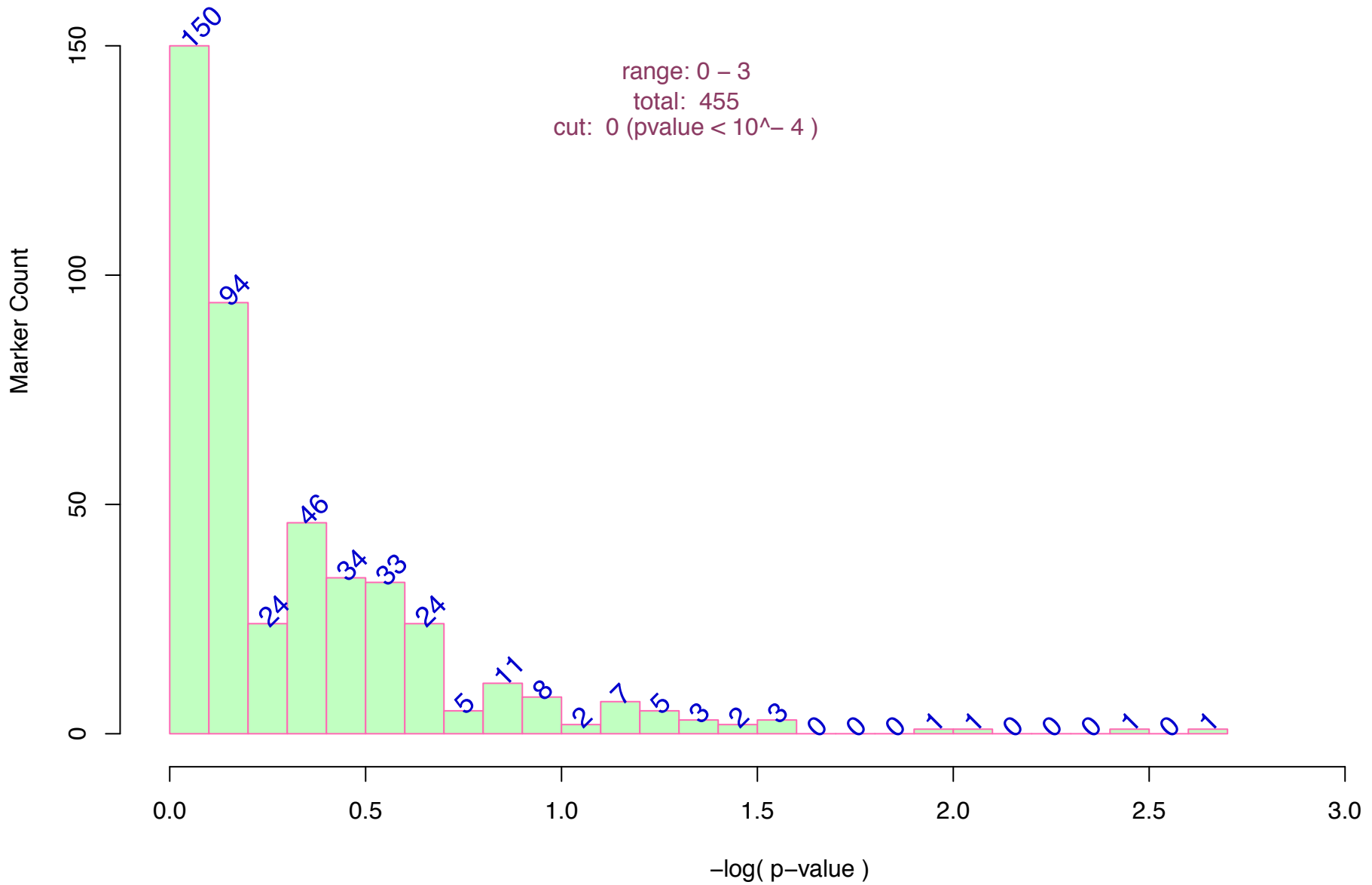


## GENOTYPE PROCESSING METHODS.

A number of data processing steps were taken to regularize the HGI genotypic data separately for each subset, imposing reference map marker order throughout. PEDSTATS (1) was run to identify and drop markers with Hardy-Weinberg equilibrium (HWE) p-value  $< 0.0001$ . In-house scripts were used to remove markers missing  $> 10\%$  and zero individuals missing  $> 20\%$  of genotypes. Study 7 was only partially genotyped on the X chromosome, and therefore chromosome X markers were retained in this sample with up to 14% missingness. RELCHECK (2) was then run to verify and adjust pedigree structures. Mendel errors were assessed using MENDEL (3), including total number of errors per family and per marker, and the genotypes of all individuals in a family for a marker giving an error were set to missing. Checks for Mendel errors, missingness, and HWE were repeated (in this order) in case any new markers or individuals crossed our criteria for removal during the previous steps. At each step, histograms (below) were visually inspected to check that cutoffs were appropriate given the observed distributions (e.g., that they successfully captured outliers). Inspection of the histograms led us to raise the missingness thresholds to 25% for individuals and 15% for markers during the final iteration. Additional checks at this stage included looking for duplicated individuals across pedigrees using RELCHECK and verifying concordance between recorded sex and X-chromosome genotypes. When an individual's sex could not be resolved by genotyped mate and offspring (without RELCHECK problems), the X-genotypes were zeroed and the phenotype of the individual was set to unknown to protect against undetectable sample swaps. Finally, we constructed subset-specific maps based on the data as processed to this point, to allow for differing inter-marker genetic distances and marker allele frequencies across studies and, in particular, across population groups. For the microsatellite data, marker-to-marker genetic distances were estimated using KELVIN (4), and marker allele frequencies were estimated by maximum likelihood estimation using MENDEL for all data sets, including those for which frequencies had been included in the HGI files. For the SNP data, marker allele frequencies were similarly estimated, but physical locations were interpolated directly onto the reference map.

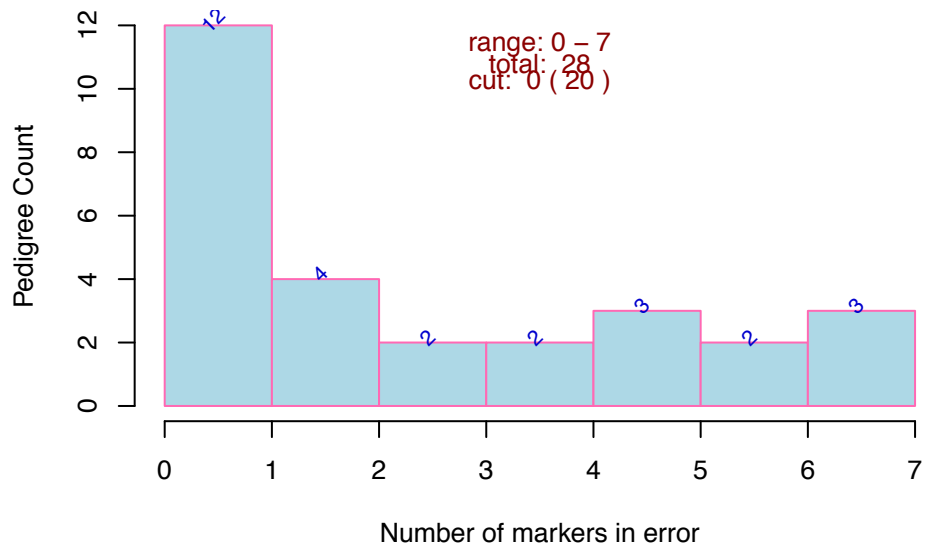
1. Wigginton JE, Abecasis GR. PEDSTATS: descriptive statistics, graphics and quality assessment for gene mapping data. *Bioinformatics*. 2005;21(16):3445-7.
2. Broman KW, Murray JC, Sheffield VC, White RL, Weber JL. Comprehensive human genetic maps: Individual and sex-specific variation in recombination. *Am J Hum Genet*. 1998;63:861-689.
3. Lange K, Cantor RM, Horvath S, Perola M, Sabatti C, Sinsheimer J, et al. Mendel version 4.0: A complete package for the exact genetic analysis of discrete traits in pedigree and population data sets. *Am J Hum Genet*. 2001;69(suppl):A1886.
4. Vieland VJ, Huang Y, Seok SC, Burian J, Catalyurek U, O'Connell J, et al. KELVIN: a software package for rigorous measurement of statistical evidence in human genetics. *Hum Hered*. 2011;72(4):276-88.

