Exploring the statistical nature of independent chromosome segments

Mary Kate Hollifield[1], M. Bermann[1], D. Lourenco[1], I. Misztal[1]

[1]Department of Animal and Dairy Science, University of Georgia, Athens, GA

ASAS Annual Meeting - June 28th, 2022

Animal Breeding and Genetics Group
College of Agricultural & Environmental Sciences
UNIVERSITY OF GEORGIA

ASAS-CSAS ANNUAL MEETING
OKLAHOMA CITY, OKLAHOMA
JUNE 26-30, 2022

UNIVERSITY OF GEORGIA

Chromosome segments in a formula for accuracy approximation

Animal Breeding and Genetics Group
College of Agricultural & Environmental Sciences
UNIVERSITY OF GEORGIA

N = number of animals with phenotypes

M_e = number of independent chromosome segments

Number of independent segments 1k-20k

$$r = \sqrt{\frac{Nh^2}{Nh^2 + M_e}}$$

Daetwyler et al., 2008

Chromosome segments in APY

Animal Breeding and Genetics Group
College of Agricultural & Environmental Sciences
UNIVERSITY OF GEORGIA

Breeding values → $\mathbf{u} = \mathbf{Ts}$ ← M_e chromosome segments

Choose any N_e animals called "founder": $\mathbf{u}_f$

$\mathbf{s} = \mathbf{Q}\mathbf{u}_f + \boldsymbol{\varepsilon}_f$ Segments as a linear function of N_e founder animals

$\mathbf{u}_n = \mathbf{T}_n\mathbf{s} = \mathbf{P}_{nf}\mathbf{u}_f + \boldsymbol{\varepsilon}_n$ Non-founder animals as linear functions of founder animals

Accuracy of APY algorithm high when $M_e = 4N_eL$
(Stam, 1980; Pocrnic et. al, 2016)

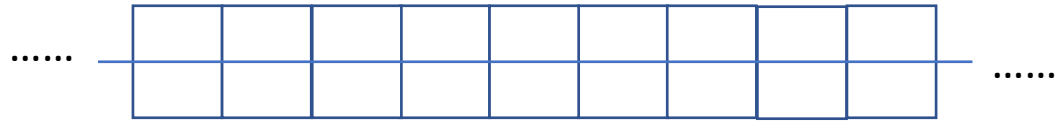
What are the chromosome segments physically?

Animal Breeding and Genetics Group
College of Agricultural & Environmental Sciences
UNIVERSITY OF GEORGIA

Hypothesis 1: equal size blocks

Average segment size 1/4M
With L=10, 40 segments per animal
$4N_eL$ equal size independent segments

Hypothesis 2: equal size haplotypes



Average haplotype size 1/2M
With L=10, 20 haplotypes per animal
$2N_eL$ equal size independent haplotypes

Materials and Methods: Data

Animal Breeding and Genetics Group
College of Agricultural & Environmental Sciences
UNIVERSITY OF GEORGIA

- QMSim (Sargolzaei & Schenkel, 2009)
  - Ten generations: 2,000 animals each
  - One polygenic trait, $h^2 = 0.3$
  - $N_e = 20$
  - Last 3 generations genotyped
    - L = 10 M, 50k SNP, 1,660 QTL

Models and Computations

Animal Breeding and Genetics Group
College of Agricultural & Environmental Sciences
UNIVERSITY OF GEORGIA

$$\mathbf{y} = \mathbf{1}'\mu + \mathbf{Z_s}\boldsymbol{s} + \boldsymbol{e}$$

- Validation:

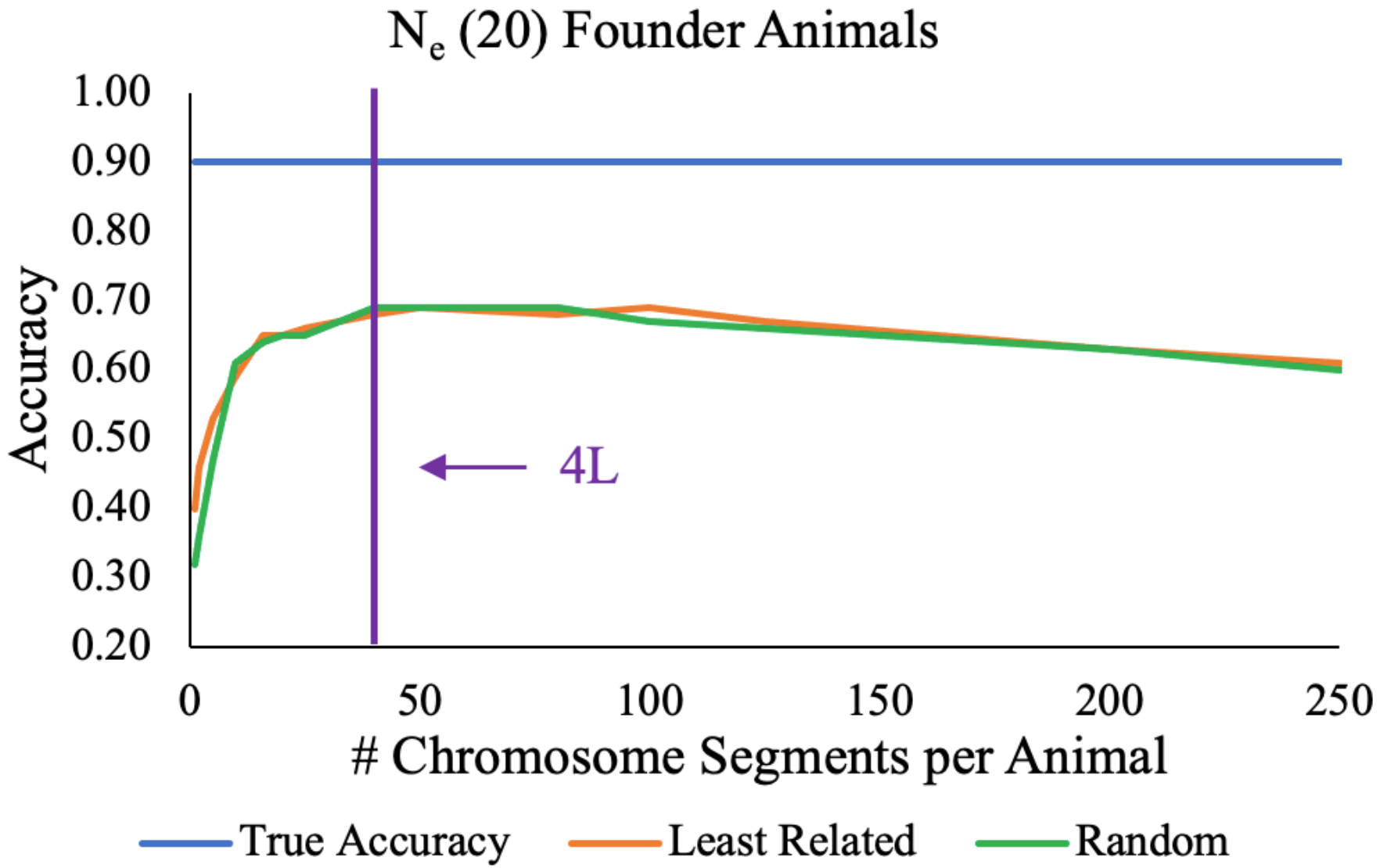
  $cor(\mathbf{TBV}, \mathbf{Z_s}\hat{\boldsymbol{s}})$

  $cor(\mathbf{TBV}, \mathbf{GEBV})$

- GEBV computed using GBLUP in BLUPF90 software suite
- $\hat{\boldsymbol{s}}$ and $\mathbf{Z_s}\hat{\boldsymbol{s}}$ computed using in-house software written in Fortran

