

MEETING ABSTRACT

Open Access

# Comprehensive characterization of an individual's LINE-1 insertion profile using next generation sequencing data

Alexander Kemper<sup>1</sup>, Musa Hindi<sup>2\*</sup>, Kenneth S Ramos<sup>1</sup>, Theodore S Kalbfleisch<sup>1</sup>

From 11th Annual UT-ORNL-KBRIN Bioinformatics Summit 2012  
Louisville, KY, USA. 30 March - 1 April 2012

## Background

The Long Interspersed Nuclear Element-1 (LINE-1) is an active retrotransposon capable of affecting cellular phenotype in terms of repressing or stimulating the expression of neighboring genes, and even modifying genetic structure via retrotransposition. New LINE-1 insertion events are occurring constantly in somatic cells. They also occur in germline cells resulting in an estimated 1 new LINE-1 element for every 30 live births. As a result of this activity, in each individual, there are estimated to be several thousand LINE-1 elements, with on the order of 1/3 of these elements differing in genotype, homozygous present, homozygous absent, or heterozygous between any two individuals. Given its potential impact on cellular phenotype and function, it is essential to fully characterize an individual's LINE-1 insertion profile in terms of number of insertion sites, and the sequence and genotype of each insertion.

## Results

By using a next generation sequencing protocol that sequences the ends of genomic DNA fragments 5000 bases in length, we are developing a methodology that can be used to both sequence an individual's genome, and comprehensively characterize that individual's LINE-1 insertion profile. We have developed a workflow that will both map fragments to the reference genome for standard variant analysis, and isolate those sequence reads that provide sequence data for the multitude of LINE-1 insertions in the individual's genome. Once isolated, we are able to assemble the reads and build contigs

that allow us to measure the genotype, and predict the phenotypic activity of each LINE-1 element in an individual's genome. Our method is price competitive with the sequencing of smaller fragments, and provides much more comprehensive information.

## Conclusion

We are developing a method that will allow researchers and clinicians to comprehensively characterize the LINE-1 retrotransposon insertion profile in individuals. Current Next Generation Sequence mapping algorithms will neither identify non-reference LINE-1s, nor will they provide LINE-1 genotypes resulting from an individual's whole genome sequence analysis. Since LINE-1 elements can greatly impact cellular function and human health, it is essential that we incorporate algorithms such as this to characterize an individual's LINE-1 insertion profile as we endeavor use genetic information in the treatment of patients.

## Author details

<sup>1</sup>Department of Biochemistry and Molecular Biology, University of Louisville, Louisville, KY 40292, USA. <sup>2</sup>Department of Computer Engineering and Computer Science, University of Louisville, Louisville, KY 40292, USA.

Published: 31 July 2012

doi:10.1186/1471-2105-13-S12-A2

**Cite this article as:** Kemper et al.: Comprehensive characterization of an individual's LINE-1 insertion profile using next generation sequencing data. *BMC Bioinformatics* 2012 **13**(Suppl 12):A2.

\* Correspondence: mmhind01@louisville.edu

<sup>2</sup>Department of Computer Engineering and Computer Science, University of Louisville, Louisville, KY 40292, USA

Full list of author information is available at the end of the article