New approach to calculate inbreeding effective population size from runs of homozygosity

Gregor Gorjanc, John Hickey & Ino Curik

ROSLIN

BBSRC
bioscience for the future

THE UNIVERSITY of EDINBURGH
Royal (Dick) School of Veterinary Studies

AGRONOMSKI FAKULTET ZAGREB

Effective population size (NeF$_{ROH}$)→ new idea !

Aims of the study

1 To present theoretical background/concept behind "new" inbreeding effective population size - NeF$_{ROH}$

2 To present NeF$_{ROH}$ calculated/obtained in empirical cattle populations

3 To analyse behaviour of NeF$_{ROH}$ in computer simulations

The effective population size (N_e) of a real population X is the size of a hypothetical ideal population (Wright-Fisher) that will result in the ***same amount of genetic drift** as in the real (actual) population considered.

1

Effective inbreeding population size (N_{eF}):
same change in inbreeding level

$\approx$ **but** $\neq$

Effective variance population size (N_{eV}):
*same change in allele frequencies
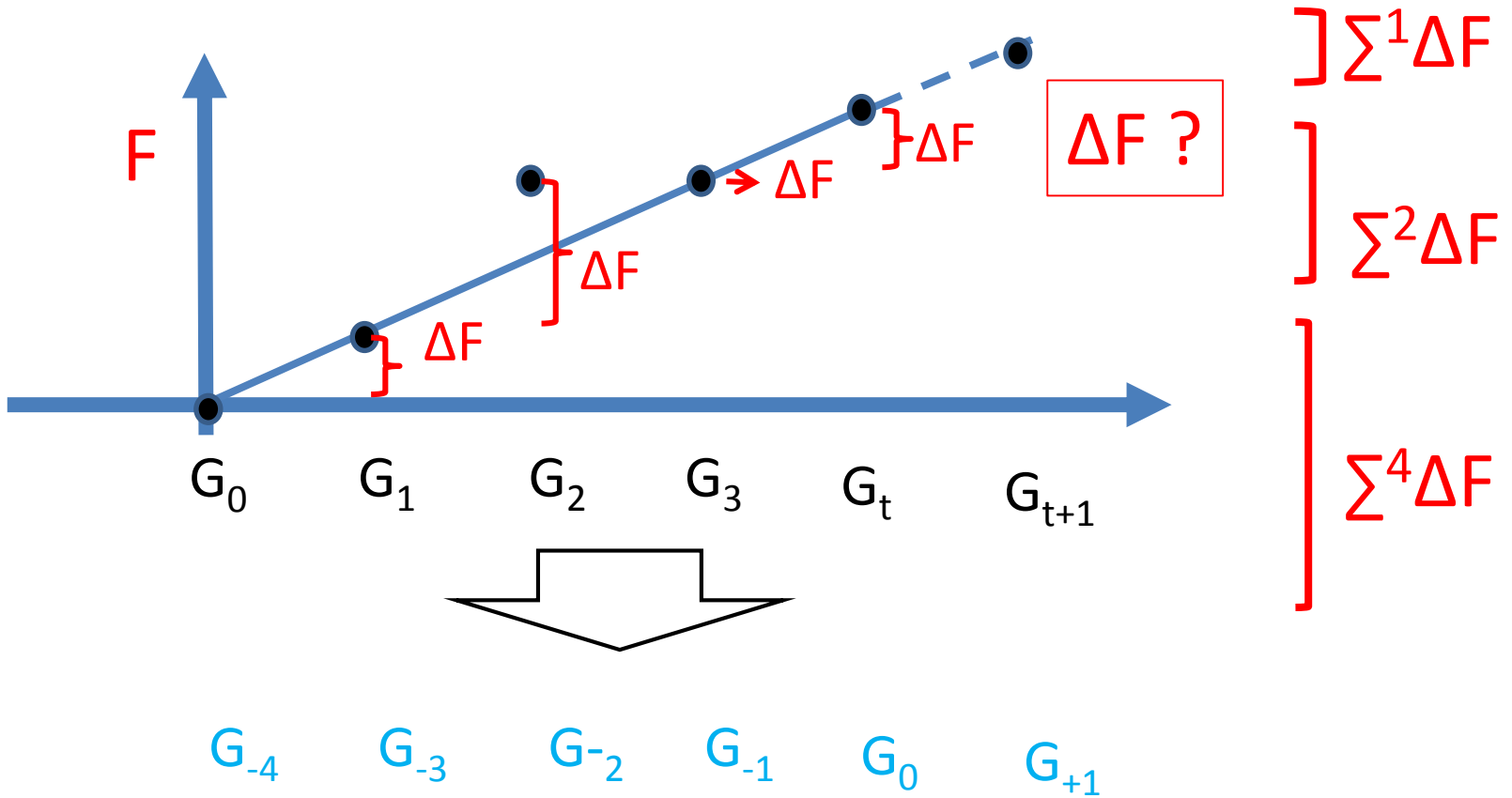
Effective eigenvalue population size (N_{eE}):
*same long term rate at which genetic variants are lost

