

LINE-1 and Alu Promote Breast Tumor Progression Via Forming Chimeric Transcripts

Jian Li

Sun Yat-Sen Memorial Hospital <https://orcid.org/0000-0002-6091-9882>

Yin Zhang

Sun Yat-Sen Memorial Hospital

Yiting Zhou

Sun Yat-Sen Memorial Hospital

Jian-you Liao

Sun Yat-Sen Memorial Hospital

Yabin Guo (✉ guoyb9@sysu.edu.cn)

Sun Yat-Sen Memorial Hospital <https://orcid.org/0000-0001-8316-8527>

Research

Keywords: Alu, breast tumor, lncRNA, LINE-1, retrotransposon

Posted Date: March 24th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-326833/v1>

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

[Read Full License](#)

Abstract

Background LINE-1 (L1) and Alu were reported to regulate tumor development and progression by constituting chimeric transcripts with diverse sequences. The landscape of L1 or Alu chimeric transcripts in breast tumor has not been reported yet, so breast tumor L1 and Alu chimeric transcripts were investigated in this study by analyzing more than 50 billion RNA sequences of breast tumor tissues and adjacent normal tissues from TCGA database.

Results The expression of L1 and Alu in breast tumor tissues were significantly higher than that in the non-tumor tissues. High expression of L1 and Alu in breast tumor predicted poor prognosis. Further exploration demonstrated that L1 and Alu were extensively chimerically expressed with adjacent transcripts. 651 L1-mRNA chimeric transcripts and 1525 Alu-mRNA chimeric transcripts showed significantly different frequencies between non-tumor and tumor tissues. 1009 L1-lncRNA chimeric transcripts and 2575 Alu-lncRNA chimeric transcripts had significantly different frequencies between non-tumor and tumor tissues. Function cluster analysis demonstrated that these differently chimerically expressed genes were involved in multi facets of tumorigenesis, including metabolism, signal transduction, immune reaction, cell cycle and apoptosis, etc.

Conclusions Chimeric expression with different sequences might be an important mechanism for L1 and Alu to promote breast tumor progression.

Background

Transposons are DNA elements that can move within (or between) genomes. Transposons can be divided into DNA transposons and retrotransposons according to their transposition manners [1, 2]. Long Interspersed Element-1 (LINE-1, L1) and Alu are two of the most active retrotransposons in the primates [3, 4]. L1 is the only active autonomous mobile DNA in human, constituting ~ 17% of the human genome. The intact L1s are about 6 kb. Most L1s are inactivated due to truncations, rearrangements or mutations, and only about 100 copies remain active in typical human genome. The typical Alu elements are ~ 300 bp. There are more than one million copies of Alu elements, accounting for ~ 11% of the human genome. Alu is a non autonomous transposon, whose transposition relies on the proteins expressed by other autonomous transposons, such as L1. The activations of L1 and Alu not only cause genome instability, but also induce abnormal gene expression [5], which are two of the most basic characteristics for tumors [6, 7]. L1 and Alu are hypomethylated and actively transcribed in most tumors [8–10]. Moreover, their activations are often accompanied by abnormal expression of tumor related genes [11, 12].

Breast cancer is among the top five malignant tumors and the second leading cause for cancer related death in women [13]. L1 and Alu are hypomethylated in breast tumor, and can regulate the expressions of adjacent genes, including forming chimeric transcripts with them to promote tumorigenesis [14–16]. Yet, there has not been a comprehensive report on the profile of L1 or Alu chimeric expressions in breast cancer. Through analyzing breast cancer transcriptome data from the Cancer Genome Atlas (TCGA)

database, we found that L1 and Alu were high expressed in breast tumor tissue and the high expression of L1 and Alu in breast tumor tissue predicted poor prognosis; further exploration demonstrated that L1 and Alu formed extensively chimeric transcripts in breast tumor.

Methods

Data resources

The raw transcriptome sequencing data (fastq format) of 606 breast tumors and 59 adjacent normal tissues, with more than 50 billion pairs of sequences were obtained from TCGA database (Table S1). The mRNA sequence dataset and the sequences of L1 and Alu were downloaded from NCBI database; lncRNA dataset were downloaded from Noncode Database (<http://www.noncode.org/index.php>). The mRNA expression dataset of breast cancers was obtained from the Dashboard Stddata database of broad research institute (<https://conference.broadinstitute.org/display/gdac/dashboard-stddata>). The lncRNA expression dataset of breast cancers were obtained from Tanric Database of TCGA [31].

Data analyses

The RNA sequencing data were mapped against the index of L1 and Alu using Bowtie2 software and the mapping outputs were filtered. For the single-end matches, the sequences of the other end were extracted from the original sequence files, and then mapped against the index of mRNAs and lncRNAs to obtain the chimeric mRNA and lncRNA datasets, and multiple statistical analyses were performed. All computer programs were written in Perl language. The violin plot and survival analysis were performed with Graphpad software. Cluster analysis were performed with consensuspathdb [32] and plots were generated with RStudio.

Results

Overview of L1 and Alu expressions in breast tumor

We mapped the transcriptome sequence reads against the index of L1 or Alu full-length sequences with Bowtie2 software (Fig. 1). The statistical result shows that the expressions of L1 and Alu in women of white race are significantly higher than those in black or Asian race. Moreover, the diversity of expression in white race is also much higher than those in black or Asian race (Fig. 1A), which is a very interesting observation. To our knowledge, there was no similar report so far, and it is worthy of further study in the future. Figure 1B shows the coverage of RNA sequencing on both retrotransposons. Obviously, the coverage of 3'-end is significantly higher than that of 5'-end in L1, which is due to the large proportion of 5'-truncation L1 copies in the genome. The highest coverage of Alu is at 20–50 bp for the similar reason. This result also shows that the data quality of the RNA sequencing meets the requirement for analyzing these two retrotransposons.

