GBSBC

Genetics, Bioinformatics, & Systems Biology Colloquium

presents

Benjamin Langmead, PHD

Associate Professor of Computer Science

Johns Hopkins University

THURSDAY
NOVEMBER 16TH

12PM

LEICHTAG
AUDITORIUM

ZOOM
JOIN

PAN-GENOMIC ADVANCES FOR FIGHTING REFERENCE BIAS

Sequencing data analysis often begins with aligning reads to a reference genome, where the reference takes the form of a linear string of bases. But linearity leads to reference bias, a tendency to miss or misreport alignments containing non-reference alleles, which can confound downstream statistical and biological results. This is a major concern in human genomics; we don't want to live in a world where diagnostics and therapeutics are differentially effective depending whether and where our genetic variants happen to match the reference.

Fortunately, computer science and bioinformatics are meeting the moment. We can now index and align sequencing reads to references that include many population variants. I will present some of the major and insights that have shaped this journey from the early days of efficient genome indexing -- especially the Burrows-Wheeler Transform -- continuing through recent methods for indexing graph-shaped references and references that include many genomes. I will emphasize recent results that show how to optimize simple and complex pan-genome representations for effective avoidance of reference bias. Finally, I will outline promising methods for the bias, including new ideas for how to measure bias, new proposals in compressed indexing, and new workflows that integrate genotype imputation to improve reference bias.

Much of this work is collaborative with Travis Gagie, Christina Boucher, Alan Kuhnle and others.

PRESENTED BY:

IGM Institute for Genomic Medicine
UC San Diego Bioinformatics & Systems Biology
UC San Diego School of Medicine

FOR MORE INFORMATION PLEASE VISIT GENOMIC.WEEBLY.COM