# Study 1 African American Hardy-Weinberg Errors

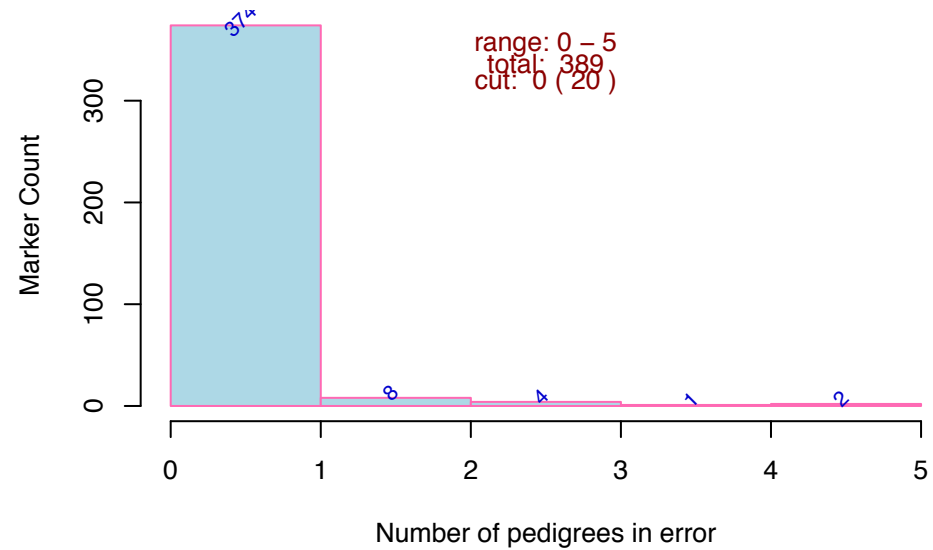


# Study 1 African American Initial Mendel Errors

## Pedigrees

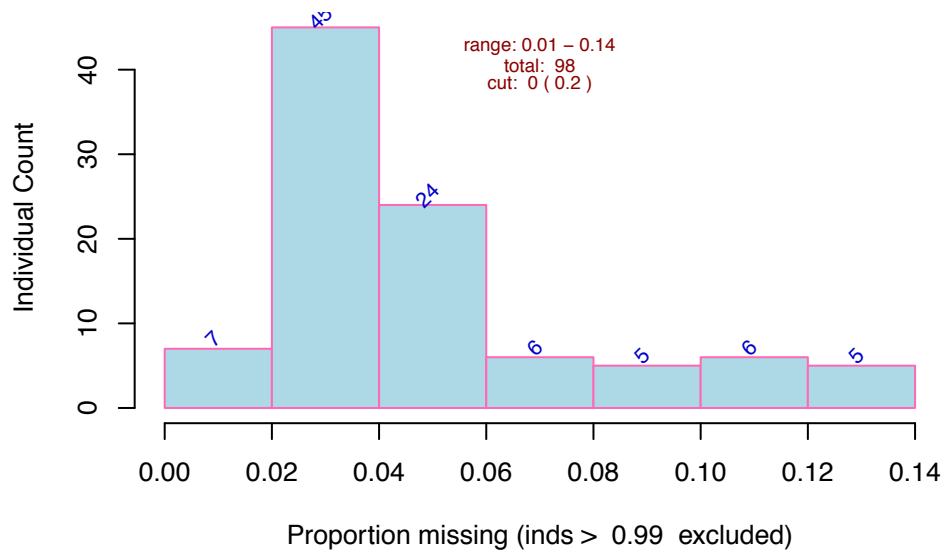


## Markers

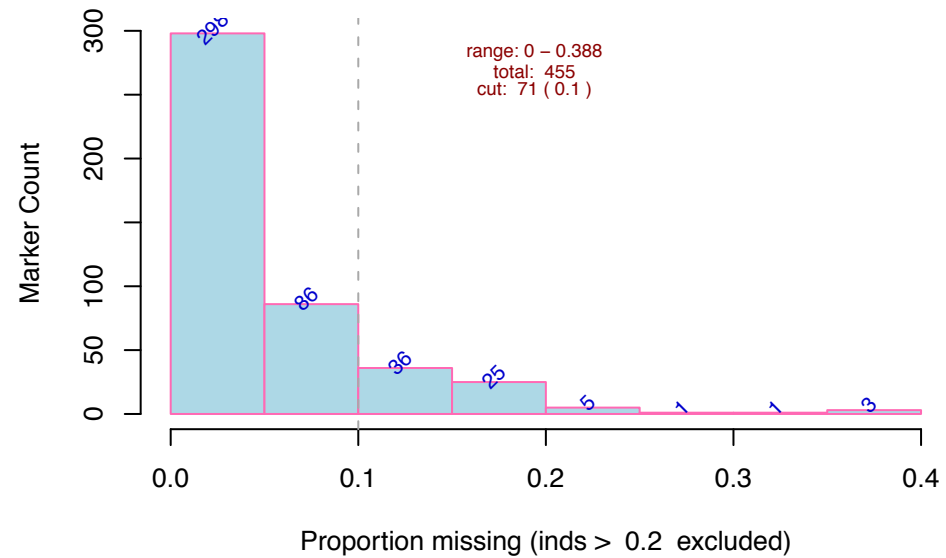


# Study 1 African American Initial Missingness

## Individuals

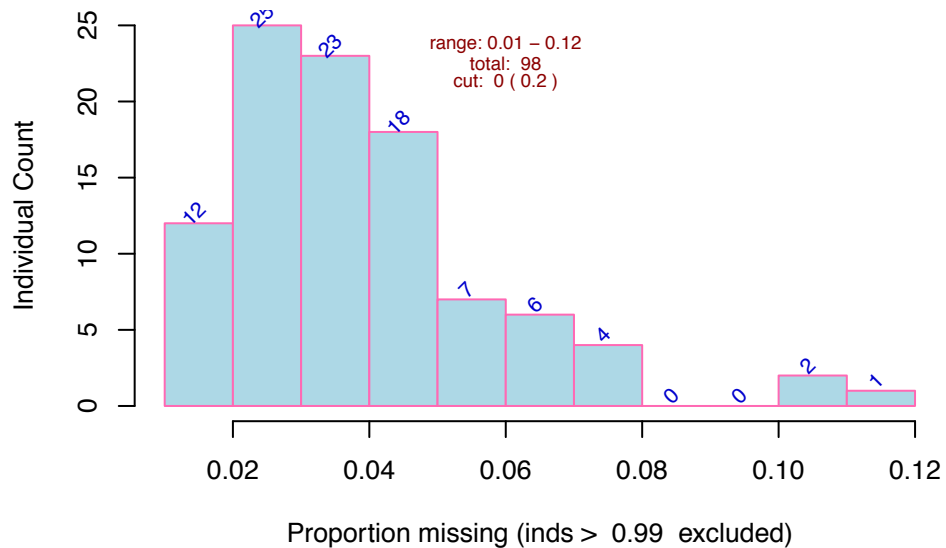


## Markers

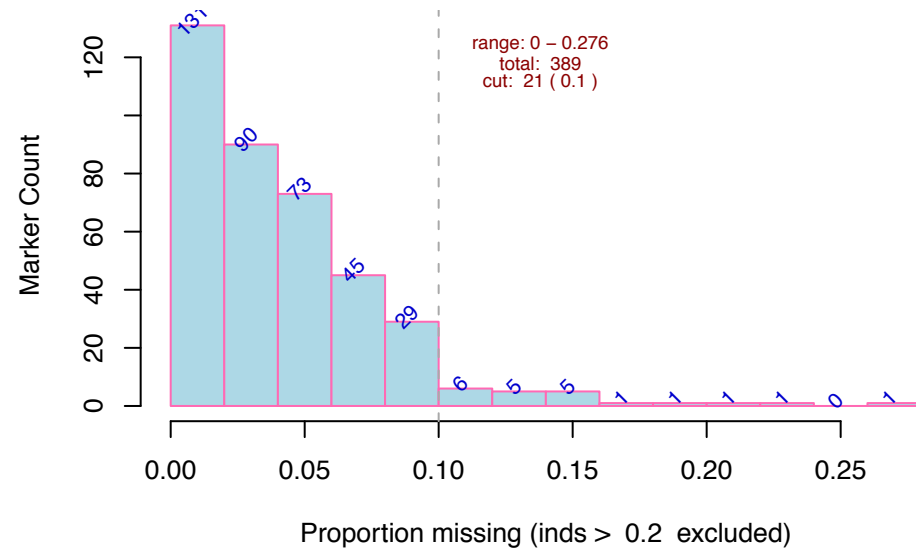


# Study 1 African American Repeat Missingness

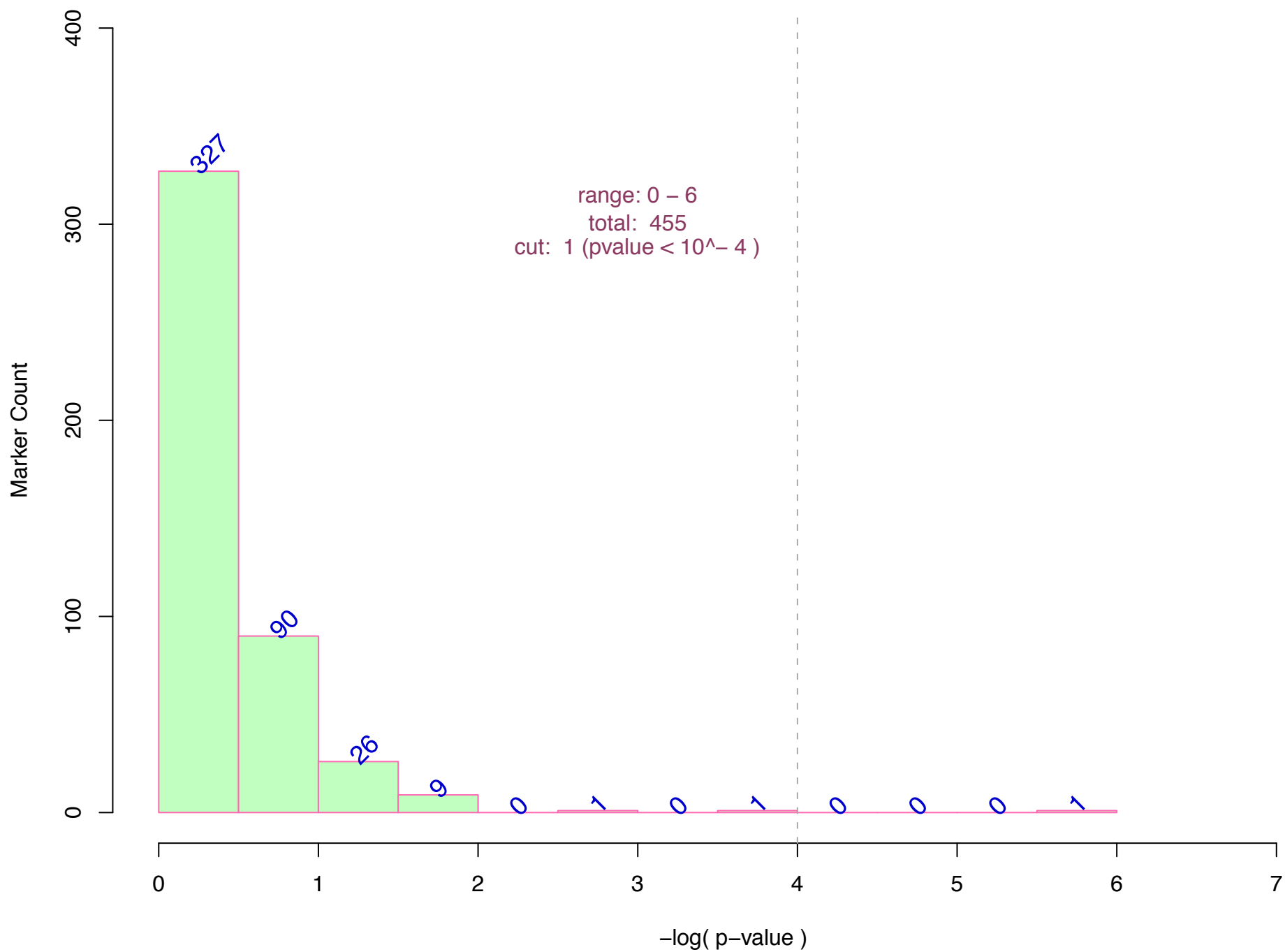
## Individuals



## Markers

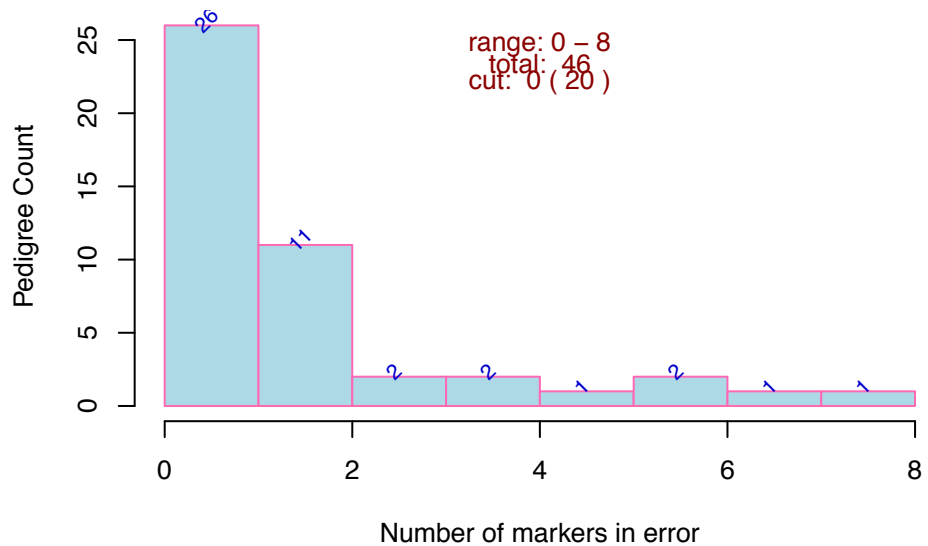


# Study 1 European American Hardy–Weinberg Errors

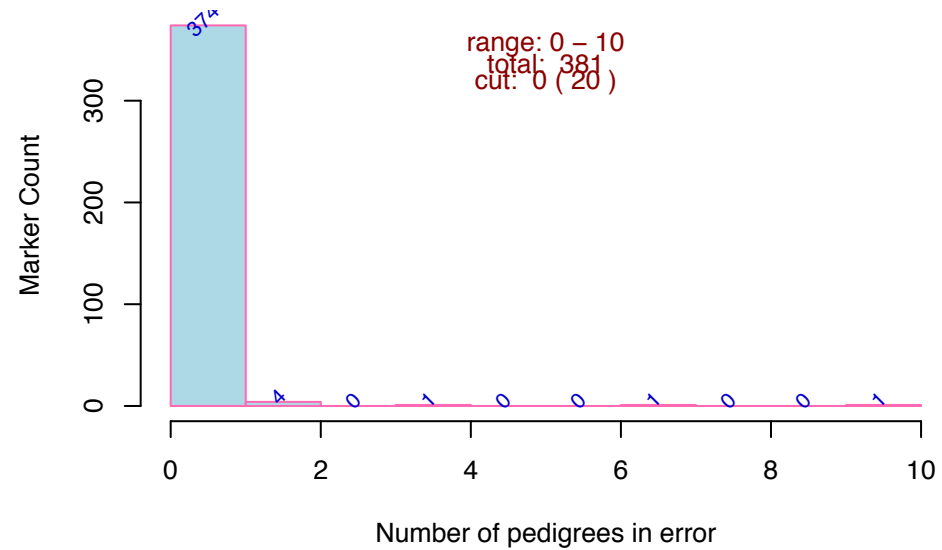