High expression of L1 and Alu in breast tumor predicted poor prognosis

Next, we compared the expressions of L1 and Alu between tumor tissues and adjacent normal tissues based on the mapping results. Results demonstrate that the expressions of both L1 and Alu in cancer tissues are significantly higher than those in adjacent tissues (Fig. 2A, Fig. 2B, Student's T-Test). Compared with differences of means or medians between cancer and normal tissues, the extraordinarily large variances in cancer tissues are even more remarkable, i.e., the expression levels of these two retrotransposons are similar in different normal tissues, whereas, the expression levels vary drastically in different cancer tissues, highlighting the complex regulations of transposon expression in tumorigenesis. Patients were further divided into high or low group according to L1 or (and) Alu expression in breast tumor tissue, and overall prognostic analysis demonstrates that patients with high L1 or (and) Alu expression have poorer prognosis than those with low expression (Fig. 2C-2F).

Widespread chimeric expression of mRNAs with L1 or Alu in breast tumor

To uncover how they promote breast tumor progression, we detected the correlation of L1 and Alu with transcriptome expression in breast tumor firstly. It shows that correlation between them is in median or low level with the highest absolute value of correlation coefficient less than 0.7 (Figure S1), which implies that L1 and Alu affect breast tumor progression not by acting on sole gene. As L1 and Alu often affect their adjacent gene expression, we scanned their chimeric expression with adjacent genes. We identified overall 101,846 mRNA sequences chimerically expressed with L1, 651 of which demonstrated different expression between in tumor tissues and in adjacent tissues ($P < 0.05$, chi-square test) (Fig. 3A, Fig. 4A, Table S2). For example, the chimeric expressions of *ESR1*, *SPC25*, *MET*, *FAM111B*, *CASP8*, etc. with L1 in tumor tissues are significantly higher than those in adjacent tissues. Whereas, the chimeric expressions of *LPP*, *CFLAR*, *ABCC9*, *ABCA9*, *LEP*, *BAG1*, *GBP5*, *OXCT1*, etc. with L1 in tumor tissues are significantly lower than those in adjacent tissues. Similarly, we identified 129,462 chimeric transcripts of mRNAs and Alu, 1,525 of which were differently expressed between in tumor tissues and in adjacent tissues ($P < 0.05$, chi-square test) (Fig. 3B, Fig. 4B, Table S3). The chimeric expression of *CD24*, *GPI*, *BRCA1*, *B3GALNT2*, *ERBB2*, *TP53*, *FLT1*, *CWF19L1*, *UNSA1*, etc. with Alu are significantly higher in tumor tissues than in adjacent tissues. And the chimeric expression *MCAM*, *PIGR*, *CLDN11*, *CCL28*, *ADIPOQ*, *KIAA1324*, *SLC2A3*, *ECI2*, *EP300*, etc. with Alu are significantly lower in tumor tissues than in adjacent tissues. Interestingly, there are more chimeric transcripts that are downregulated than those are upregulated (Fig. 3E, Fig. 3F), indicating that a lot of L1/Alu chimeric transcripts may play roles in keeping the homeostasis under physical condition and their downregulations may contribute for tumorigenesis.

The chimeric expression of lncRNAs with L1 or Alu in breast tumor

Besides the coding genes, long non-coding RNAs (lncRNAs) are increasingly recognized to play important roles in evolution and disease progression [17]. Therefore, we analyzed the chimeric expressions of lncRNAs with L1 or Alu. 16,817 lncRNA sequences are chimerically expressed with L1, 1,009 of which are differently expressed between in tumor tissues and adjacent tissues ($P < 0.05$, chi square test) (Fig. 3C, Fig. 4C, Table S4). The chimeric expressions of LOC286437 (NONHSAT138089.2), KCNQ10T1 (NONHSAT017523.2), LINC00393 (NONHSAT034309.2), LINC01410 (NONHSAT131526.2), LINC01087

(NONHSAT074491.2), BHLHE40-AS1 (NONHSAT246212.1), etc. with L1 in tumor tissues are significantly higher than that in adjacent tissues. And the chimeric expressions of LCAT_00100281 (NONHSAT174269.1), LCAT_00113317 (NONHSAT176719.1), CECR7 (NONHSAT191948.1), LINC01088 (NONHSAT097104.29), etc. with L1 in tumor tissues are significantly lower than that in adjacent tissues. Similarly, we detected 46,457 lncRNA genes chimerically expressed with Alu, 2,575 of which were differently expressed between in tumor and adjacent tissues ($P < 0.05$, chi-square test) (Fig. 3D, Fig. 4D, Table S5). For example, the chimeric expressions of LINC01296 (NONHSAT035480.2), SNHG16 (NONHSAT056037.2), LINC00514 (NONHSAT147902.2), LINC01441 (NONHSAT080425.2), etc. with Alu are higher in tumor tissues than in adjacent tissues. And the chimeric expressions of LINC00894 (NONHSAT138981.2), LINC00702 (NONHSAT156543.1), LINC02202 (NONHSAT104813.2), LCAT_00037989 (NONHSAT158544.1), NORAD (NONHSAT079548.2), XIST (NONHSAT137541.2), etc. with Alu are significantly lower in tumor tissues than in adjacent tissues.

L1 and Alu differently chimerically expressed genes involved in multi facets of tumorigenesis

Tumorigenesis is a multi-step process and many elements are involved in this progression. We found a large number of genes chimerically expressed with L1 and Alu in breast tumor in this study. So, we further performed functional cluster analysis of them and found that the functions of these chimeric genes were related to various aspects of tumorigenesis, including pathway related to metabolism, metabolism of protein, signal transduction, gene transcription, immune system, cell cycle and apoptosis etc. (Fig. 5A-5G). Results imply that L1 and Alu might promote breast cancer progression via influencing multi facets of tumorigenesis.

Comparison of the chimeric expression levels and the overall expression levels of related genes

Theoretically, there are two potential reasons for the change of chimeric expressions with L1 and Alu in cancer cells: 1) L1 or Alu is adjacent to these genes and the expression changes of these genes cause the changes of the chimeric expressions; 2) the changes of L1 and Alu activity cause the changes of chimeric expression. If 1) is the major cause, the chimeric expressions should display a strong correlation with the overall expressions of related genes, otherwise, the correlation should be weak. We downloaded gene expression data from previous transcriptome studies of breast cancer, and compared it with the chimeric expression data in current study. Figure 6 shows that there is no correlation between the mRNA or lncRNA chimeric expressions with L1 or Alu and total expressions of these genes, which indicates that the expression difference of these chimeric RNAs is not due to the "passenger effect" of the overall expressions, rather, they should be regulated by an independent mechanism.

