doi 10.18699/bgrs2024-1.1-27

Identification of small RNAs stably associated with chromosome in the metaphase of 3T3 cell line

Ye B.[1, 2†], Meng Y.[3†], Zhang L.[2†], Xinhui Li[2]*, Hua Li[1, 2]**

[1] *Jiangsu Province Engineering Research Center of Molecular Target Therapy and Companion Diagnostics in Oncology, Suzhou Vocational Health College, Suzhou, China*

[2] *State Key Laboratory of Systems Medicine for Cancer, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China*

[3] *Shanghai Key Laboratory of Embryo Original Diseases, The International Peace Maternity and Child Health Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China*

†These authors contributed equally to this work.

** xhli@sjtu.edu.cn; ** kaikaixinxin@sjtu.edu.cn*

Key words: small RNAs; mouse 3T3 Cells; mitotic chromosomes; RNA sequencing

Motivation and Aim: During the different stages of mitosis, various types of noncoding RNAs (ncRNAs) that are associated with chromatin must maintain strict proportions and structures. These ncRNAs form an RNA infrastructure that is involved not only in processing other RNAs, such as mRNAs, tRNAs, and rRNAs, but also in gene regulation by targeting mRNAs and chromatin. Previous studies have revealed the lncRNA composition of mitotic chromosomes, suggesting the functional significance of these RNAs [1, 2]. However, the profile of chromosome-associated small RNAs in mammals has not yet been systematically investigated. In this study, we performed deep sequencing of chromosome-associated small RNAs in the metaphase of mouse 3T3 cell lines under three different conditions (buffer A, low salt, and high salt), as previous research indicates that different concentrations of salt solutions can extract various types of small RNAs associated with the metaphase chromosomes while maintaining the stability of the core chromosome structure. We aim to identify small RNAs tightly bound to the metaphase chromosomes across the three conditions, which indicates the functional significance of these RNAs during mitosis and possibly in disease mechanisms [3].

Methods and materials: Mouse 3T3 cells were cultured in DMEM supplemented with 10 % heat inactivated fetal calf serum and penicillin-streptomycin. $5X10^7$ Cells were harvested when they achieved 70-80 % confluence and were treated with colcemid (100 ng/ml) for 12 hours to synchronized at the metaphase.

The mitotic chromosomes were isolated as previously described in Meng et al 2016 [1]. In addition to the low salt (LS) and high salt (HS) samples, a third set of chromosomes (buffer A (BA)sample) was treated with Buffer A (15 mM Tris–HCl (pH 7.5), 80 mM KCl, 2 mM EDTA, 2 mM spermine, 5 mM spermidine, 0.1 mM PMSF, and 0.05 % digitonin) on ice for 3 minutes. The mixture was centrifuged at 1750g for 6 minutes and then resuspended in the PA buffer before loading onto the sucrose gradient. [interphase nuclear miRNA extraction] Total RNA was extracted from each sample with TRIzol according to the manufacturer's instruction (Thermo Fisher Scientific). The small RNA library was prepared with TruSeq® Small RNA Library Preparation Kits following the Illumina's protocol. Deep sequencing of the library was performed on Illumina HiSeq platform using single end 1x50 as the sequencing mode.

Raw sequencing data were first processed with FASTX-Toolkit to remove sequencing adaptors and low-quality reads (http://hannonlab.cshl.edu/fastx_toolkit/). Reads longer than 30 bp accounted for a very small part of all reads due to the size selection of gel purification in the library preparation and were removed from downstream analysis. Bowtie was then used to map the clean reads generated from above steps to the mouse genome (version mm10) with default parameters except -v -k and -m all set to 1. Only uniquely mapped reads were used for downstream analysis.

Reads were grouped into the same cluster if the distance between their 5' ends were less than 15 bp, which is the minimal length of annotated mouse miRNAs. In this study, the highest-frequency read in each read cluster was used to defined the position of the read cluster.

For each condition (BA, LS or HS), read clusters were ranked by their RPM and a cut-off of RPM was set such that the total number of reads in clusters with RPM above the cut-off accounted for >90 % of all reads in the condition. The read clusters above the cut-off were then compared with the mouse genomic features (version mm10). The microRNA (miRNA) annotation data were retrieved from miRBase (miRBase v21; http://www.mirbase.org/) [4]. The Piwi-interacting RNA (piRNA) and transcription start site (TSS) annotation data were retrieved from UCSC genome browser mouse mm10. Read clusters were identified as miRNAs or piRNAs if they fell into the annotated genomic regions of miRNA or piRNAs, respectively.

miREval2.0 (http://mimirna.centenary.org.au/mireval/) were employed to predict novel miRNAs from read clusters above the RPM cut-off that does not match any annotated small RNA [5]. The express sequence tag (EST) annotation used to support the predicted novel miRNAs were also retrieved from UCSC genome browser mouse mm10.