# Study 1 European American Initial Mendel Errors

## Pedigrees

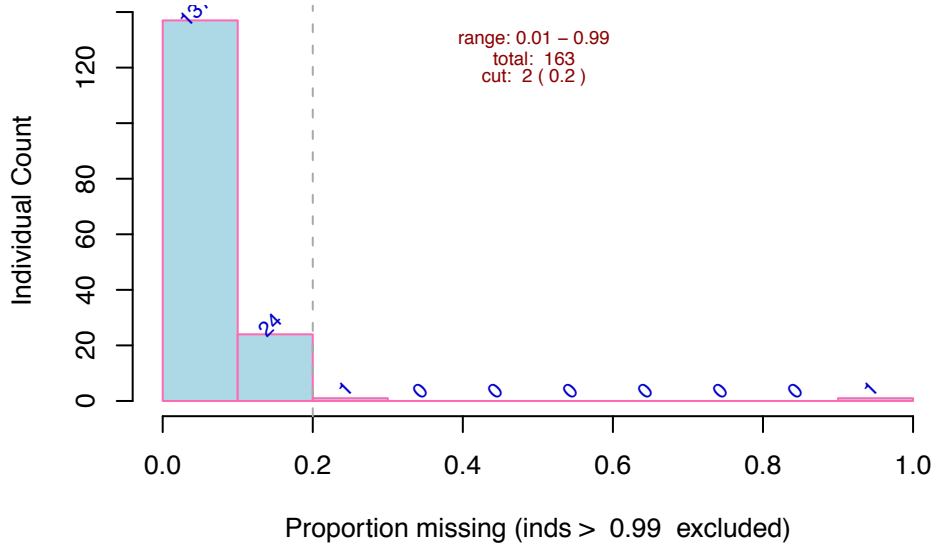


## Markers

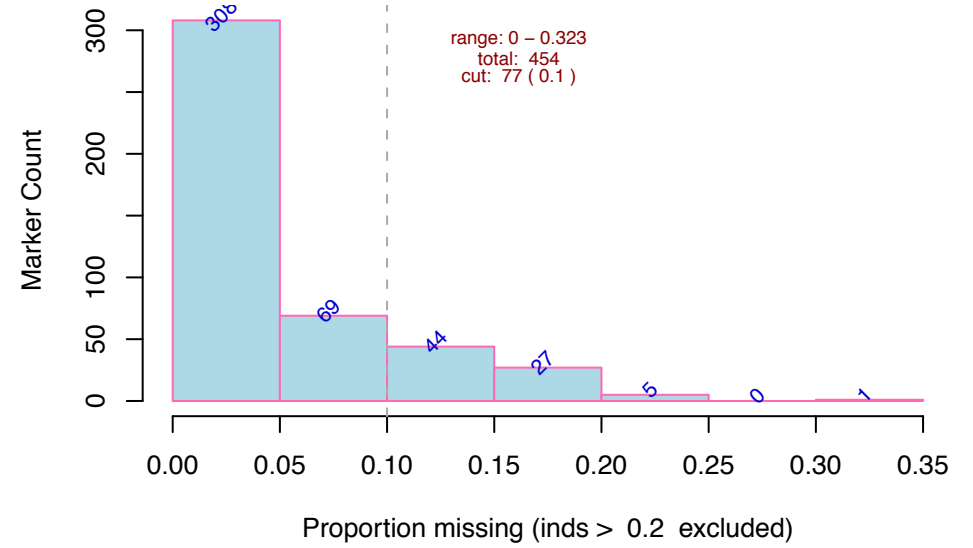


# Study 1 European American Initial Missingness

## Individuals



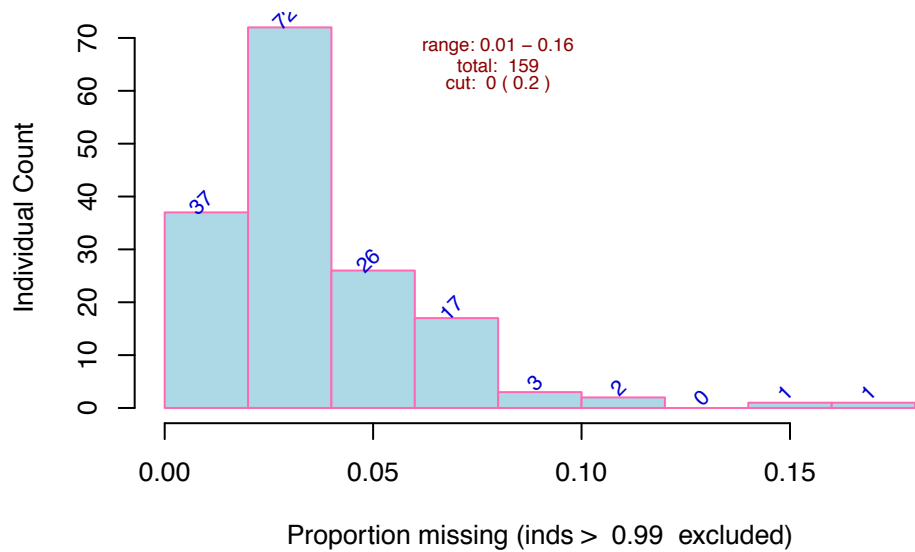
## Markers



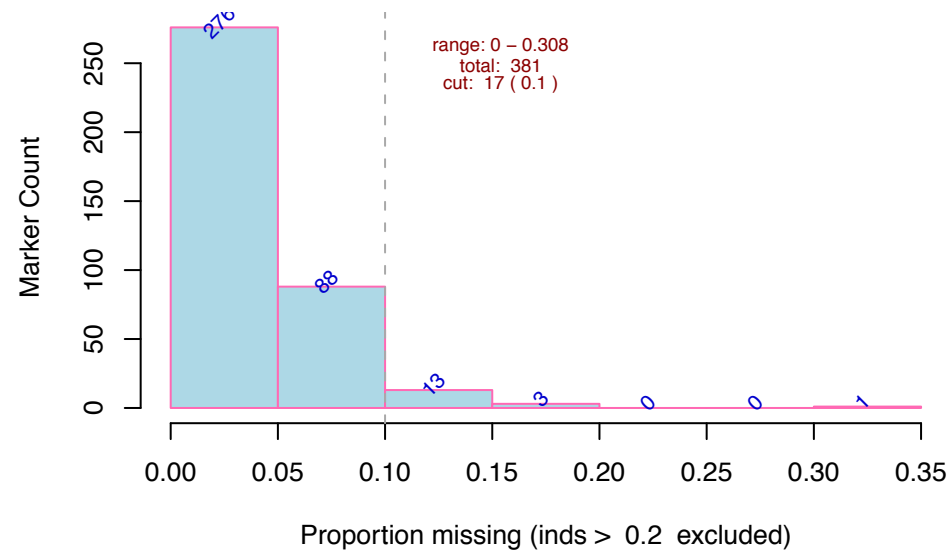


# Study 1 European American Repeat Missingness

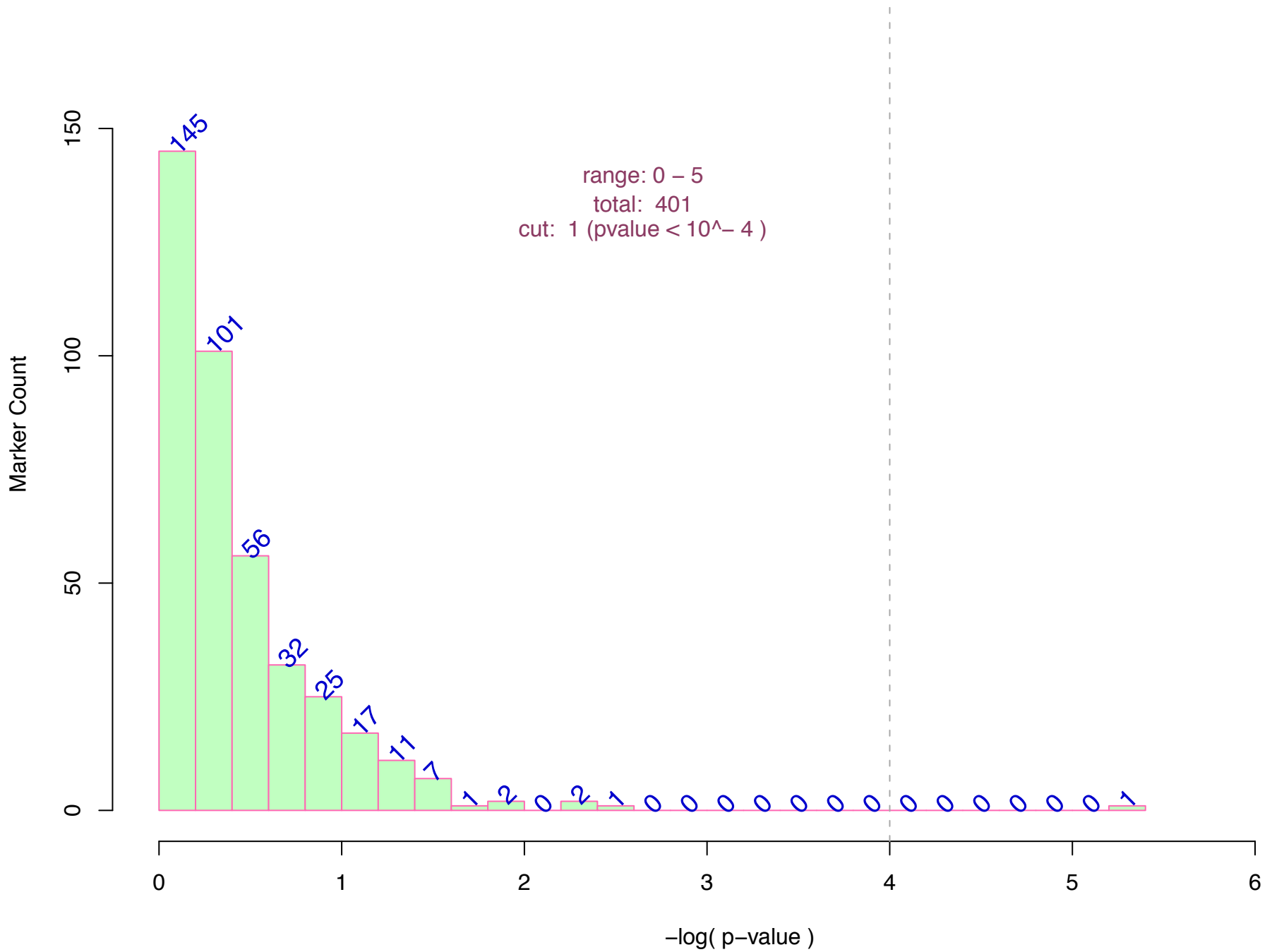
## Individuals



## Markers

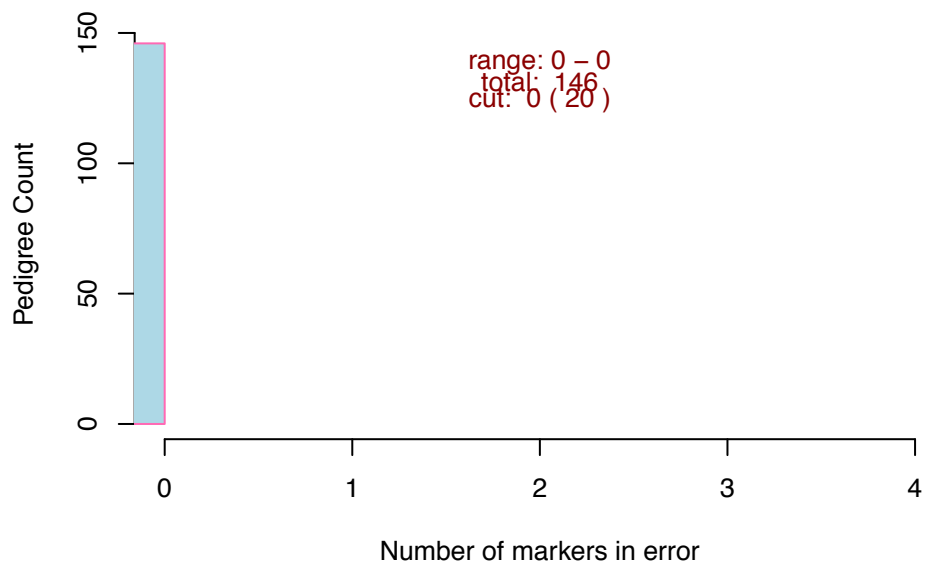


# Study 2 African American Hardy-Weinberg Errors

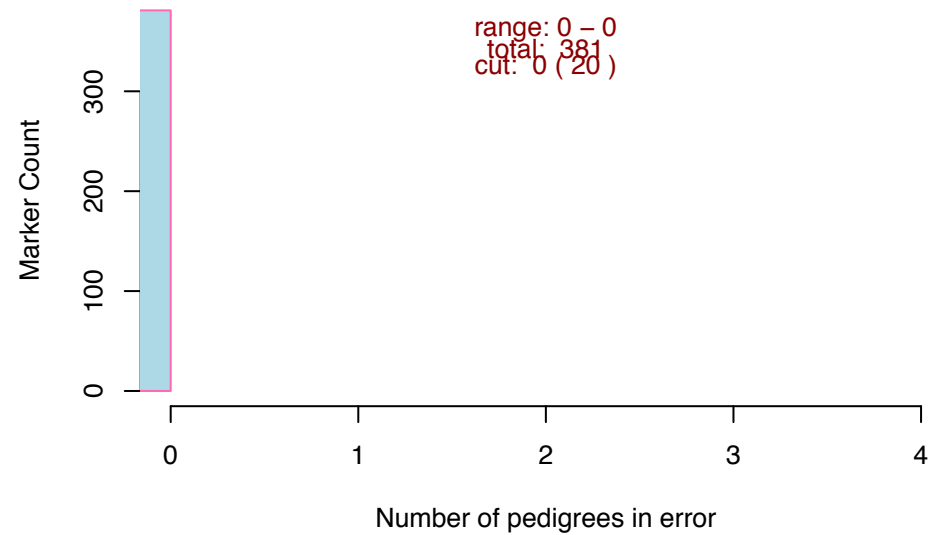


# Study 2 African American Initial Mendel Errors

## Pedigrees

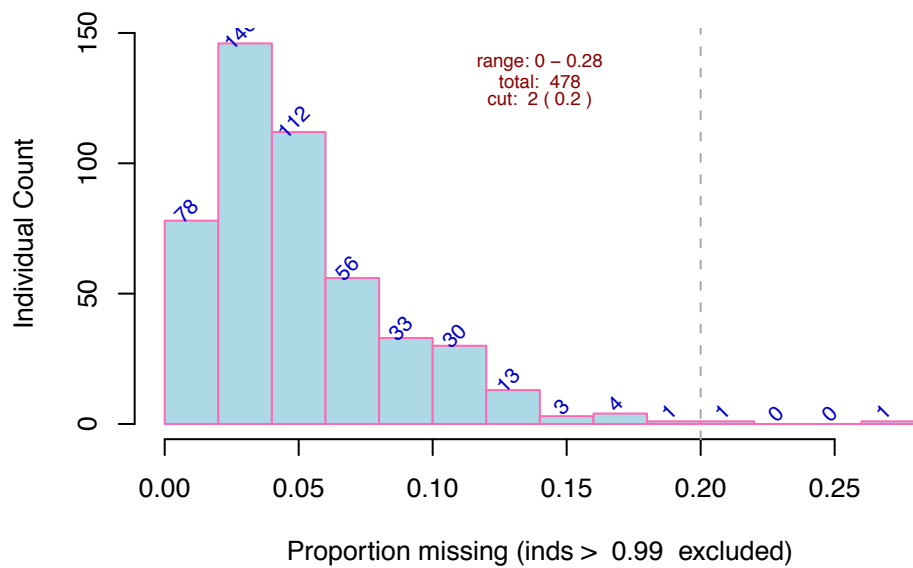


## Markers

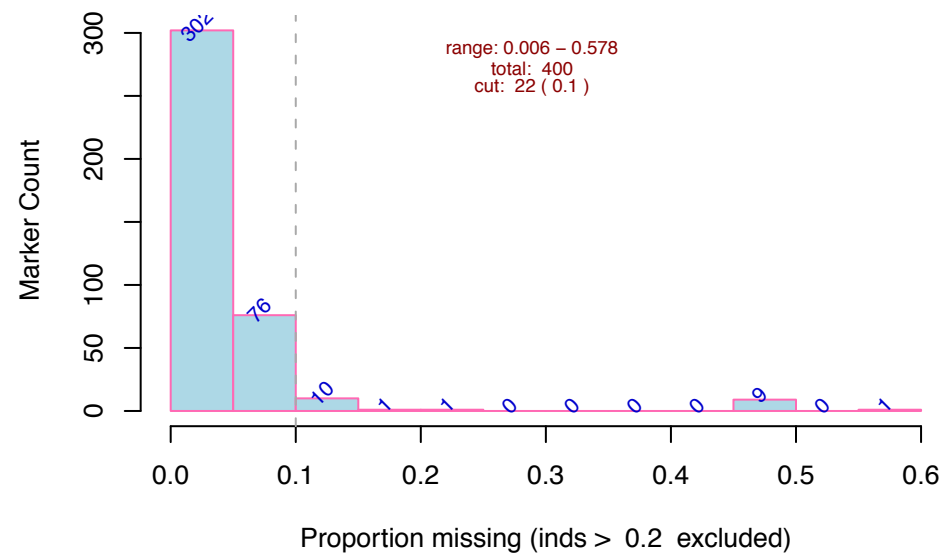