Discussion

By retrieving the RNA sequencing data of hundreds of breast tumors and dozens of adjacent normal tissues, we found that the expression of L1 or (and) Alu in tumor was significantly higher than that in normal tissue, and higher expression of L1 or (and) Alu in breast tumor predicted poorer prognosis, which was consistent with previous reports on the hypomethylation and upregulation of L1 in breast cancers

[14, 18, 19]. In addition to the upregulation of average expression level, the expression in tumor tissues showed extraordinarily high diversity, highlighting the heterogeneity of the activation and regulation of retrotransposons in tumors.

To explore the mechanism of L1 and Alu promote breast cancer progression, we investigated the expression of L1 and Alu chimeric transcripts in breast tumor. We found that there were significant differences in the L1 or Alu chimeric expressions with a large number of protein coding genes and long noncoding RNA genes between in tumor and normal tissues. In addition to the change of expression levels, the formation of chimeric RNAs with L1 or Alu may also endow genes with new functions, and influence tumor progression. Miglio U et al. reported that the chimeric expression of L1 and *MET* in certain breast tumor greatly enhanced the invasion of tumor cells [16]. We also found that the chimeric expressions of *MET* with L1 in tumor were significantly higher than those in adjacent normal tissue. *BRCA1* gene mutation is one of the most notorious poor prediction factors for breast cancer. It was reported that high frequency of Alu insertions was the main cause for *BRCA1* mutation and abnormal insertion of Alu not only greatly increased the incidence of breast cancer, but also induced chemotherapy resistance [20, 21]. *TP53* is one of the most important tumor suppressor gene. It was reported that *TP53* was mutated and inactivated due to Alu insertion in most tumors. *TP53* mutation caused by Alu high-frequency insertion was not only related to the occurrence, but also related to the formation of secondary malignant tumor and chemotherapy resistant phenotype in breast cancer [22–24]. Through large sample, high-throughput study, we have confirmed that previous reports about frequent insertion of Alu leading to *BRCA1* or *TP53* are common in breast cancer. These results suggest that more attention should be paid to the role of retrotransposons in breast cancer.

In addition, many lncRNAs demonstrating significantly different chimeric expressions with L1 or Alu between cancer and adjacent normal tissue have also been proved to be involved in the incidence and progression of breast cancer. For example, Cai C. et al. found that SNHG16 (NONHSAT056037.2) promoted the migration of breast cancer cells by competitive binding miR-98 and E2F5 [25]. Jiang M. et al. reported that LINC01296 (NONHSAT035480.2) was highly expressed in breast cancer, and its over expression was related to poor prognosis [26]. Tan BS et al. discovered that the expression of NORAD (NONHSAT079548.2) was low in breast tumor, and its knockdown could promote breast tumor metastasis [27]. Xing F. et al. reported that XIST (NONHSAT137541.2) was low expressed in breast tumors, and its knockout promoted breast tumor brain metastasis [28]. All these lncRNAs showed significantly different chimeric expressions with Alu between tumor tissue and normal tissue in our study. According to the report of Feng W. et al., KCNQ10T1 (NONHSAT017523.2) was highly expressed in breast cancer and promoted tumor growth [29]. DeVaux RS. et al. reported that the expression of BHLHE40-AS1 (NONHSAT246212.1) was increased in breast cancer, which could modulate interleukin-6/STAT3 activity to promote breast tumor progression [30]. We also found that chimeric expressions of these two lncRNAs with L1 increased significantly in breast tumor tissues.

In addition to the above examples, there are also many well-known genes playing important roles in the genesis and development of breast tumor, such as *ESR1*, *LPP*, *CFLAR*, *LEP*, *BAG1*, *GBP5*, *OXCT1*, *CD24*,

GPI, B3GALNT2, ERBB2, FLT1, CWF19I1, NUSAP1, MCAM, KIAA1324, SLC2A3, ECI2, etc., have significantly different chimeric expression with L1 or Alu between in cancer and normal tissues. Moreover, by comparing L1 and Alu chimeric expressions with the total expressions of related genes in breast cancer, we found that there was no correlation between them, indicating that these chimeric expressions were not the "passenger effect" of the overall expression changes. Instead, the chimeric expressions are more likely to be regulated by independent mechanisms, which may also endow new functions to these genes.

Although the expression of L1 and Alu are significantly upregulated in breast cancer, there are more downregulated chimeric mRNAs than upregulated ones (Fig. 3E, 3F), which indicates that many tumor suppressors may function as chimeric transcripts with L1 or Alu under physical condition. Therefore, the changes of L1 and Alu that contribute for tumorigenesis include not only activations, but also suppressions.

Furthermore, the functions of many genes identified in this study that are differentially expressed with L1 or Alu between cancer and normal tissue are still elusive (especially a large number of lncRNA genes), suggesting the potential role of these genes in the genesis of breast cancer, which is worth further investigations.

Conclusions

Through analyzing a large number of RNA sequencing data, we found that L1 and Alu were high expressed in breast tumor and their high expression predicted poor prognosis for patients. We also found that gene chimeric expression with L1 or Alu in breast cancer is a very common phenomenon, but not just some special cases. The chimeric expression of these genes may be one of the reasons for promoting breast cancer formation and progression. Function cluster analysis of L1 and Alu chimeric genes demonstrated that functions of these chimeric transcripts were involved in various aspects of tumorigenesis including metabolism, immune reaction, signaling pathway, cell cycle and apoptosis etc. Generally, there are not enough studies on the relationship between human endogenous transposons and cancer yet, and the role of transposons in cancer may be largely underestimated. We believe that in-depth researches in this field will provide new ideas and targets for cancer diagnosis and treatment.

Abbreviations

LINE-1 or L1: Long Interspersed Element-1, TCGA: the Cancer Genome Atlas, lncRNAs: long non-coding RNAs

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

They are available in the article.

Competing interests

The authors declare that they have no conflicts of interest with the contents of this article.

Funding

This work was supported by grants from Guangdong Natural Science Foundation (2018A030313819 to Y. G.), National Natural Science Foundation of China (81872295 to Y. G.), Guangdong Science and Technology Department (2017B030314026), Guangzhou Science and Technology Program (201904020008).

