Characterization of copy number variants in a large multi-breed population of beef and dairy cattle using high-density single nucleotide polymorphism genotype data

P. Rafter[*†], D. C. Purfield[*], D. P. Berry[*2], A. C. Parnell[†], I.C. Gormley[†], J.F. Kearney[#], M.P. Coffey[¥], T. R. Carthy[*]

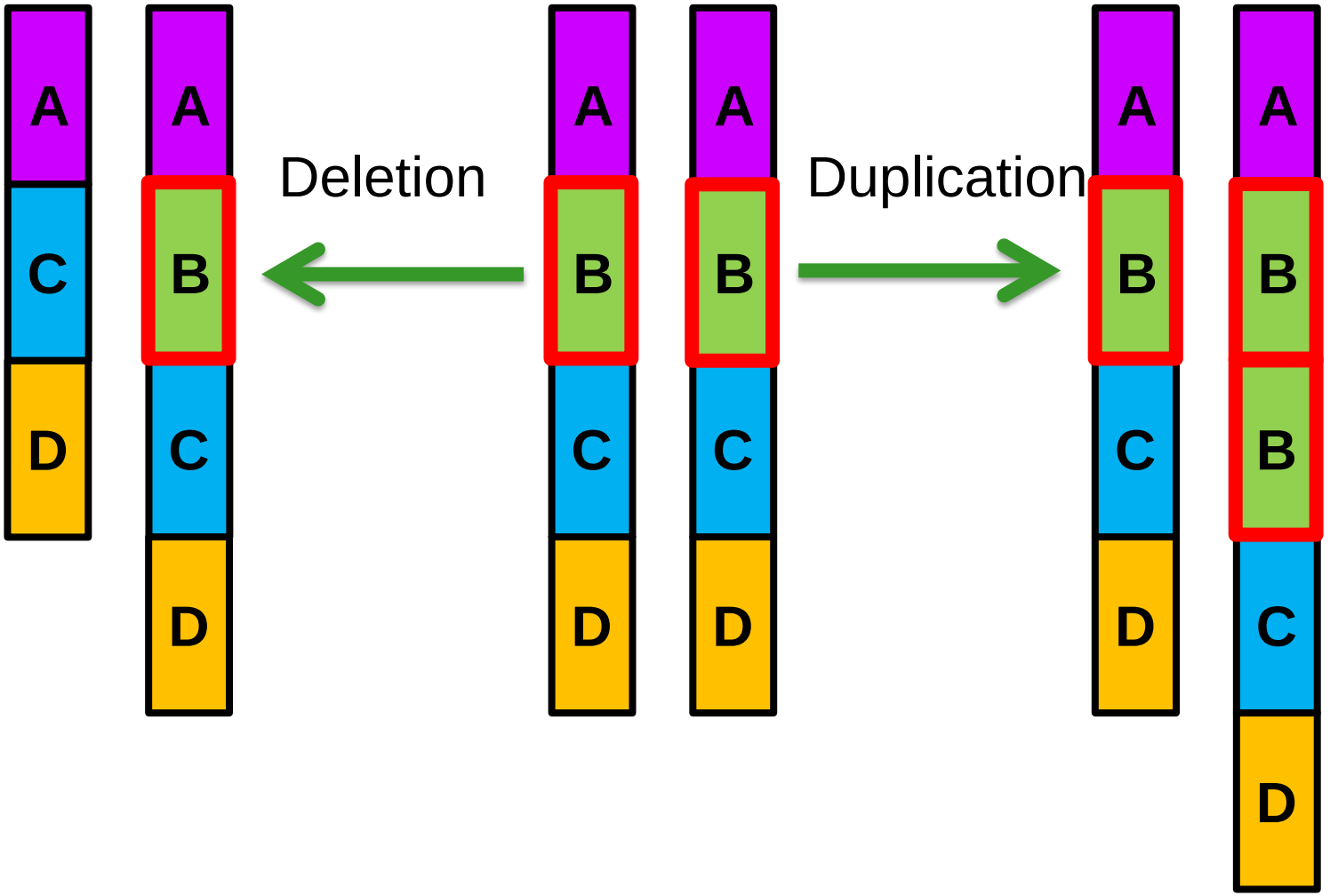
*Teagasc, Animal & Grassland Research and Innovation Centre, Moorepark, Fermoy, Co. Cork, Ireland.

† UCD School of Mathematics and Statistics, Insight Centre for Data Analytics, University College Dublin, Belfield, Dublin 4, Ireland.

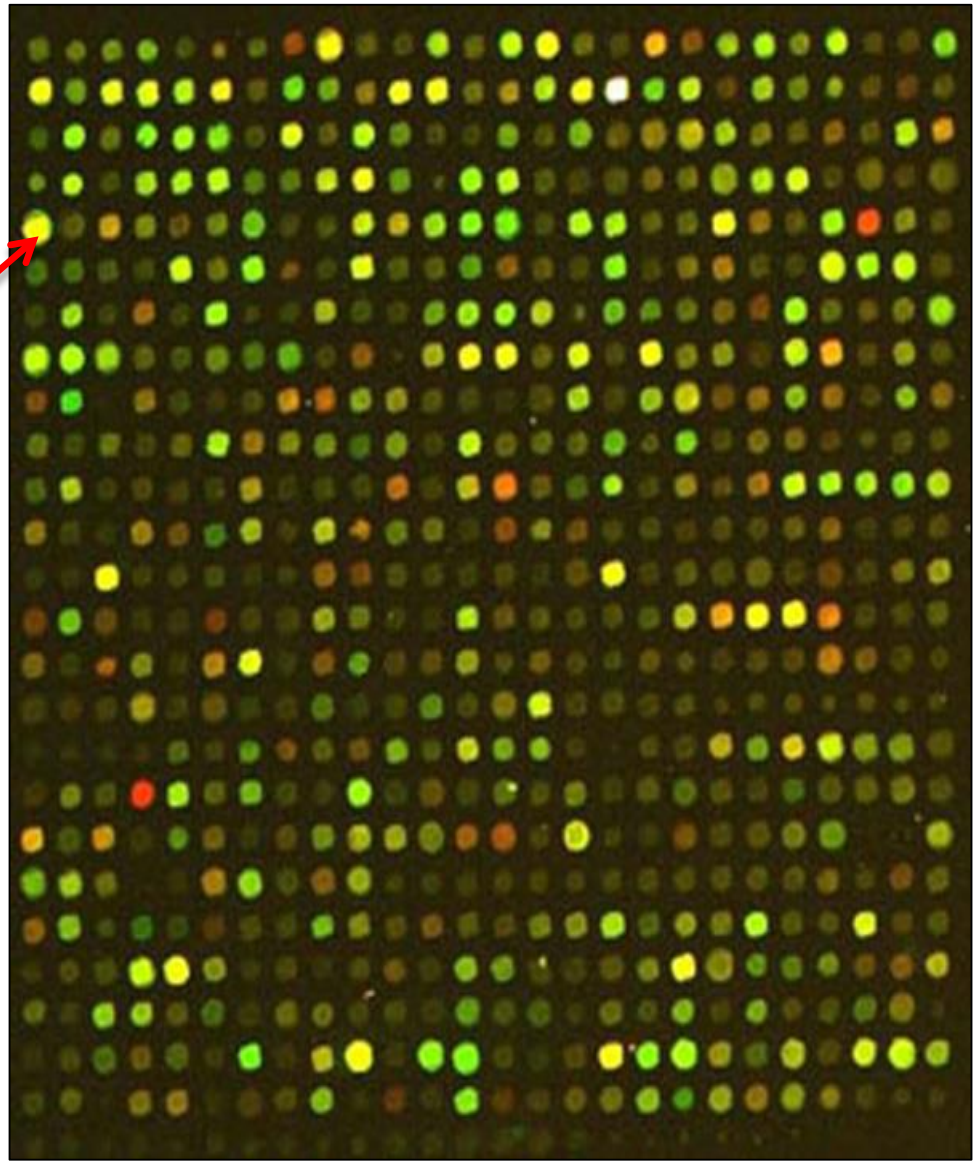
¥ Animal & Veterinary Sciences, SRUC, Roslin Institute Building, Easter Bush, Midlothian, Scotland. EH25 9RG.

ICBF, Highfield House, Shinagh, Bandon, Co.Cork, Ireland.

