685 | low   | TARGET.40.PAVEC |
| TARGET.40.PANGR\ | 9.254794521 | 0 | 0.362090864 | low   | TARGET.40.PAVEC |
| TARGET.40.PASKZZ | 1.484931507 | 0 | 0.26325869 | low   | TARGET.40.PASEFS |
| TARGET.40.PASFCV | 5.649315068 | 0 | 0.216036584 | low   | TARGET.40.PASEFS |
| TARGET.40.PASEFS | 5.78630137 | 0 | 0.550370725 | low   | TARGET.40.PASEFS |
| TARGET.40.PAKZKK | 1.238356164 | 1 | 0.948795119 | high  | TARGET.40.PAKZKK |
| TARGET.40.PAPKWI | 0.526027397 | 1 | 0.545873282 | low   | TARGET.40.PAPKWI |
| TARGET.40.0A4I42 | 0.961643836 | 1 | 3.621506191 | high  | TARGET.40.0A4I42 |
| TARGET.40.PATEEM | 2.632876712 | 1 | 1.405639929 | high  | TARGET.40.PATEEM |
| TARGET.40.PAUBIT | 2.819178082 | 0 | 0.35277244 | low   | TARGET.40.PAUBIT |
| TARGET.40.0A4I0W | 1.871232877 | 0 | 0.391739553 | low   | TARGET.40.0A4I0W |
| TARGET.40.0A4I3W | 1.17260274 | 1 | 1.790747411 | high  | TARGET.40.0A4I3W |
| TARGET.40.PASSLM | 5.109589041 | 1 | 1.114250168 | high  | TARGET.40.PASSLM |
| TARGET.40.PATAW\ | 4.663013699 | 0 | 0.25668101 | low   | TARGET.40.PATAW\ |
| TARGET.40.PAVCLP | 1.4 | 0 | 0.412967449 | low   | TARGET.40.PAVCLP |
| TARGET.40.NAASJI | 1.063013699 | 1 | 0.488103249 | low   | TARGET.40.NAASJI |
| TARGET.40.PAVEC | 1.304109589 | 0 | 0.184612911 | low   | TARGET.40.PAVEC |
| Target | Value | Correlation | Type |
|--------|-------|-------------|------|
| TARGET.40.0A4I5B 1.695890411 | 0 | 1.298829953 high |
| TARGET.40.PAPWW 5.295890411 | 1 | 1.856946121 high |
| TARGET.40.PAUXP2 1.147945205 | 1 | 0.807902333 high |
| TARGET.40.PANPUN 0.126027397 | 1 | 0.654474881 high |
| TARGET.40.PALKDF 8.991780822 | 0 | 0.951840767 high |
| TARGET.40.PANZH0 2.224657534 | 1 | 0.333253294 low |
| TARGET.40.PAMLK! 0.194520548 | 1 | 0.95774747 high |
| TARGET.40.NAASJK 2.071232877 | 1 | 0.555732606 low |
| TARGET.40.NAAEDI 0.750684932 | 1 | 2.385116775 high |
| TARGET.40.PAUTYF 1.443835616 | 1 | 1.742805011 high |
| TARGET.40.PASUUF 4.452054795 | 0 | 0.460305239 low |
| TARGET.40.0A4I40 0.643835616 | 1 | 0.641749782 low |
| TARGET.40.PARDAY 0.287671233 | 1 | 3.598854801 high |
| futime | fustat | riskScore | risk   |
|--------|--------|-----------|--------|
| 1.002739726 | 1     | 0.377895934 | low    |
| 1.583561644  | 1     | 4.872046206 | high   |
| 5.24109589   | 0     | 0.721059995 | high   |
| 4.336986301  | 0     | 0.48202024  | low    |
| 7.150684932  | 0     | 1.922694365 | high   |
| 4.328767123  | 0     | 0.313995867 | low    |
| 1.032876712  | 0     | 0.662114722 | high   |
| 0.17260274   | 1     | 1.107690583 | high   |
| 0.783561644  | 0     | 0.67084173  | high   |
| 1.326027397  | 1     | 1.784315131 | high   |
| 1.054794521  | 1     | 2.758579026 | high   |
| 5.81369863   | 0     | 0.216719604 | low    |
| 0.649315068  | 1     | 1.89097957  | high   |
| 4.78630137   | 0     | 2.781805901 | high   |
| 1.638356164  | 1     | 1.035356219 | high   |
| 8.356164384  | 0     | 0.837265853 | high   |
| 4.21369863   | 0     | 0.330867254 | low    |
| 0.115068493  | 1     | 1.056769019 | high   |
| 1.397260274  | 1     | 3.760075495 | high   |
| 0.153424658  | 1     | 1.018767749 | high   |
| 1.58630137   | 0     | 0.219512948 | low    |
| 5.123287671  | 0     | 0.312471124 | low    |
| 1.052054795  | 1     | 4.148320027 | high   |
| 1.402739726  | 1     | 1.220095913 | high   |
| 6.180821918  | 0     | 0.858452021 | high   |
| 5.052054795  | 0     | 0.457345369 | low    |
| 0.98630137   | 1     | 1.046199264 | high   |
| 8.479452055  | 0     | 0.411161115 | low    |
| 4.42739726   | 0     | 0.914022871 | high   |
| 10.8109589   | 0     | 1.681179065 | high   |
| 1.939726027  | 0     | 0.724223444 | high   |
| IncRNA       | mRNA   | cor   | pvalue         | IncRNA       | function                      |
|--------------|--------|-------|----------------|--------------|-------------------------------|