Author contributions

Y. G. designed the project. Yin Z. and J.Y. L. collected the original sequence data. J. L., Yin Z, Yiting Z. and Y. G. analyzed the data. J. L., Yin Z. and Y. G. wrote the manuscript.

References

1. Cordaux R, Batzer MA. The impact of retrotransposons on human genome evolution. *Nat Rev Genet.* 2009;10(10):691-703.
2. Huang CR, et al. Mobile interspersed repeats are major structural variants in the human genome. *Cell.* 2010;141(7):1171-82.
3. Beck CR, et al. LINE-1 retrotransposition activity in human genomes. *Cell.* 2010;141(7):1159-70.
4. Lander ES, et al. Initial sequencing and analysis of the human genome. *Nature.* 2001;409(6822):860-921.
5. Elbarbary RA, Lucas BA, Maquat LE. Retrotransposons as regulators of gene expression. *Science.* 2016;351(6274):aac7247. doi:10.1126/science.aac7247.
6. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011;144(5):646-74.
7. Campbell BB, et al. Comprehensive Analysis of Hypermutation in Human Cancer. *Cell.* 2017;171(5):1042-56 e10.
8. Daskalos A, et al. Hypomethylation of retrotransposable elements correlates with genomic instability in non-small cell lung cancer. *Int J Cancer.* 2009;124(1):81-7.
9. Howard G, et al. Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. *Oncogene.* 2008;27(3):404-8.

10. Lin CH, et al. Genome-wide hypomethylation in hepatocellular carcinogenesis. *Cancer Res.* 2001;61(10):4238-43.
11. Tubio JMC, et al. Mobile DNA in cancer. Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes. *Science.* 2014;345(6196):1251343. doi: 10.1126/science.1251343.
12. Jang HS, et al. Transposable elements drive widespread expression of oncogenes in human cancers. *Nat Genet.* 2019;51(4):611-7.
13. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. *CA Cancer J Clin.* 2017;67:7–30.
14. Hon GC, et al. Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. *Genome Res.* 2012;22(2):246-58.
15. Park SY, et al. Alu and LINE-1 Hypomethylation Is Associated with HER2 Enriched Subtype of Breast Cancer. *PLoS ONE.* 2014;9(6): e100429. doi:10.1371/journal.pone.0100429.
16. Miglio U, et al. The expression of LINE1-MET chimeric transcript identifies a subgroup of aggressive breast cancers. *Int J Cancer.* 2018;143(11):2838-48.
17. Anastasiadou E, Jacob LS, Slack FJ. Non-coding RNA networks in cancer. *Nat Rev Cancer.* 2018;18(1):5-18.
18. Deroo LA, et al. Global DNA methylation and one-carbon metabolism gene polymorphisms and the risk of breast cancer in the Sister Study. *Carcinogenesis.* 2014;35(2):333-8.
19. Patnala R, et al. Inhibition of LINE-1 retrotransposon-encoded reverse transcriptase modulates the expression of cell differentiation genes in breast cancer cells. *Breast Cancer Res Treat.* 2014;143(2):239-53.
20. Petrij-Bosch A, et al. BRCA1 genomic deletions are major founder mutations in Dutch breast cancer patients. *Nat Genet.* 1997;17(3):341-5.
21. Wang Y, et al. BRCA1 intronic Alu elements drive gene rearrangements and PARP inhibitor resistance. *Nat Commun.* 2019;10(1):5661. doi: 10.1038/s41467-019-13530-6.
22. Niederacher D, et al. Patterns of allelic loss on chromosome 17 in sporadic breast carcinomas detected by fluorescent-labeled microsatellite analysis. *Genes Chromosomes Cancer.* 1997;18(3):181-92.
23. Cui F, Sirotnin MV, Zhurkin VB. Impact of Alu repeats on the evolution of human p53 binding sites. *Biol Direct.* 2011;6:2. doi:10.1186/1745-6150-6-2.
24. Kamat N, et al. High incidence of microsatellite instability and loss of heterozygosity in three loci in breast cancer patients receiving chemotherapy: a prospective study. *BMC Cancer.* 2012;12:373. doi:10.1186/1471-2407-12-373.
25. Cai C, et al. SNHG16 contributes to breast cancer cell migration by competitively binding miR-98 with E2F5. *Biochem Biophys Res Commun.* 2017;485(2):272-8.
26. Jiang M, et al. Overexpression of long noncoding RNA LINC01296 indicates an unfavorable prognosis and promotes tumorigenesis in breast cancer. *Gene.* 2018;675:217-24.

27. Tan BS, et al. LncRNA NORAD is repressed by the YAP pathway and suppresses lung and breast cancer metastasis by sequestering S100P. *Oncogene*. 2019;38(28):5612-26.
28. Xing F, et al. Loss of XIST in Breast Cancer Activates MSN-c-Met and Reprograms Microglia via Exosomal miRNA to Promote Brain Metastasis. *Cancer Res*. 2018;78(15):4316-30.
29. Feng W, et al. The Dysregulated Expression of KCNQ10T1 and Its Interaction with Downstream Factors miR-145/CCNE2 in Breast Cancer Cells. *Cell Physiol Biochem*. 2018;49(2):432-46.
30. DeVaux RS, et al. Long noncoding RNA BHLHE40-AS1 promotes early breast cancer progression through modulating IL-6/STAT3 signaling. *J Cell Biochem*. 2020; 121(7):3465-78.
31. Li J, et al. TANRIC: An interactive open platform to explore the function of lncRNAs in cancer. *Cancer Res*. 2015;75(18): 3728-37.
32. Kamburov A, Stelzl U, Lehrach H, et al. The ConsensusPathDB interaction database: 2013 update. *Nucleic Acids Res*. 2013;41:D793-D800.

