Functional analysis of repeat regions in the eukaryotic genomes

Functional analysis of repetitive DNA derived from transposable elements in the human genome

Lu Zeng

A thesis submitted for the degree of Master of Philosophy

Discipline of Genetics

School of Molecular and Biomedical Science

The University of Adelaide

July 2013

Table of Contents

Abstract

Nearly half of the human genome is made up of transposable elements (TEs). With the rapid progress of sequencing technologies, we are now much better able to systematically analyze these TEs. We have used multiple types of omics data, including the genomic sequences, epigenetic data and transcriptomic data, to investigate the potential functions of TEs across the entire human genome. Comparative analysis revealed that a large proportion of potentially functional transposable elements were located in introns, and they were mainly associated with gene repression. Functional classification from GO enrichment showed that different functions were enriched in protein coding regions containing TEs compared to non-protein coding regions. For example, protein coding genes with Alus in non-coding regions are enriched with respect to intracellular membrane-bounded organelles, while protein coding genes with Alus in coding regions are more enriched at intracellular non-membrane-bounded organelles. Significantly, transcriptome data showed that the genes with TEs had lower expression levels compared with genes without TEs, revealing a novel aspect of the impact of TEs on the human genome. In addition, genome wide analysis of repeats with regulatory elements showed that MIR and L2 repeats were more probable to be active regulators while L1 repeats were less probable to be regulators. In conclusion, the role of TEs is significant across the genome. Repeats reduce or repress the expression of related gene, either through the proximal promoter, 5'UTR or 3'UTR or perhaps as components of lincRNA exons.

Declaration

This work contains no material which has been accepted for the award of any other degree or diploma in any university of other tertiary institution to Lu Zeng, and to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright holder(s) of these works contained within this thesis (as listed below) resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue, and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signed…… Date…….

Acknowledgements

I would like to express my sincere gratitude to the following people: Dr. Dave Adelson, gave me the opportunity and support to join the Master by Research program in the University of Adelaide. I am so lucky to have such a great supervisor. Dave, the knowledge that I learned from you is not just valuable for my Master, but will support me for my whole academic career.

Dr. Chaochun Wei, my supervisor in SHANGHAI JIAOTONG University, introduced me to a completely new field of bioinformatics. Over the year I stayed in Adelaide, he was always in contact with Dave and me to co-advise my research. Furthermore, he gave me this opportunity to do the Master in Adelaide.

Dan Kortschak, my co-supervisor in the University of Adelaide, helped me do research and always provide valuable suggestions.

Joy Raison, helped me generate nice figures. Sim Lim Lin, helped me sort out the tough problems I encountered in my research, and kept encouraging me. Zhipeng Qu, made both my life and research easier and fluent. Reuben Buckley, helped me practice English and providing me valuable advices. All other members of the Adelson lab, past and present, made it such a supportive and enjoyable place to work.

I would like to also specially thank SHANGHAI JIAOTONG University and The University of Adelaide, who provided me this opportunity to come here for my research.