# Study 2 African American Initial Missingness

## Individuals

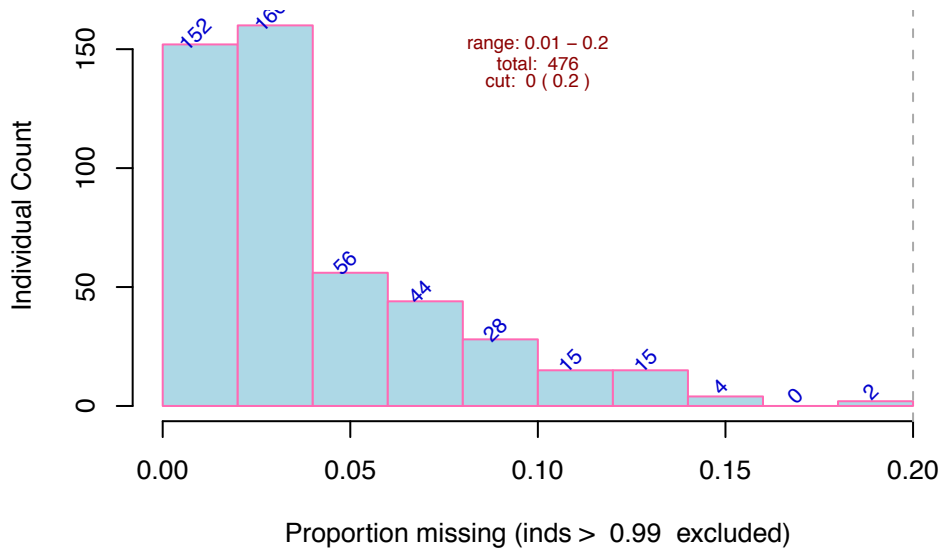


## Markers

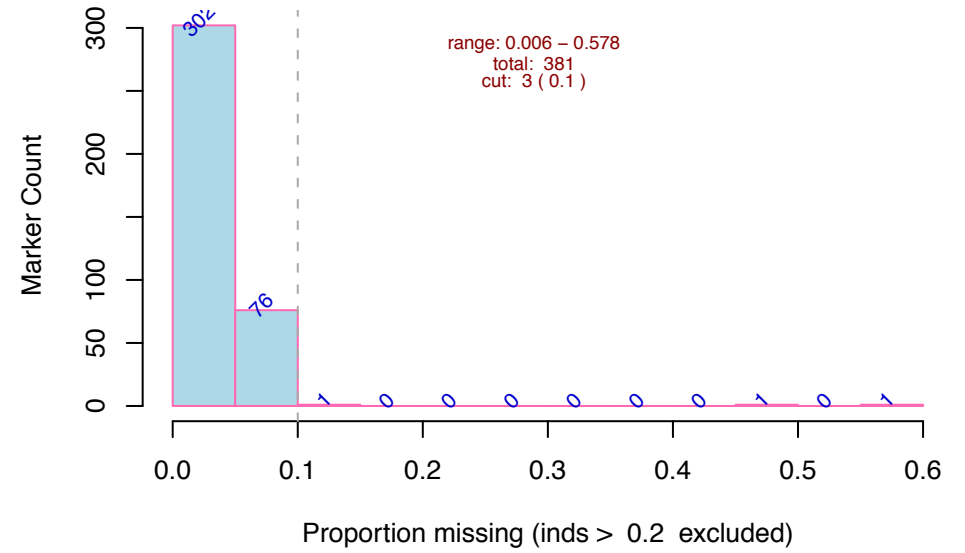


# Study 2 African American Repeat Missingness

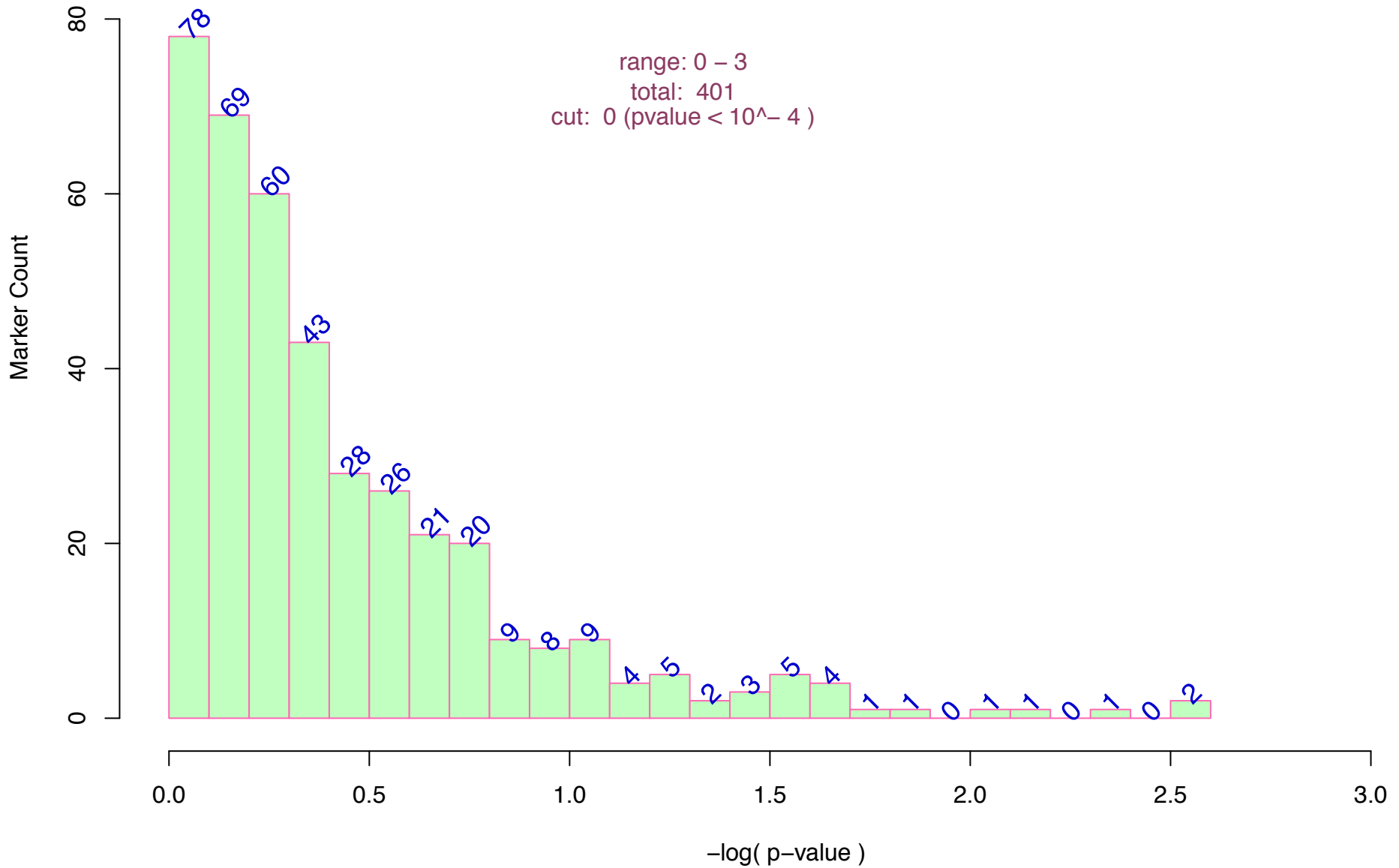
## Individuals



## Markers

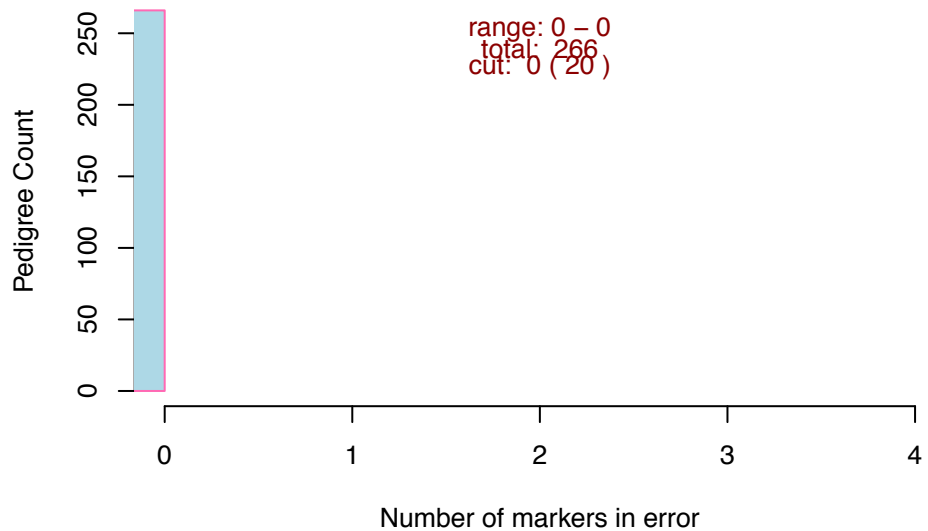


## Study 2 European American Hardy–Weinberg Errors

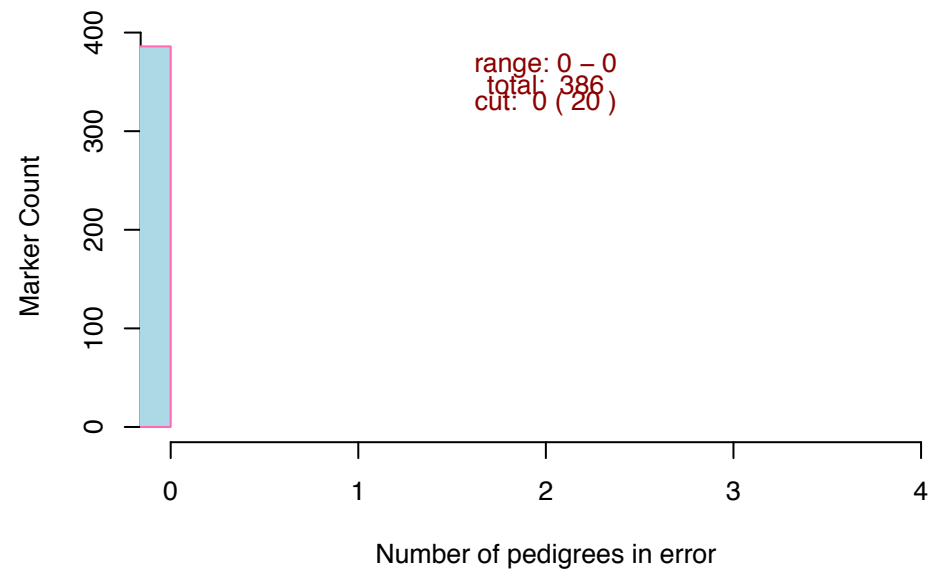


# Study 2 European American Initial Mendel Errors

## Pedigrees

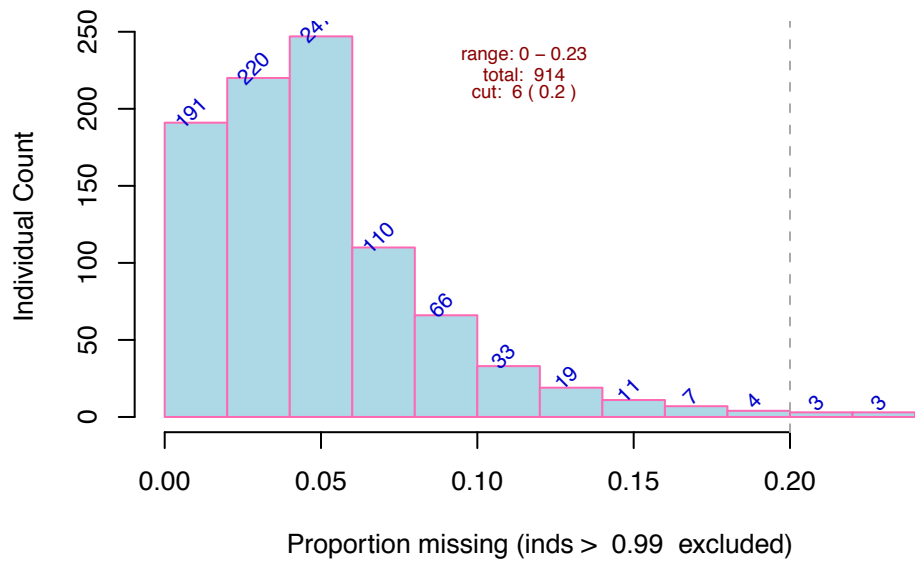


## Markers

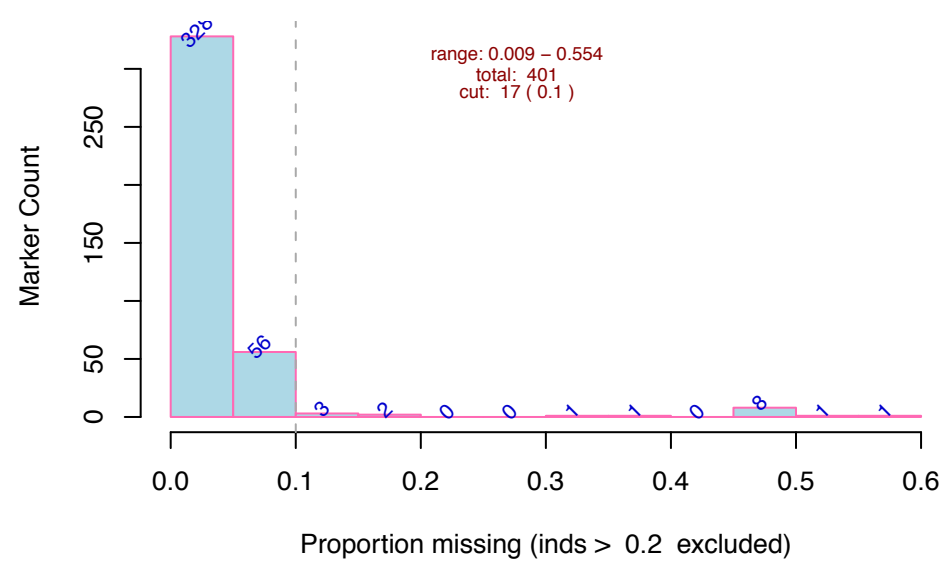


# Study 2 European American Initial Missingness

## Individuals



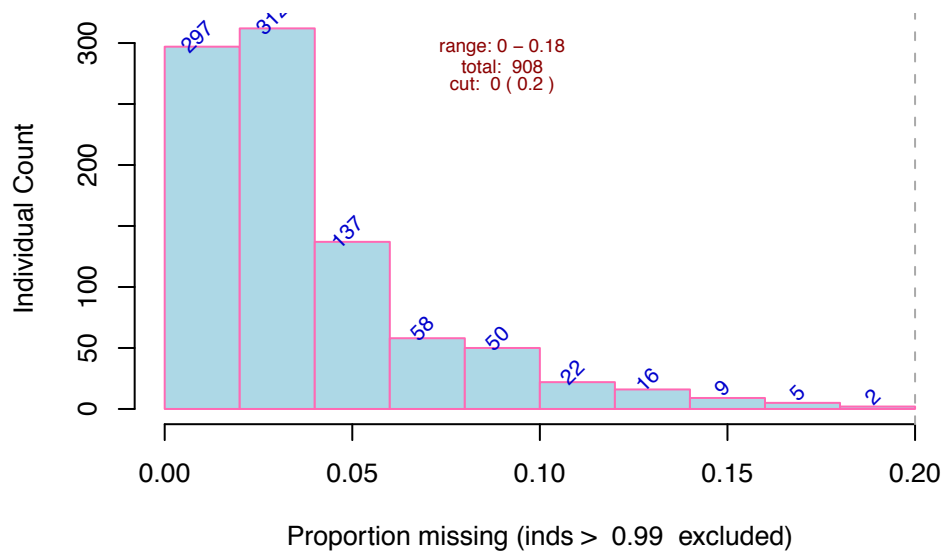