Figures

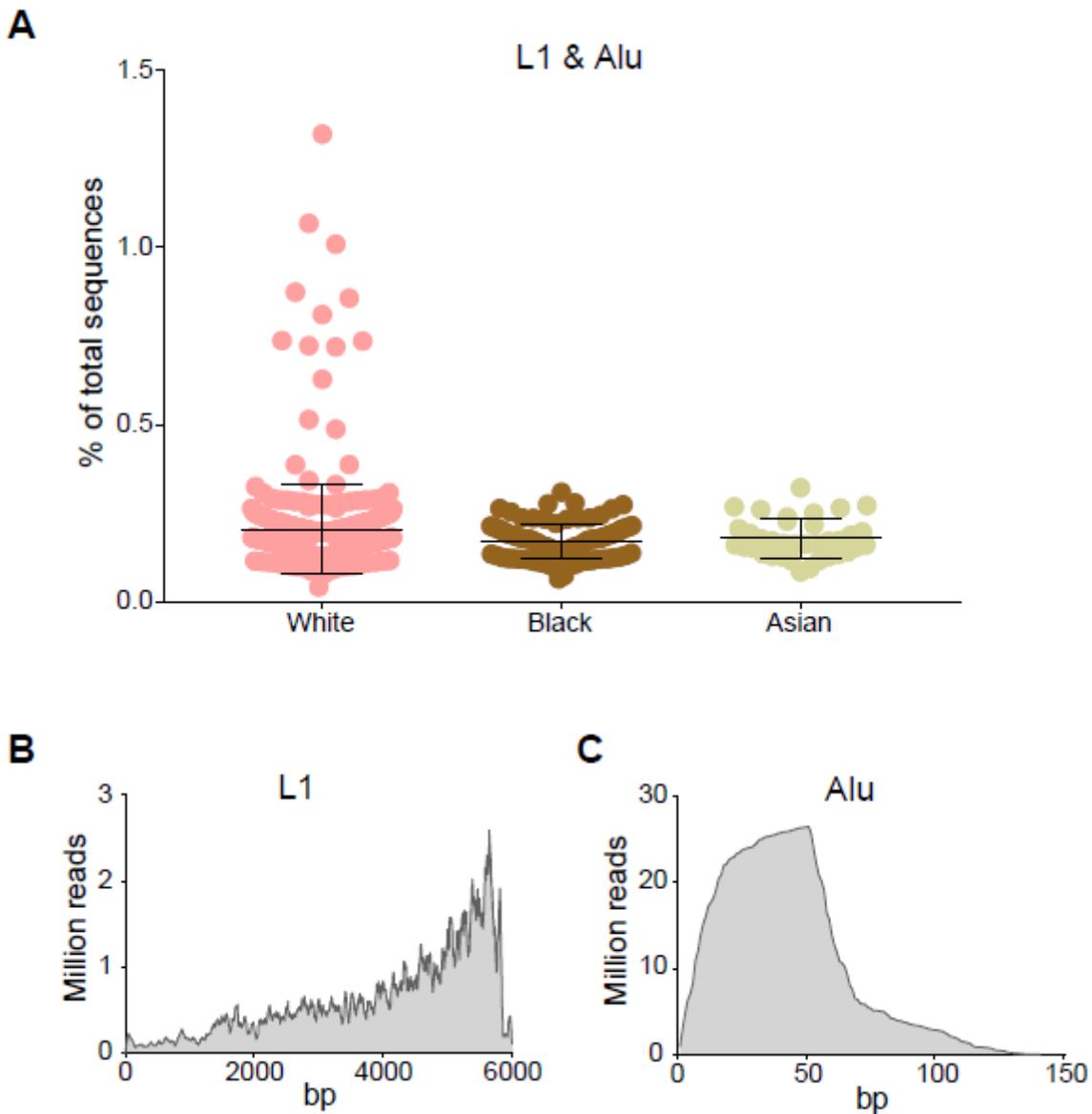


Figure 1

The overall expression of L1 and Alu. A. The distribution of L1 and Alu expression levels in White, Black and East Asian populations; B-C, The sequence read numbers at the full length of L1 (B) and Alu (C) in the mapping result under single nucleotide resolution.

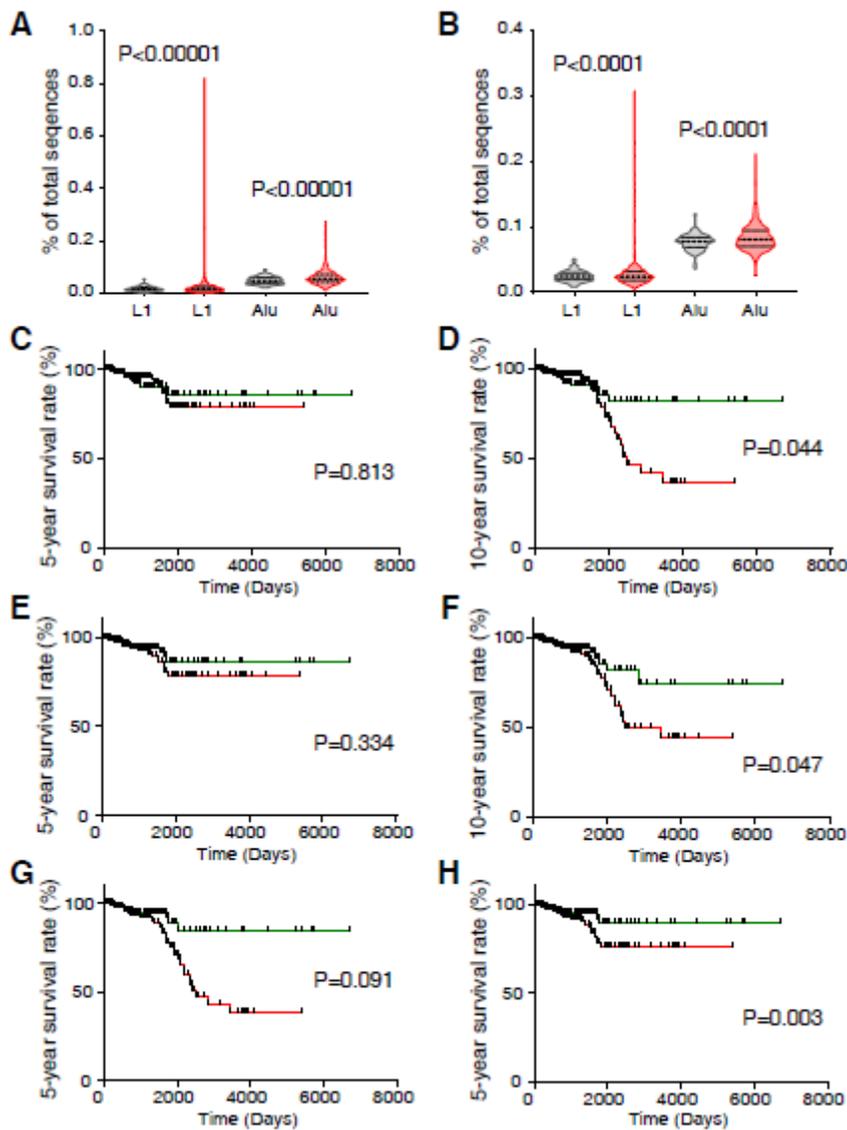


Figure 2

High expression of L1 and Alu in breast tumor predicted poor prognosis. A-B, The different expression levels of L1 and Alu in breast cancer and adjacent tissues under paired end (A) or single end (B) match model, the red dots indicate tumor tissue expression, and the black dots indicate normal tissues expression; C, Comparing of 5-year survival rate between Alu high (red) and low (green) breast cancer patients; D, Comparing of 10-year survival rate between Alu high (red) and low (green) breast cancer patients; E, Comparing of 5-year survival rate between L1 high (red) and low (green) breast cancer patients; F, Comparing of 10-year survival rate between L1 high (red) and low (green) breast cancer patients; G, Comparing of 5-year survival rate between Alu&L1 high (red) and low (green) breast cancer patients; H, Comparing of 10-year survival rate between Alu&L1 high (red) and low (green) breast cancer patients.

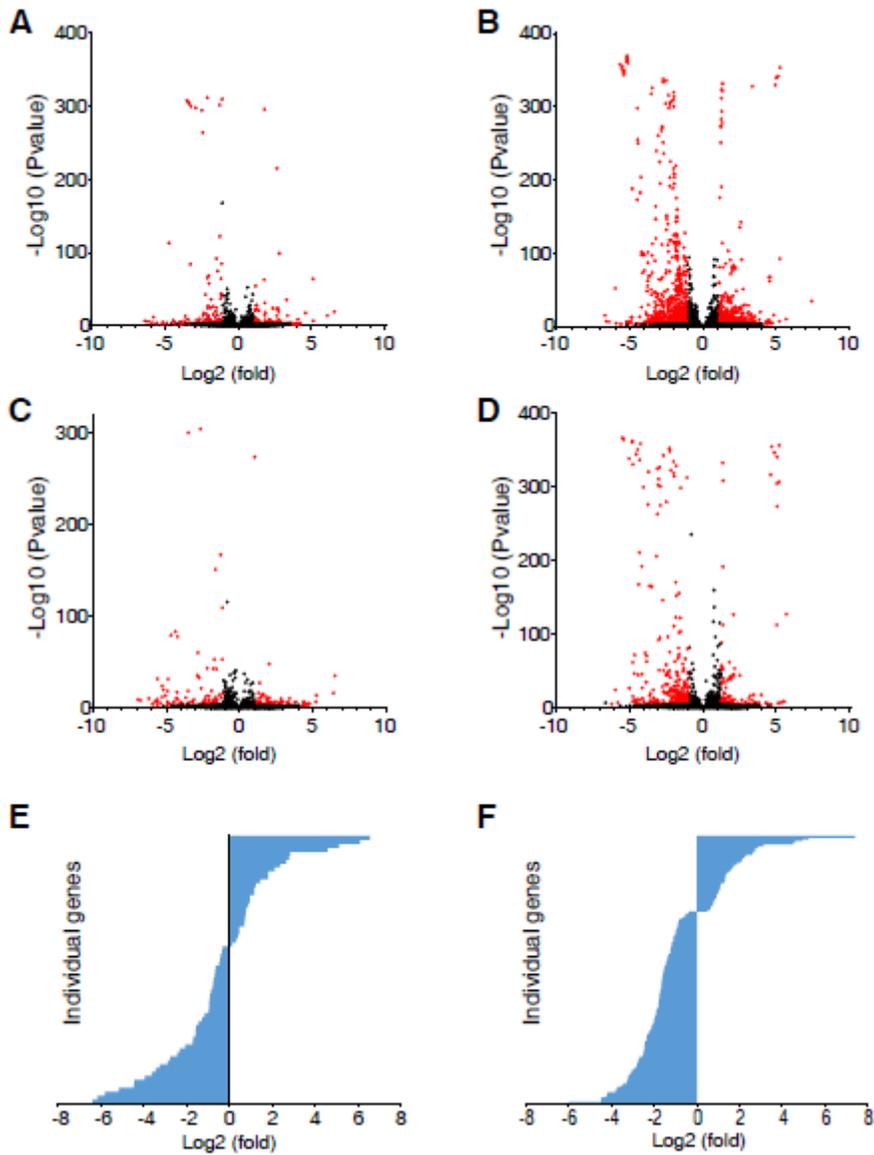


Figure 3

The expression difference of gene chimerically expressed with L1 or Alu in breast cancer and adjacent tissues. The reds dot on the left side of the volcano plots indicate that expressions decrease by > 2 folds ($P < 0.05$), the red dots on the right side indicate that expressions increase by > 2 folds ($P < 0.05$), and the black dots indicate changes are < 2 folds or not significant. A, the expressions of L1-mRNA chimeric transcripts; B, the expressions of Alu-mRNA chimeric transcripts; C, the expressions of L1-lncRNA chimeric transcripts; D, the expressions of Alu-lncRNA chimeric transcripts; E-F, The expression level distribution of chimeric mRNA with L1 (E) or Alu (F) in cancer and adjacent tissues ($P < 0.05$).