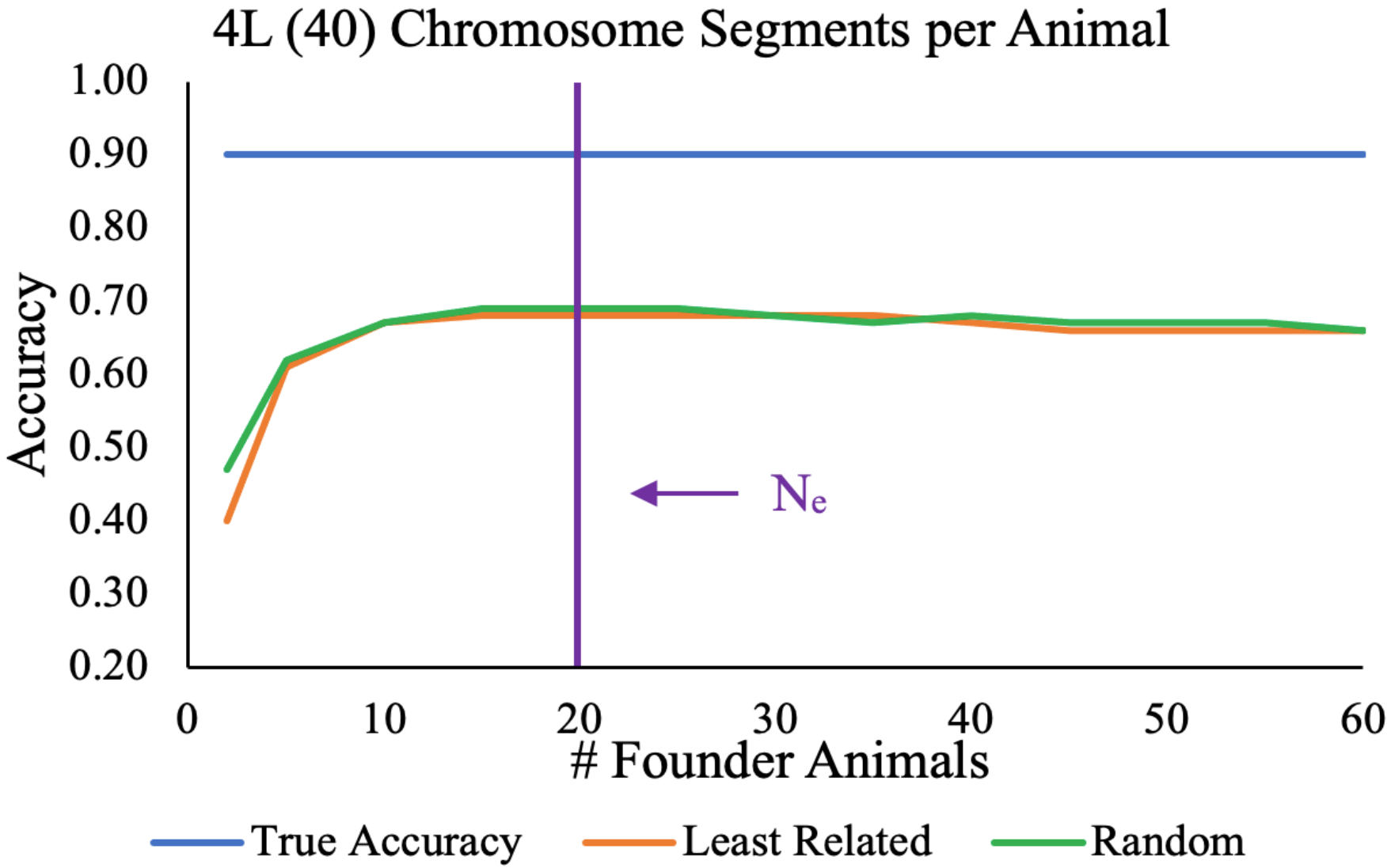
Results

Animal Breeding and Genetics Group
College of Agricultural & Environmental Sciences
UNIVERSITY OF GEORGIA



Results

Animal Breeding and Genetics Group
College of Agricultural & Environmental Sciences
UNIVERSITY OF GEORGIA



Why accuracies are not as high as GBLUP?

Animal Breeding and Genetics Group
College of Agricultural & Environmental Sciences
UNIVERSITY OF GEORGIA

- Wrong hypothesis
- Used pseudo-founders, not founders
- Used most similar segment of pseudo-founders, not a combination

Thank You!

UNIVERSITY OF GEORGIA

Animal Breeding and Genetics Group
College of Agricultural & Environmental Sciences
UNIVERSITY OF GEORGIA

marykate.hollifield@uga.edu