What is a copy number variant (CNV)?



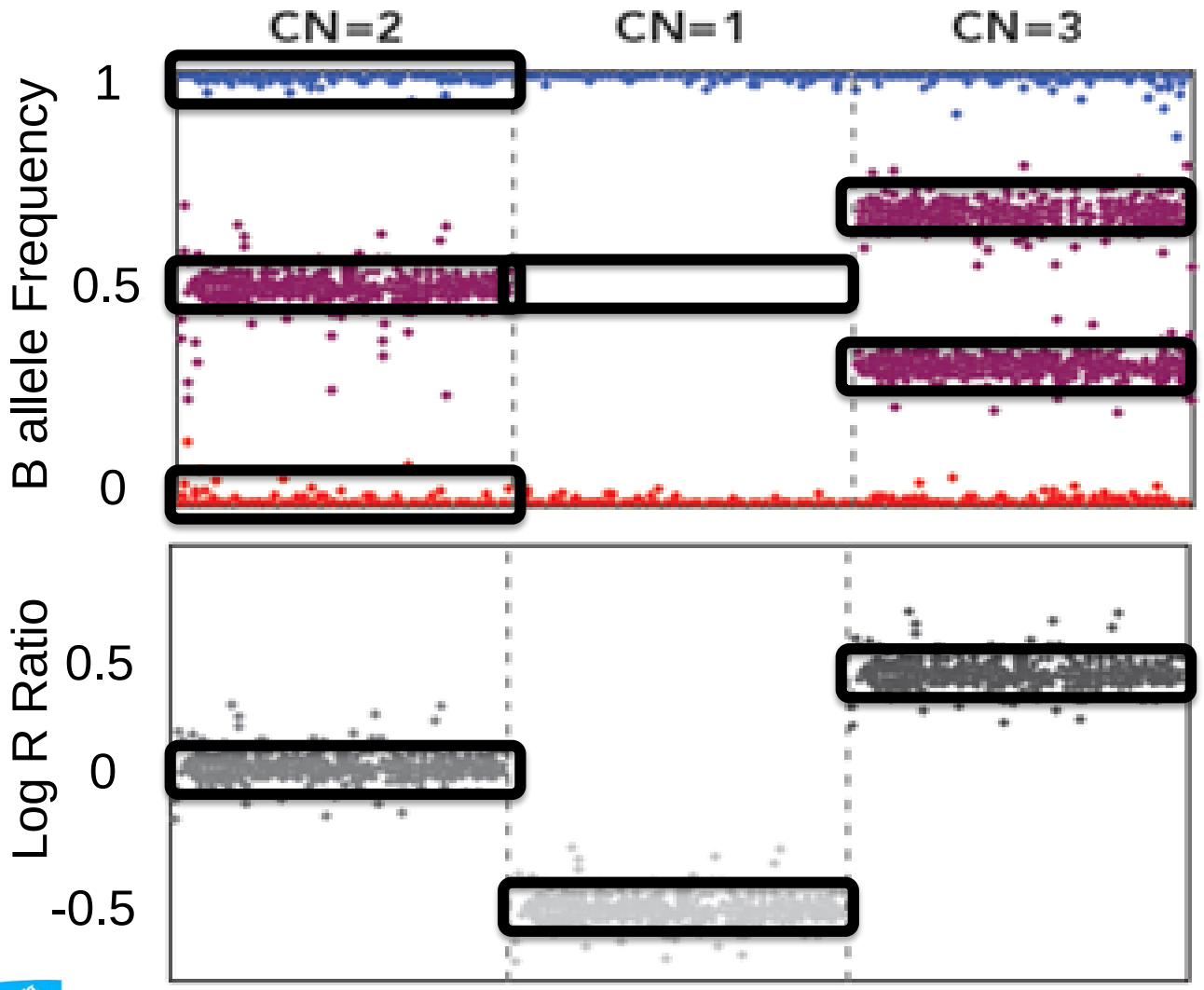
$$\log\left(\frac{Observed\ Fluorescence}{Expected\ Fluorescence}\right)$$



UCD DUBLIN

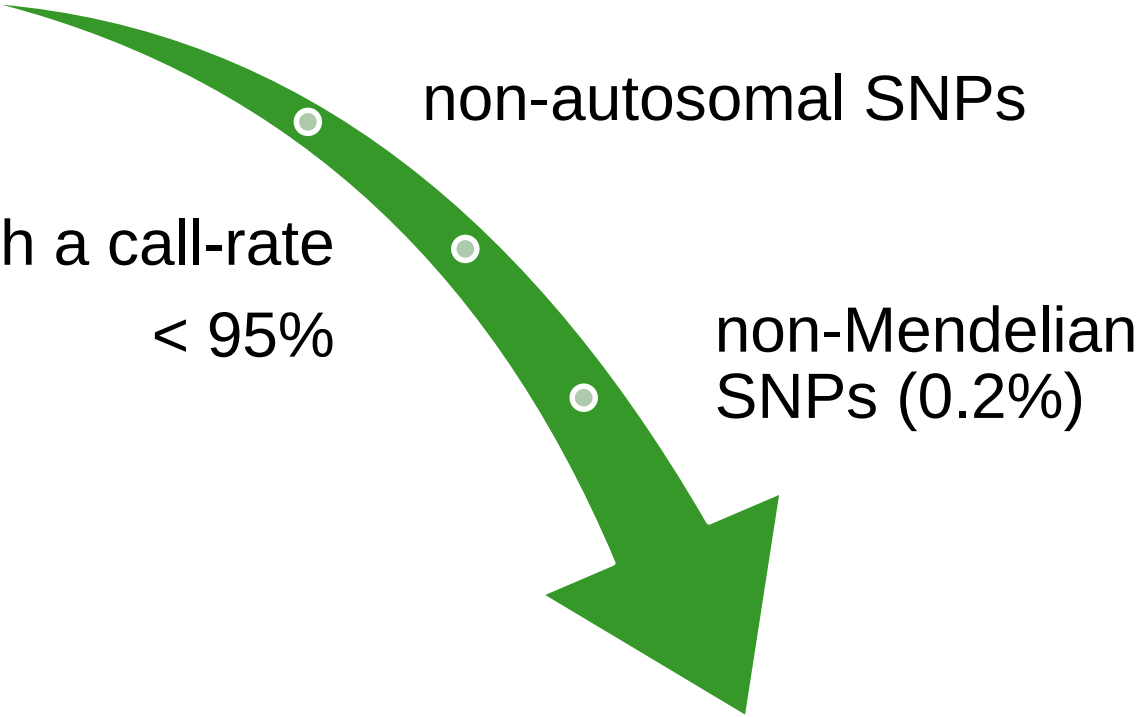
teagasc
Agriculture and Food Development Authority

