

Title:

Comprehensive survey of LINE-1 transcriptional activity in human cell lines, healthy somatic tissue, and tumors

Authors:

Fabio CP Navarro 1,2; Jacob Hoops 1,2; Lauren Bellfy 4; Eliza Cerveira 4; Qihui Zhu 4; Chengsheng Zhang 4; Charles Lee 4,5; Mark B. Gerstein 1,2,3

Affiliations:

1 Program in Computational Biology and Bioinformatics, 2 Department of Molecular Biophysics and Biochemistry, and 3 Department of Computer Science, Yale University, Bass 432, 266 Whitney Avenue, New Haven, CT; 4 The Jackson Laboratory for Genomic Medicine, Farmington, CT; 5 Department of Life Sciences, Ewha Womans University, Seoul, Korea

Abstract

Long interspersed nuclear element 1 (LINE-1) is a main source of variation in humans and other mammals. However, LINE-1 activity is difficult to study because of its highly repetitive nature and the effects of pervasive transcription. We developed and validated a method to gauge LINE-1 transcriptional activity accurately by removing the effects of pervasive transcription. We evaluated and validated our method in human cell-lines using the ENCODE dataset and found that most L1 transcription signal derives, as expected, from the cytoplasm. Our method also allowed us to perform comprehensive, uniform, and unbiased measurements of LINE-1 activity across healthy somatic cells and tumor cells. Previously, LINE-1 had been shown to be active in human germline and tumor cells but not in healthy somatic tissue, with the exception of some activity in the human brain. In contrast, we found that LINE-1 activity was limited in the central nervous system, but present in some normal somatic cells and tumor cells. Interestingly, the amount of LINE-1 activity was associated with the amount of cell turnover and, in tumor cells, with the amount of genomic instability. Our results suggest a mechanism in which LINE-1 activity

gives rise to insertions and deletions overlapping LINE-1 target sites, potentially contributing to the mutagenic landscape in tumors.