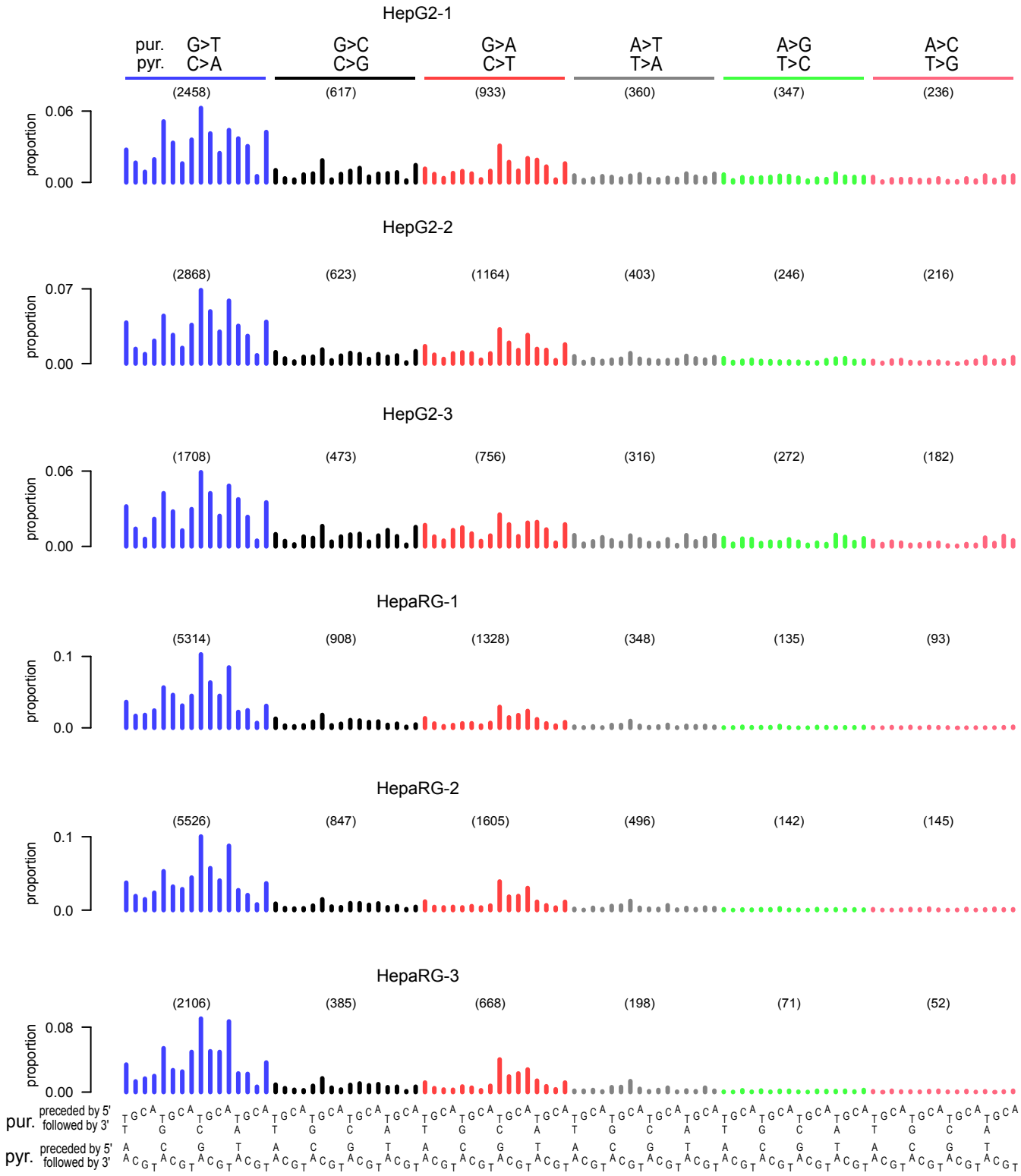
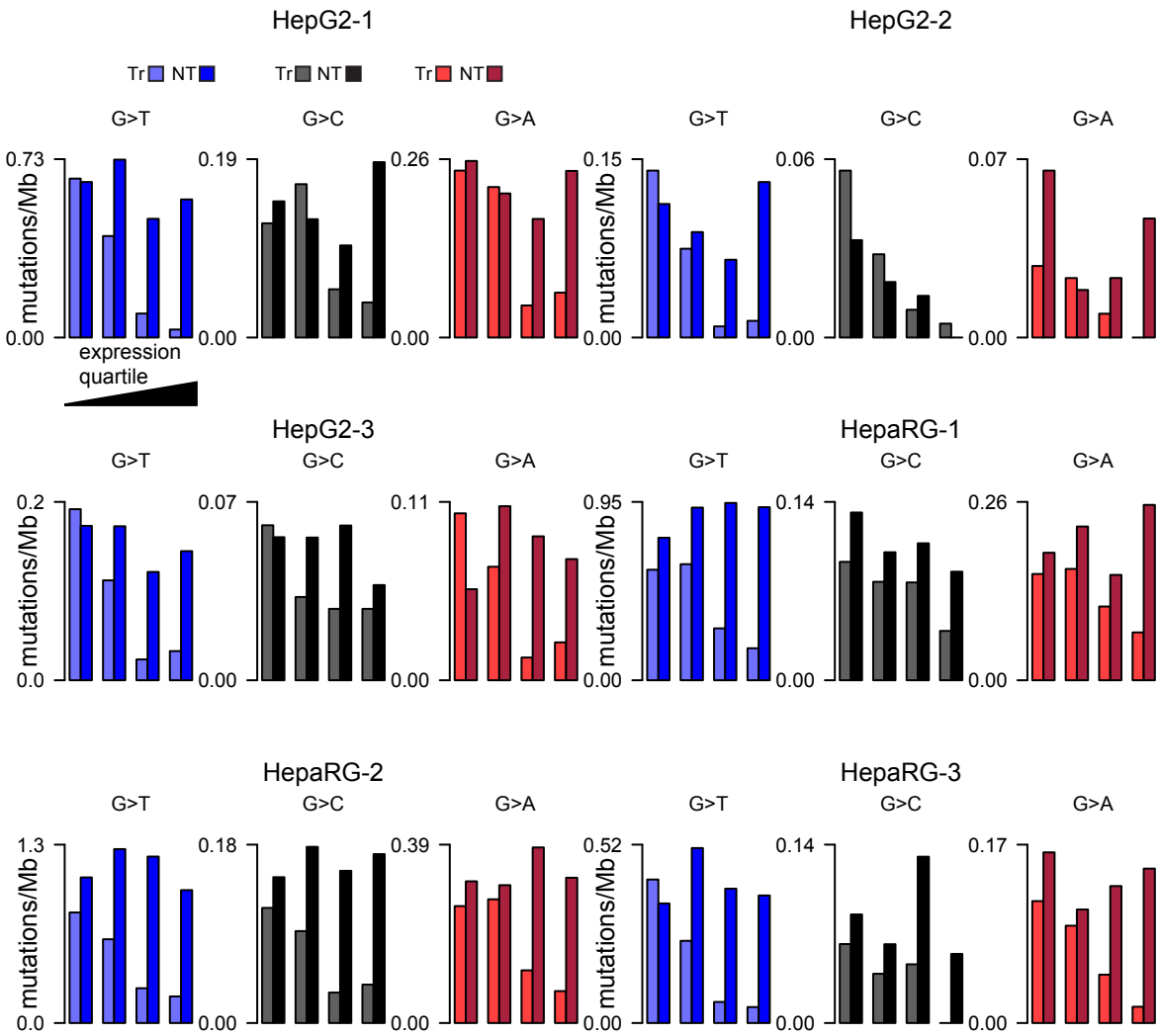
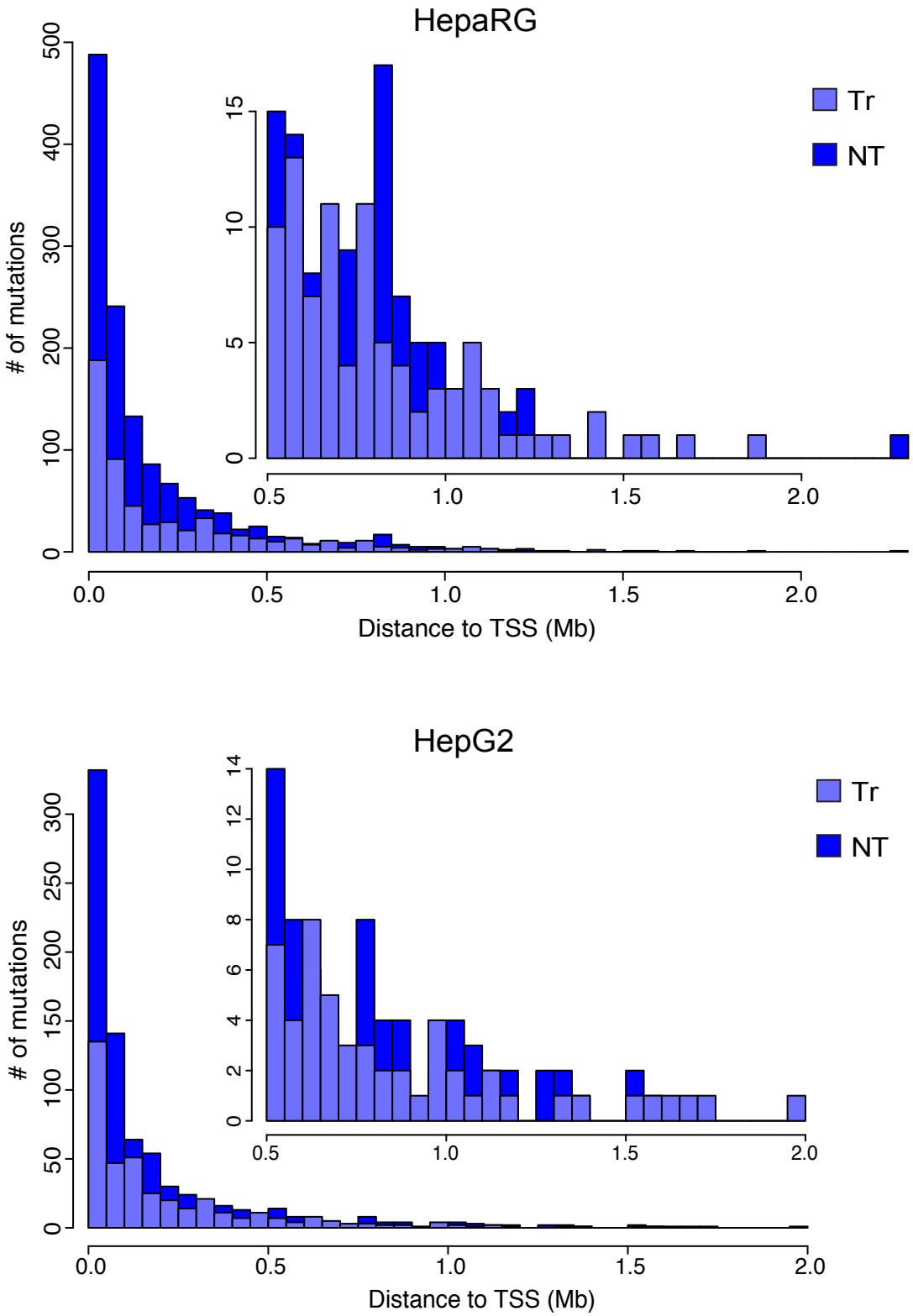