Effective linkage disequilibrium population size (N_{eLD}):
*same change in gametic phase/linkage disequilibrium

$\Delta F = 1/(2Ne) \rightarrow Ne = 1/(2\Delta F)$

Setting up the basic population based on ROH length!
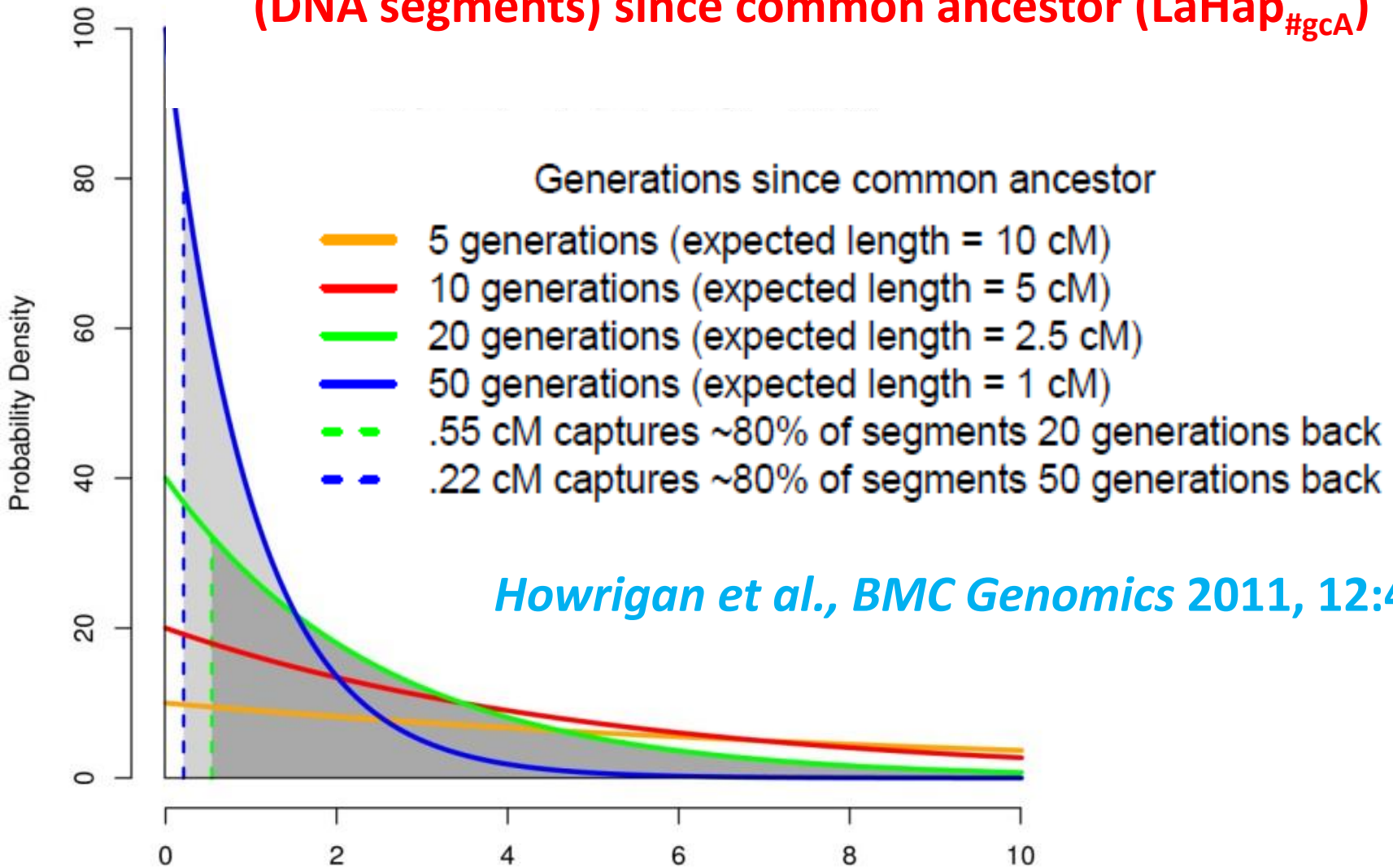
1