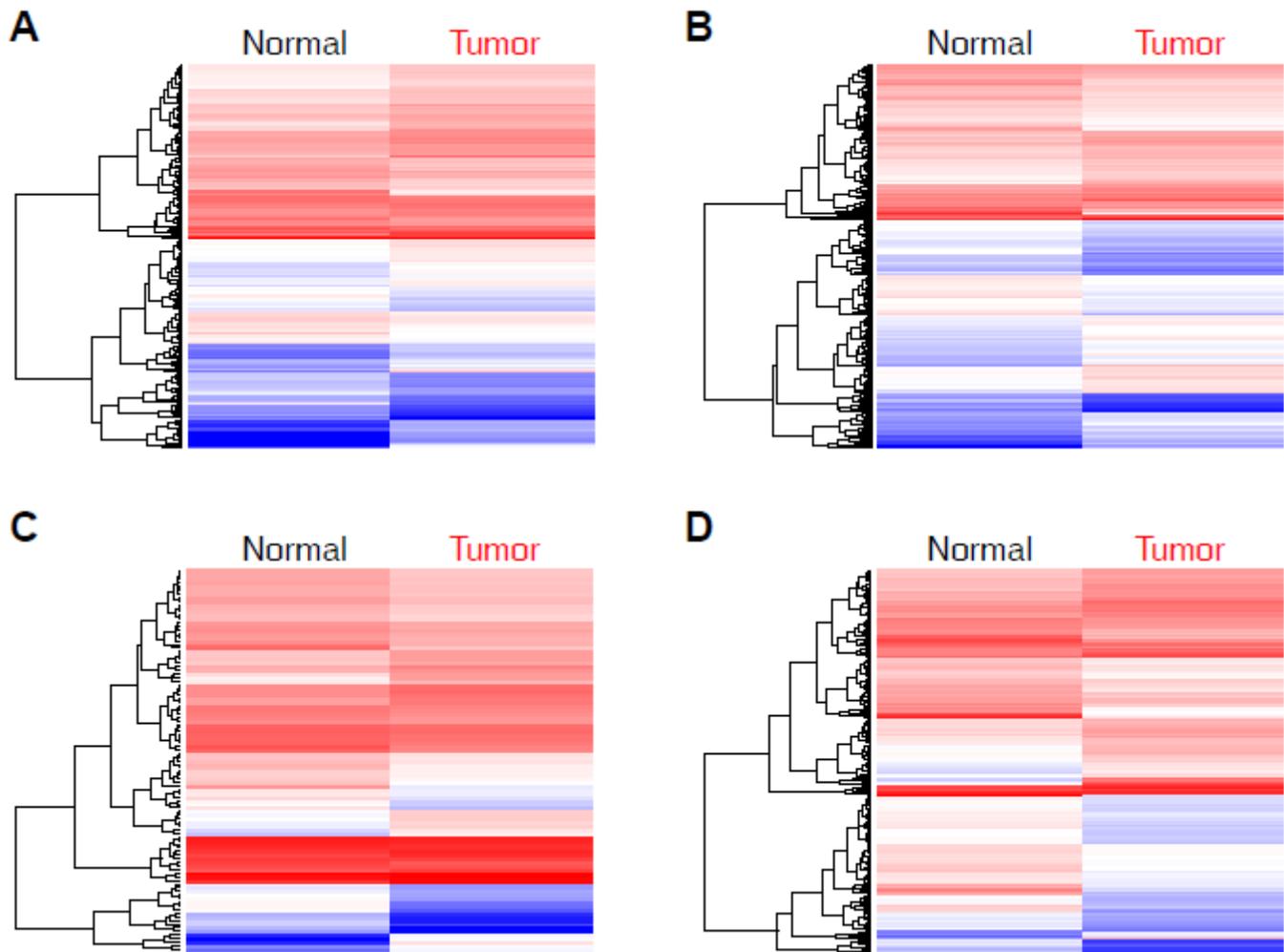


Figure 4

Cluster analysis of L1 or Alu chimeric expressions in breast tumor tissues ($P < 0.05$). A, L1-mRNA chimeric transcripts; B, Alu-mRNA chimeric transcripts; C, L1-lncRNA chimeric transcripts; D, Alu-lncRNA chimeric transcripts;