## Markers



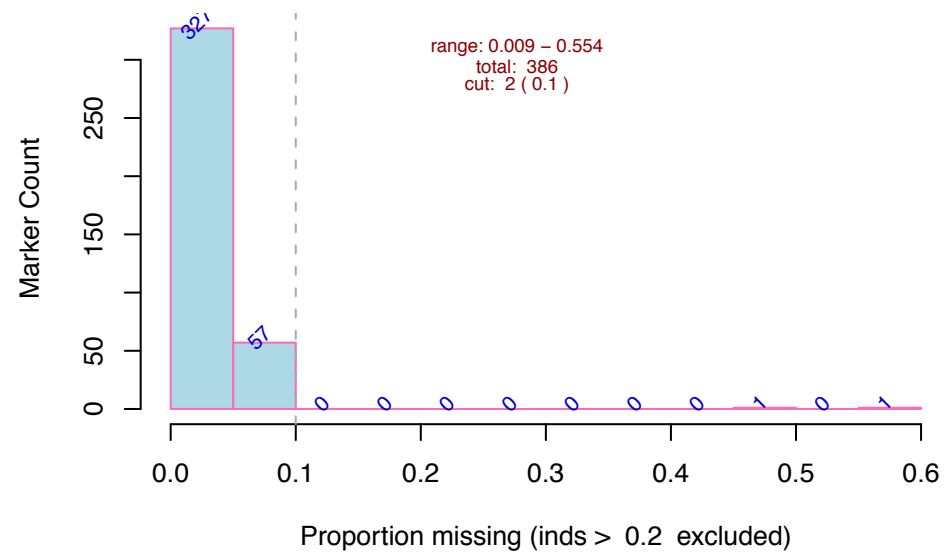


# Study 2 European American Repeat Missingness

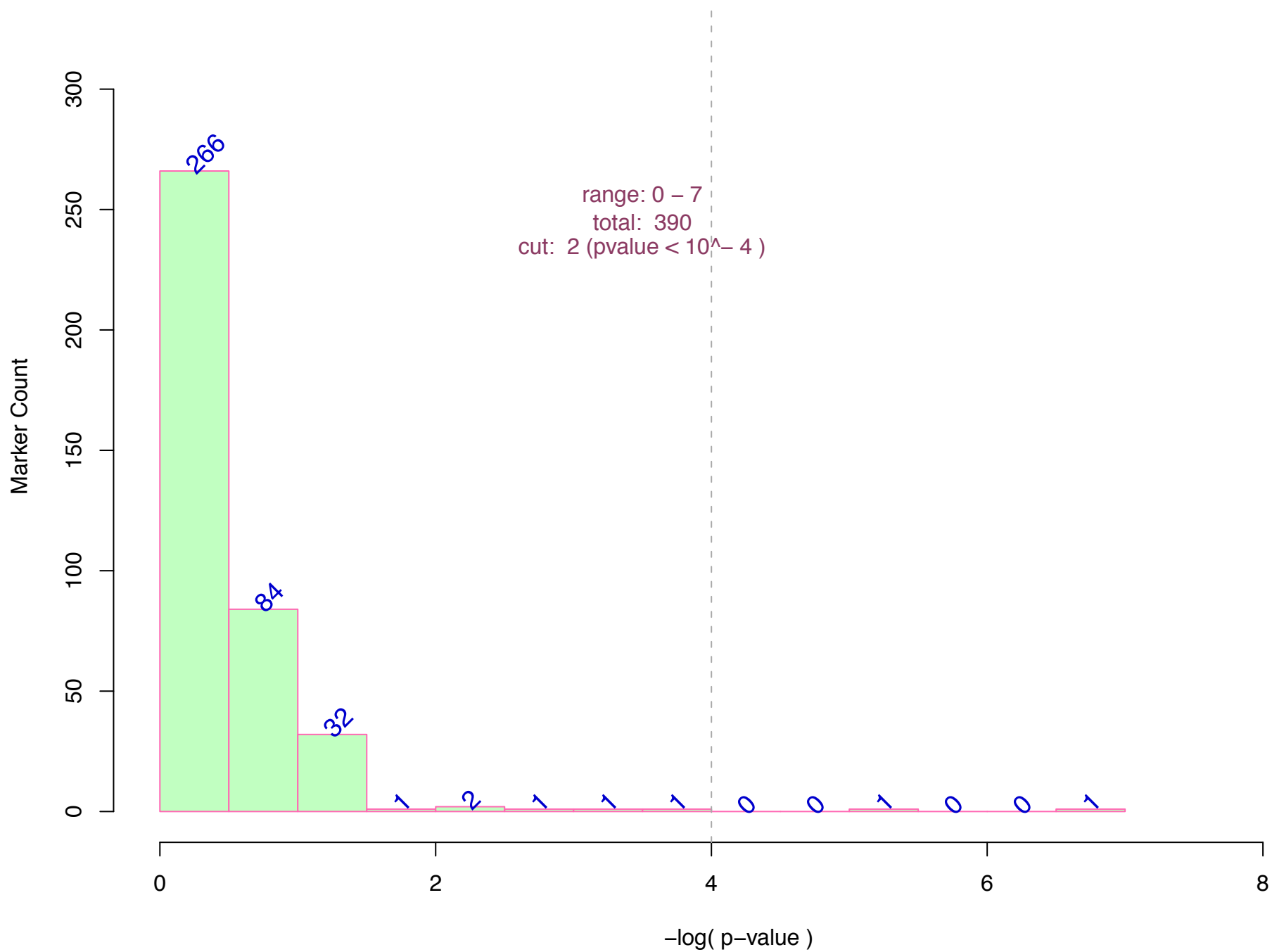
## Individuals



## Markers

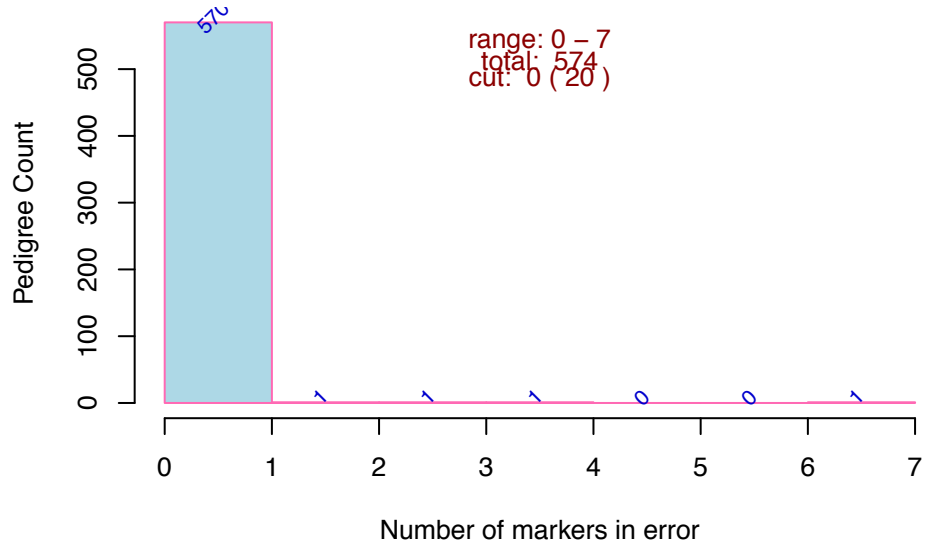


### Study 3 Han Chinese Hardy-Weinberg Errors

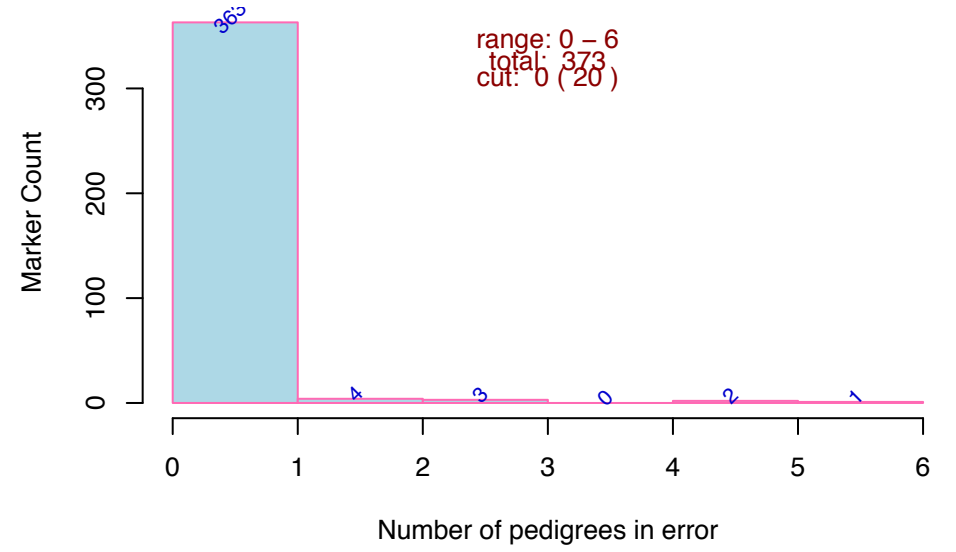


# Study 3 Han Chinese Initial Mendel Errors

## Pedigrees

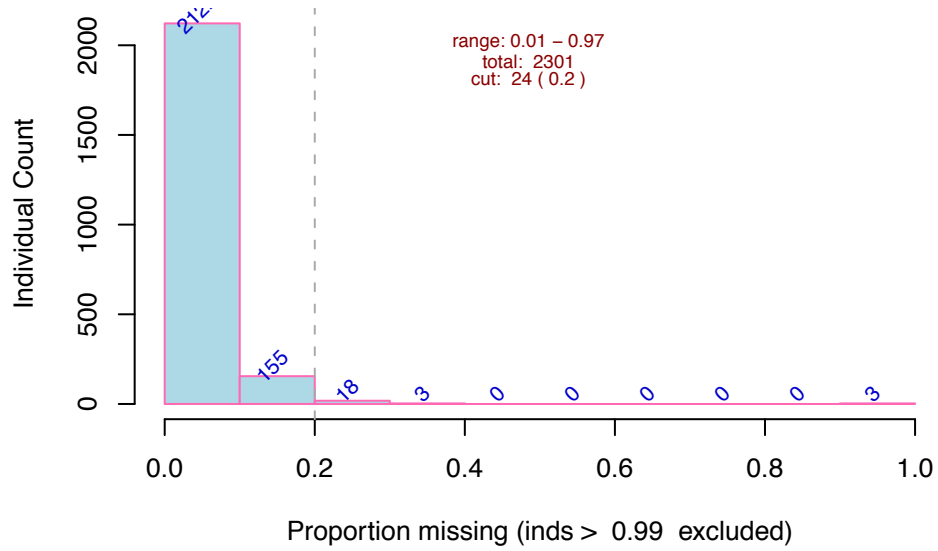


## Markers

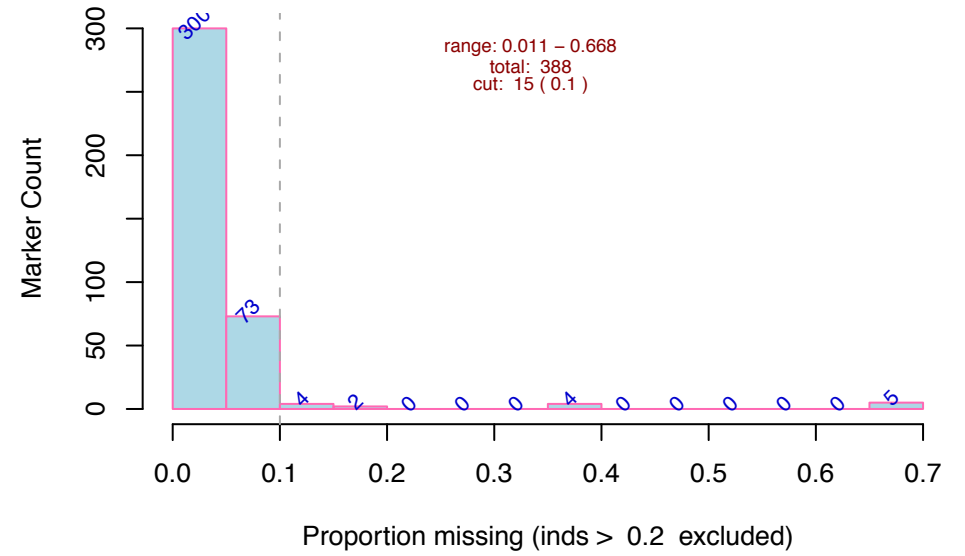


# Study 3 Han Chinese Initial Missingness

## Individuals

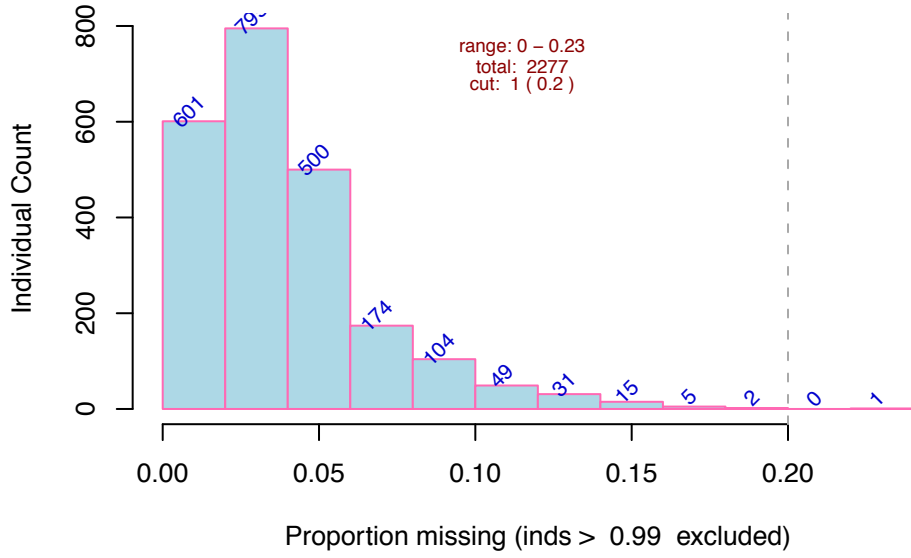


## Markers

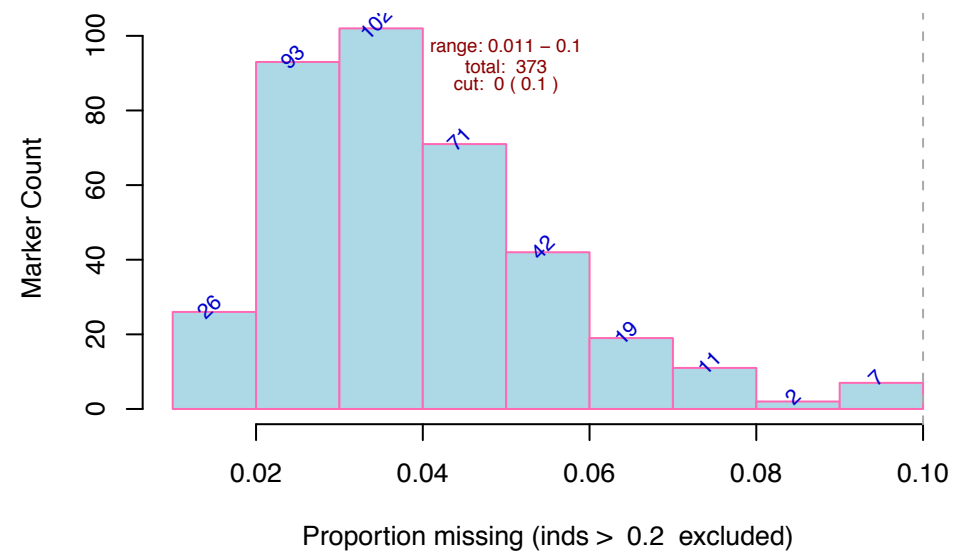


# Study 3 Han Chinese Repeat Missingness

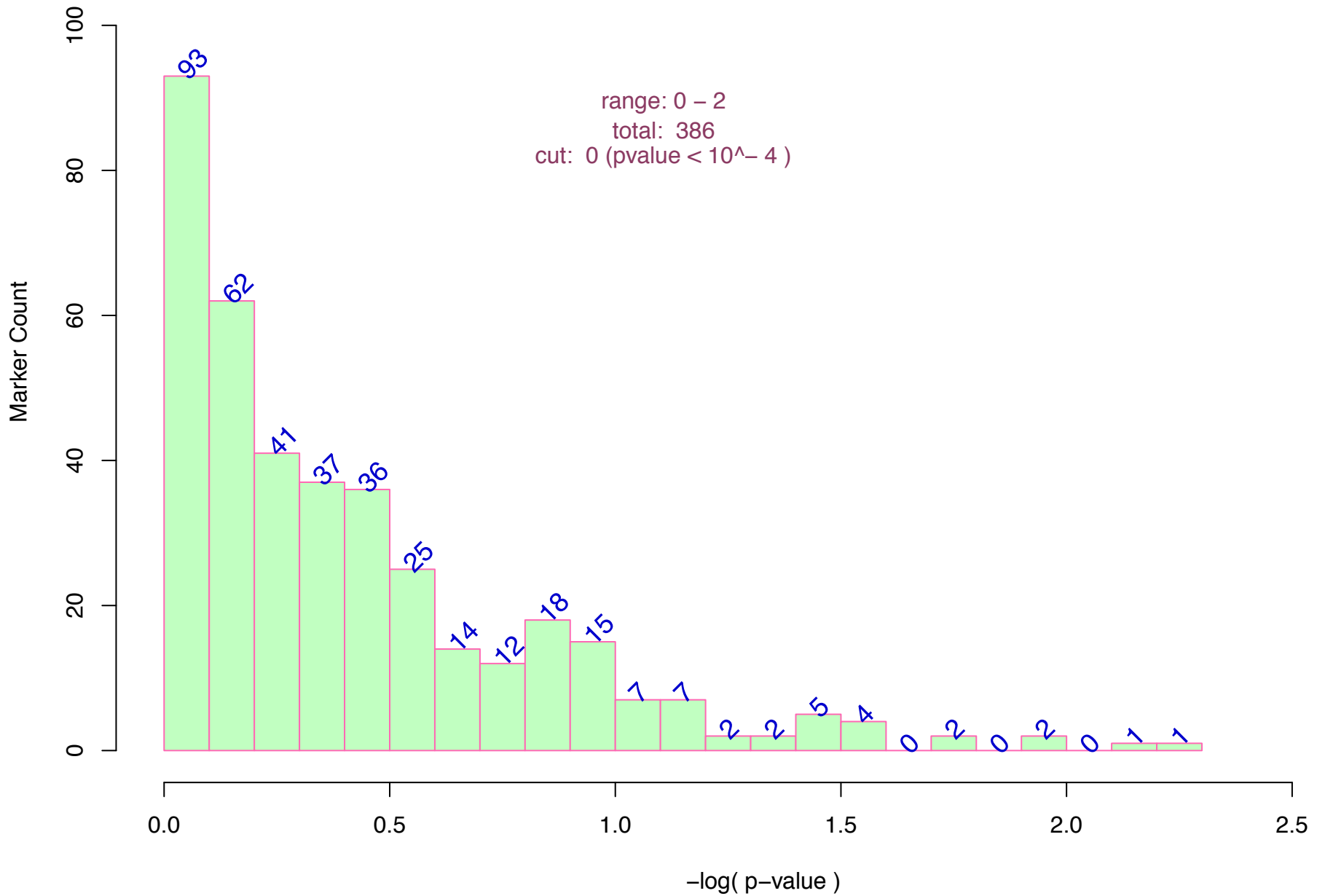