$$F_{ROH} = \frac{\sum_{j=1}^{n} L_{ROH_j}}{L_{total}}$$

F_{ROH>10Mb}

F_{ROH>5Mb}

F_{ROH>2Mb}

<25   **Generations**   <10   <5.0   0

>2Mb   **ROH length**   >5Mb   >10Mb
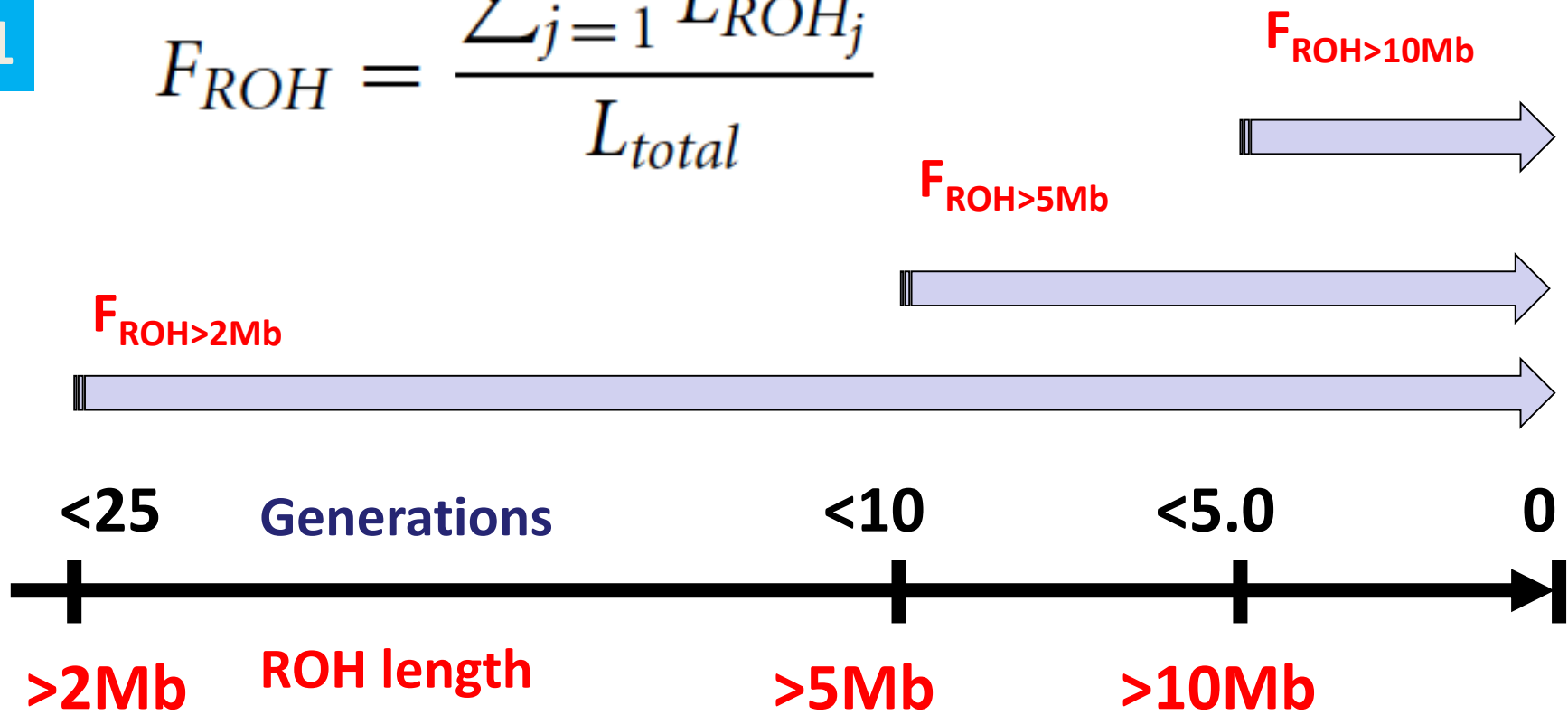
Distribution of expected lengths of autozygous haplotypes (DNA segments) since common ancestor (LaHap$_{\#gcA}$)

Generations since common ancestor

— 5 generations (expected length = 10 cM)
— 10 generations (expected length = 5 cM)
— 20 generations (expected length = 2.5 cM)
— 50 generations (expected length = 1 cM)
- - .55 cM captures ~80% of segments 20 generations back
- - .22 cM captures ~80% of segments 50 generations back