Copy number analysis



Genotype data

Illumina BovineHD, 777,962 SNPs

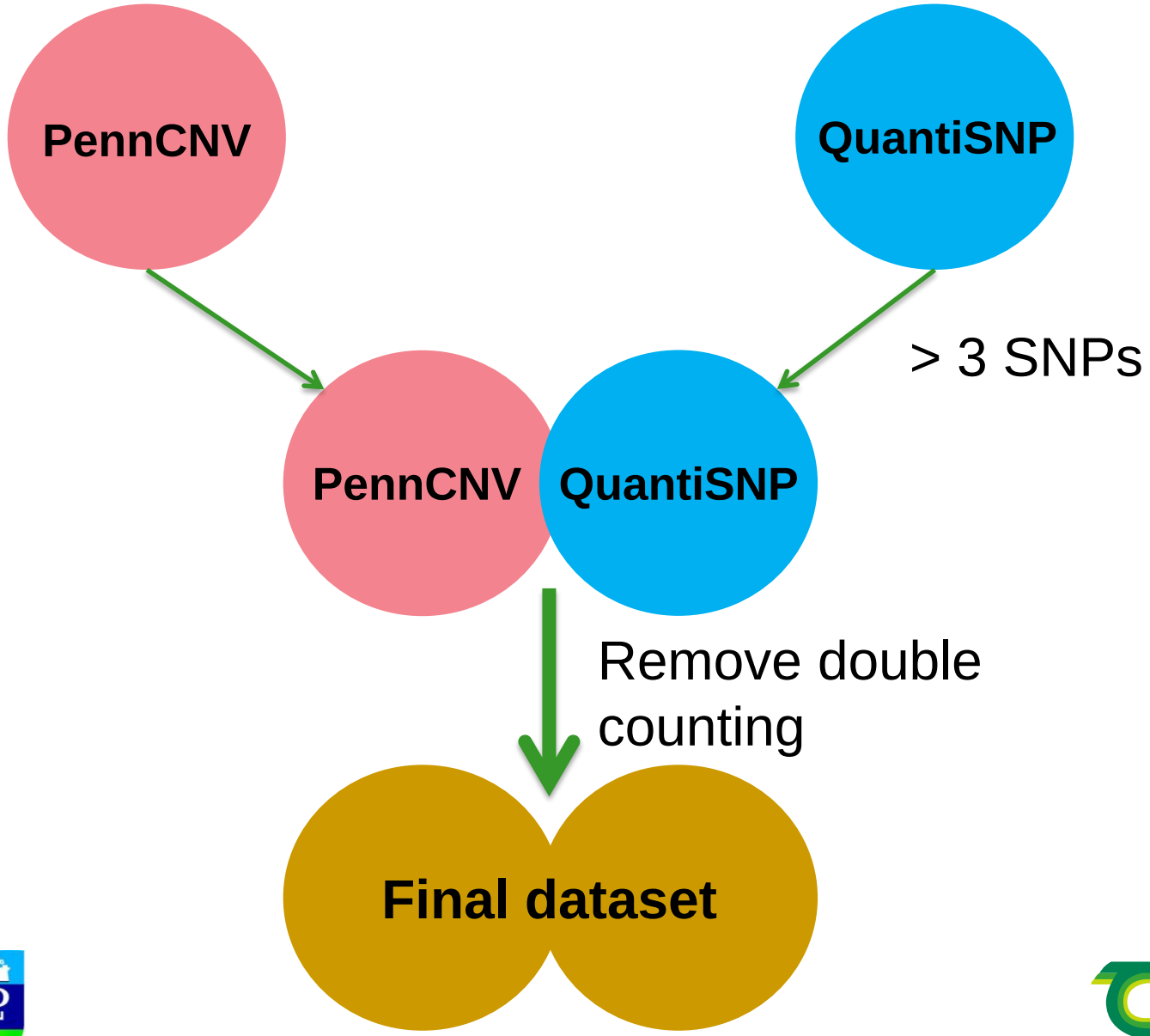


713,162 SNPs

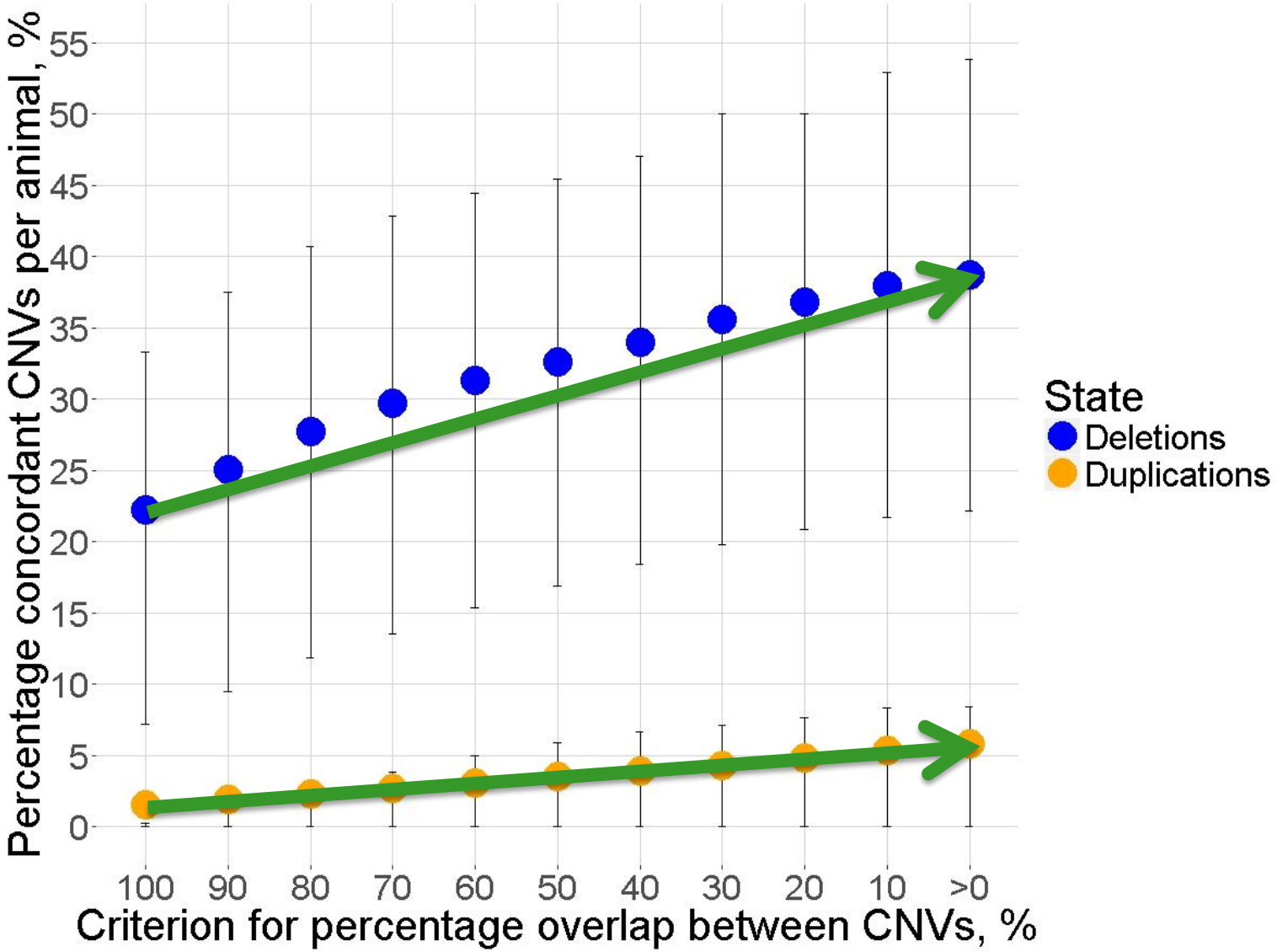
UCD DUBLIN

teagasc
Agriculture and Food Development Authority

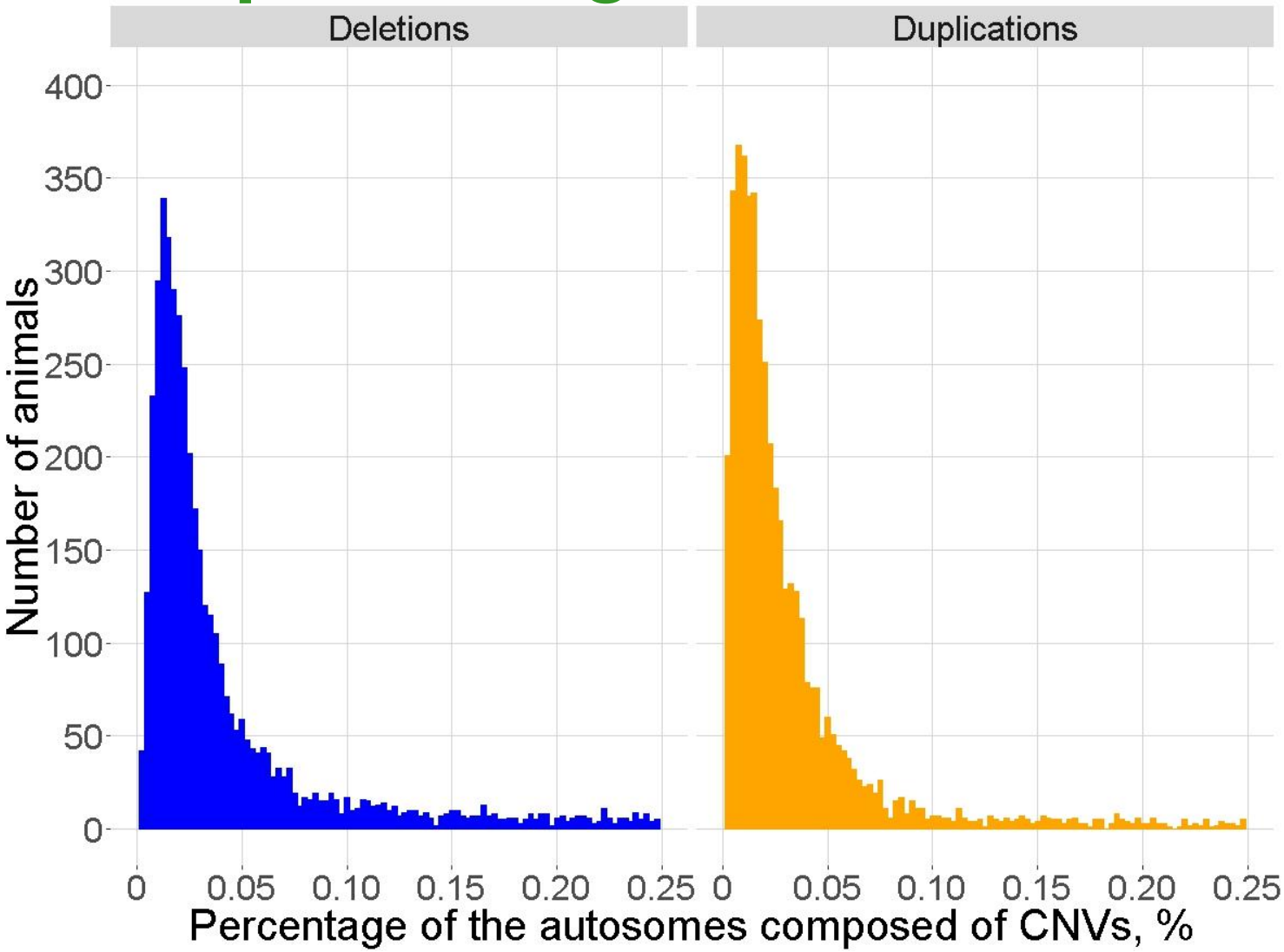
The CNV dataset