## Individuals



## Markers

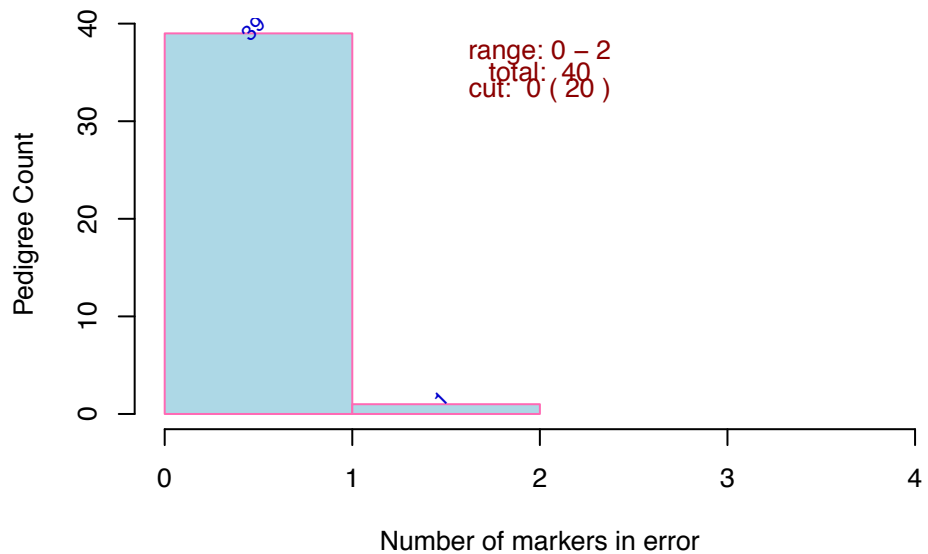


# Study 4 European American Hardy–Weinberg Errors

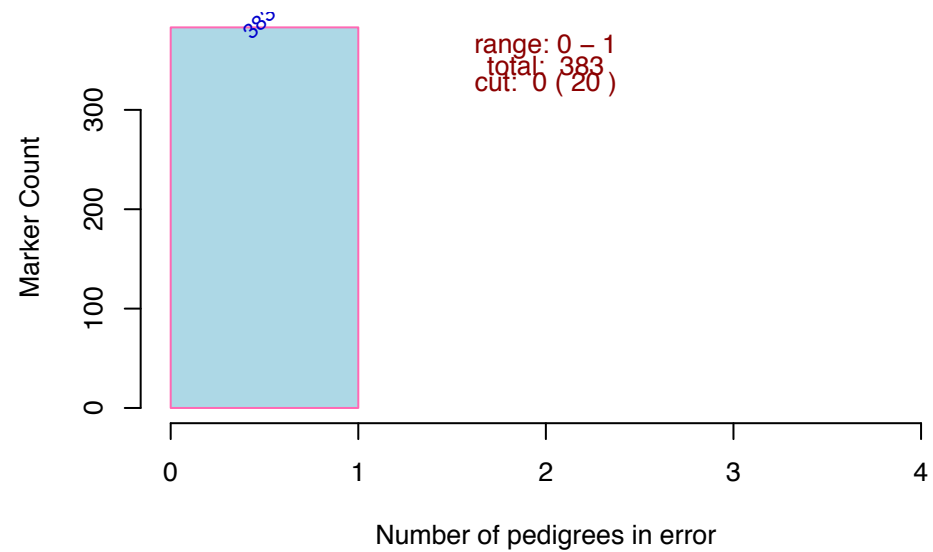


# Study 4 European American Initial Mendel Errors

## Pedigrees

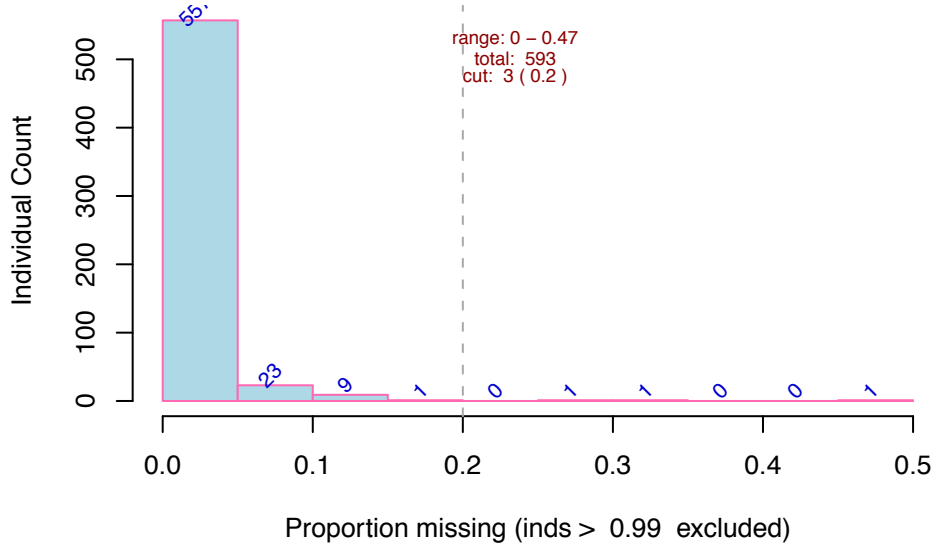


## Markers

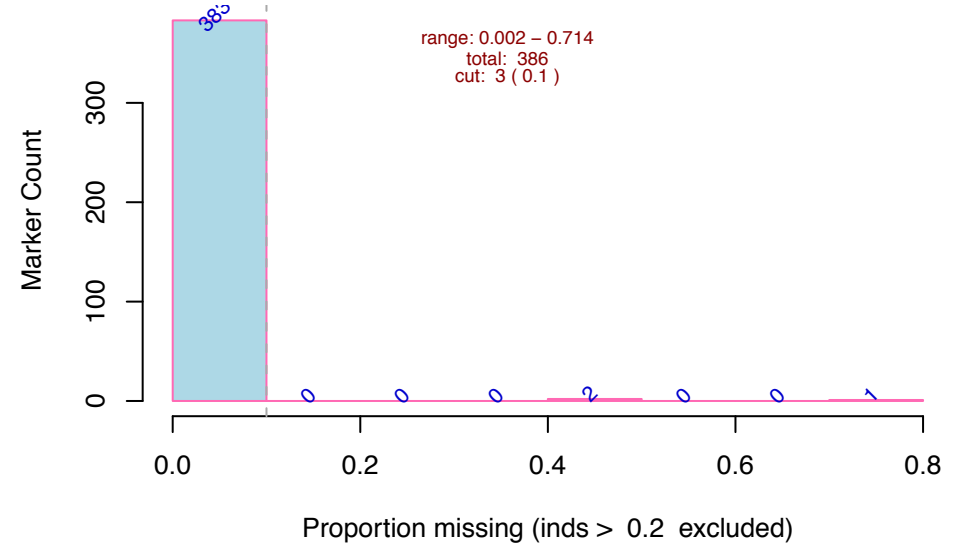


# Study 4 European American Initial Missingness

## Individuals



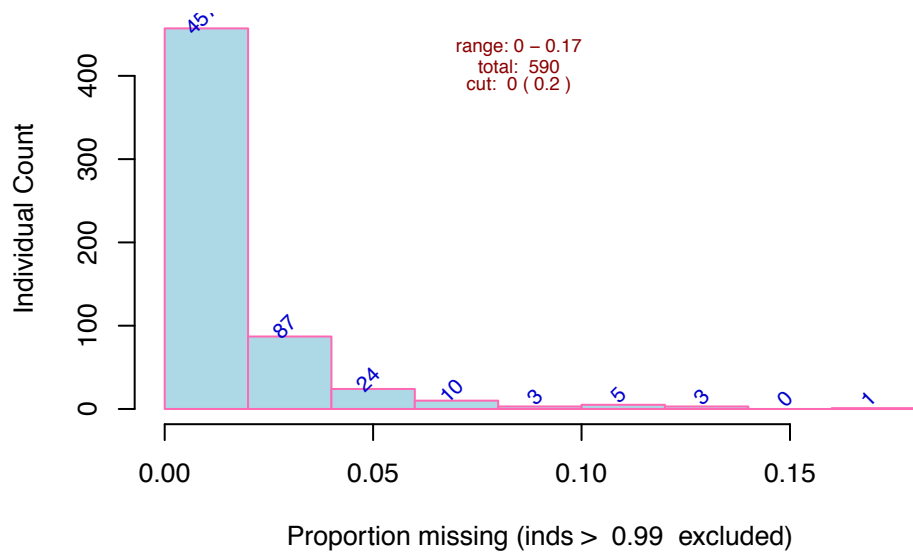
## Markers



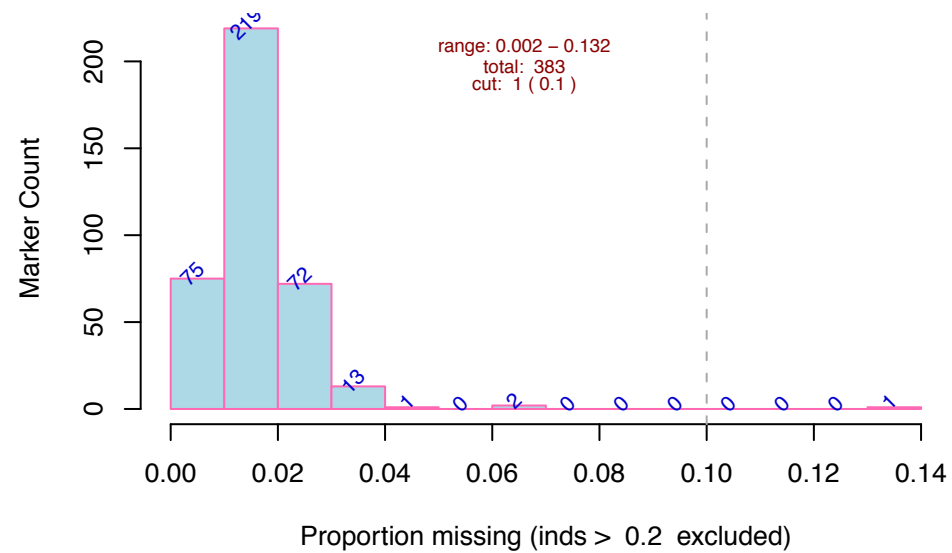


# Study 4 European American Repeat Missingness

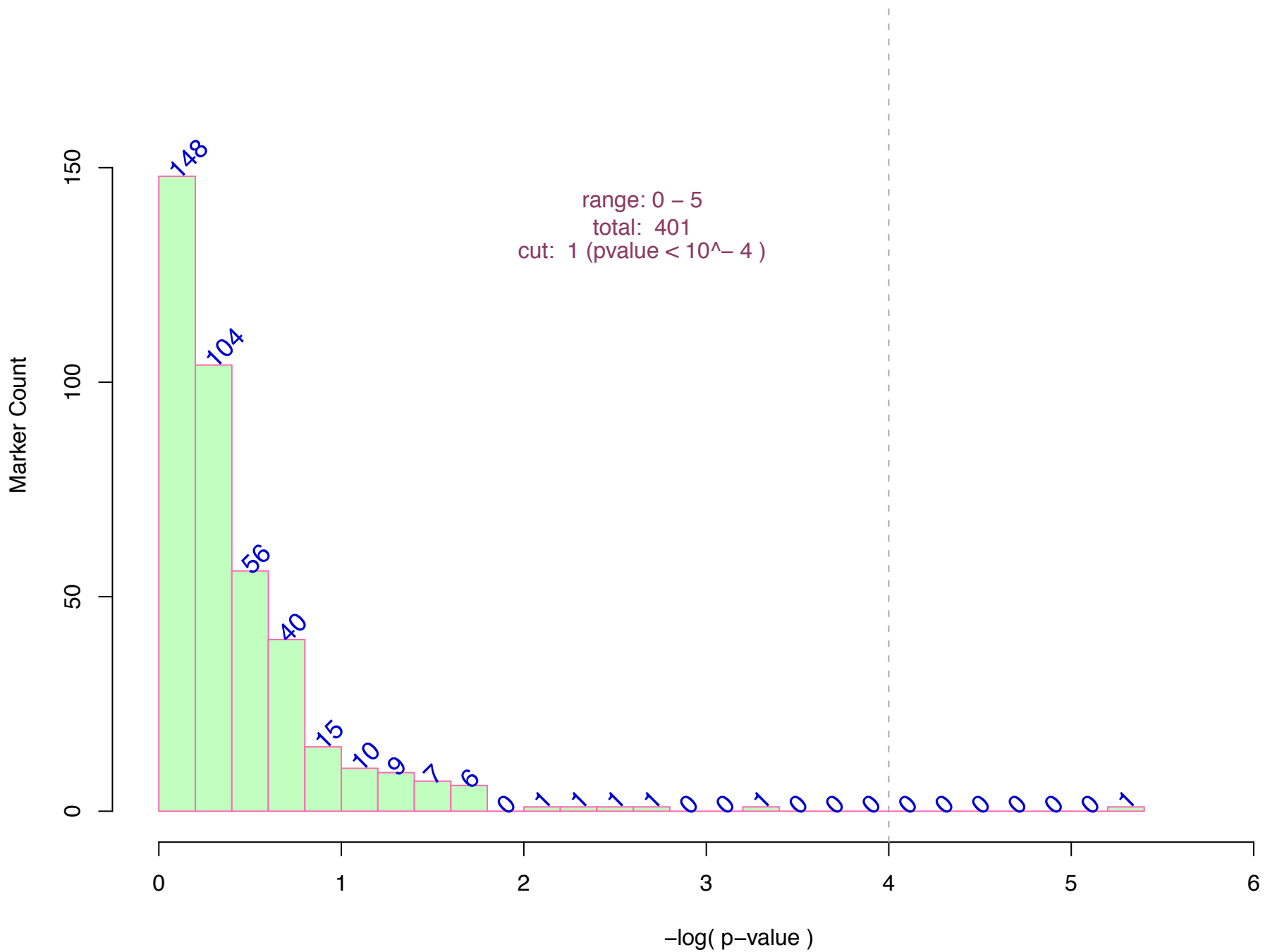
## Individuals



## Markers

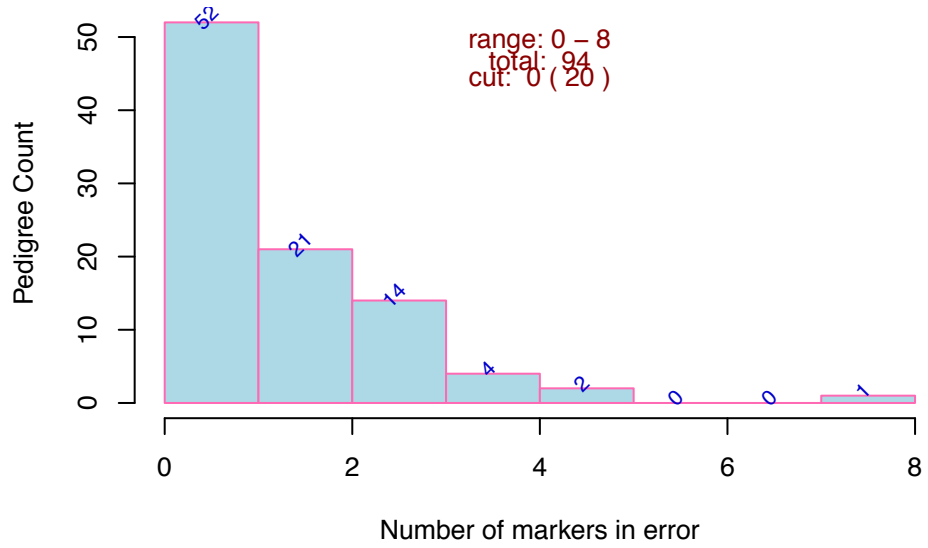


# Study 5 Hispanic Hardy-Weinberg Errors

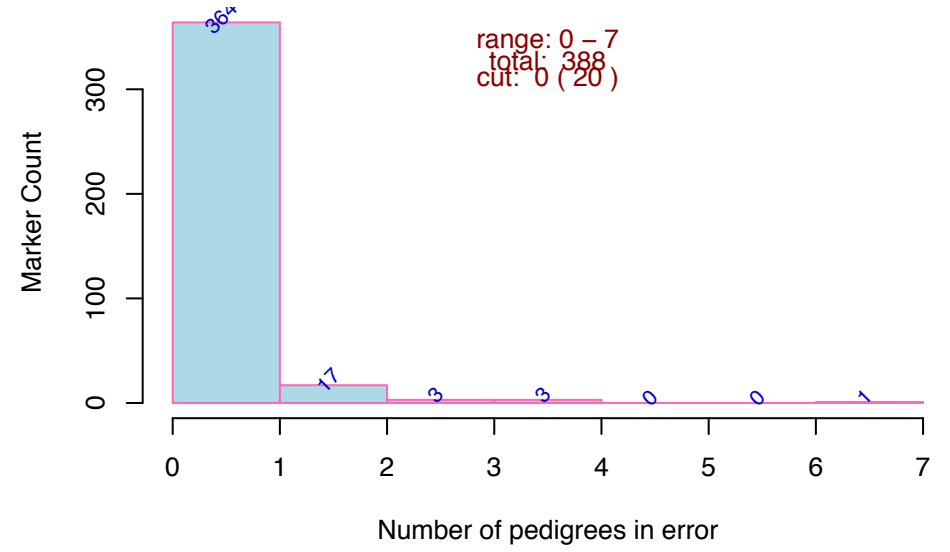