Howrigan et al., BMC Genomics 2011, 12:460

Probability Density

cM distance (1 cM ~ 1 Mb)

$$\Delta F = 1/(2Ne) \quad \rightarrow \quad Ne = 1/(2\Delta F)$$

$$Ne_{FROH > LMb} = \frac{\text{\# generations as f(LROH)}}{2 \sum_{gt0}^{gtb} \Delta F}$$

$$\sum^{5g} \Delta F$$

$$\sum^{10g} \Delta F$$

$$\sum^{25g} \Delta F$$

1

<25   **Generations**   **<10**   **<5**   **0**

>2Mb   **ROH length**   **>5Mb**   **>10Mb**
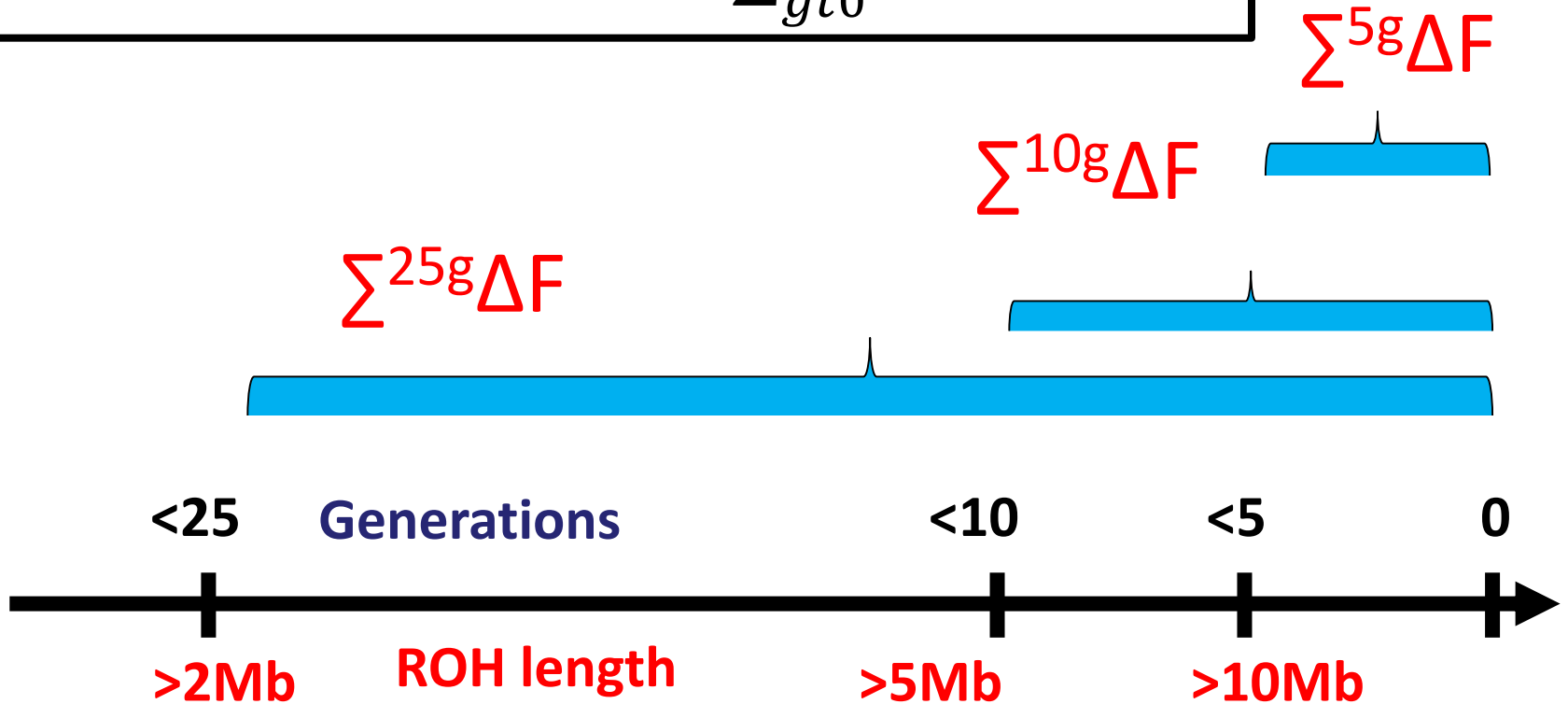
Realized (accumulated) inbreeding during t generations

