
**RELACIÓN ENTRE MARCAS EPIGENÉTICAS Y ACTIVIDAD
TRANSCRIPCIONAL EN ELEMENTOS TRANSPONIBLES****MATÍAS EZEKIEL VÉJAR REYES
INGENIERO CIVIL EN BIOINFORMÁTICA****RESUMEN**

Los elementos transponibles (TE) son segmentos de ADN que tienen la capacidad de moverse de una región a otra, en un genoma. Éstos se transponen, teniendo efectos desconocidos en la expresión de los genes. El proyecto ENCODE ha mostrado que sobre el 70% del genoma puede transcribirse. También se sabe que cerca de la mitad del genoma humano corresponde a TEs, por lo que se deduce que cerca de un 35% del genoma corresponde únicamente a TEs que están siendo transcritos. La célula, intentando defenderse de la propagación de los TEs, pone en marcha mecanismos de regulación a nivel epigenético que actúan a nivel transcripcional. Las regulaciones epigenéticas son modificaciones químicas de las bases nitrogenadas o de las histonas, que no cambian la secuencia del ADN, y que permiten regular la expresión de genes y TEs. Se esperaría, para ciertas marcas epigenéticas, una relación proporcional entre el número de marcas en un TE predicho como activo y su expresión. Aunque se desconoce la relación que existe entre la actividad transcripcional de los TEs con la regulación epigenética, se han asociado algunas marcas con la actividad del ADN. En este trabajo, se propone buscar esta correlación entre la expresión diferencial de TEs activos y las mencionadas tres marcas epigenéticas. Como la mayoría de los TEs son fósiles inactivos y fragmentos, es difícil identificar el locus del elemento del cual fue transcrito debido a que existen múltiples parálogos en el genoma; Entretanto, hay 2 herramientas computacionales, TEcandidates y SQUIRE, que permiten predecir la expresión de TEs transcripcionalmente activos en un genoma sin perder su ubicación. Se usarán esas herramientas para seleccionar un conjunto de TEs activos de diferentes superfamilias. Tras predecir los TEs activos se observó como esperado que gran parte de los TEs predichos como activos no presentan marcas epigenéticas. Entretanto, ningún TE sobre-expresado presentó marcas epigenéticas tampoco. La única de las 3 marcas en presentar unión con TEs fue

H3K27ac. Sin embargo, todos los TEs con la marca H3K27ac, no fueron activos, lo que sugiere que esta marca podría tener efectos represivos en TEs. Además, esta marca pareciera tener una preferencia por las subclases LTR y no-LTR, en retrotransposones. Considerando los resultados, se puede establecer que hay indicios de una relación entre la marca H3K27ac y retrotransposones predichos como activos en cáncer de próstata en humano. Además, ninguna de las otras marcas presentó ninguna dependencia con TEs. Se sugiere para futuros estudio, para tener un mejor entendimiento de la relación entre TEs y marcas epigenéticas, la utilización de un mayor repertorio de marcas. Además, sería deseable sets de datos de RNA-Seq en estudios sobre diferentes condiciones y especies.

ABSTRACT

Transposable elements (TEs) are DNA segments that can move from one region to another in a genome. They transpose, having unknown effects on gene expression. The ENCODE project has shown that about 70% of the genome can be transcribed. It is also known that about half of the human genome corresponds to TEs, so it is deduced that about 35% of the genome corresponds only to TEs that are being transcribed. The cell, trying to defend itself against the spread of TEs, sets in motion regulatory mechanisms at the epigenetic level. Epigenetic regulations are a set of chemical modifications of nitrogenous bases or histones, which do not change the DNA sequence, and which allow the expression of genes, including TEs, to be regulated. Thus, one might expect to find, for certain epigenetic marks, a proportional relationship between the number of marks in a predicted active TE and its expression. Although the relationship between the transcriptional activity of TEs with epigenetic regulation is unknown, some marks have been associated with DNA activity. In this work, we propose to search for this relationship between differential expression of active TEs and the three epigenetic marks. As most TEs are inactive fossils and fragments, it is difficult to identify the locus of the element from which it was transcribed because there are multiple paralogs in the genome; meanwhile, there are 2 computational tools, TEcandidates and SQUIRE, that allow predicting the expression of transcriptionally active TEs in a genome without losing their location. We will use these tools to select a set of active TEs from different superfamilies. After predicting transcriptionally active TEs, it was observed that as expected, most of the TEs that were predicted to be active did not exhibit epigenetic marks. However, none of the overexpressed TEs showed epigenetic marks either. The only one of the 3 marks to show binding to TEs was H3K27ac. However, all TEs with the H3K27ac mark were not active, suggesting that this epigenetic mark might have repressive effects on TEs. In addition, this mark appeared to prefer LTR and non-LTR subclasses in retrotransposons. Considering the results obtained, it can be established that there are signs of a relationship between H3K27ac marks and retrotransposons predicted to be active

in human prostate cancer. In addition, none of the other marks showed any dependence on TEs. It is suggested for future studies, in order to have a better understanding of the relationship between TEs and epigenetic marks, the use of a larger repertoire of marks. In addition, additional RNA-Seq dataset on different conditions and organisms would be desirable.