Identification of intragenic retrocopies in chimaric transcripts in humans

Rafael Luiz Vieira Mercuri, Helena Beatriz da Conceição, Pedro Alexandre Favoretto Galante

Instituto de Ensino e Pesquisa, Hospital Sírio-Libanês

Abstract

The evolution of a species occurs by the emergence of new genes. These can be originated through LINE1 (L1) mediated gene duplication, a phenomenon often encountered in the human genome. Although sometimes these genes are not functional, recent studies have shown that some retrocopies are transcribed and functional. Among all the retrocopies genes encoding a genome, those inserted into the intronic regions of (other) coding genes deserve attention. Because they are in a gene region, they may influence the transcription and post-transcriptional processing of the "host gene." It is currently known that there are retrocopies present in the human genome that are located in introns of coding genes. However, very little is known about the influence and contribution of these retrocopies in relation to their host genes. The aim of this work was to elaborate a systematic study of the retrocopies inserted in introns and exons of human coding genes (intragenic). The RCPedia (https://www.bioinfo.mochsl.org.br/rcpedia/) was used as a database to identify intragenic retrocopies in humans. Subsequently a comparison was made with GTEx (https://gtexportal.org/) to verify the expression profile in 53 human tissues of host genes and their intragenic retrocopies. First, we found 2499 intragenic retrocopies (990 in the same transcription strand and 1509 in the opposite strand to their host genes) and of these, 65% (1630 retrocopies) had their expression confirmed by GTEx. Testis presented the highest number of expressed retrocopies and bladder the lowest. Interestingly, , some of expressed retrocopies are located into genes involved in pathological pathways as Ras Homolog Family G (RHOG). Thus, our results bring important knowledge and can contribute to a better understanding of the origin of new genes and genetic novelties.

Funding: FAPESP