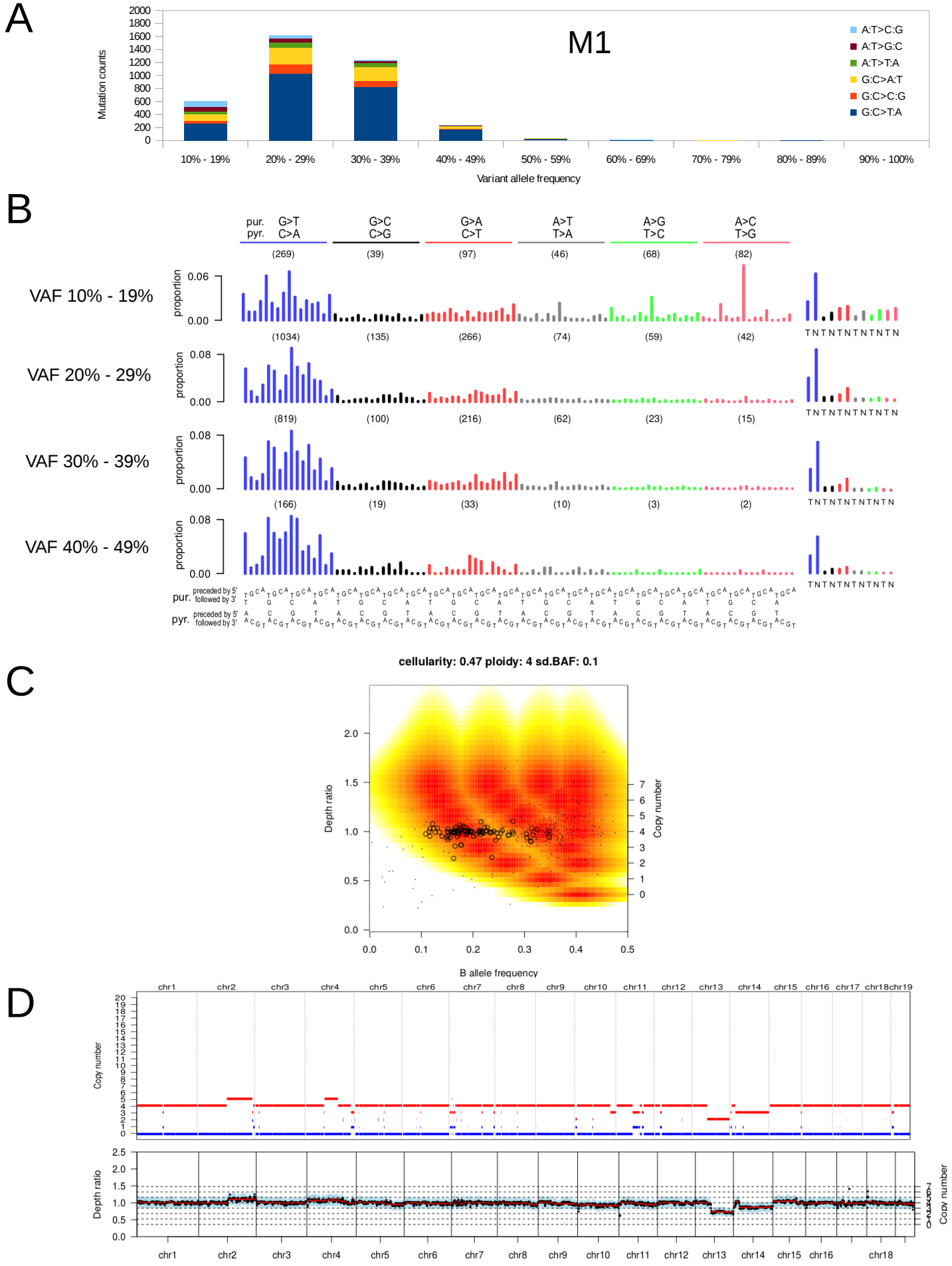
Supplemental Fig. S1 Mutation spectra of human cell lines treated with AFB1



Supplemental Fig. S2 Transcription strand bias by expression quartile for human cell lines treated with AFB1