Results: Here we only used the uniquely mapped reads for downstream data analysis. We grouped reads into the same cluster if the distance between their 5' ends were less than 15 bp (Methods and Materials). By doing this, we obtained 122940, 77275, 75730, 96720 read clusters for BA, LS, HS and NC, respectively. Here we selected read clusters with large read numbers and made comparisons between the three conditions. For each condition, we first set a cut-off for the read numbers such that the total number of reads in clusters above the cut-off accounted for at least 90% of all reads in this condition. Furthermore, to make quantitative comparisons between conditions, we used reads per million (RPM) to normalize the read numbers of each read cluster (Methods and Materials). In the following paragraphs, we used RPM to measure the amount of small RNAs associated with chromosomes.

For those read clusters above the cut-offs, we compared the RPM distributions between the three conditions. We found that, in BA and LS, the average RPM was very close to each other, while the average RPM in HS was significantly lower than in the other two conditions. This result shows that small RNAs of high concentration are more likely to be washed away from chromosomes in high-salt environment.

The read clusters above the cut-offs are highly consistent among the three conditions. We picked the top-ranked (by RPM) 148 read clusters in each condition and found 128 read clusters in common, demonstrating a statistically significant overlap among them [6].

The three sets have a common set of 132 read clusters whose association with metaphase chromosomes are both high and stable (i.e., not significantly influenced by salt concentration), suggesting their potential functional importance in mitosis. By contrast, most of the 132 clusters have a relatively large variation in RPM between condition

HS/LS/BA and condition NC (i.e., the interphase cell nucleus of the mouse 3T3 cell line). Specifically, some of the 132 clusters have very low RPMs in condition NC, indicating their special roles in metaphase.

Based on the experimental protocols and small RNA annotations, we classified the majority of 132 read clusters mentioned above into three categories: microRNA (miRNA), Piwi-interacting RNA and TSS-associated (TSSa) RNA. We first compared the 132 read clusters with annotated mouse miRNAs and found that 72 of 132 could perfectly match the miRNA genomic locations. Thus, the 72 read clusters were classified into annotated mouse miRNAs. As we expected, in all three conditions, the total reads of the 72 miRNAs accounted for more than half of the total reads of the 132 clusters, but the average read numbers showed no significant difference between the 72 miRNAs and the rest.

Piwi-interacting RNA (piRNA) constitutes the largest category of non-coding small RNAs in mammals. We compared the remaining 60 read clusters (out of 132) with annotated mouse piRNAs and found 6 read clusters to be piRNAs. For the rest (54) of the 132 read clusters, there were 6 of them located in the vicinity of transcription start sites (TSSs), making it highly possible that the 6 read clusters belonged to the category of TSSa RNA. The rest (48) of the 132 read clusters could not be mapped to any known small RNA; since they all had very high abundance (comparable with the 72 miRNAs), they were likely to be unknown chromosome-associated small RNAs.

Conclusion: In this study, we performed deep sequencing of chromosome-associated small RNAs in the metaphase of mouse 3T3 cell lines under three different conditions: buffer A, low salt, and high salt. The small RNAs identified with high abundance were consistent across the three conditions, indicating the functional significance of these stably attached RNAs. Most of these small RNAs can be mapped to known features such as miRNAs and piRNAs, while a considerable amount (36.3 %) could not be annotated. Additionally, some unannotated small RNAs are not present in interphase 3T3 cell lines, suggesting their possible mitosis-specific functions. Analyzing these different types of small RNAs helps us better understand mitosis, leading to new insights into the mechanisms of cell division and possible role in pathogenesis.

Funding: This work was supported by grants from the National Key R&D Program of China (2020YFA0908100), the National Natural Science Foundation of China (32370572) and the K.C. Wong Education Foundation (H.K.).

References

1. Meng Y., Yi X., Li X., Hu C., Wang J., Bai L., Czajkowsky D. M., Shao Z. The non-coding RNA composition of the mitotic chromosome by 5'-tag sequencing. *Nucleic Acids Res*. 2016;44(10):4934-4946. doi 10.1093/nar/gkw195
2. Zhang L., Hu C., Xu Z., Li H., Ye B., Li X., Czajkowsky D.M., Shao Z. Quantitative catalogue of mammalian mitotic chromosome-associated RNAs. *Sci Data*. 2024;11(1):43. doi 10.1038/s41597-023-02884-8
3. Orlov Y.L., Chen W.L., Sekacheva M.I., Cai G., Li H. Editorial: high-throughput sequencing-based investigation of chronic disease Markers and mechanisms. *Front Genet.* 2022;13:922206. doi 10.3389/fgene.2022.922206
4. Kozomara A., Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* 2014;42(D1):D68-D73. doi 10.1093/nar/gkt1181
5. Gao D., Middleton R., Rasko J.E., Ritchie W. miREval 2.0: a web tool for simple microRNA prediction in genome sequences. *Bioinformatics.* 2013;29(24):3225-3226. doi 10.1093/bioinformatics/btt545
6. Li H., Tong P., Gallegos J., Dimmer E., Cai G., Molldrem J.J., Liang S. PAND: A distribution to identify functional linkage from networks with preferential attachment property. *PLoS One.* 2015;10(7): e0127968. doi 10.1371/journal.pone.0127968