| RP11-69E11.4 | ADGRD1 | 0.454 | 5.52059E-06   | RP11-69E11.4 | ossification                  |
| RP11-69E11.4 | BMP8A  | 0.882 | 3.86934E-31   | RP11-69E11.4 | extracellular region          |
| RP11-69E11.4 | BMP8B  | 0.786 | 1.85071E-20   | RP11-69E11.4 | proteaceous extracellular     |
| RP11-69E11.4 | CA3    | 0.458 | 4.30594E-06   | RP11-69E11.4 | skeletal system development   |
| RP11-69E11.4 | CGREF1 | 0.498 | 4.26573E-07   | RP11-69E11.4 | cartilage development         |
| RP11-69E11.4 | CHMP4C | 0.607 | 1.43974E-10   | RP11-69E11.4 | cell junction                  |
| RP11-69E11.4 | CHST13 | 0.539 | 3.05304E-08   | RP11-69E11.4 | collagen trimer               |
| RP11-69E11.4 | COL11A2| 0.424 | 2.54689E-05   | RP11-69E11.4 | calcium ion binding           |
| RP11-69E11.4 | BMP8A  | 0.882 | 3.86934E-31   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | BMP8B  | 0.786 | 1.85071E-20   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | CA3    | 0.458 | 4.30594E-06   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | CHMP4C | 0.607 | 1.43974E-10   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | CHST13 | 0.539 | 3.05304E-08   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | COL11A2| 0.424 | 2.54689E-05   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | BMP8A  | 0.882 | 3.86934E-31   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | BMP8B  | 0.786 | 1.85071E-20   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | CA3    | 0.458 | 4.30594E-06   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | CHMP4C | 0.607 | 1.43974E-10   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | CHST13 | 0.539 | 3.05304E-08   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | COL11A2| 0.424 | 2.54689E-05   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | BMP8A  | 0.882 | 3.86934E-31   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | BMP8B  | 0.786 | 1.85071E-20   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | CA3    | 0.458 | 4.30594E-06   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | CHMP4C | 0.607 | 1.43974E-10   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | CHST13 | 0.539 | 3.05304E-08   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | COL11A2| 0.424 | 2.54689E-05   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | BMP8A  | 0.882 | 3.86934E-31   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | BMP8B  | 0.786 | 1.85071E-20   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | CA3    | 0.458 | 4.30594E-06   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | CHMP4C | 0.607 | 1.43974E-10   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | CHST13 | 0.539 | 3.05304E-08   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | COL11A2| 0.424 | 2.54689E-05   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | BMP8A  | 0.882 | 3.86934E-31   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | BMP8B  | 0.786 | 1.85071E-20   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | CA3    | 0.458 | 4.30594E-06   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | CHMP4C | 0.607 | 1.43974E-10   | RP11-69E11.4 | extracellular matrix          |
| RP11-69E11.4 | CHST13 | 0.539 | 3.05304E-08   | RP11-69E11.4 | extracellular matrix          |
| Gene      | Description                                      |
|-----------|--------------------------------------------------|
| EPCAM     | G-protein coupled receptor                       |
| GNG4      | Integral component of G-protein coupled receptor |
| GRAMD1    | Integral component of plasma membrane            |
| ISM1      | Integral component of plasma membrane            |
| LGR6      | Integral component of plasma membrane            |
| MEPE      | Cytosol                                          |
| MFAP4     | Cytoplasm                                        |
| MSC       | Nucleus                                          |
