

MIT Open Access Articles

BIC-seq: a fast algorithm for detection of copy number alterations based on high-throughput sequencing data

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Xi et al. "BIC-seq: a fast algorithm for detection of copy number alterations based on high-throughput sequencing data". Genome Biology 2010 11(Suppl 1):010.

As Published: <http://dx.doi.org/10.1186/gb-2010-11-S1-010>

Publisher: BioMed Central Ltd.

Persistent URL: <http://hdl.handle.net/1721.1/69934>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of use: Creative Commons Attribution



SELECTED ORAL PRESENTATION

Open Access

BIC-seq: a fast algorithm for detection of copy number alterations based on high-throughput sequencing data

Ruibin Xi¹, Joe Luquette¹, Angela Hadjipanayis², Tae-Min Kim¹, Peter J Park^{1,3,4*}

From Beyond the Genome: The true gene count, human evolution and disease genomics
Boston, MA, USA. 11-13 October 2010

DNA copy number alterations (CNA), which are amplifications and deletions of certain regions in the genome, play an important role in the pathogenesis of cancer and have been shown to be associated with other diseases such as autism, schizophrenia and obesity. Next-generation sequencing technologies provide an opportunity to identify CNA regions with unprecedented accuracy. We developed a CNA detection algorithm based on single-end whole-genome sequencing data for samples with matched controls. This algorithm, called BIC-seq, can accurately and efficiently identify the CNAs via minimizing the Bayesian information criterion (BIC). We applied BIC-seq on a glioblastoma multiforme (GBM) tumor genome from the Cancer Genome Atlas (TCGA) project and identified hundreds of CNVs, some were as small as 10 bp. We compared these CNAs with those detected using the array Comparative Genomic Hybridization (CGH) platforms and found that about one third were 'missed' by the array-CGH platforms, most of which were CNAs less than 10 kb. We selected 17 of the CNAs not detected by the array-based platforms for validation, ranging from 110 bp to 14 kb, and found that 15 of them are true CNAs. We further extended BIC-seq to the multi-sample case to identify recurrent CNAs in across multiple tumor genomes.

Author details

¹Center for Biomedical Informatics, Harvard Medical School, 10 Shattuck St, Boston, Massachusetts 02115, USA. ²Harvard Medical School, 77 Louis Pasteur Avenue, Boston, Massachusetts 02115, USA. ³Department of Medicine, Brigham and Women's Hospital, 77 Louis Pasteur Avenue, Boston,

¹Center for Biomedical Informatics, Harvard Medical School, 10 Shattuck St, Boston, Massachusetts 02115, USA

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

Massachusetts 02115, USA. ⁴Harvard-MIT Health Sciences and Technology Informatics Program at Children's Hospital, 300 Longwood Ave., Boston, Massachusetts 02115, USA.

Published: 11 October 2010

doi:10.1186/gb-2010-11-S1-O10

Cite this article as: Xi et al.: BIC-seq: a fast algorithm for detection of copy number alterations based on high-throughput sequencing data. *Genome Biology* 2010 **11**(Suppl 1):O10.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