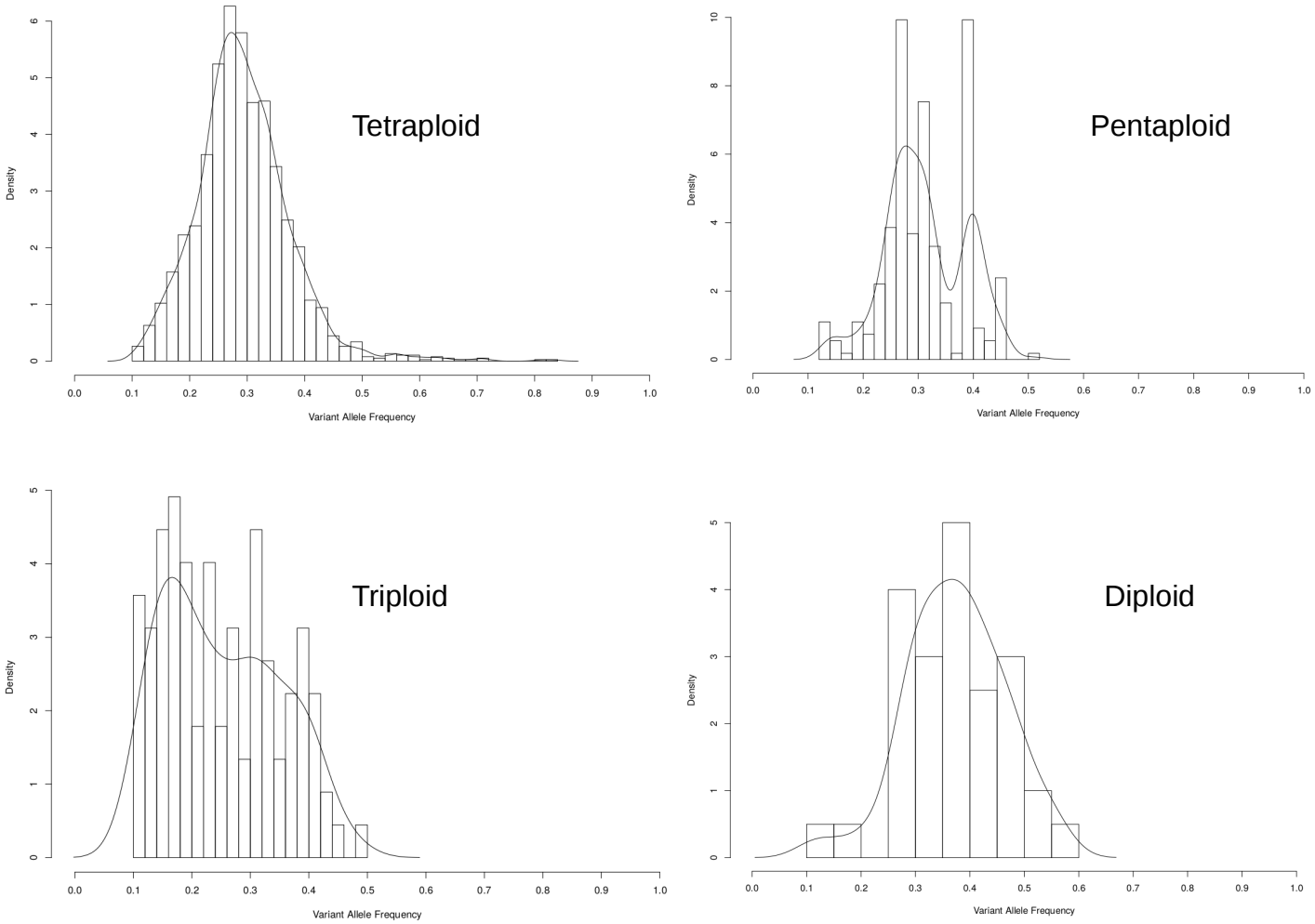
Supplemental Fig. S3 Transcription strand bias along the transcripts for human cell lines treated with AFB1.



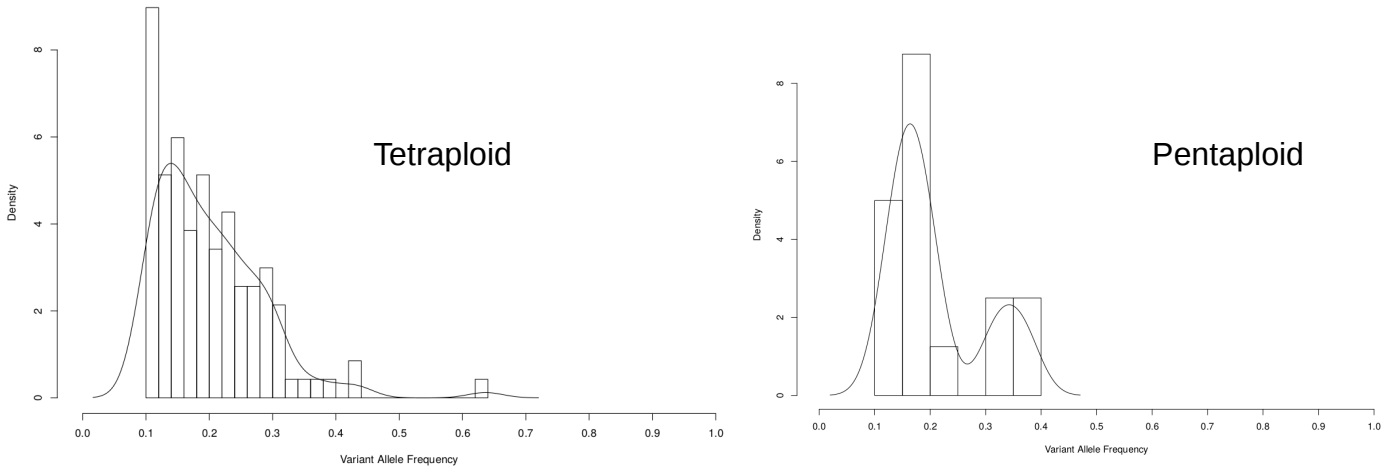
Supplemental Fig. S4 Analysis of variant allel e frequency (VAF) effects on mutational signatures in mouse tumor M1. SNV (A) and mutation spectra (B) distribution across different VAFs. T, transcribed; N, non-transcribed (C) Tumor content (cellularity) and ploidy (D) Copy number and depth ratio across the genome.

A

M1



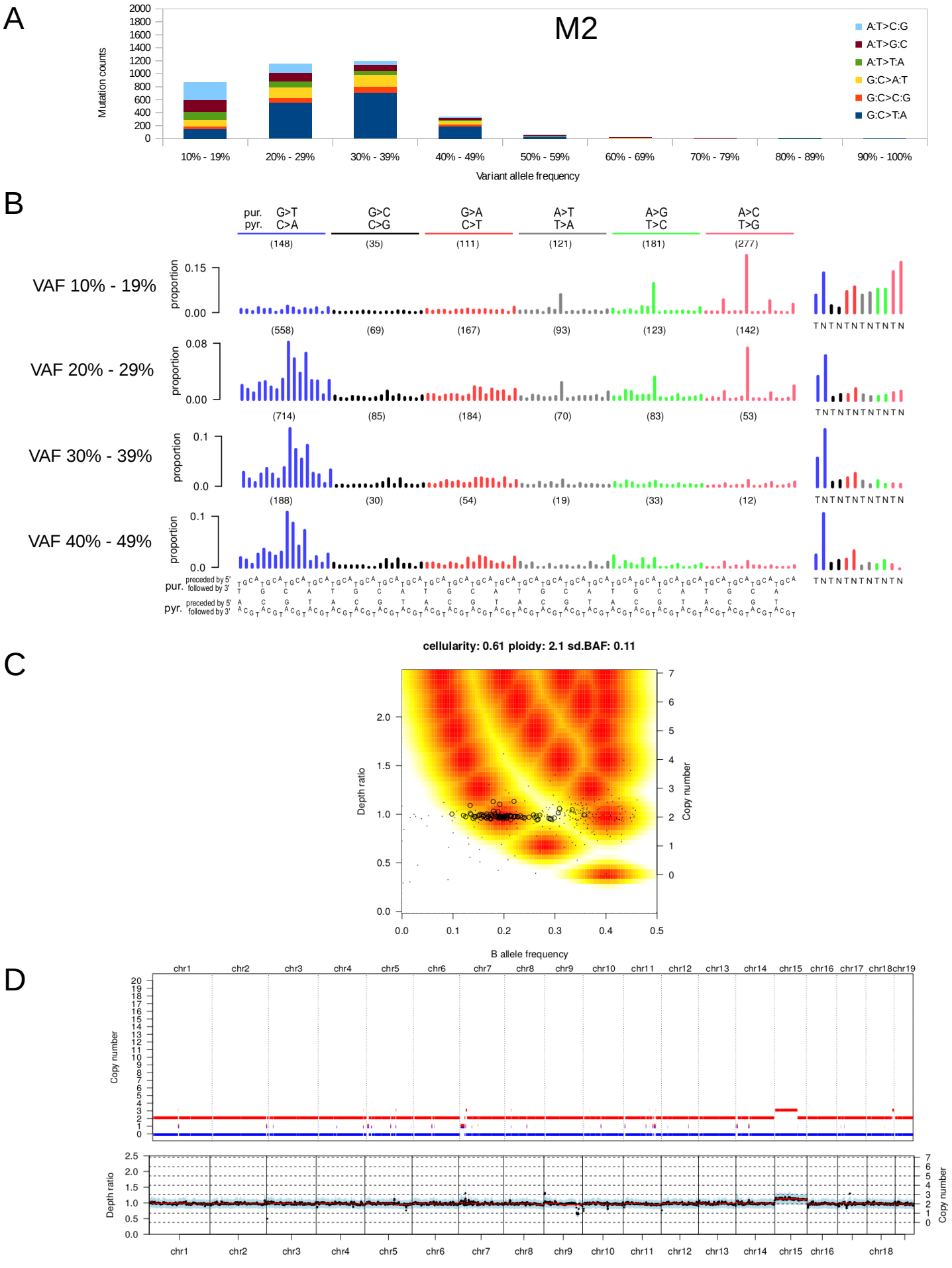
B



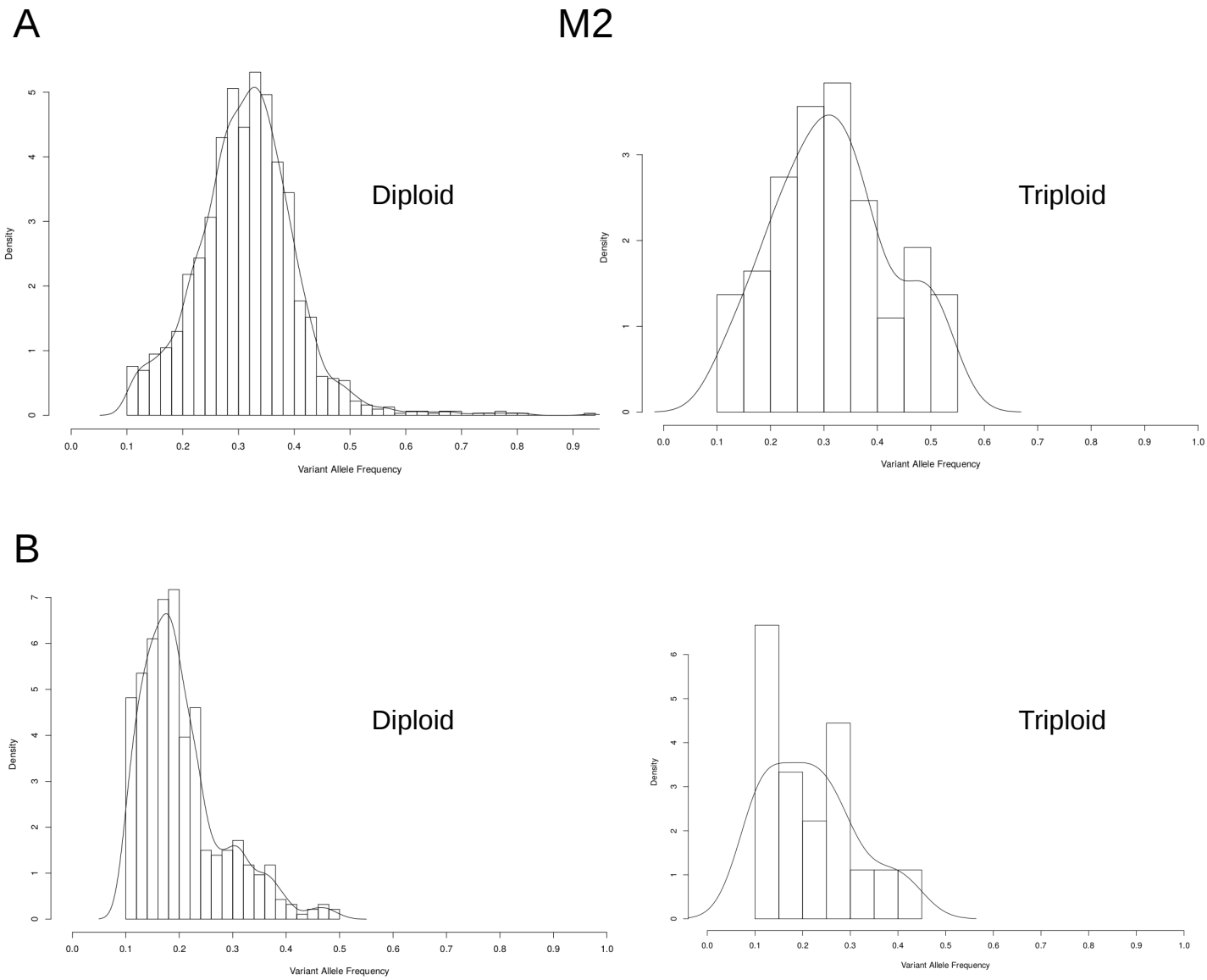
Triploid region only contains five somatic mutations with VAF of 0.316, 0.291. 0.233, 0.178 and 0.154 respectively

Diploid region only contains four somatic mutations with VAF of 0.212, 0.194, 0.156 and 0.119 respectively

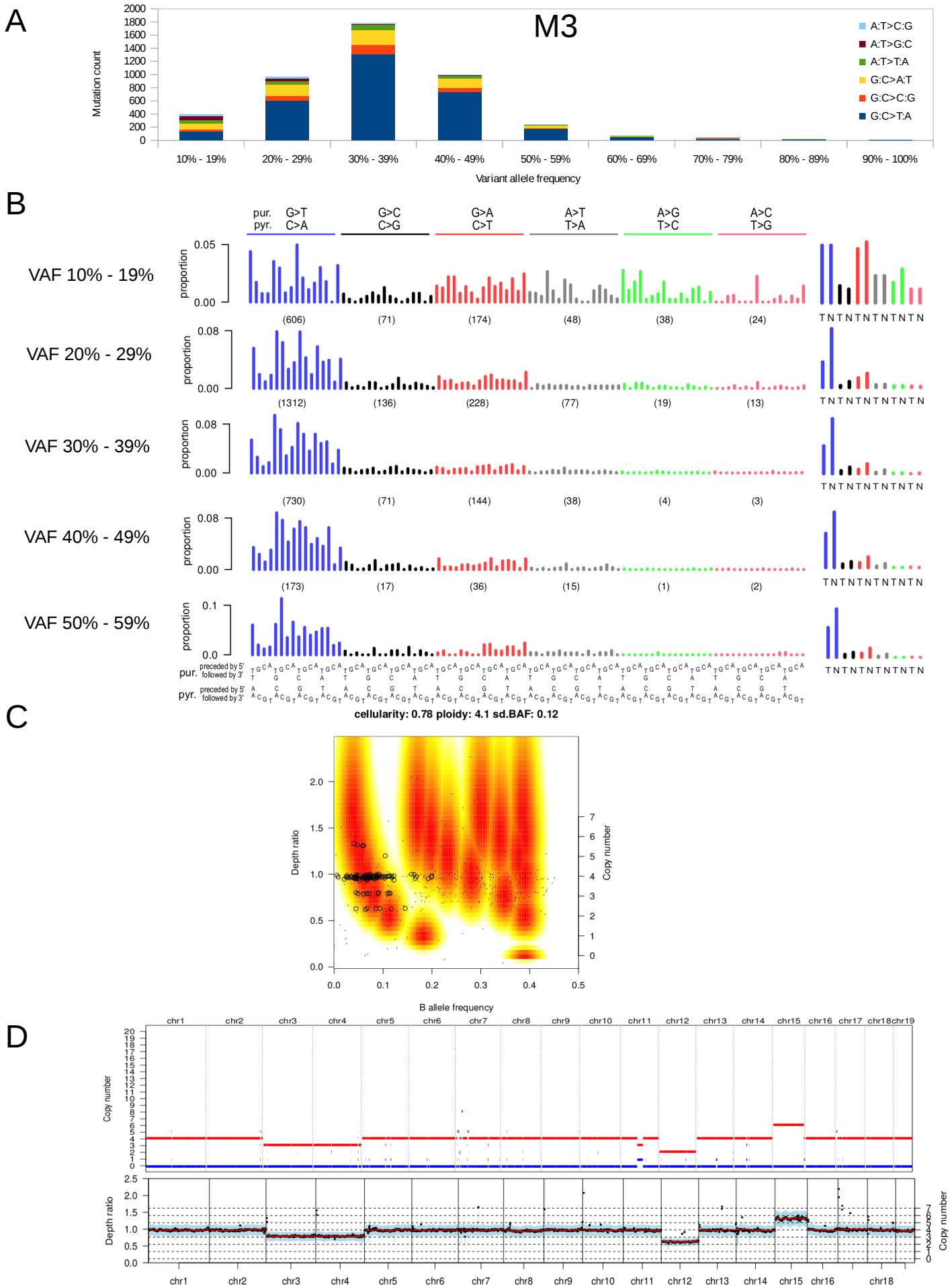
Supplemental Fig. S5 Probability densities of G:C > T:A **(A)** and A:T > C:G **(B)** mutation counts by variant allele frequency (VAF) in different copy number regions of the genome of mouse tumor M1. Probability density histograms at were created using the R function hist() with probability=T; probability density lines were estimated using the R density() function.



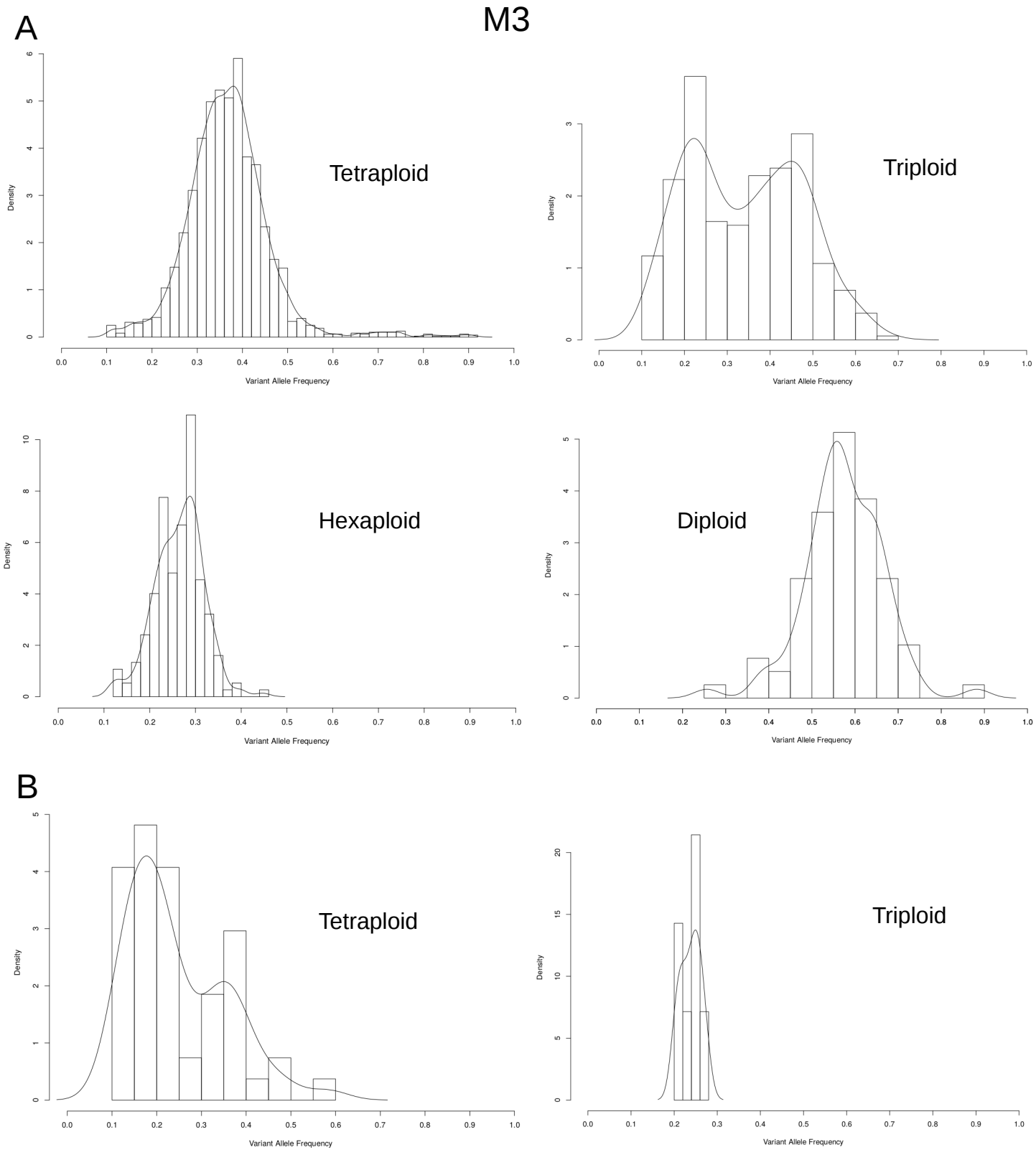
Supplemental Fig. S6 Analysis of variant allele frequency (VAF) effects on mutational signatures in mouse tumor M2. SNV **(A)** and mutation spectra **(B)** distribution across different VAFs. T, transcribed; N, non-transcribed **(C)** Tumor content (cellularity) and ploidy **(D)** Copy number and depth ratio across the genome.



Supplemental Fig. S7 Probability densities of G:C > T:A **(A)** and A:T > C:G **(B)** mutation counts by variant allele frequency (VAF) in different copy number regions of the genome of mouse tumor M2. Probability density histograms at were created using the R function hist() with probability=T; probability density lines were estimated using the R density() function.



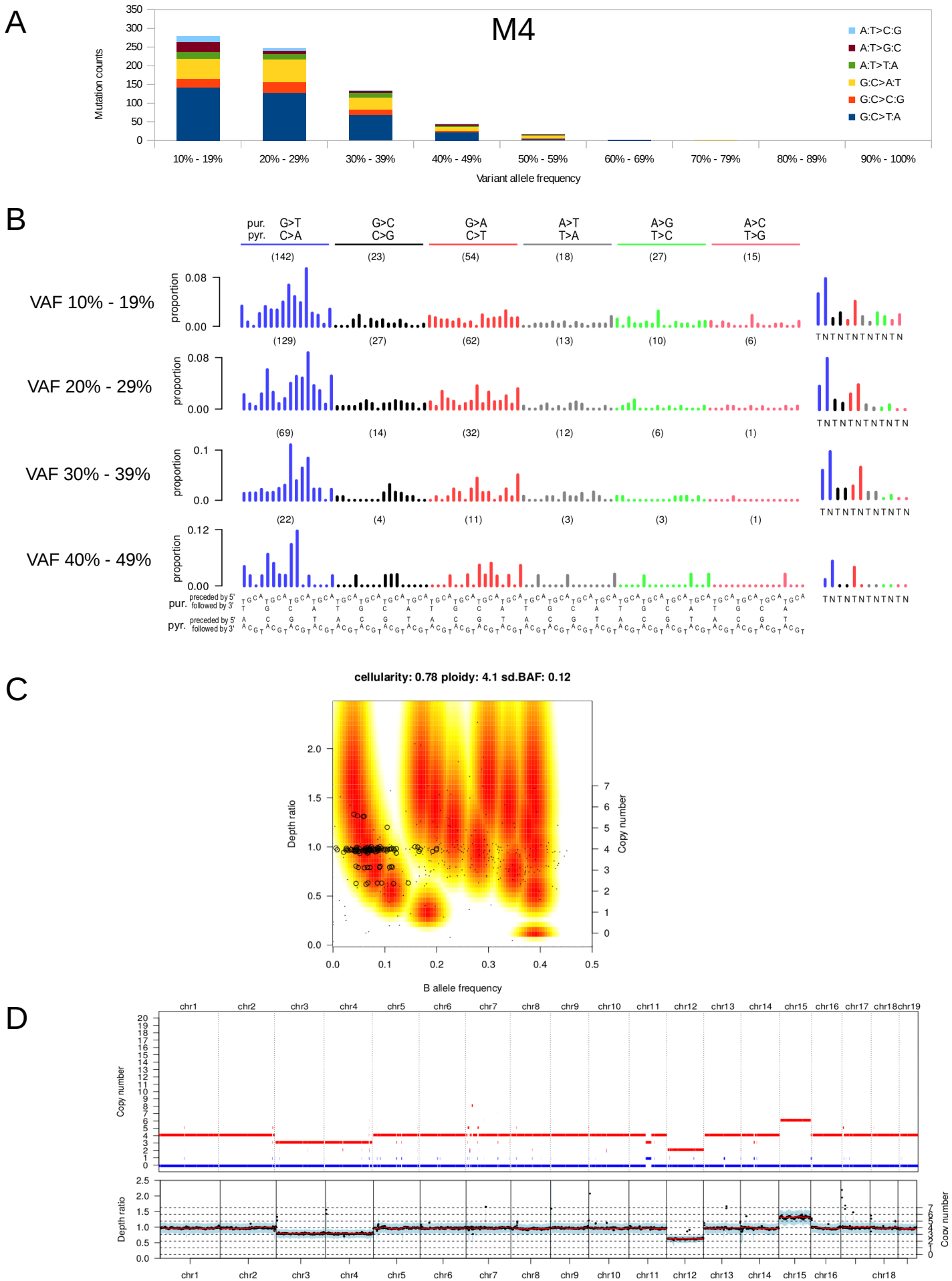
Supplemental Fig. S8 Analysis of variant allele frequency (VAF) effects on mutational signatures in mouse tumor M3. SNV **(A)** and mutation spectra **(B)** distribution across different VAFs. T, transcribed; N, non-transcribed **(C)** Tumor content (cellularity) and ploidy **(D)** Copy number and depth ratio across the genome.



Hexaploid region only contain two somatic mutations with VAF of 0.229 and 0.107 respectively

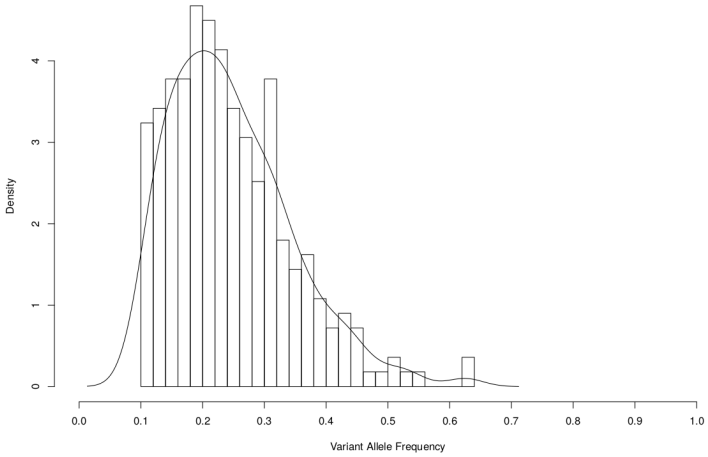
Diploid region only contain three somatic mutations with VAF of 0.290, 0.231 and 0.52 respectively

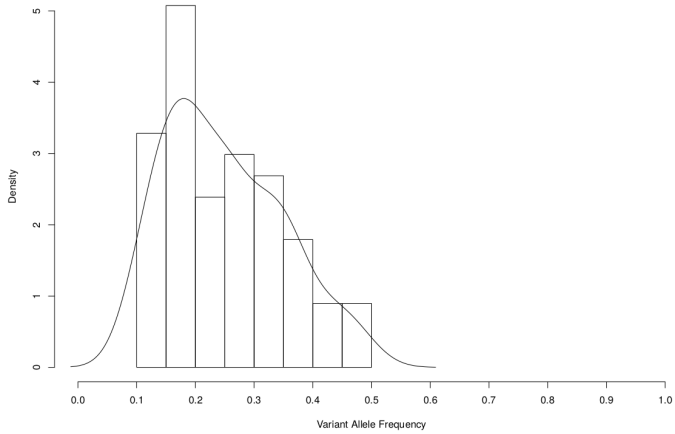
Supplemental Fig. S9 Probability densities of G:C > T:A **(A)** and A:T > C:G **(B)** mutation counts by variant allele frequency (VAF) in different copy number regions of the genome of mouse tumor M3. Probability density histograms at were created using the R function hist() with probability=T; probability density lines were estimated using the R density() function.

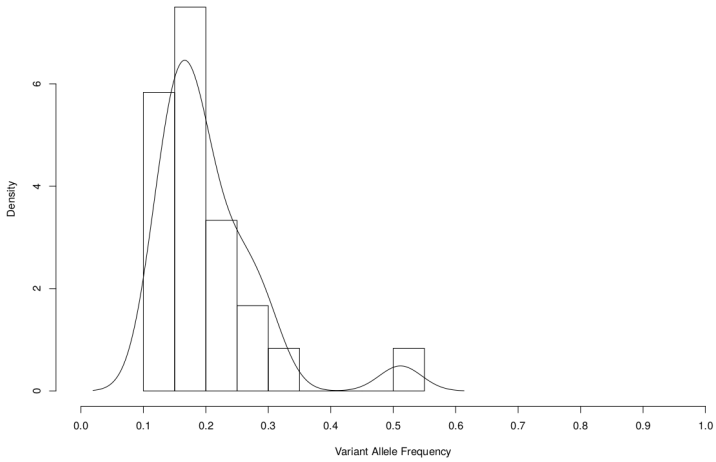


Supplemental Fig. S10 Analysis of variant allele frequency (VAF) effects on mutational signatures in mouse tumor M4. SNV **(A)** and mutation spectra **(B)** distribution across different VAFs. T, transcribed; N, non-transcribed **(C)** Tumor content (cellularity) and ploidy **(D)** Copy number and depth ratio across the genome.

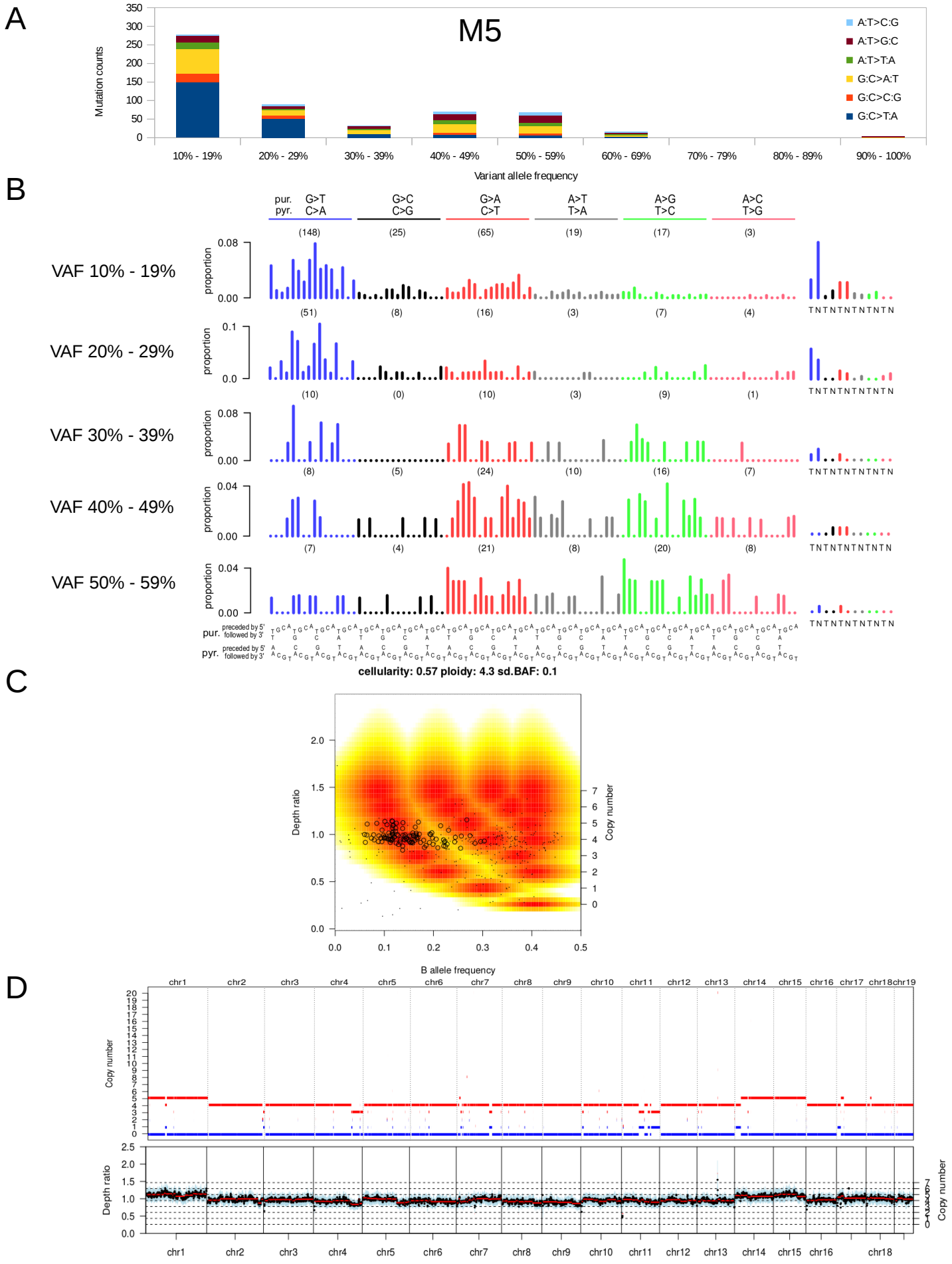
M4





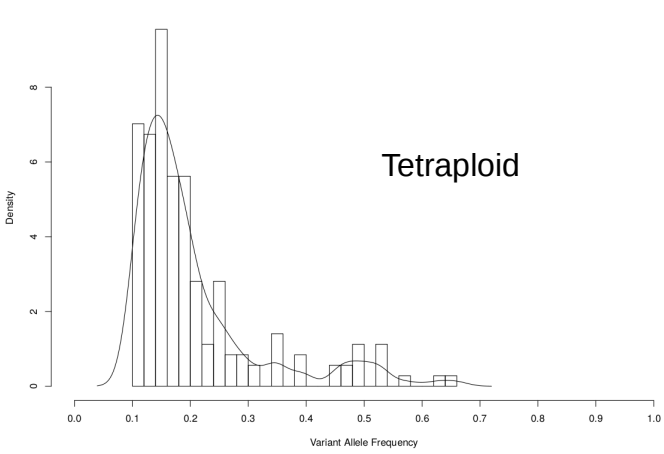


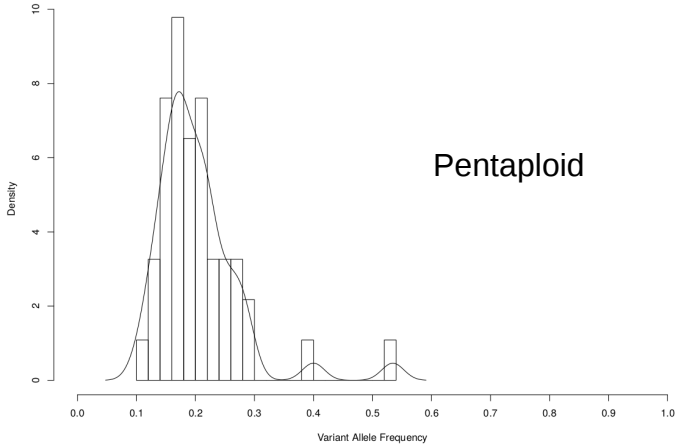
Supplemental Fig. S11 Probability densities of G:C > T:A mutation counts by variant allele frequency (VAF) in different copy number regions of the genome of mouse tumor M4. Probability density histograms at were created using the R function hist() with probability=T; probability density lines were estimated using the R density() function.



Supplemental Fig. S12 Analysis of variant allele frequency (VAF) effects on mutational signatures in mouse tumor M5. SNV **(A)** and mutation spectra **(B)** distribution across different VAFs. T, transcribed; N, non-transcribed **(C)** Tumor content (cellularity) and ploidy **(D)** Copy number and depth ratio across the genome.

