

# Model-based genotype and ancestry estimation for potential hybrids with mixed-ploidy

Vivaswat Shastry<sup>1</sup>, Paula Adams<sup>2</sup>, Dorothea Lindtke<sup>3</sup>, Elizabeth Mandeville<sup>4</sup>, Thomas Parchman<sup>5</sup>, Zach Gompert<sup>6</sup>, and C. Alex Buerkle<sup>1</sup>

<sup>1</sup>University of Wyoming

<sup>2</sup>The University of Alabama System

<sup>3</sup>University of Bern

<sup>4</sup>University of Guelph

<sup>5</sup>University of Nevada Reno

<sup>6</sup>Utah State University

August 3, 2020

## Abstract

Non-random mating among individuals can lead to spatial clustering of genetically similar individuals and population stratification. This deviation from panmixia is commonly observed in natural populations. Consequently, individuals can have parentage in single populations or involving hybridization between differentiated populations. Accounting for this mixture and structure is important when mapping the genetics of traits and learning about the formative evolutionary processes that shape genetic variation among individuals and populations. Stratified genetic relatedness among individuals is commonly quantified using estimates of ancestry that are derived from a statistical model. Development of these models for polyploid and mixed-ploidy individuals and populations has lagged behind those for diploids. Here, we extend and test a hierarchical Bayesian model, called **entropy**, which can utilize low-depth sequence data to estimate genotype and ancestry parameters in autopolyploid and mixed-ploidy individuals (including sex chromosomes and autosomes within individuals). Our analysis of simulated data illustrated the trade-off between sequencing depth and genome coverage and found lower error associated with low depth sequencing across a larger fraction of the genome than with high depth sequencing across a smaller fraction of the genome. The model has high accuracy and sensitivity as verified with simulated data and through analysis of admixture among populations of diploid and tetraploid *Arabidopsis arenosa*.

## Hosted file

entropy\_ploidy.pdf available at <https://authorea.com/users/348436/articles/473747-model-based-genotype-and-ancestry-estimation-for-potential-hybrids-with-mixed-ploidy>