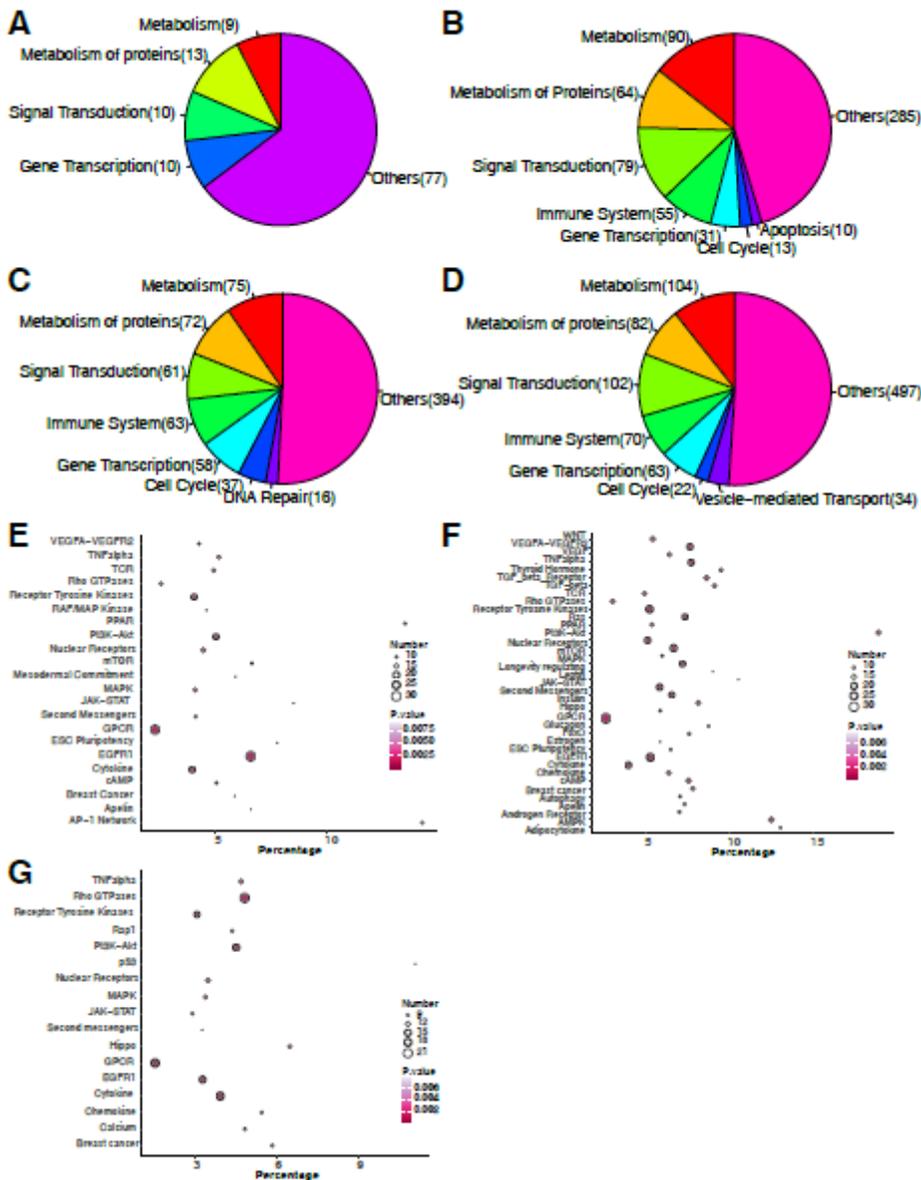


Figure 5

L1 and Alu differently chimerically expressed genes involved in multi facets of tumorigenesis. A, Major function cluster of L1-mRNA chimeric transcripts up-expressed in breast tumor tissue; B, Major function cluster of L1-mRNA chimeric transcripts down-expressed in breast tumor tissue; C, Major function cluster of Alu-mRNA chimeric transcripts up-expressed in breast tumor tissue; D, Major function cluster of Alu-mRNA chimeric transcripts down-expressed in breast tumor tissue; E, Major involved signal pathway in down-expressed L1-mRNA chimeric transcripts in breast tumor tissue; F, Major involved signal pathway in down-expressed Alu-mRNA chimeric transcripts in breast tumor tissue; G, Major involved signal pathway in up-expressed Alu-mRNA chimeric transcripts in breast tumor tissue.

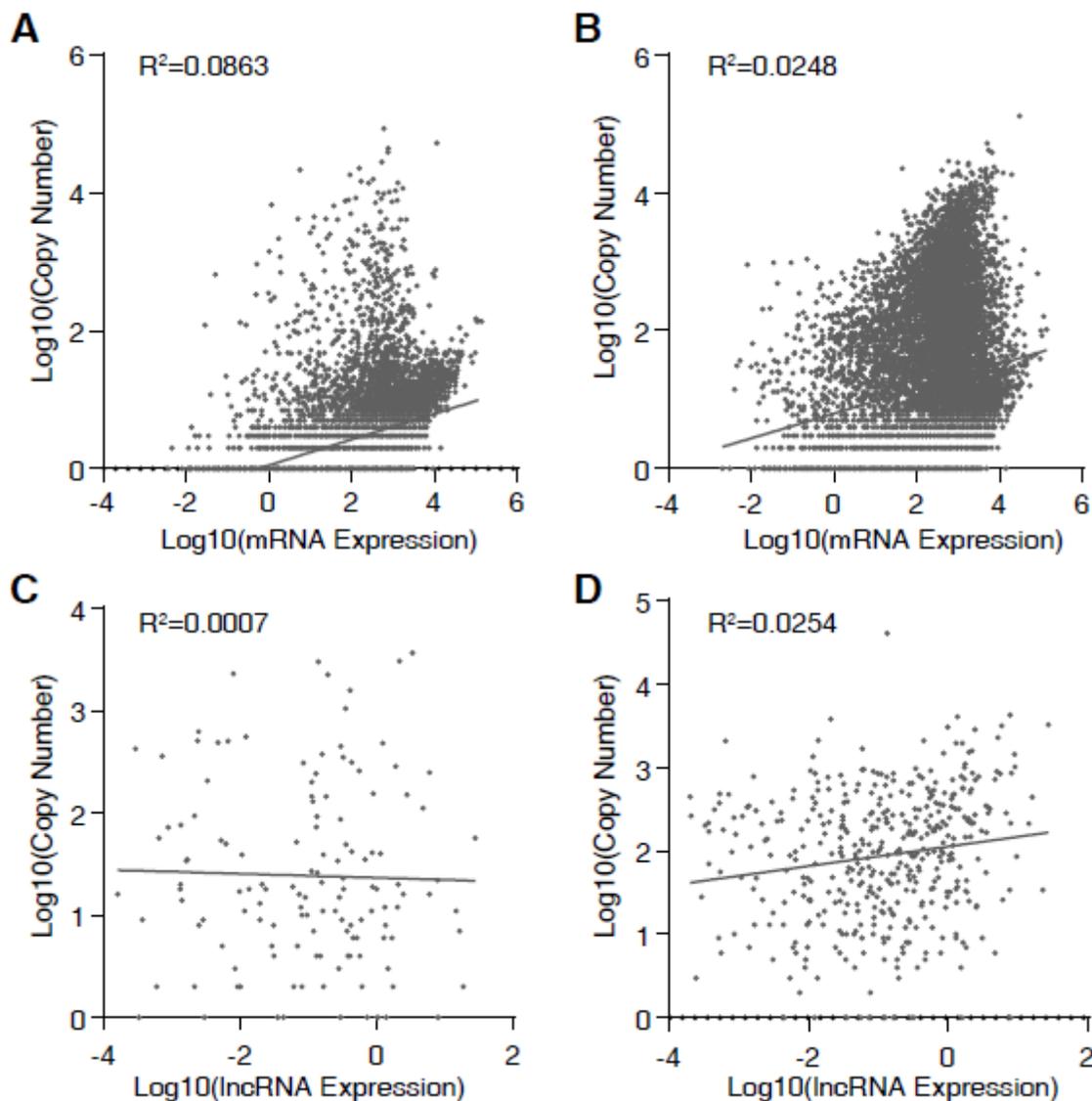


Figure 6

The relationship between chimeric expression level of the mRNA and lncRNA gene with L1 or Alu and the overall expression level of these genes in breast tumor tissues. A-D, scatter plots for the logarithm of L1 chimerically expressed mRNA (A), Alu chimerically expressed mRNA (B), L1 chimerically expressed lncRNA (C), Alu chimerically expressed lncRNA(D) copy numbers and the logarithm of the relative expressions of related genes in breast tumor tissue.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigureS1.pdf](#)
- [SupTable1TCGAID.xls](#)
- [SupTable2L1gene.xls](#)

- [SupTable3Alugene.xls](#)
- [SupTable4L1IncRNA.xls](#)
- [SupTable5AluIncRNA.xls](#)