Concordance between PennCNV and QuantiSNP



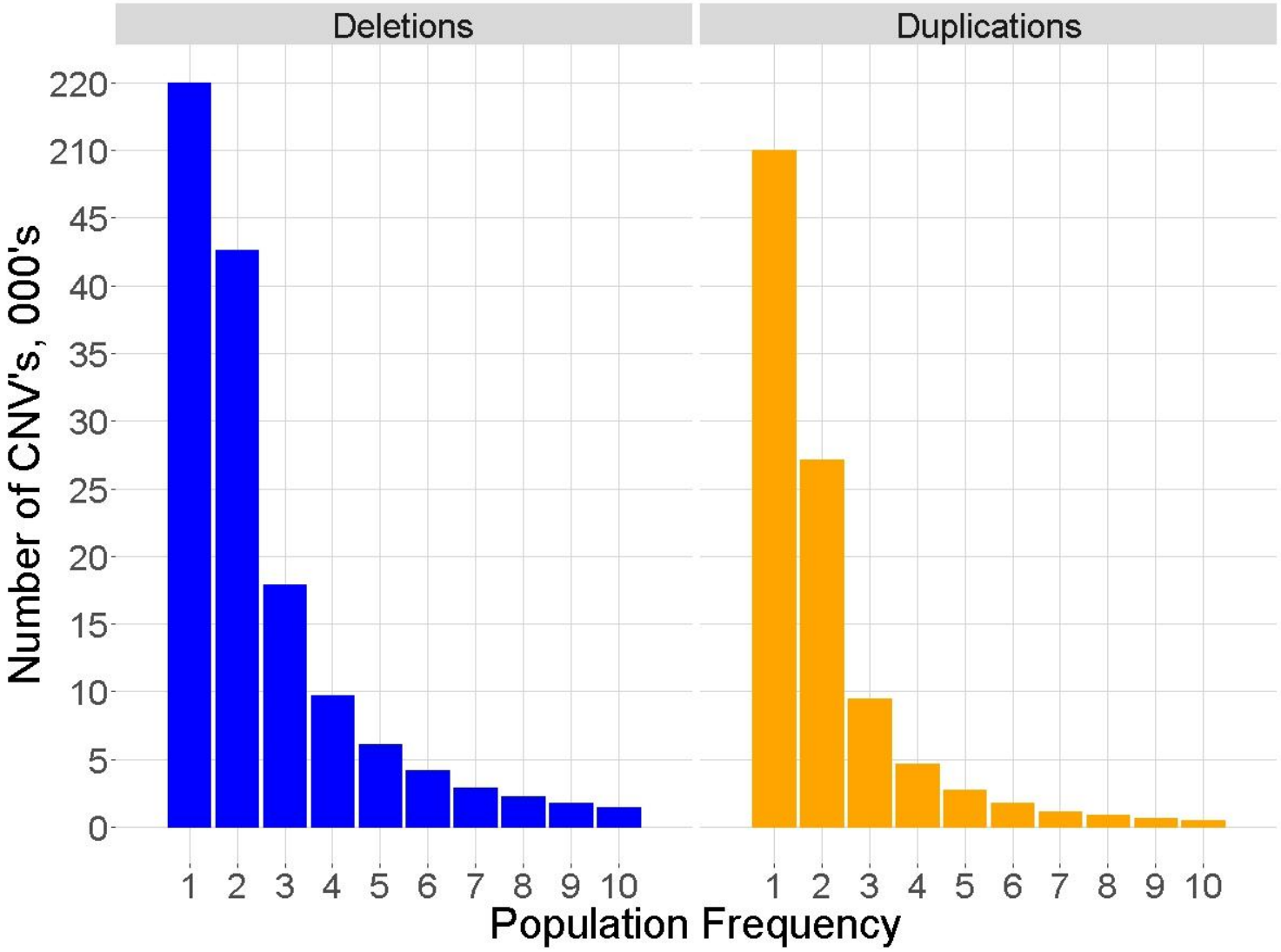
CNV percentage of autosomes



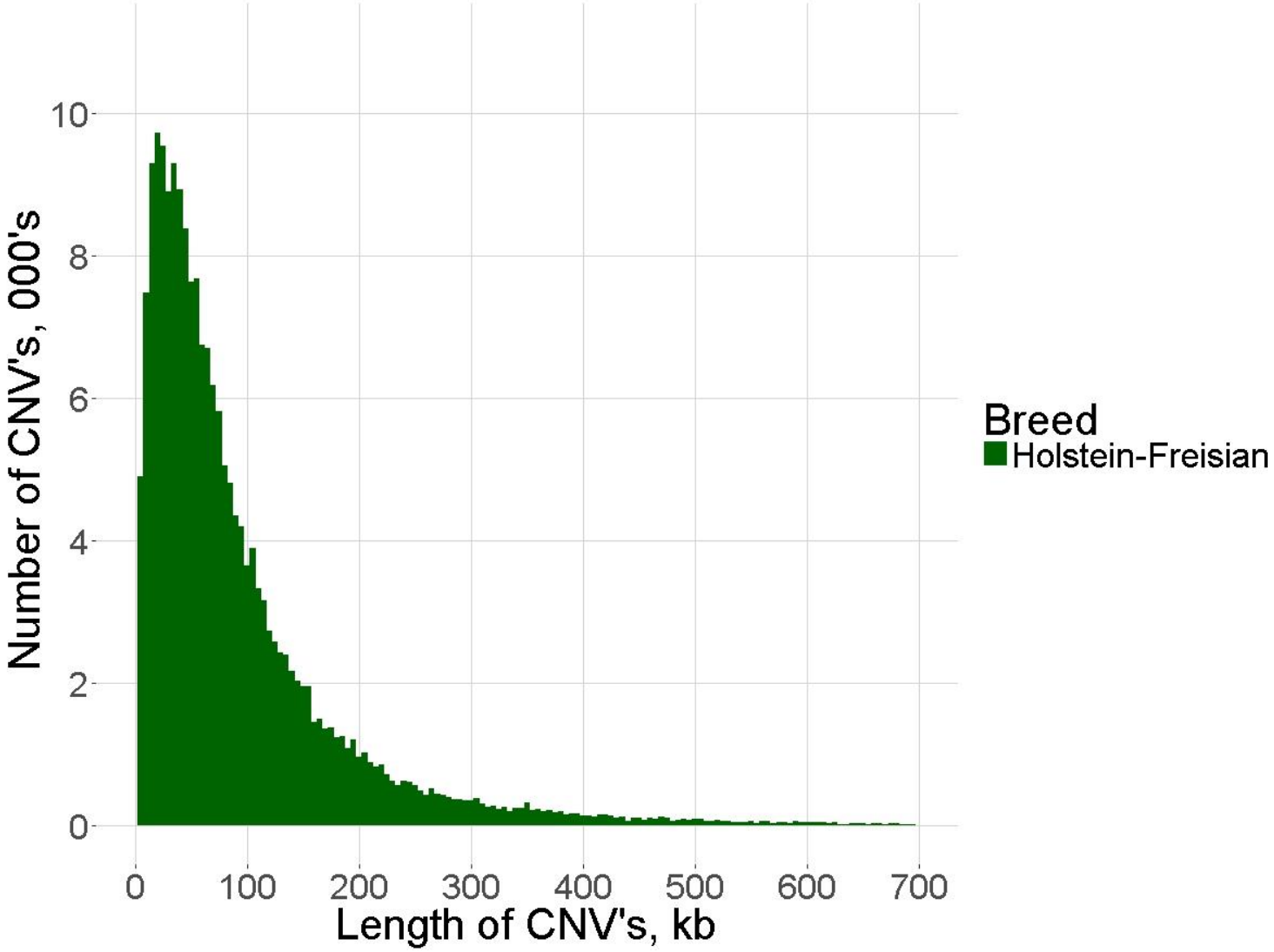
UCD DUBLIN

teagasc
Agriculture and Food Development Authority

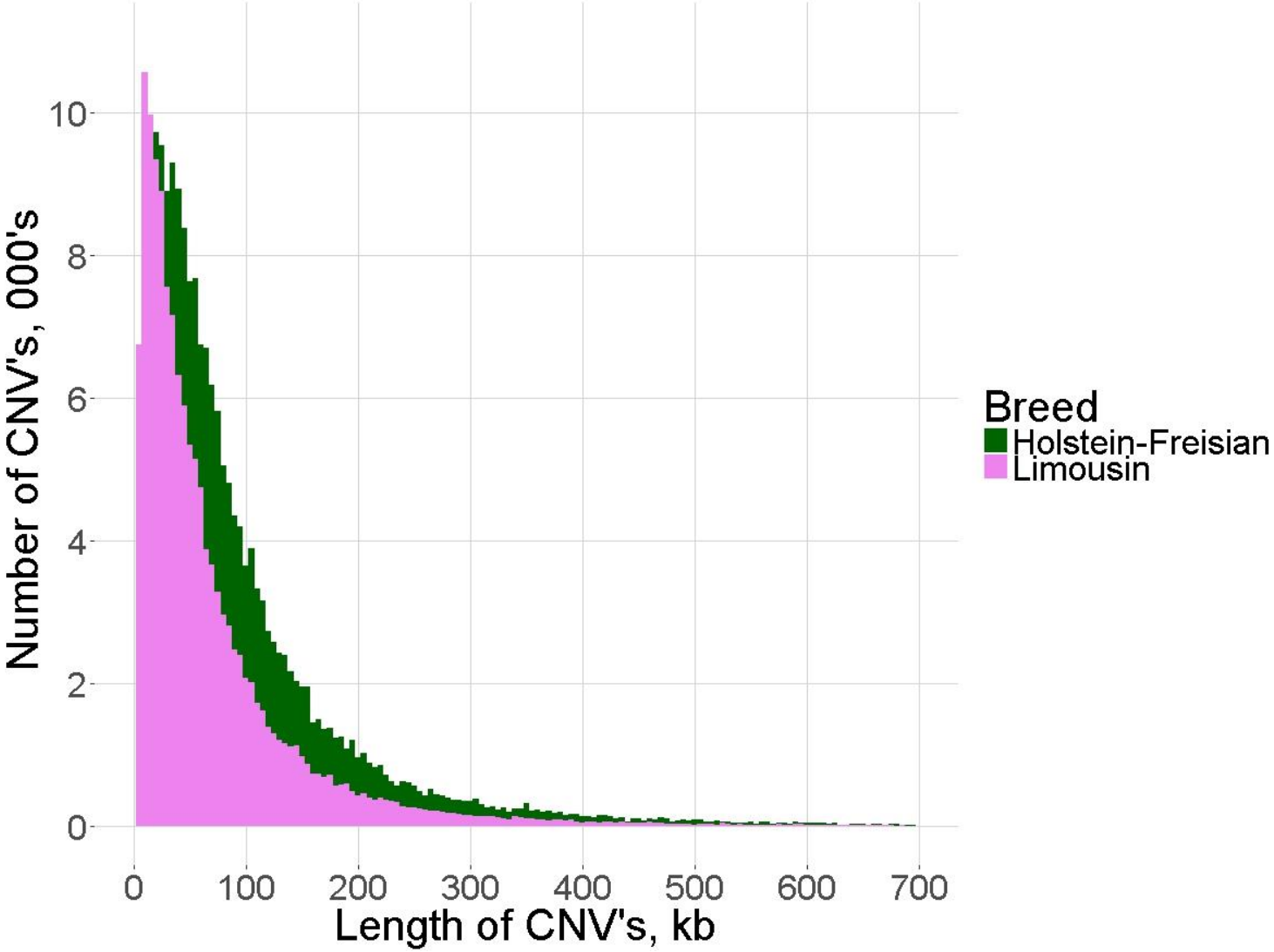
Frequency of CNVs



Breed comparison – deletion length



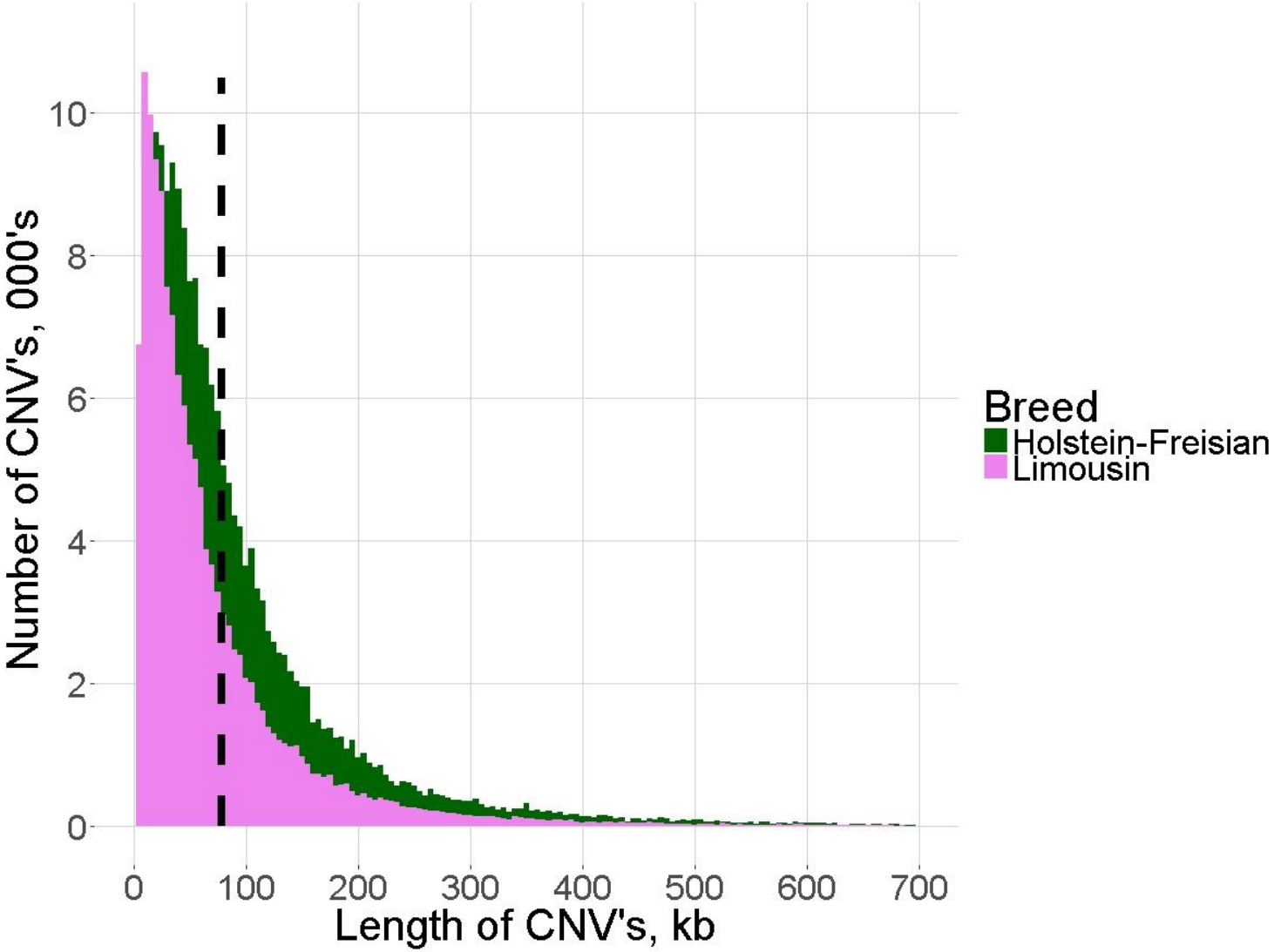
Breed comparison – deletion length



UCD DUBLIN

teagasc
Agriculture and Food Development Authority

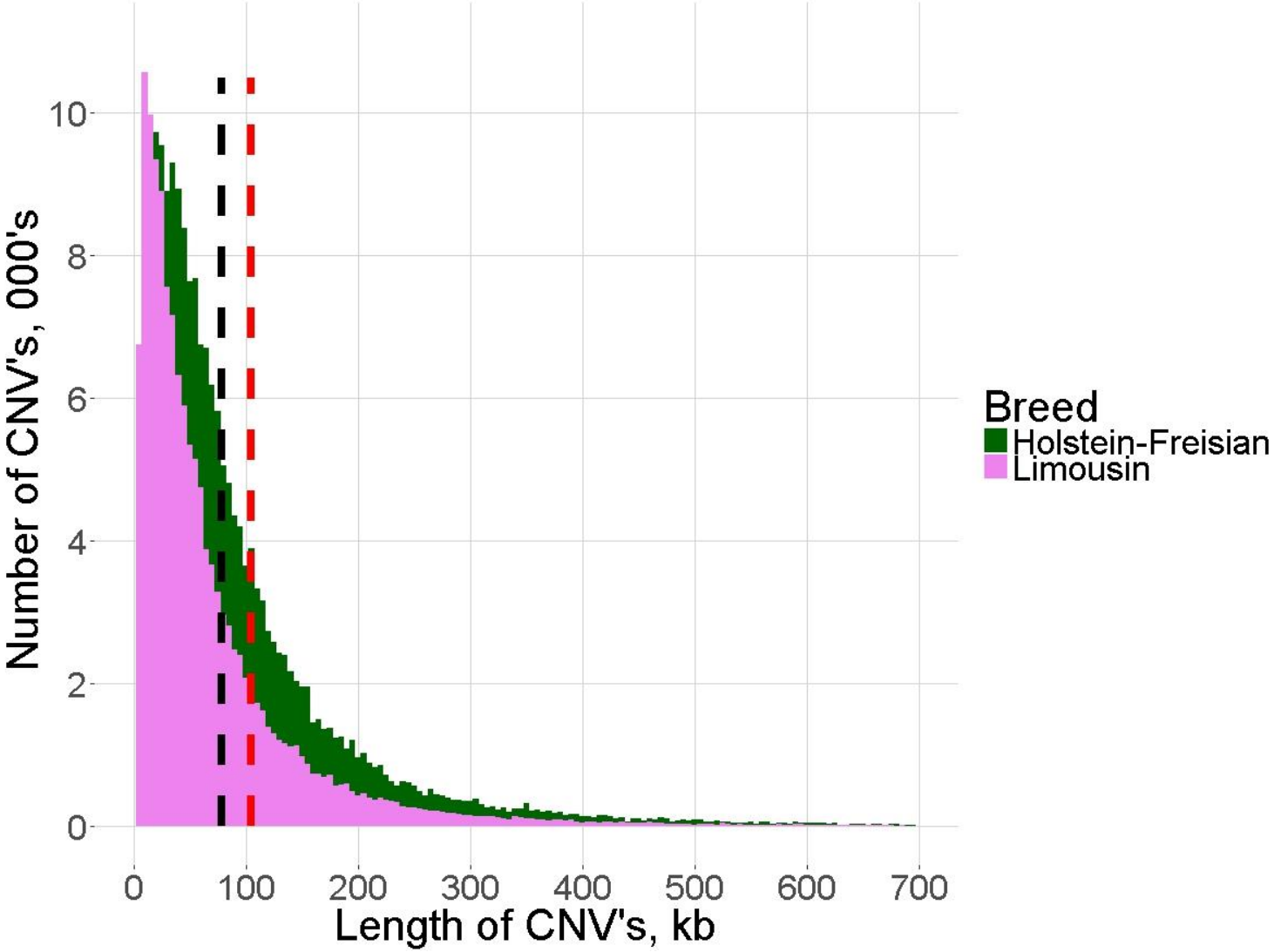
Breed comparison – deletion length



UCD DUBLIN

teagasc
Agriculture and Food Development Authority