# Study 5 Hispanic Initial Mendel Errors

## Pedigrees

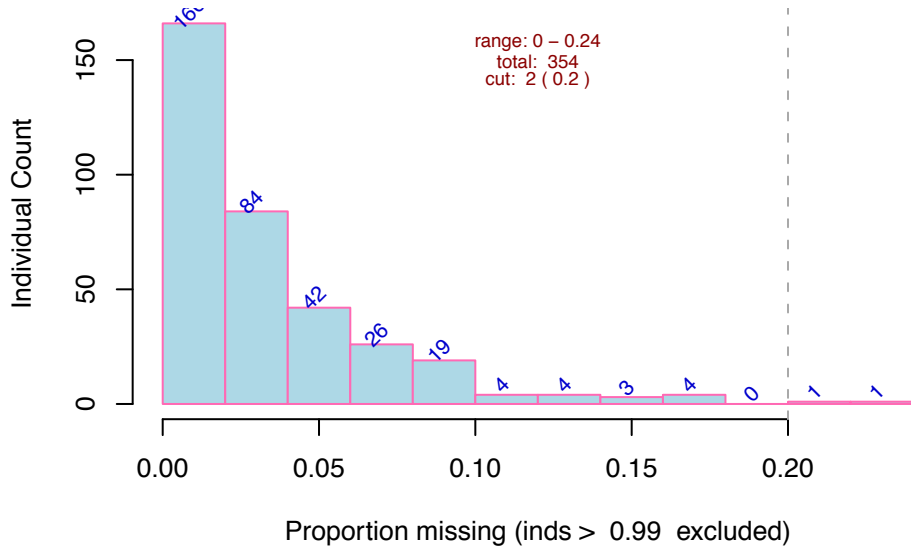


## Markers

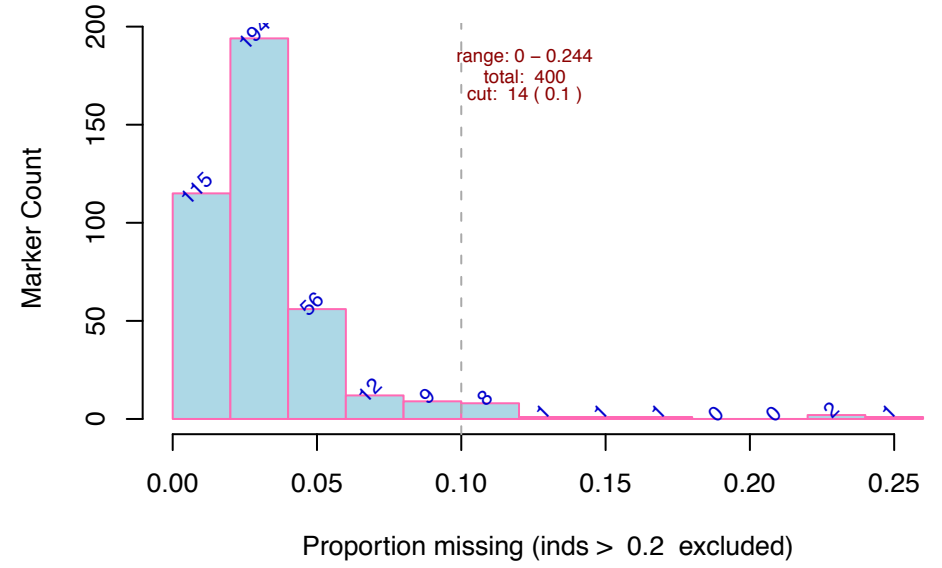


# Study 5 Hispanic Initial Missingness

## Individuals

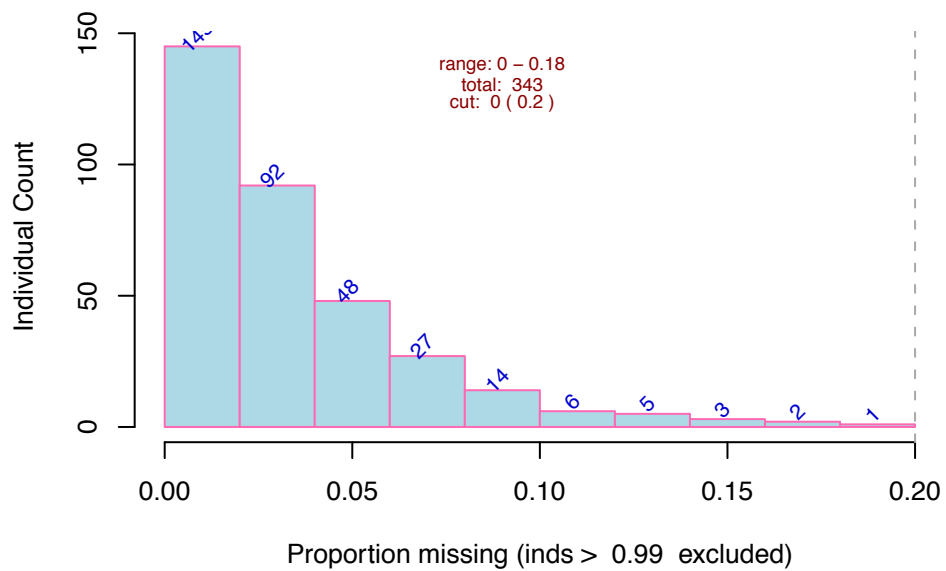


## Markers

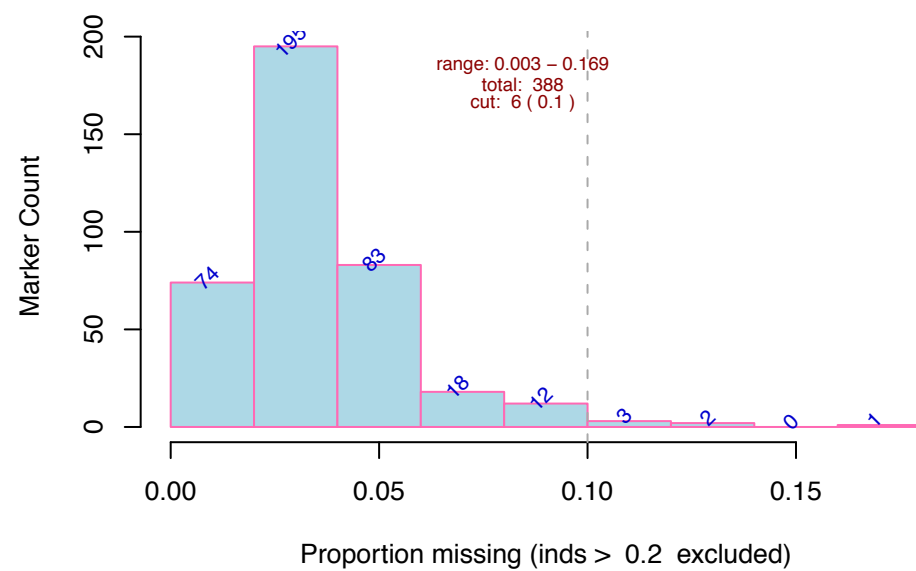


# Study 5 Hispanic Repeat Missingness

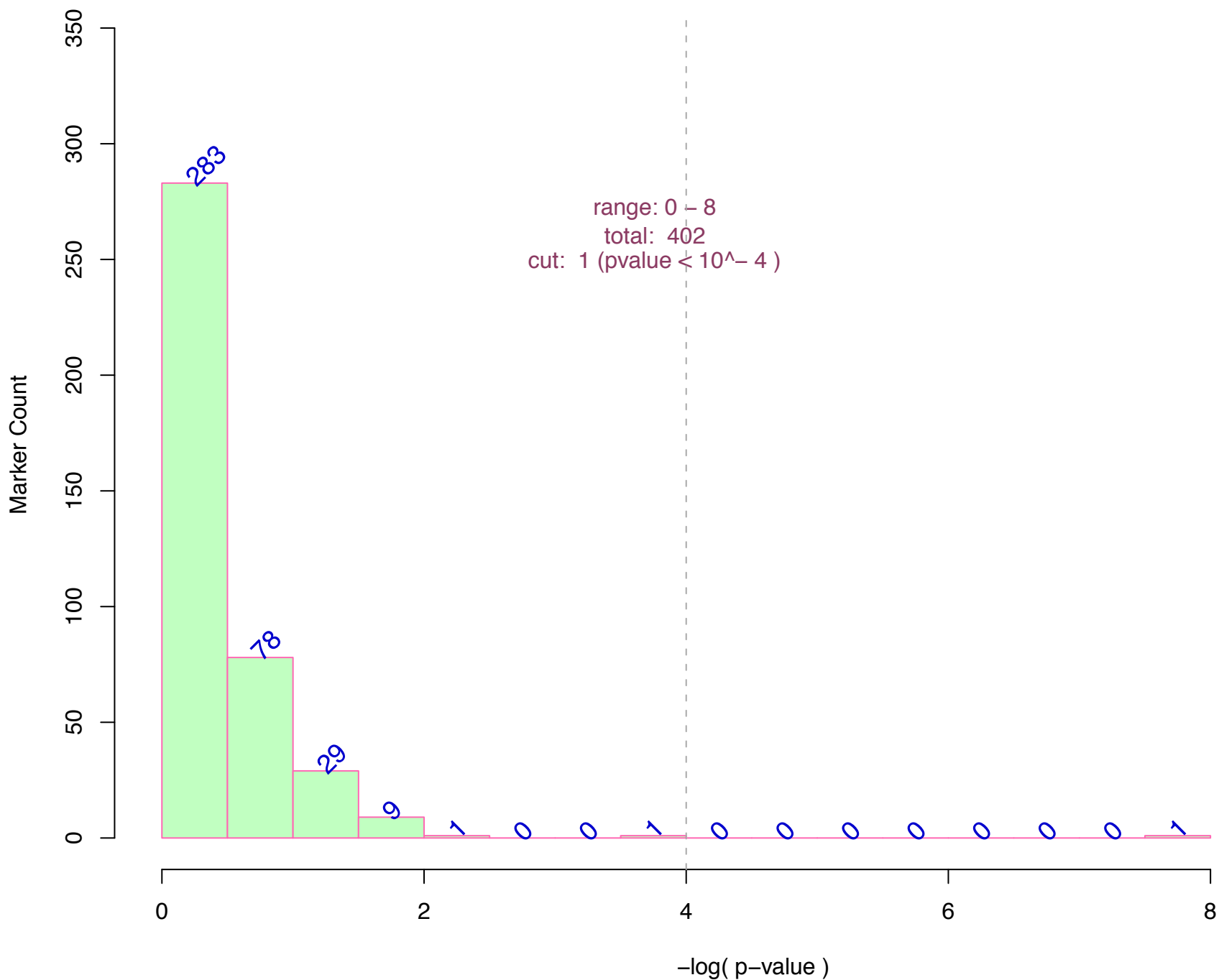
## Individuals



## Markers

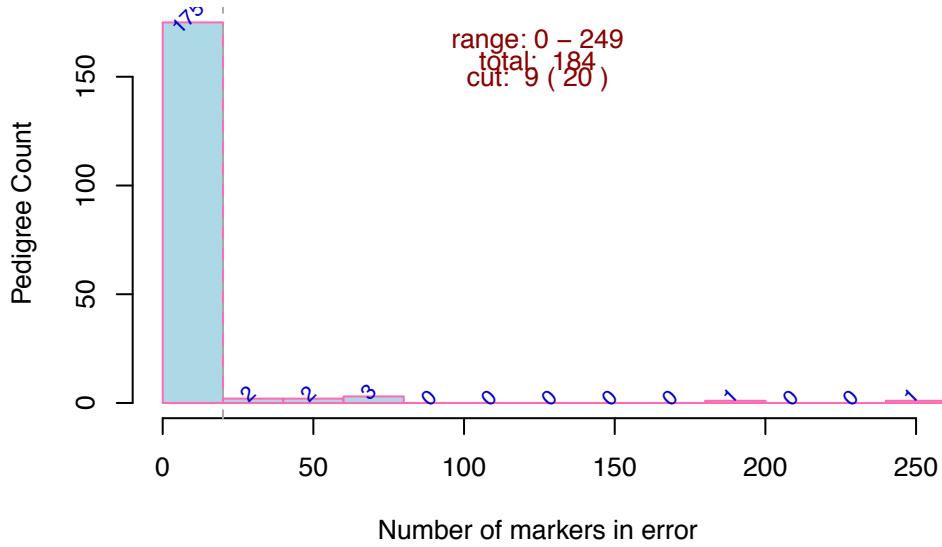


# Study 6 Hispanic Hardy-Weinberg Errors

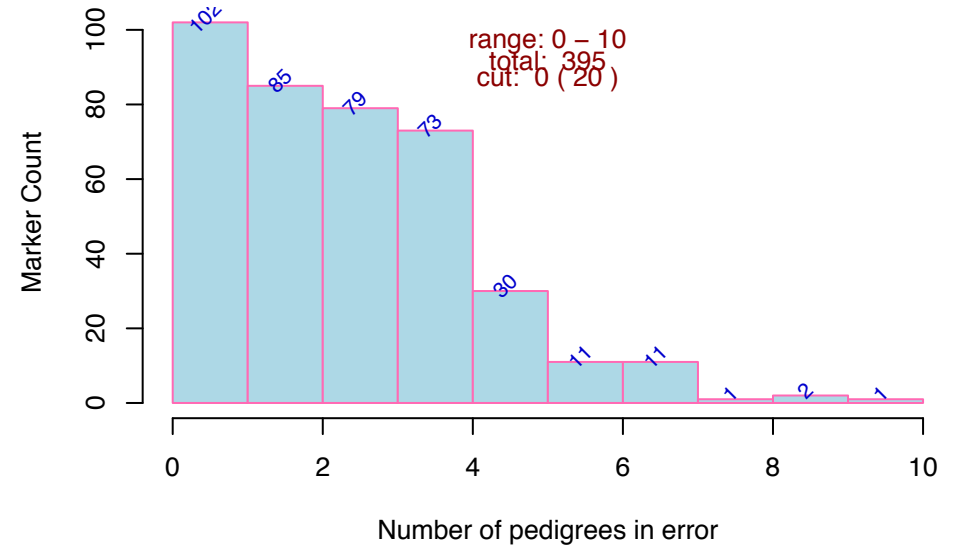


# Study 6 Hispanic Initial Mendel Errors

## Pedigrees

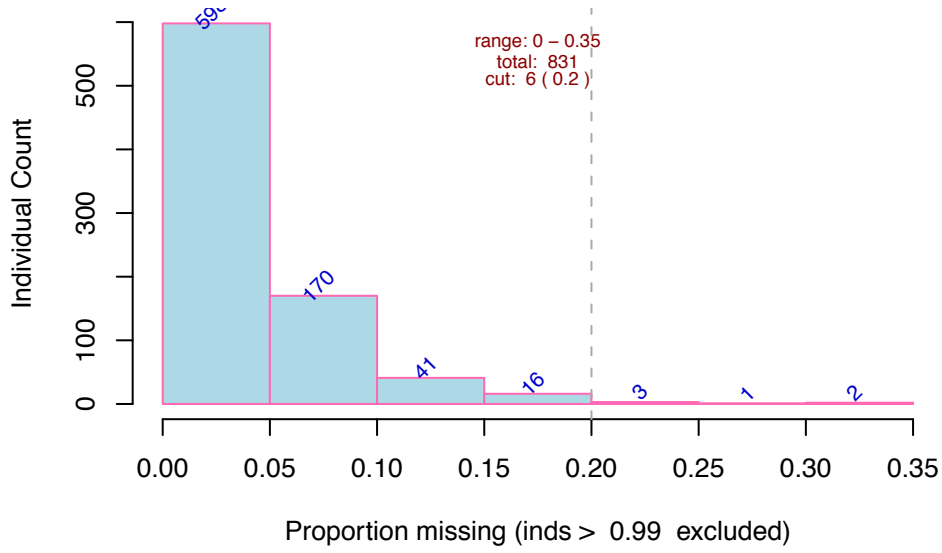


## Markers

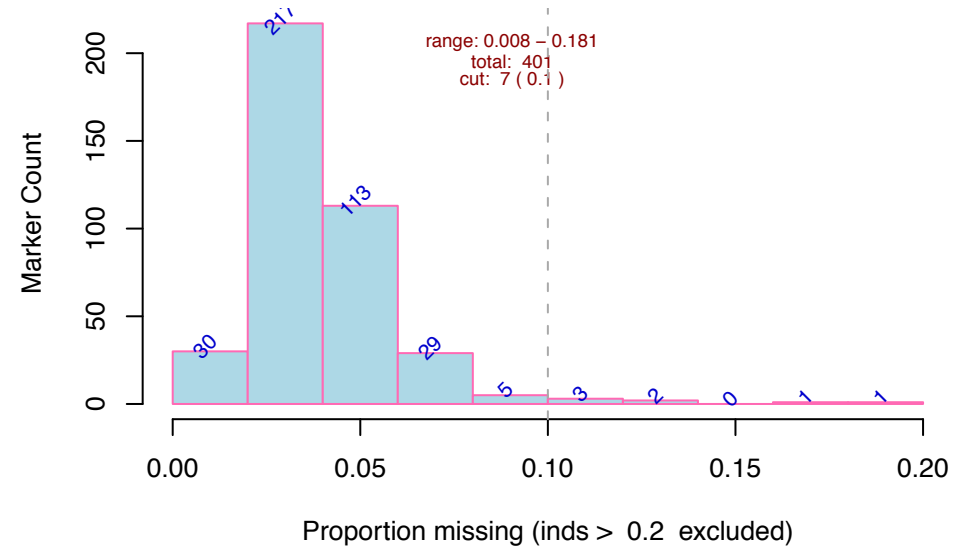