| OGN       | Protein digestion and collagen trimer            |
| PTN       | Protein digestion and collagen catabolic proc    |
| RASGRP2   | Protein digestion and PI3K-Akt signaling         |
| RHBDL2    | Protein digestion and PI3K-Akt signaling         |
| SLC8A3    | ADGRD1                                          |
| SYT12     | BMP8A                                           |
| TMCC2     | BMP8B                                           |
| TPD52     | VCAM1                                           |
| VCAM1     | ADGRD1                                          |
| C1QTNF3   | BMP8A                                           |
| CA3       | BMP8B                                           |
| CGREF1    | C1QTNF3                                         |
| CHAD      | C1QTNF3                                         |
| CHMP4C    | C1QTNF3                                         |
| CHST13    | C1QTNF3                                         |
| COL11A2   | C1QTNF3                                         |
| COL13A1   | C1QTNF3                                         |
| COL22A1   | C1QTNF3                                         |
| COL8A2    | C1QTNF3                                         |
| CORT      | C1QTNF3                                         |
| CPE       | C1QTNF3                                         |
| CPNE5     | C1QTNF3                                         |
| CYFIP2    | C1QTNF3                                         |
| DMP1      | C1QTNF3                                         |
| EFHD1     | C1QTNF3                                         |
| EPCAM     | C1QTNF3                                         |
| GNG4      | C1QTNF3                                         |
| GRAMD1    | C1QTNF3                                         |
| ISM1      | C1QTNF3                                         |
| KCNK17    | C1QTNF3                                         |
| LGR6      | C1QTNF3                                         |
| LYNX1     | C1QTNF3                                         |
| MEPE      | C1QTNF3                                         |
| MFAP4     | C1QTNF3                                         |
| MLC1      | C1QTNF3                                         |
| MSC       | C1QTNF3                                         |
| NPB       | C1QTNF3                                         |
| OGN       | C1QTNF3                                         |


| Gene          | Value  | P-value   |
|--------------|--------|-----------|
| SATB2-AS1 PHEX | 0.599  | 2.75211E-10 |
| SATB2-AS1 PTPRZ1 | 0.646  | 3.72399E-12 |
| SATB2-AS1 PYGM | 0.485  | 9.42125E-07 |
| SATB2-AS1 RASGRP2 | 0.462  | 3.55124E-06 |
| SATB2-AS1 RHBDL2 | 0.689  | 3.33546E-14 |
| SATB2-AS1 SEMA6D | 0.458  | 4.3731E-06  |
| SATB2-AS1 SLC13A5 | 0.445  | 8.93889E-06 |
| SATB2-AS1 SLC8A3 | 0.711  | 2.14899E-15 |
| SATB2-AS1 SYT12 | 0.57   | 3.12617E-09 |
| SATB2-AS1 TAC4 | 0.455  | 5.07506E-06 |
| SATB2-AS1 TMCC2 | 0.77   | 2.98786E-19 |
| SATB2-AS1 TPD52 | 0.516  | 1.38774E-07 |
| SATB2-AS1 TSPAN11 | 0.552  | 1.14287E-08 |
| SATB2-AS1 WIFI | 0.453  | 5.64743E-06 |
| Term           | Description                           |
|---------------|---------------------------------------|
| cytoplasm     |                                       |
| lumen         |                                       |
| cell adhesion |                                       |
| junction      |                                       |
| lumen         |                                       |
| cytoplasm     |                                       |
| plasma membrane|                                     |
| membrane      |                                       |
| lumen         |                                       |
| cytoplasm     |                                       |
| membrane      |                                       |
