

NAME OF PRESENTING AUTHOR: Emmanuel Sapin

EMAIL ADDRESS OF PRESENTING AUTHOR: emmanuel.sapin@colorado.edu

LOCATION OF PRESENTING AUTHOR: America (North, Central or South)

TIME ZONE OF PRESENTING AUTHOR: MDT

TYPE OF SUBMISSION: Oral paper

MEMBER STATUS: Non-member

ELIGIBLE FOR THOMPSON AWARD: Yes

ELIGIBLE FOR ROWEWARD: Yes

TITLE: Novel Approach for Parallelizing Pairwise Comparison Problems as Applied to Detecting Segments Identical By Decent in UK BioBank

FULL AUTHOR LIST: Emmanuel Sapin, Matt Keller

AFFILIATIONS: Institute for Behavioral Genetics at University of Colorado Boulder

KEYWORDS: **IBD segments**, UK BioBank, Germline, Parallelization,

ABSTRACT:

Pairwise comparison problems arise in many areas of science. In genomic with datasets already large and getting larger operations that require pairwise comparisons—either on pairs of SNPs or pairs of individuals—will be extremely computationally challenging. We propose a generic algorithm for pairwise comparison problems based on balanced incomplete block design and able to break a large pairwise comparison problem (of order n^2 comparisons) into multiple smaller pairwise comparison problems (each of order n comparisons). We demonstrated that this procedure is very efficient for calling IBD segments using GERMLINE in the large UK Biobank dataset, with a user time savings roughly 180-fold overrunning GERMLINE on the entire sample in our particular instance.

GRANT SUPPORT:
