SUPPLEMENTAL FIGURES

Figure S1. Fold difference in G4 loci coverage at different genic components (A) and non-genic functional regions (B) as compared to the genome-wide average. Red horizontal line indicates no difference as compared to the genome-wide average. NCNR—non-coding non-repetitive subgenome.

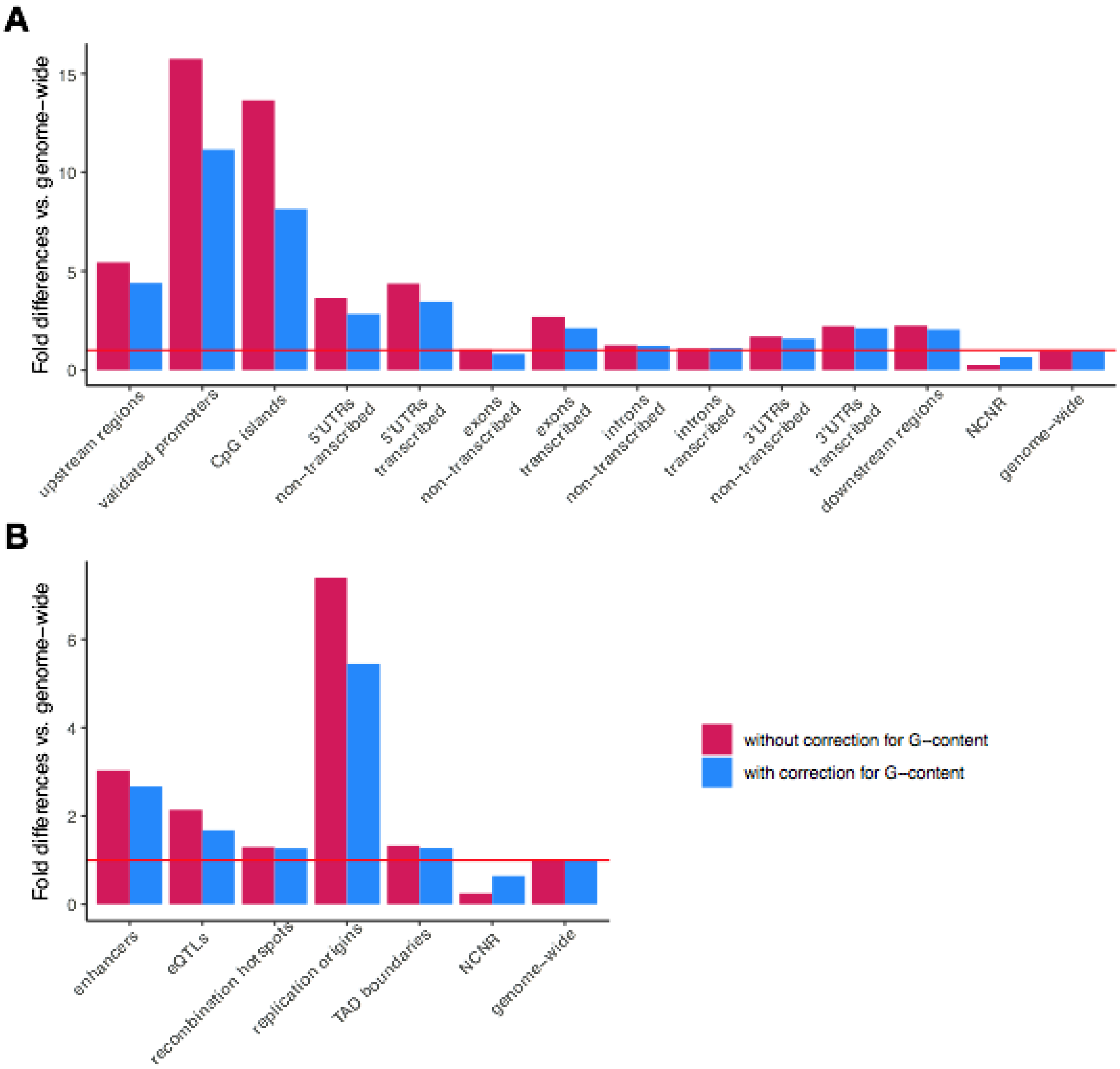


Figure S2. Distribution of predicted G4 stability scores across the genome. G4 stability scores were predicted by Quadron (Sahakyan et al. 2017). The vertical red line is the stability threshold (=19). G4 loci with a score greater than the stability threshold are predicted to be stable. G4 loci with a score lower than the stability threshold are predicted to be unstable.

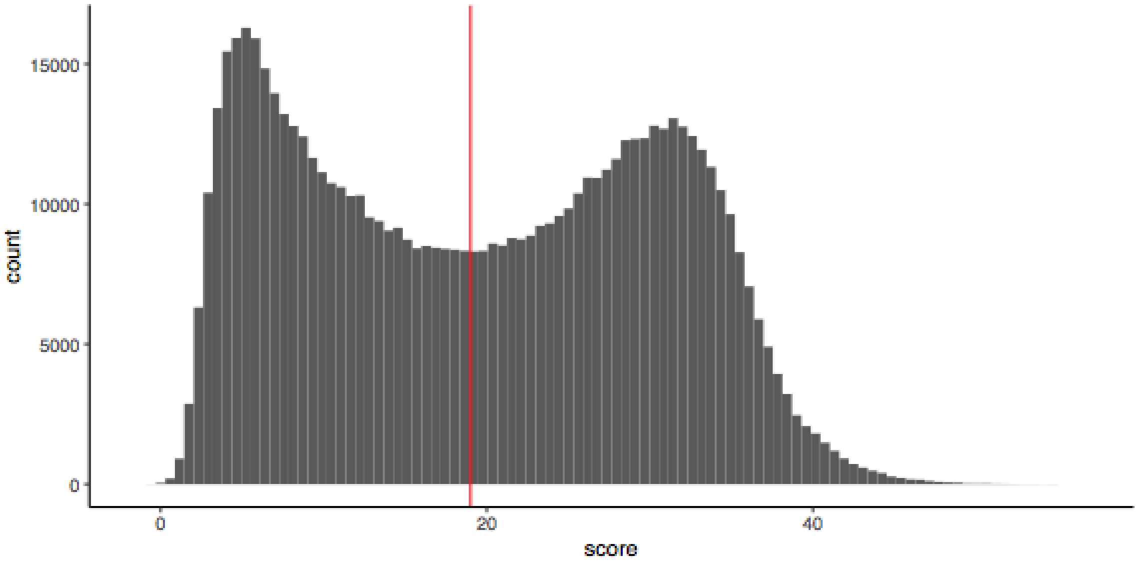


Figure S3. HKA odds ratios for FANTOM enhancers (Lizio et al. 2015; FANTOM Consortium et al. 2014; Andersson et al. 2014) and ENCODE enhancers (ENCODE Project Consortium et al. 2020), and for NCNR regions with high and low levels of recombination (top and bottom 25% of recombination rates according to (Halldorsson et al. 2019), respectively).

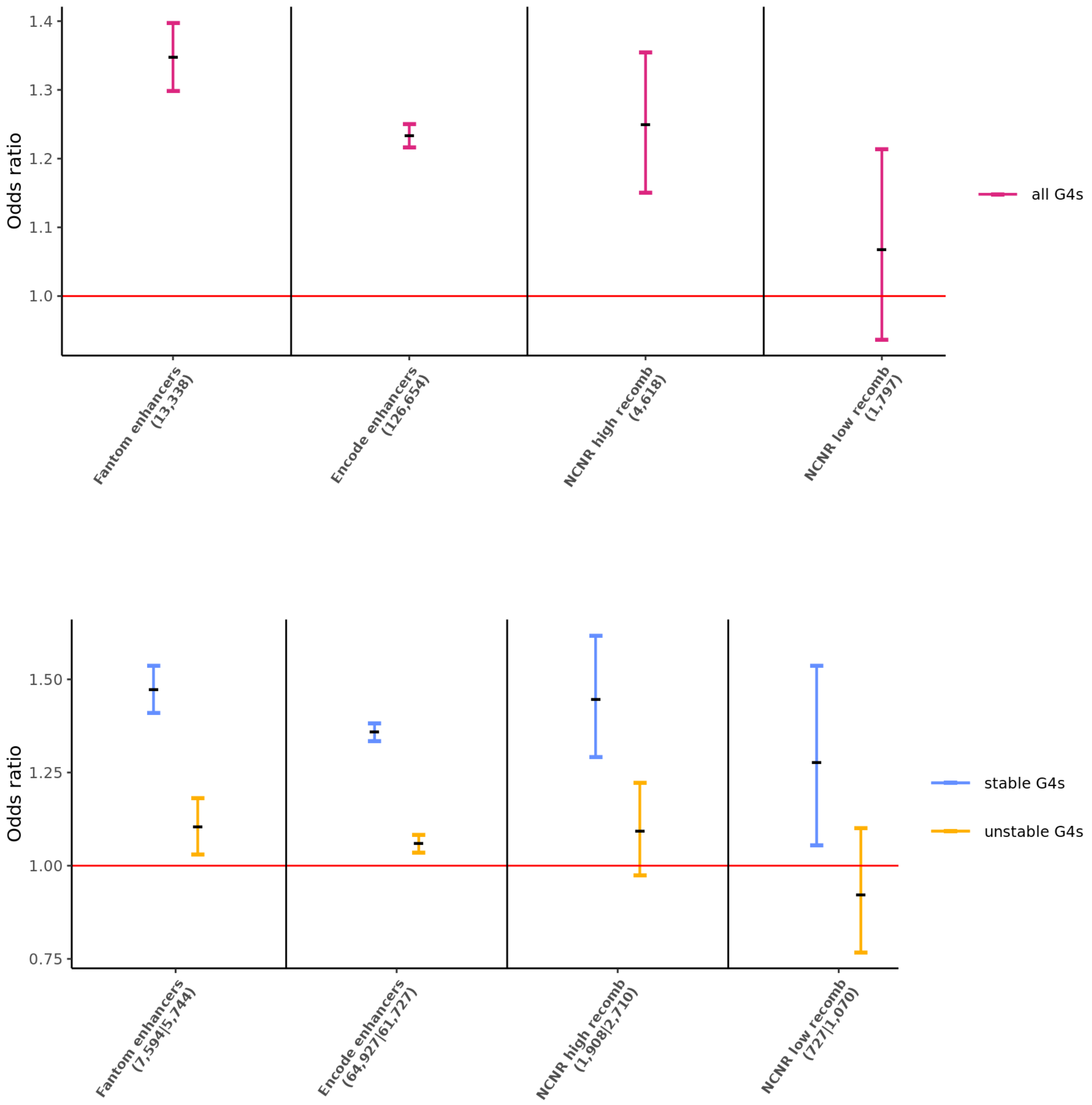
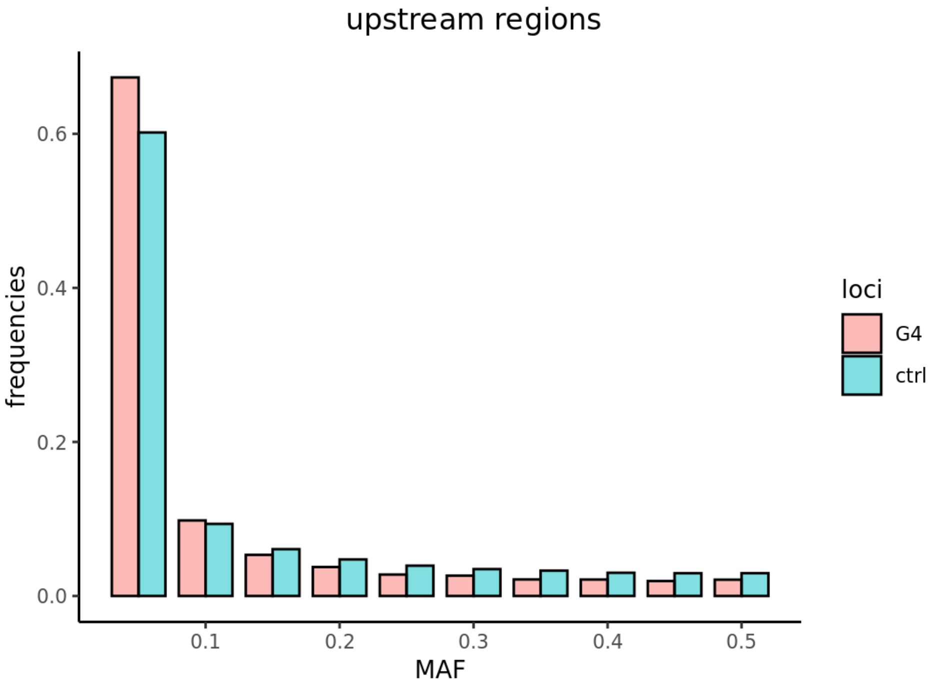
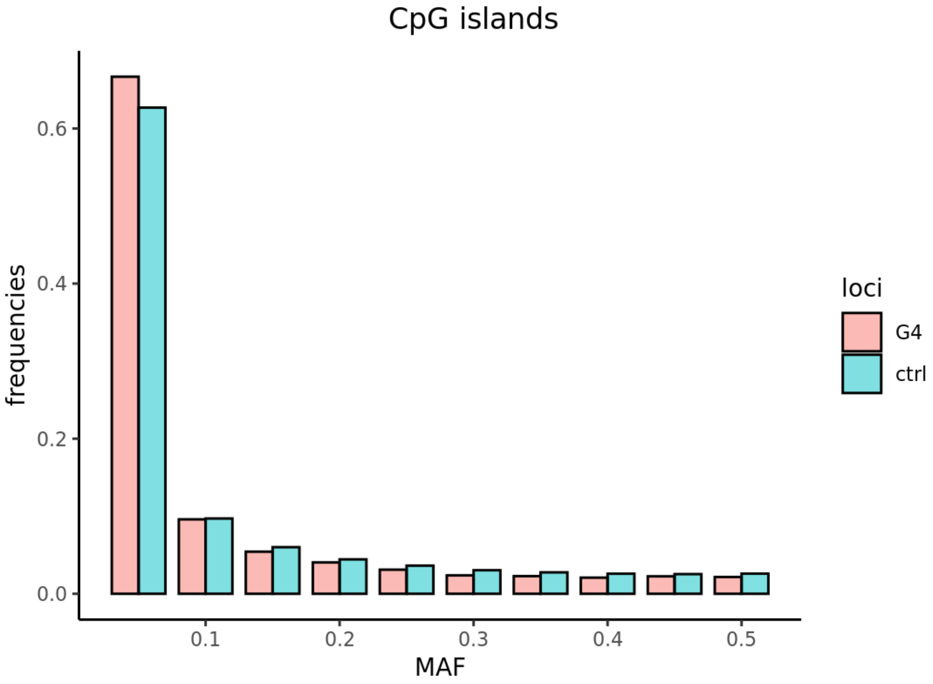


Figure S4. Site frequency spectrum. Histograms of Minor Allele Frequencies (MAFs) of SNPs from SGDP (Mallick et al. 2016) located in genic components (A-J), non-genic functional components (K-M) or the NCNR subgenome (N). SNPs are split into intersecting G4 loci or not (control). Single- and doubletons were not included.

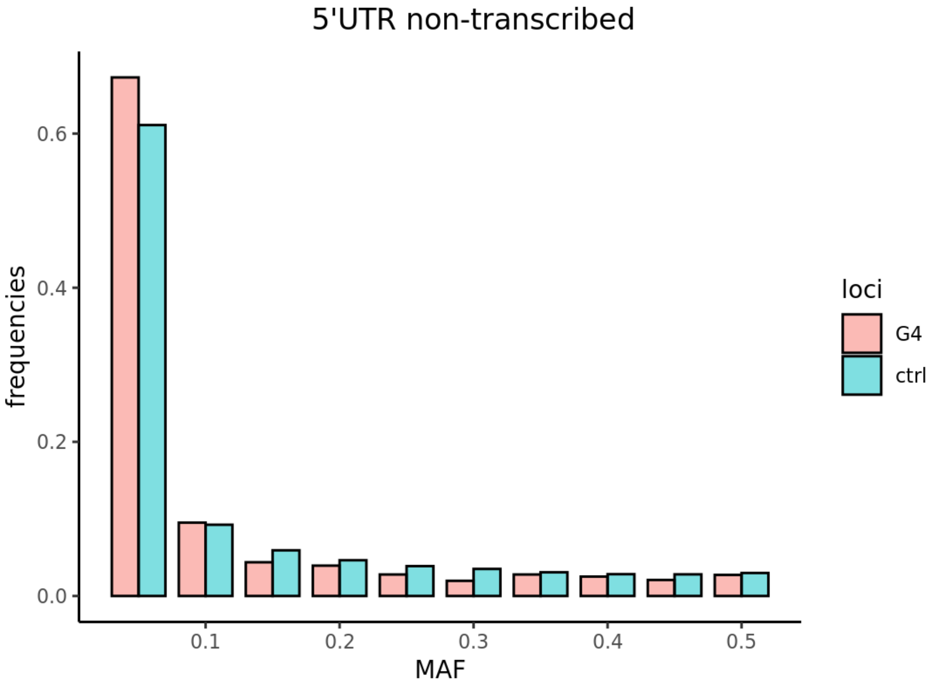
A



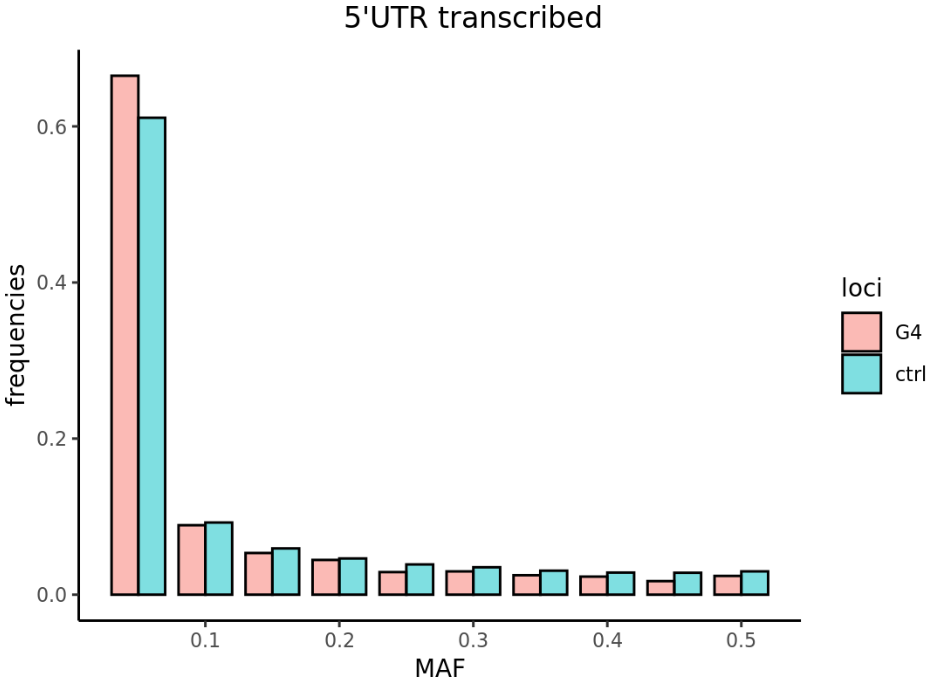
B



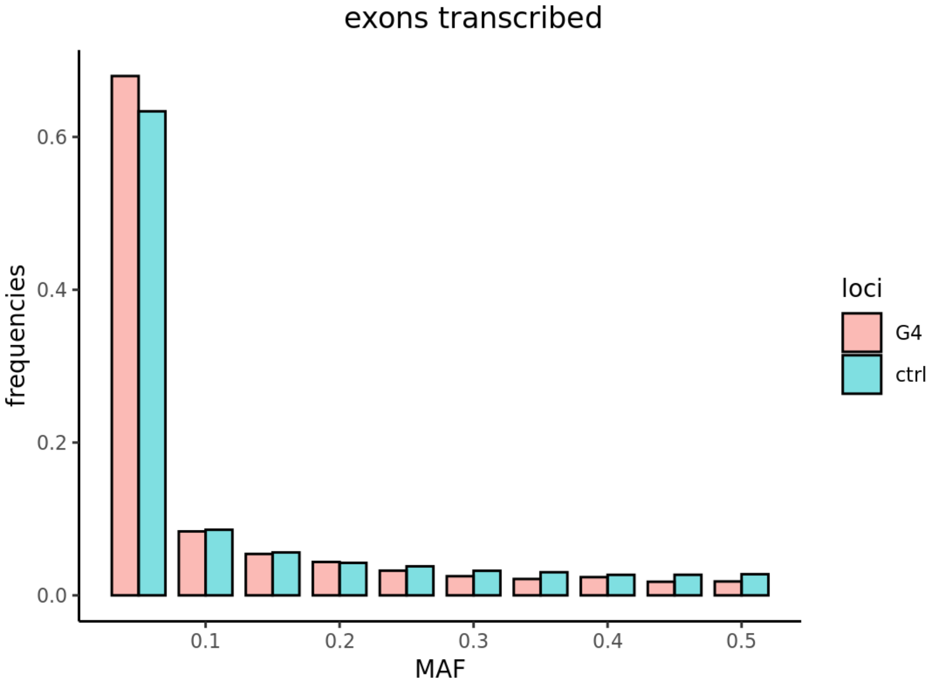
C


5'UTR non-transcribed
frequencies
0.0
0.2
0.4
0.6
MAF
0.1
0.2
0.3
0.4
0.5
loci
G4
ctrl

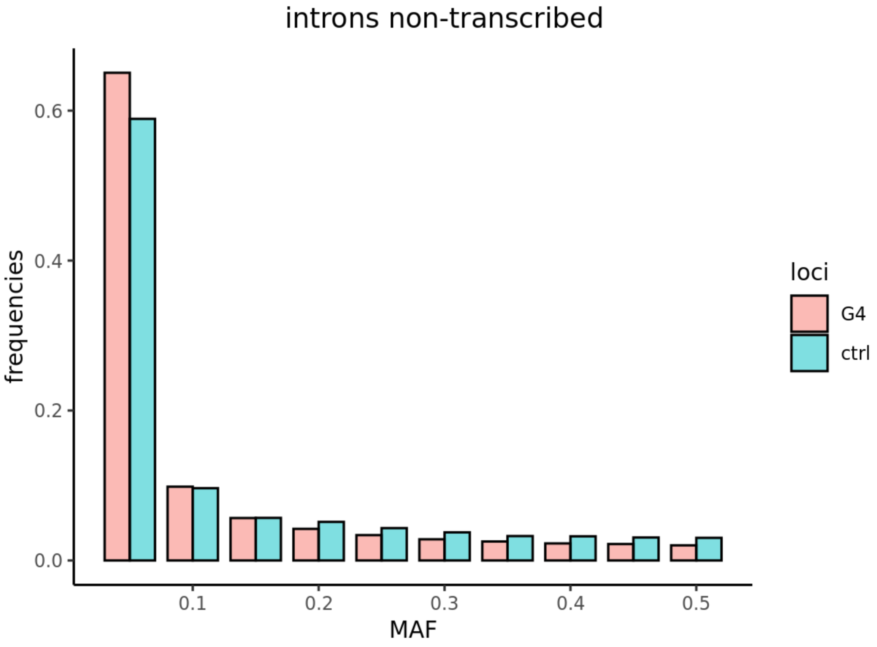
D


5'UTR transcribed
frequencies
0.0
0.2
0.4
0.6
MAF
0.1
0.2
0.3
0.4
0.5
loci
G4
ctrl

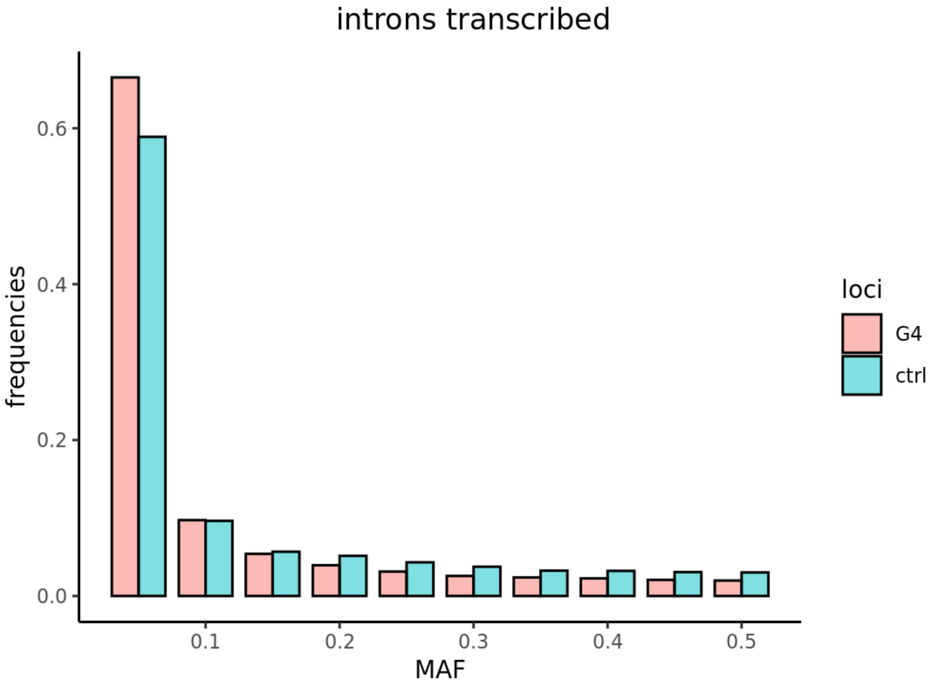
E


exons transcribed
frequencies
0.6
0.4
0.2
0.0
0.1
0.2
0.3
0.4
0.5
MAF
loci
G4
ctrl

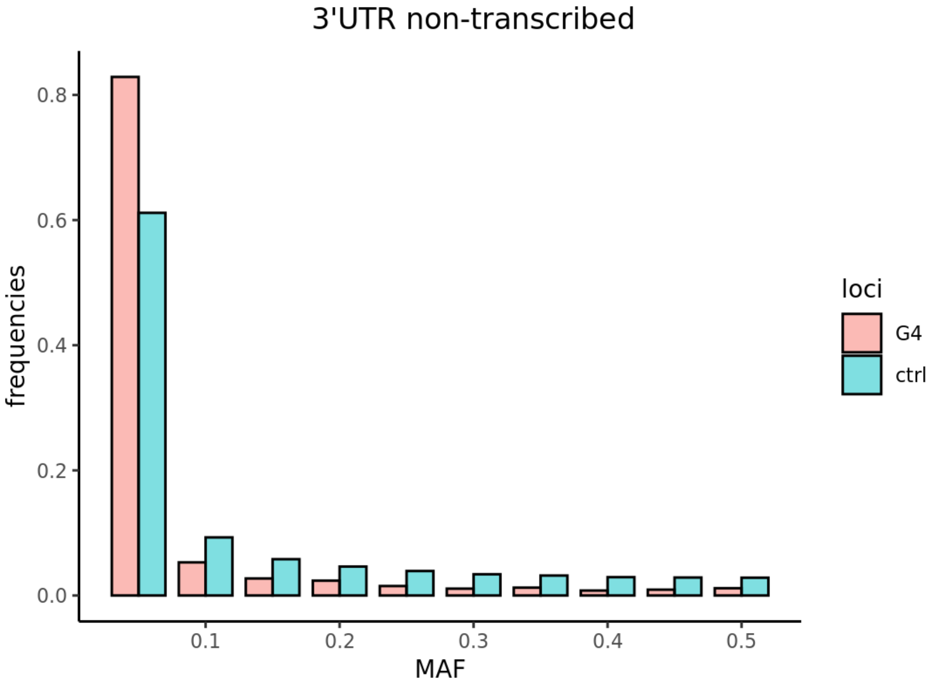
F


introns non-transcribed
frequencies
0.6
0.4
0.2
0.0
0.1
0.2
0.3
0.4
0.5
MAF
loci
G4
ctrl

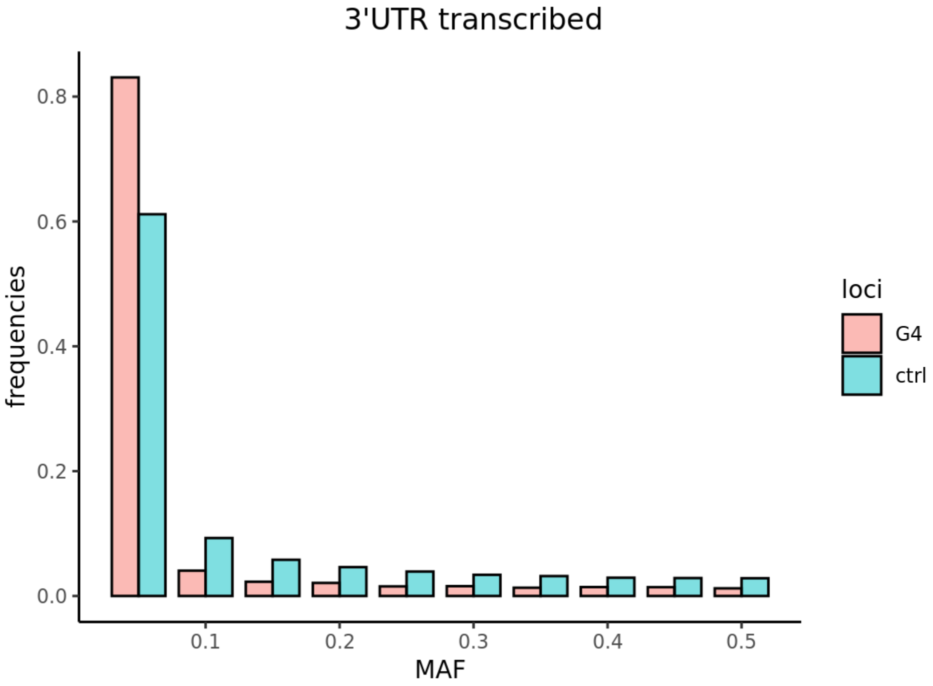
G


introns transcribed
frequencies
0.0
0.2
0.4
0.6
0.1
0.2
0.3
0.4
0.5
MAF
loci
G4
ctrl

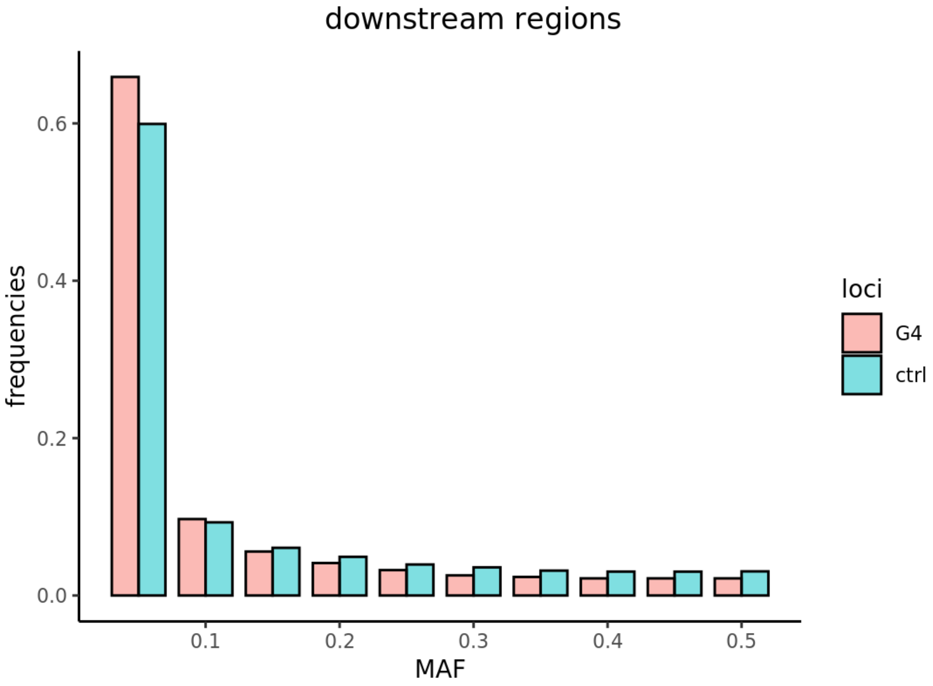
H


3'UTR non-transcribed
frequencies
0.0
0.2
0.4
0.6
0.8
0.1
0.2
0.3
0.4
0.5
MAF
loci
G4
ctrl

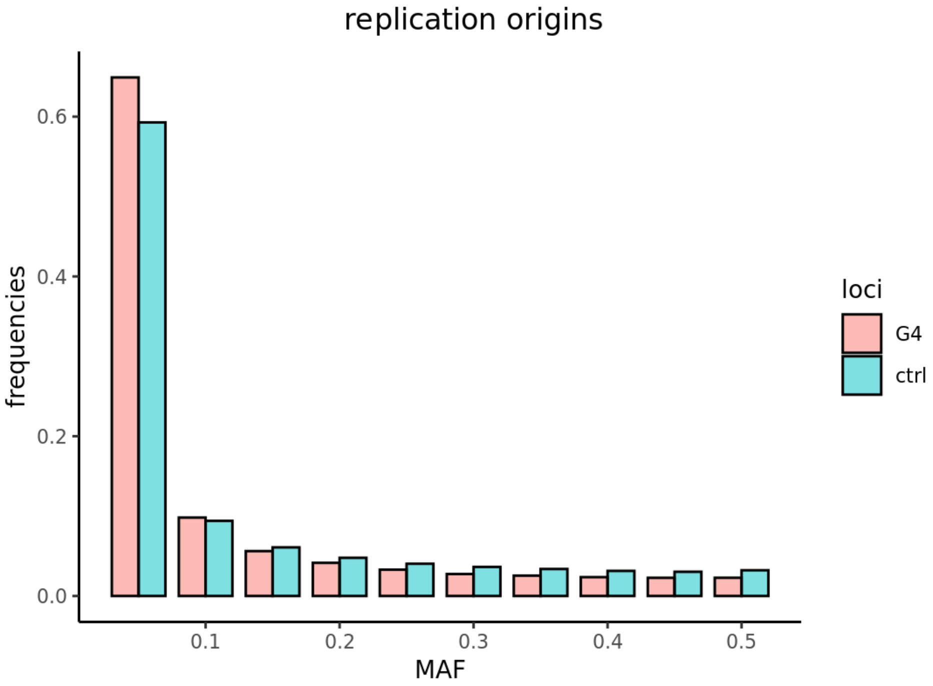
I


3'UTR transcribed
frequencies
0.8
0.6
0.4
0.2
0.0
0.1
0.2
0.3
0.4
0.5
MAF
loci
G4
ctrl

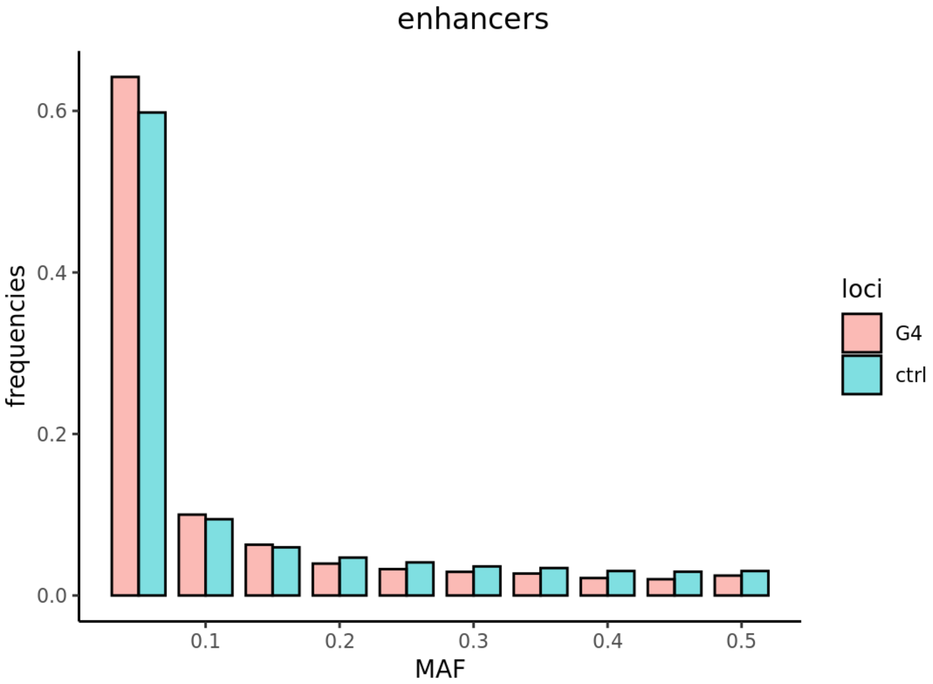
J


downstream regions
frequencies
0.6
0.4
0.2
0.0
0.1
0.2
0.3
0.4
0.5
MAF
loci
G4
ctrl

