ALIGNMENT-FREE RAPID SEQUENCE CENSUS QUANTIFICATION (KALLISTO)

Tech ID: 25009 / UC Case 2015-156-0

BRIEF DESCRIPTION

Sequence census experiments utilize next-generation sequence data to estimate the relative abundance of target sequences. Since the samples are often short DNA fragments, they must first be assigned to the correct transcripts and genes that produced them, and this alignment or mapping step currently takes up the majority of computing power and time in most expression analyses.

To avoid this costly process, UC Berkeley researchers have developed a software program (kallisto) for quantifying abundances of transcripts from RNA-Seq data, or more generally of target sequences using high-throughput sequencing reads based on pseudo-alignment for rapidly determining the compatibility of reads with targets, without the need for alignment. Pseudo-alignment of reads preserves the key information needed for quantification.

SUGGESTED USES

» RNA-Seq analysis and quantification

» Sequence census experiments for copy number variation analysis of DNA sequencing

» High-throughput sequencing

ADVANTAGES

» Hundreds of times faster than standard tools (can quantify 30 million human reads in less than 3 minutes)

» Increased accuracy, especially compared to programs that shred reads into k-mers

» Tractable to use the bootstrap to determine uncertainty on estimates and includes infrastructure for managing the large amount of data associated with bootstrapped samples

CONTACT

Terri Sale
terri.sale@berkeley.edu
tel: 510-643-4219.

INVENTORS

» Bray, Nicolas L.

» Pachter, Lior S.

» Pimentel, Harold J.

OTHER INFORMATION

KEYWORDS

RNA-Seq; Next-generation sequencing; High-throughput; sequence census

CATEGORIZED AS

» **Biotechnology**

   » Bioinformatics

   » Genomics

» **Medical**

   » Research Tools

   » Software

» **Research Tools**

   » Bioinformatics

   » Nucleic Acids/DNA/RNA

RELATED CASES

2015-156-0, 2015-121-0