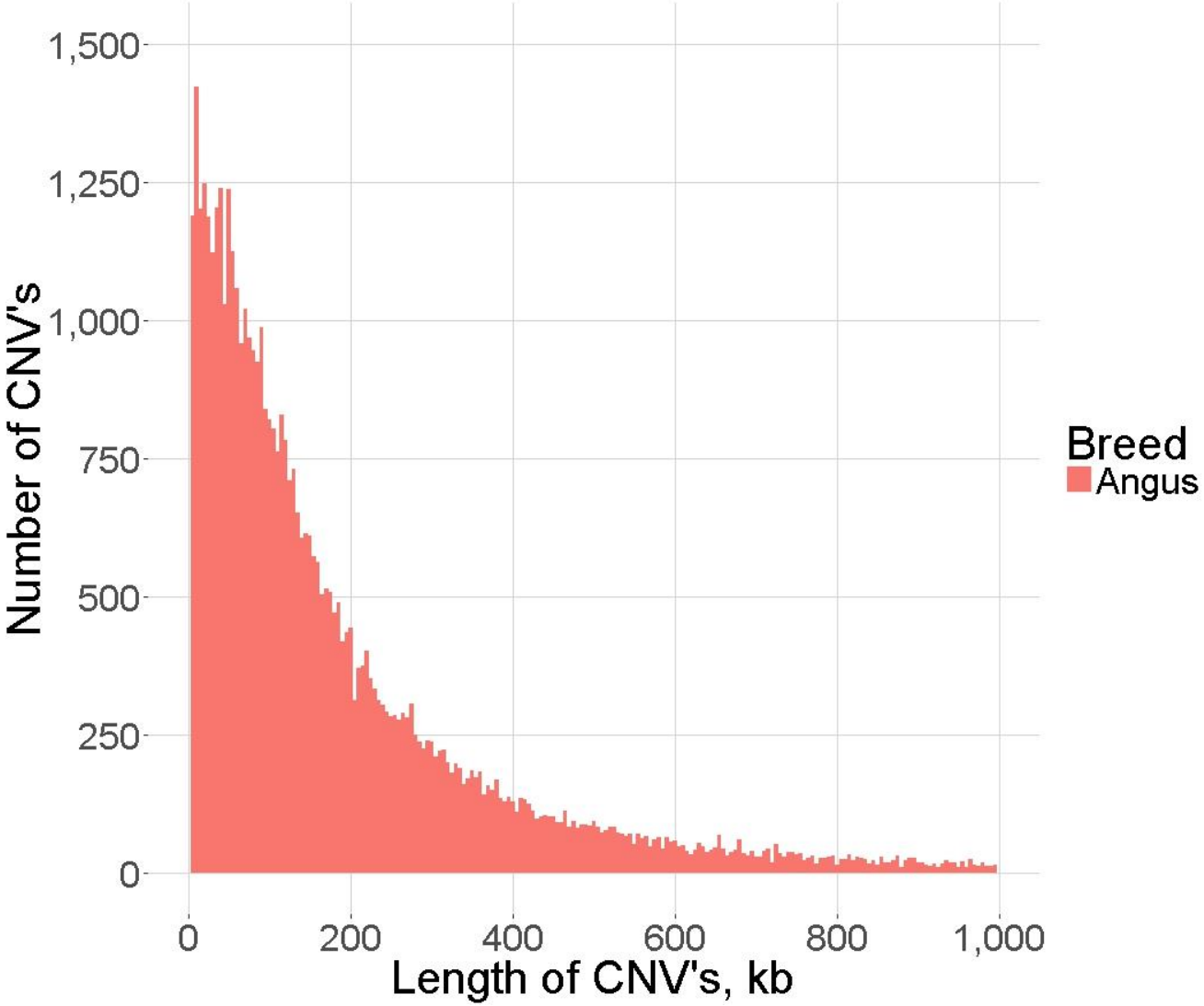
Breed comparison – deletion length



UCD DUBLIN

teagasc
Agriculture and Food Development Authority

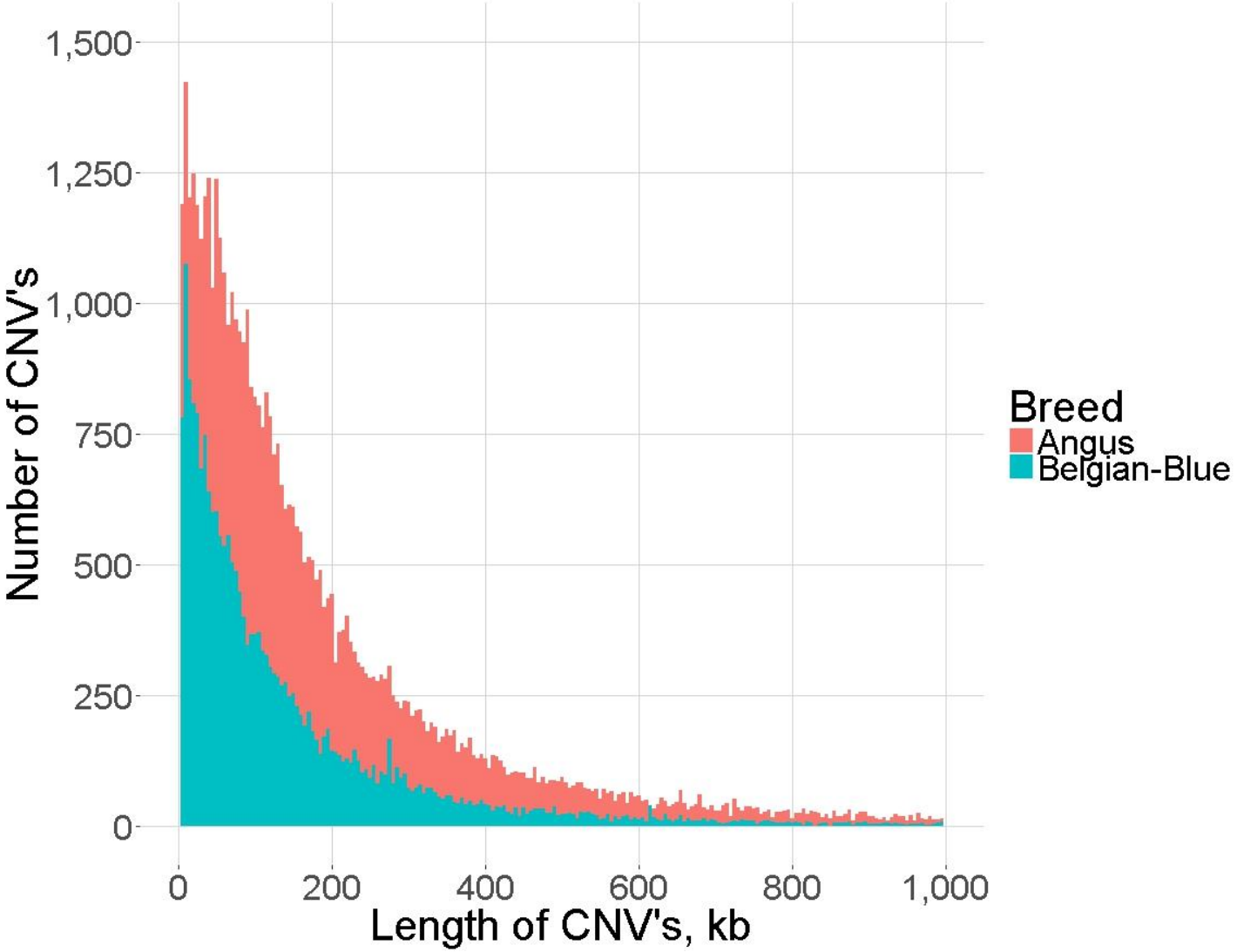
Breed comparison – duplication length



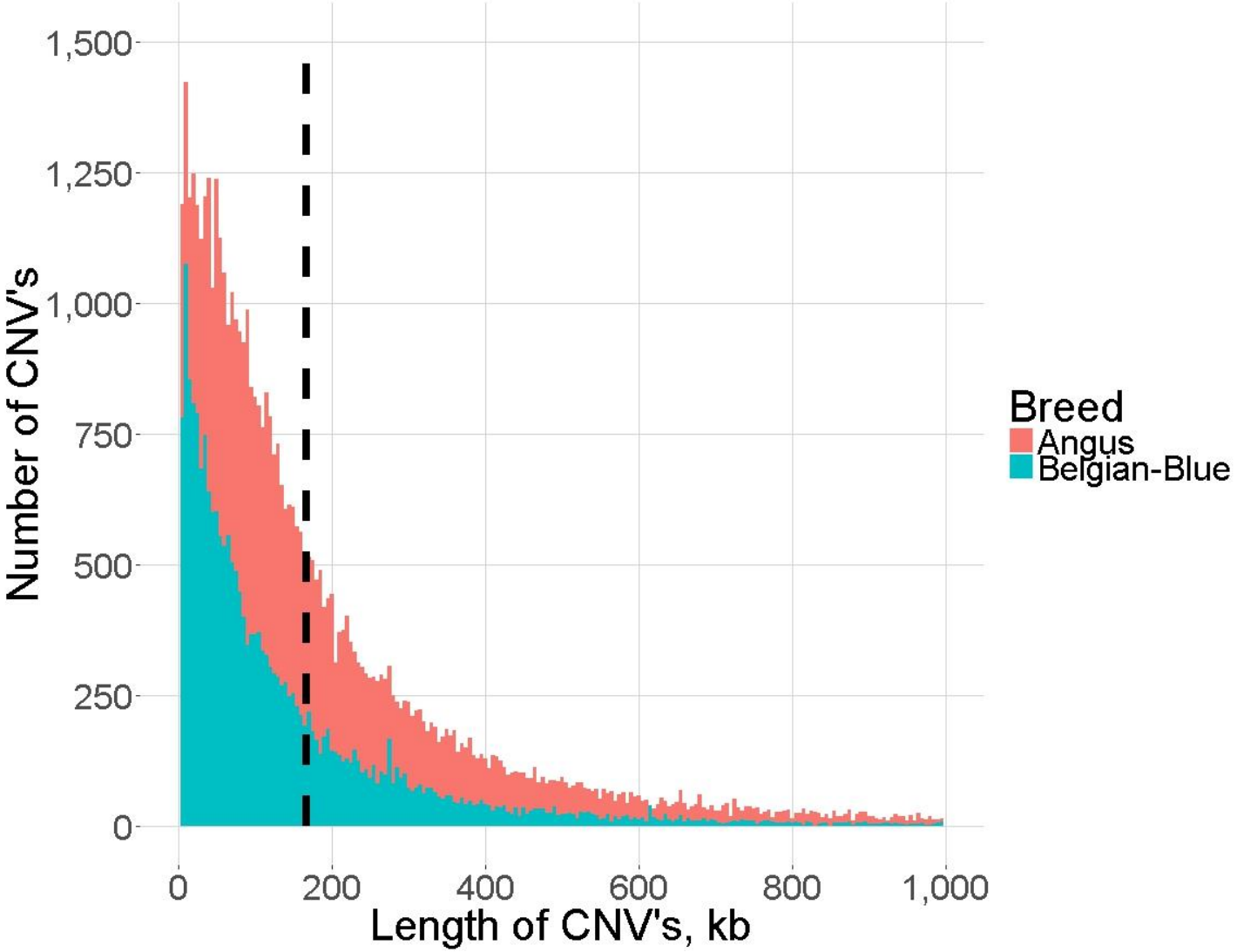
UCD DUBLIN

teagasc
Agriculture and Food Development Authority