$$F_{ROH>10Mb} = 0.05 \rightarrow \sum{}^{5g}\Delta F \rightarrow \Delta F = 2\sum{}^{5g}\Delta F /5 = 0.01 \rightarrow Ne_{FROH>10Mb} = 50$$

$$F_{ROH}v = F_{ROH}\begin{pmatrix} g - 25 \rightarrow L2.00Mb \\ g - 24 \rightarrow L2.08Mb \\ g - 23 \rightarrow L2.17Mb \\ ... \\ g - 2 \rightarrow L25.00Mb \end{pmatrix}$$

$$F_{ROH>2Mb-5Mb} = 0.25\text{-}0.10 \rightarrow {}_{25g}\sum{}^{10g}\Delta F \rightarrow \Delta F = {}_{25g}\sum{}^{10g}\Delta F /15 = 0.01$$
$$\rightarrow Ne_{FROH>2Mb-5Mb} = 50$$

$$Ne_{FROH}v = Ne_{FROH}\begin{pmatrix} F_{ROHg-25} - F_{ROHg-24} \\ F_{ROHg-24} - F_{ROHg-23} \\ F_{ROHg-23} - F_{ROHg-22} \\ ... \\ F_{ROHg-3} - F_{ROHg-2} \end{pmatrix}$$

Empirical estimates of Ne$_{FROH}$ in cattle populations

Breed	N	L$_{ROH}$/GB	F$_{ROH}$	Ne$_{FROH}$	Source
Italian Holstein	2093	8/6.25	0.051	61	Marras et al., 2015
Swiss Holstein	2568	5/10	0.053	94	Signer-Hasler et al., 2017
Italian Brown	749	8/6.25	0.068	46	Marras et al., 2015
Brown Swiss (AUT)	304	8/6.25	0.074	42	Ferenčaković et al., 2013
Brown Swiss	281	5/10	0.084	60	Signer-Hasler et al., 2017
Fleckvieh (AUT)	502	8/6.25	0.019	165	Ferenčaković et al., 2013
Italian Simmental	479	8/6.25	0.015	208	Marras et al., 2015
Swiss Flechvieh	547	5/10	0.027	185	Signer-Hasler et al., 2017

Empirical estimates of Ne_{FROH} in cattle populations

Breed	N	L_{ROH}/GB	F_{ROH}	Ne_{FROH}	Source
Norwegian Red	499	8/6.25	0.035	90	Ferenčaković et al., 2013
Tyrol Grey (AUT)	117	8/6.25	0.036	87	Ferenčaković et al., 2013
Piedmontese (ITA)	364	8/6.25	0.007	446	Marras et al., 2015
Marchigiana (ITA)	410	8/6.25	0.031	101	Marras et al., 2015
Pinzgauer (AUT)	118	8/6.25	0.027	116	Ferenčaković et al., 2013
Reggiana (ITA)	168	4/12.5	0.035	179	Mastrangelo et al., 2017
Nellore (BRA)	1278	8/6.25	0.014	223	Zavarez et al., 2015
Gyr (BRA)	2908	8/6.25	0.037	85	Peripoli et al., 2018

Unfortunately, in the first simulation a bug was found …

3

bug corrected … but … computer simulations still in progress !

Related methods

Individual increase in inbreeding allows estimating effective sizes from pedigrees

Juan Pablo Gutiérrez[1]*, Isabel Cervantes[1], Antonio Molina[2], Mercedes Valera[3], Félix Goyache[4]

$$\Delta F_i = 1 - \sqrt[t_i]{1 - F_i}$$

$$\Delta F_i^* = 1 - \sqrt[t_i - 1]{1 - F_i}$$

PLOS | GENETICS

The IBDNe method

RESEARCH ARTICLE

Ancestry-specific recent effective population size in the Americas

Sharon R. Browning[1]*, Brian L. Browning[2], Martha L. Daviglus[3], Ramon A. Durazo-Arvizu[4], Neil Schneiderman[5], Robert C. Kaplan[6], Cathy C. Laurie[1]

1. Theoretical explanation for NeF_{ROH} calculation is logical!

2. Empirical estimates of NeF_{ROH} calculated in several cattle populations are reasonable.

3. Confirmation of accuracy and bias of the proposed method as well as validation of assumptions made is required.

4. Need for the comparison with other methods.

Acknowledgments

Croatia – Slovenia bilateral scientific projects: MZO & ARRS
"Estimation of effective population size and inbreeding from high-throughput genomic information", 2012/2013

Thank you for your attention !

Questions, good suggestions & provocative chalenging comments are desirable!