# Study 6 Hispanic Initial Missingness

## Individuals



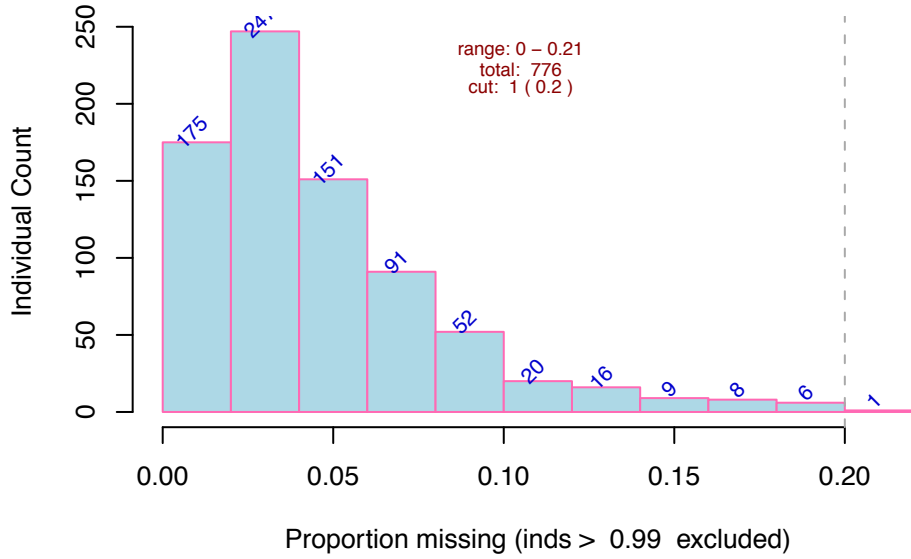
## Markers



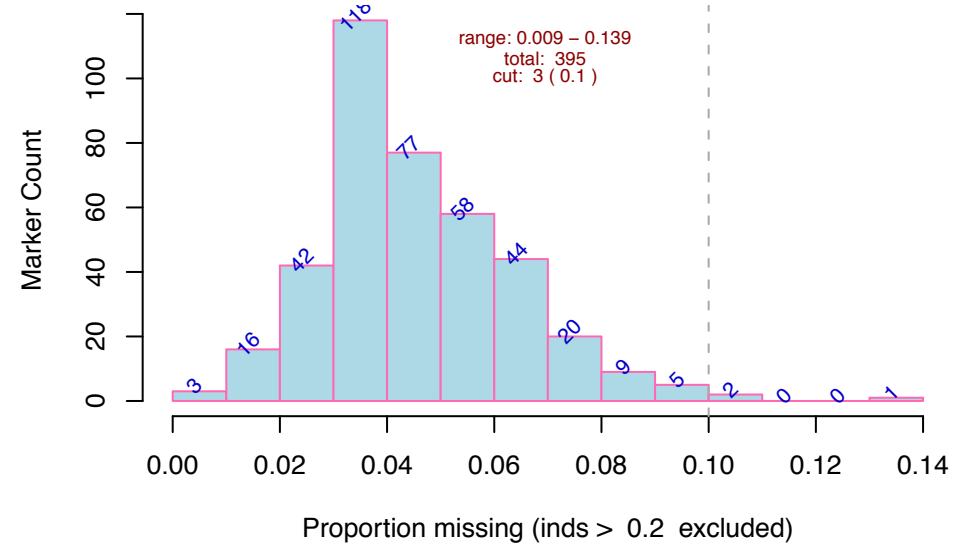


# Study 6 Hispanic Repeat Missingness

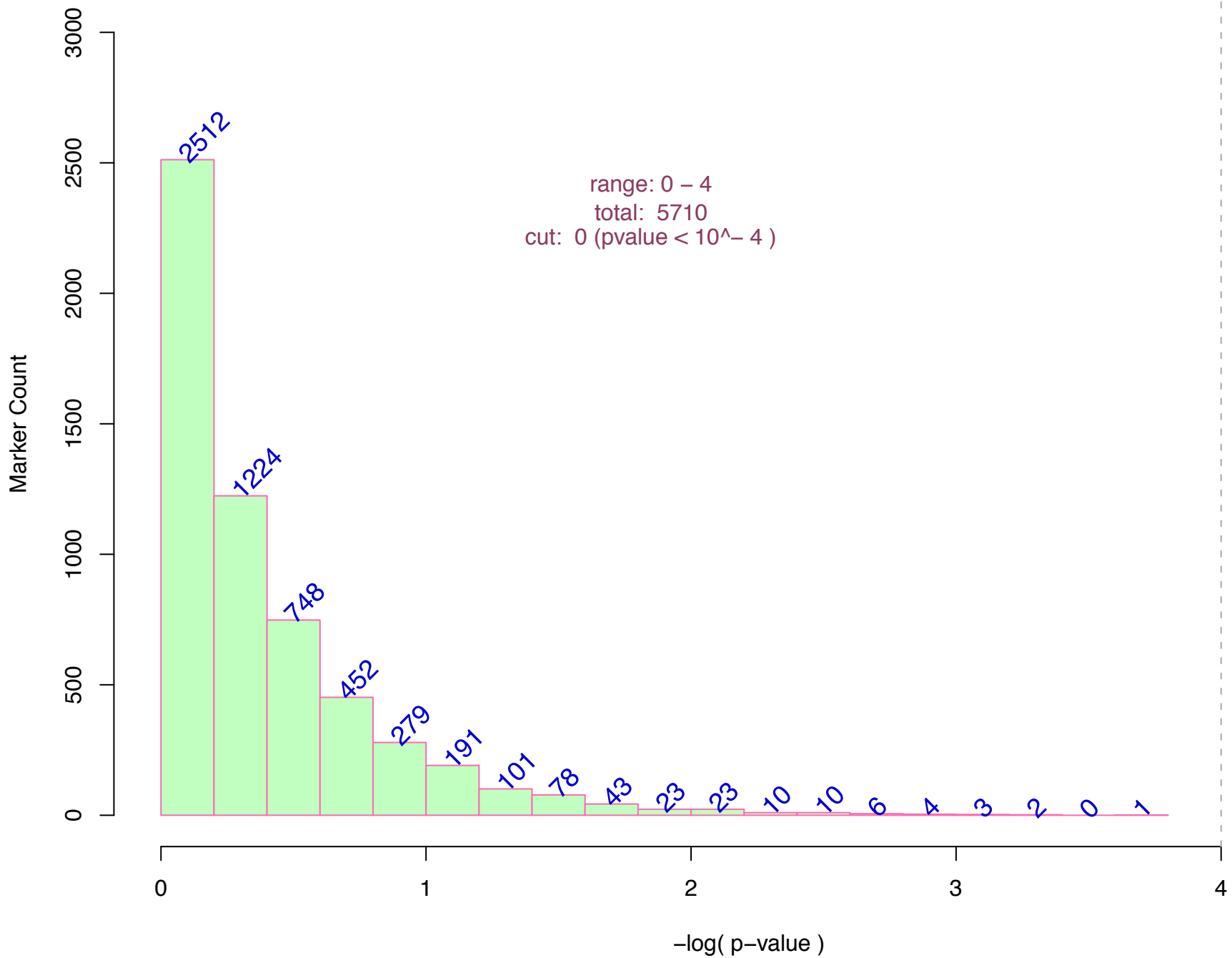
## Individuals



## Markers

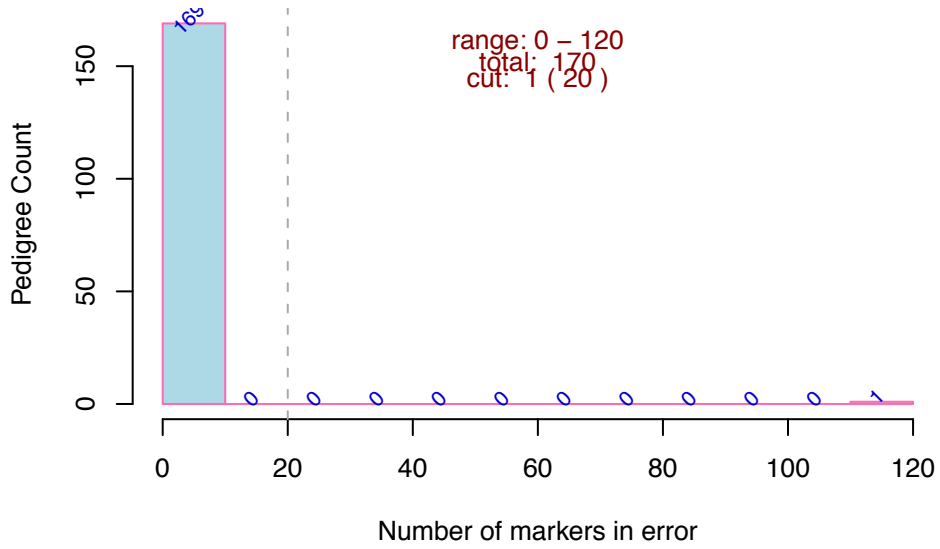


# Study 7 African American Hardy-Weinberg Errors



# Study 7 African American Initial Mendel Errors

## Pedigrees

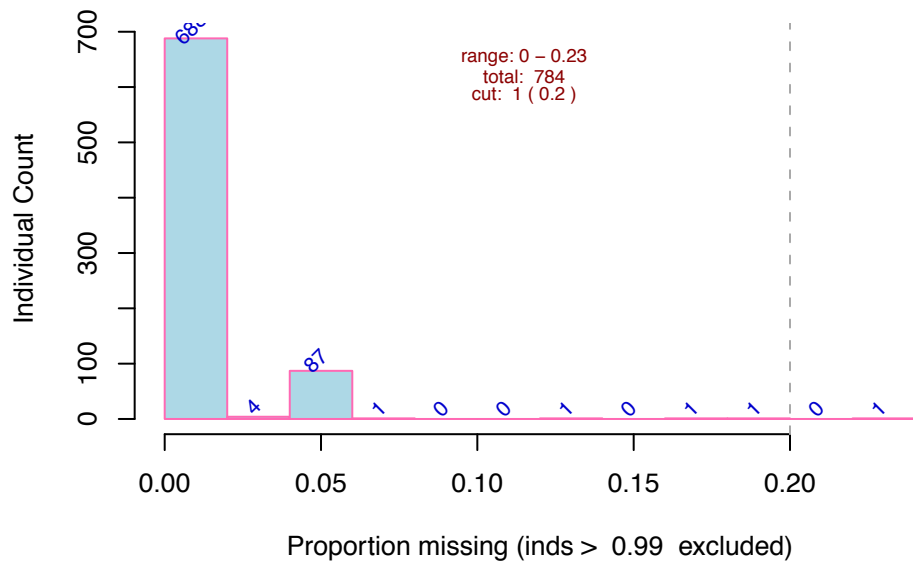


## Markers

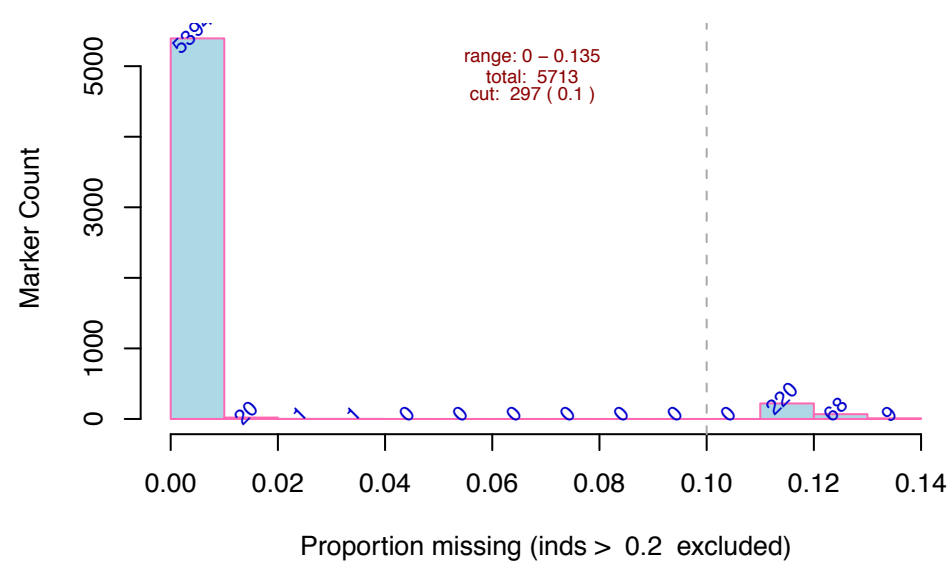


# Study 7 African American Initial Missingness

## Individuals

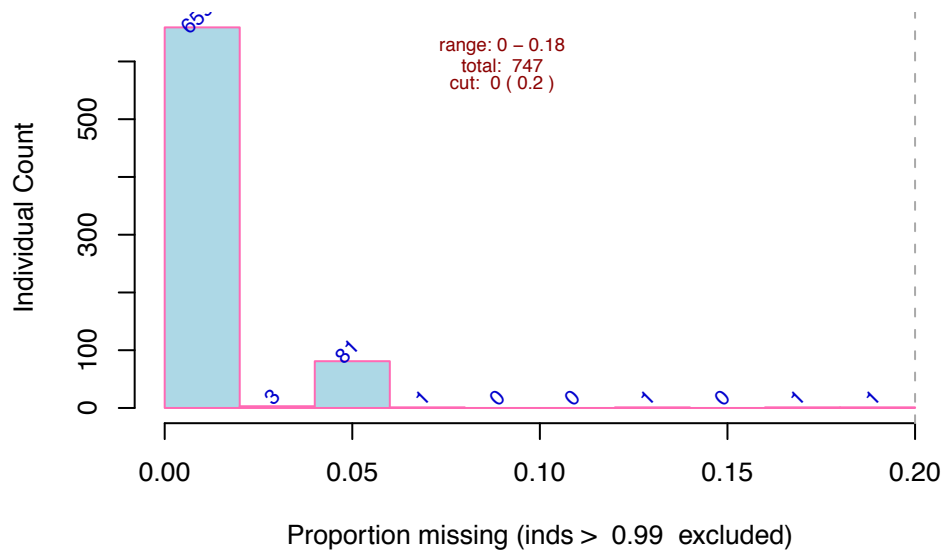


## Markers



# Study 7 African American Repeat Missingness

## Individuals



## Markers