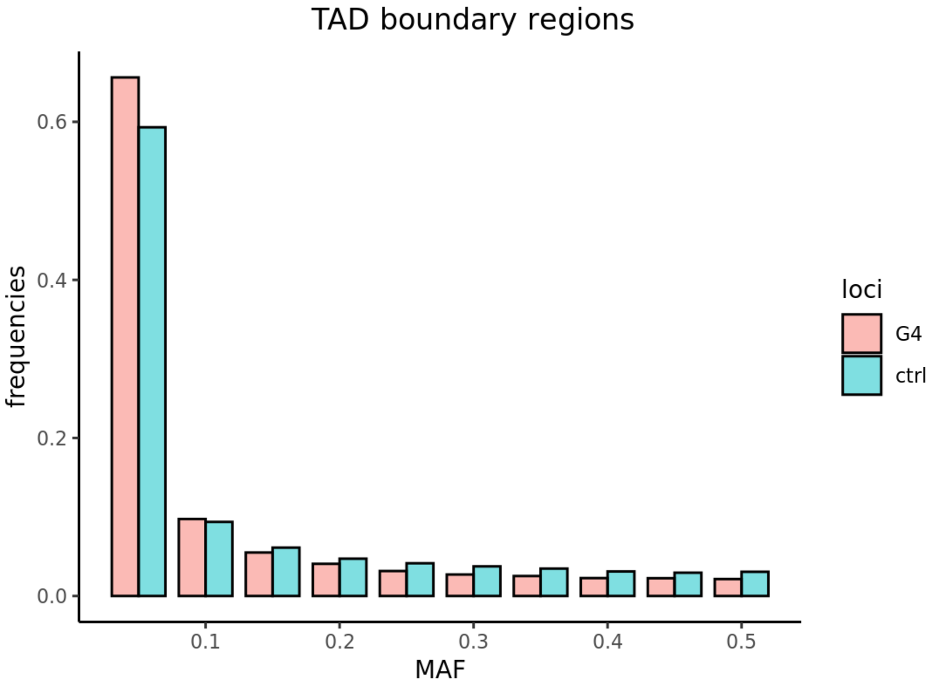
K


replication origins
frequencies
0.6
0.4
0.2
0.0
0.1
0.2
0.3
0.4
0.5
MAF
loci
G4
ctrl

L


enhancers
frequencies
0.6
0.4
0.2
0.0
0.1
0.2
0.3
0.4
0.5
MAF
loci
G4
ctrl

M


TAD boundary regions
frequencies
0.6
0.4
0.2
0.0
0.1
0.2
0.3
0.4
0.5
MAF
loci
G4
ctrl

N

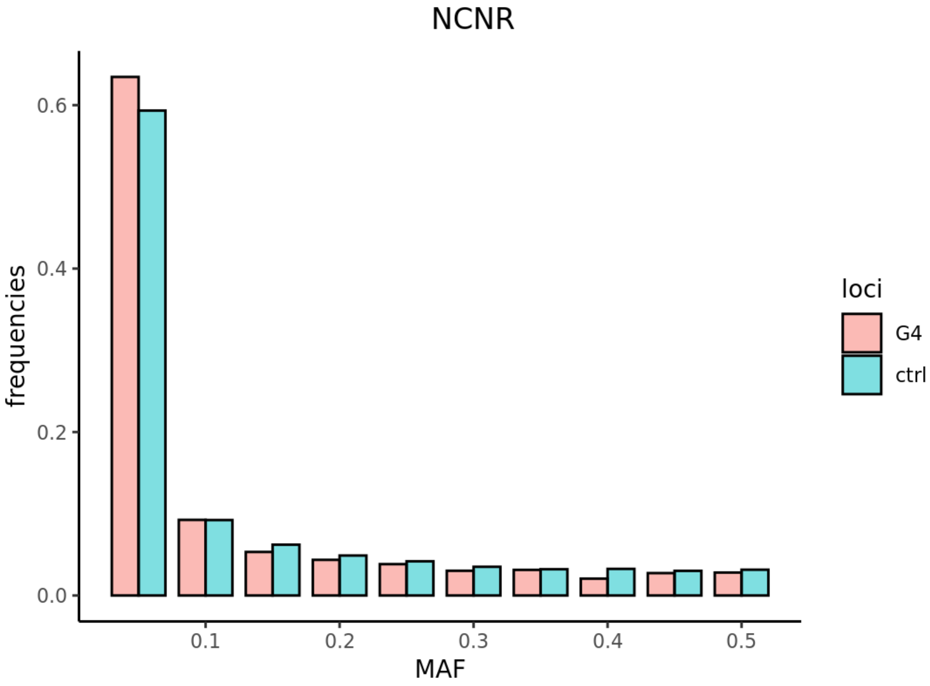

NCNR
frequencies
0.6
0.4
0.2
0.0
0.1
0.2
0.3
0.4
0.5
MAF
loci
G4
ctrl

Figure S5. Random selection of a unique annotation in an overlapping set. In this example, annotations of exons 1 to 3 form an overlapping set. To prevent the multiple uses of the same genomic coordinates, only one annotation from each such set was randomly selected.