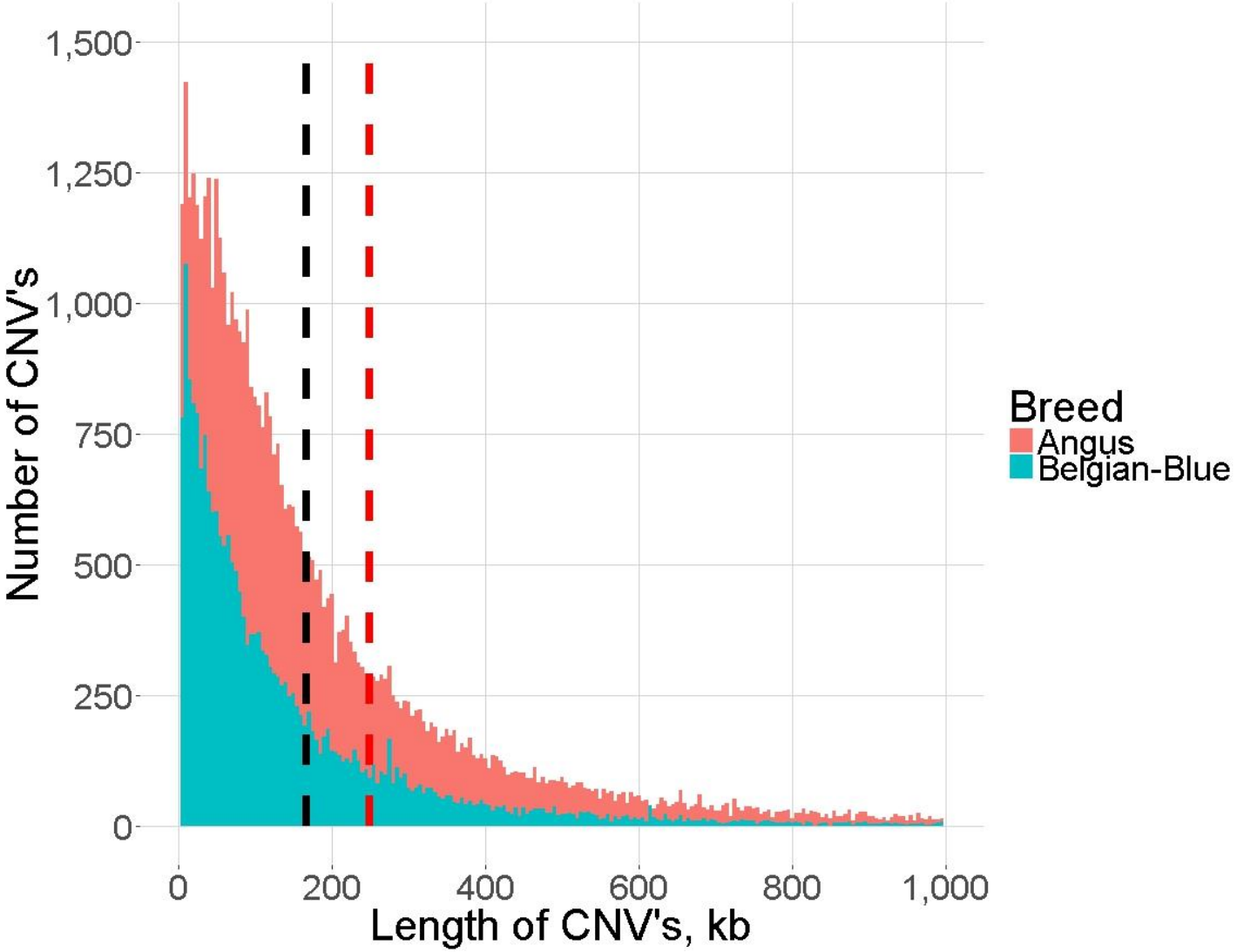
Breed comparison – duplication length



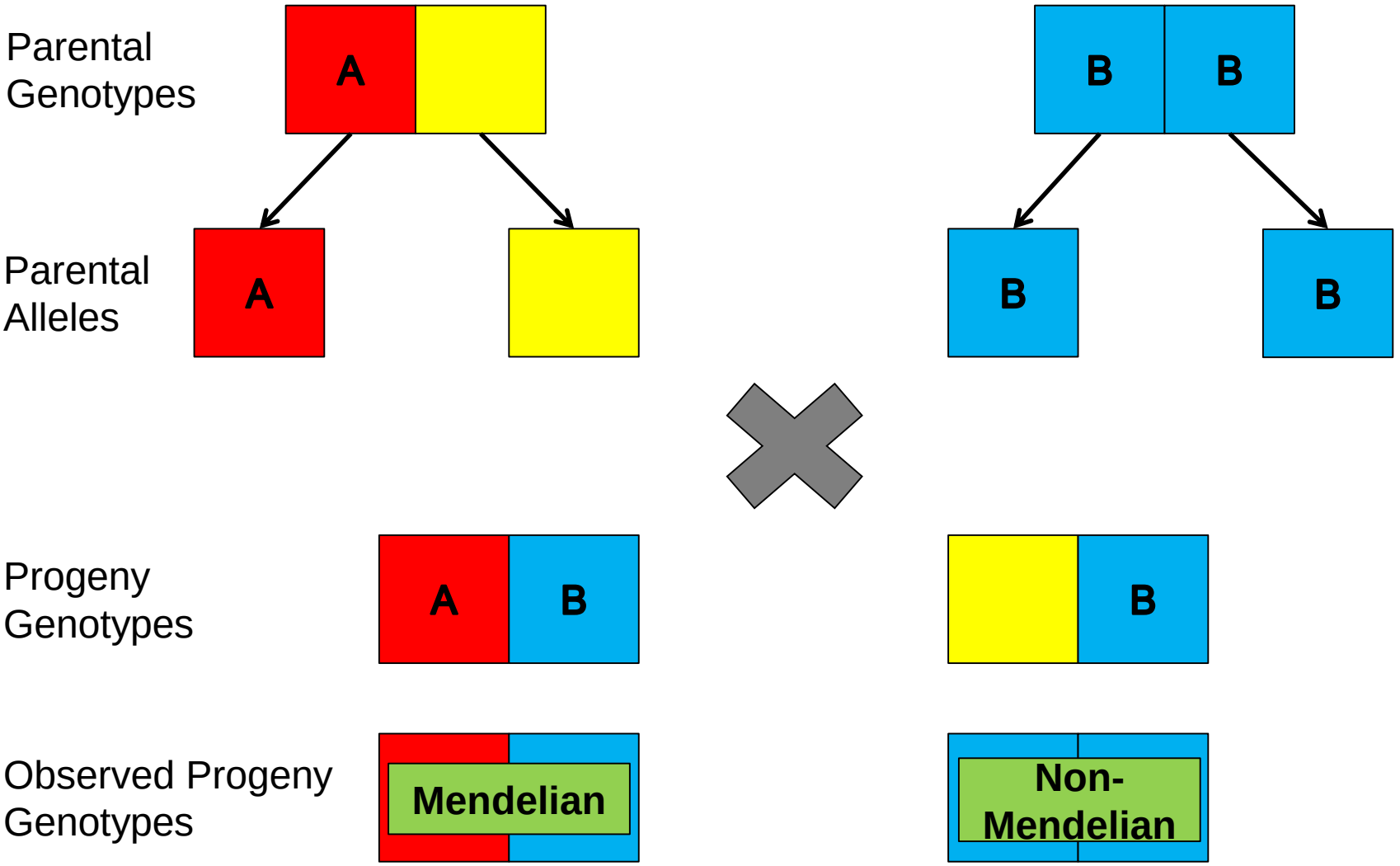
Breed comparison – duplication length



Breed comparison – duplication length



Non-Mendelian Inheritance



UCD DUBLIN

teagasc
Agriculture and Food Development Authority

Conclusions

- For most animals a small proportion of their genome were composed of CNVs.
- Small differences did exist between some breeds for mean length of CNVs, and mean number of CNVs per animal.
- The presence of CNVs can give rise to apparent Mendelian inconsistencies in SNP genotype data.