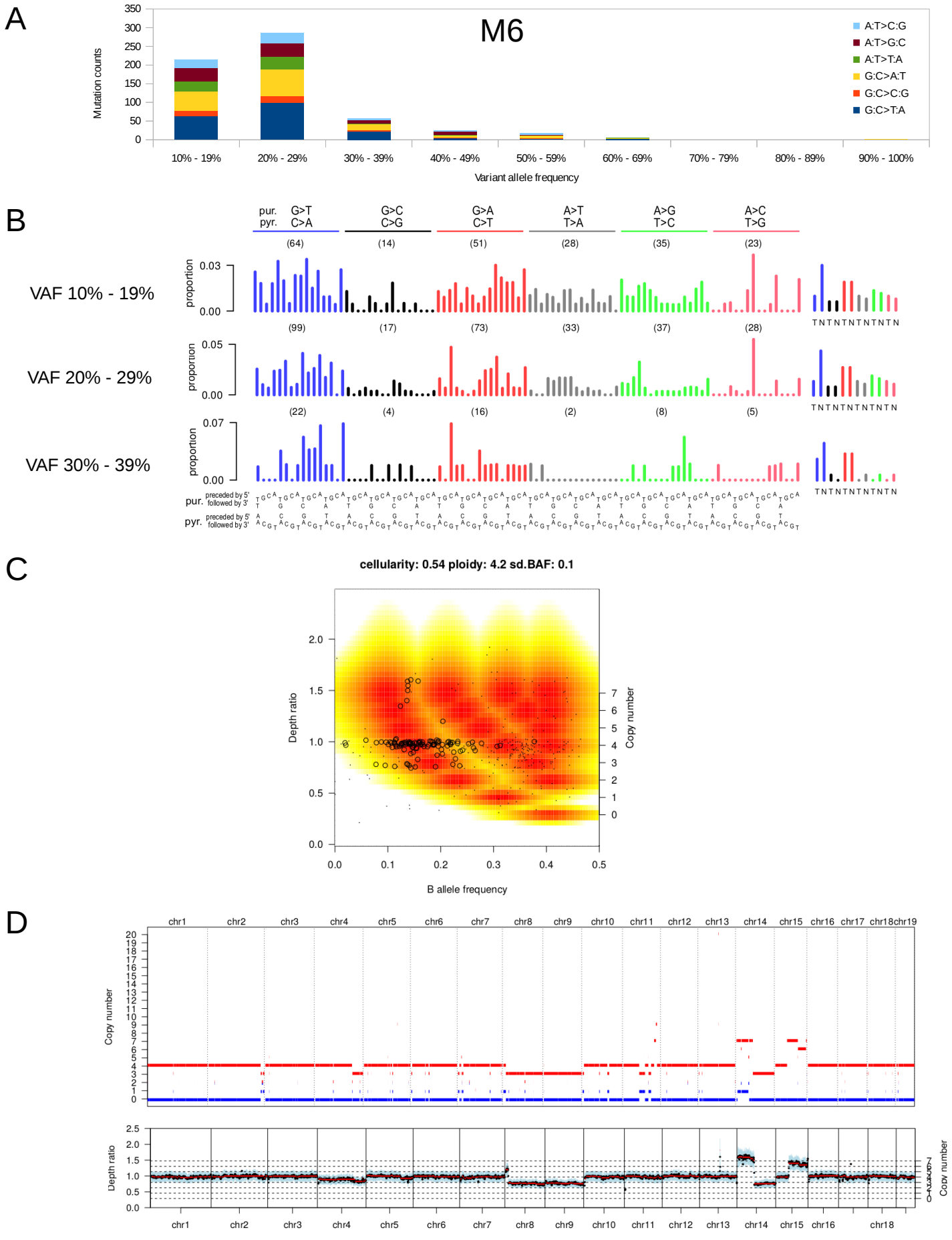
M5





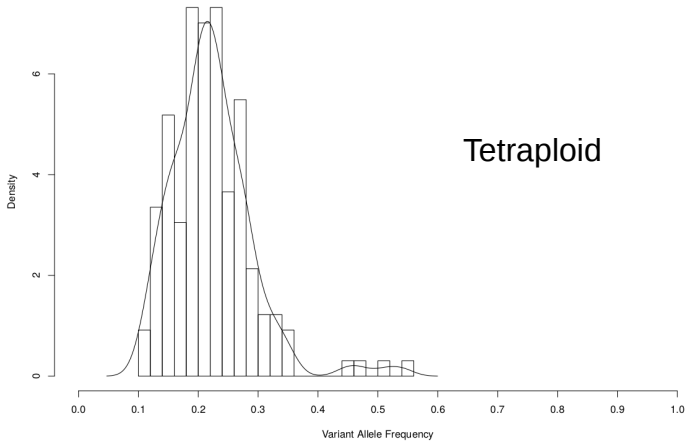
Triploid region only contain two somatic mutations with VAF of 0.150 and 0.176 respectively

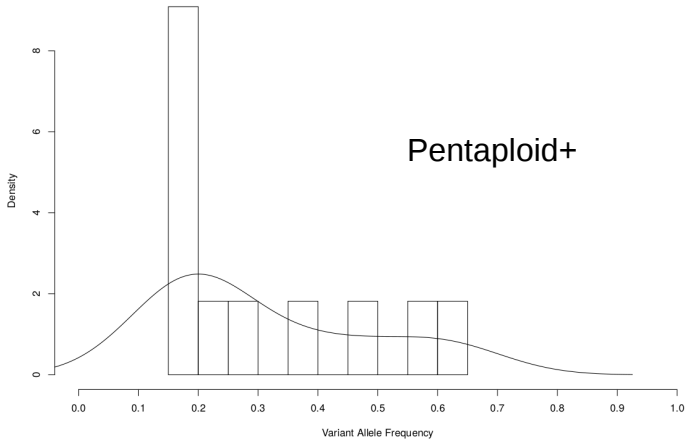
Supplemental Fig. S13 Probability densities of G:C > T:A mutation counts by variant allele frequency (VAF) in different copy number regions of the genome of mouse tumor M5. Probability density histograms at were created using the R function hist() with probability=T; probability density lines were estimated using the R density() function.

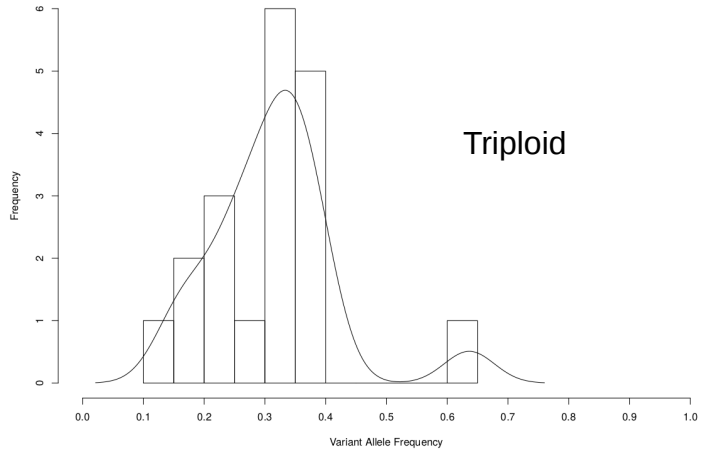


Supplemental Fig. S14 Analysis of variant allele frequency (VAF) effects on mutational signatures in mouse tumor M6. SNV **(A)** and mutation spectra **(B)** distribution across different VAFs. T, transcribed; N, non-transcribed **(C)** Tumor content (cellularity) and ploidy **(D)** Copy number and depth ratio across the genome.

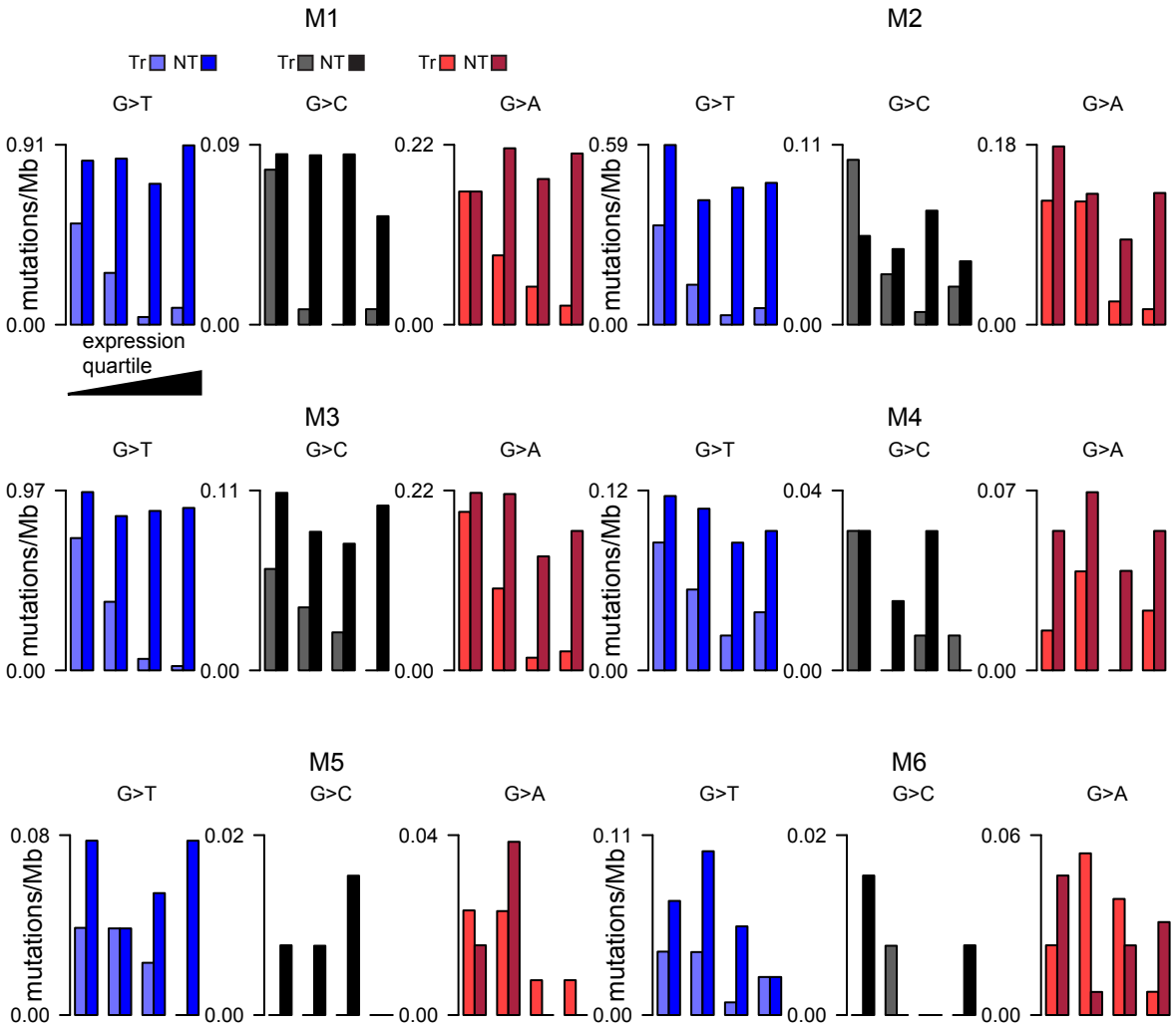
M6



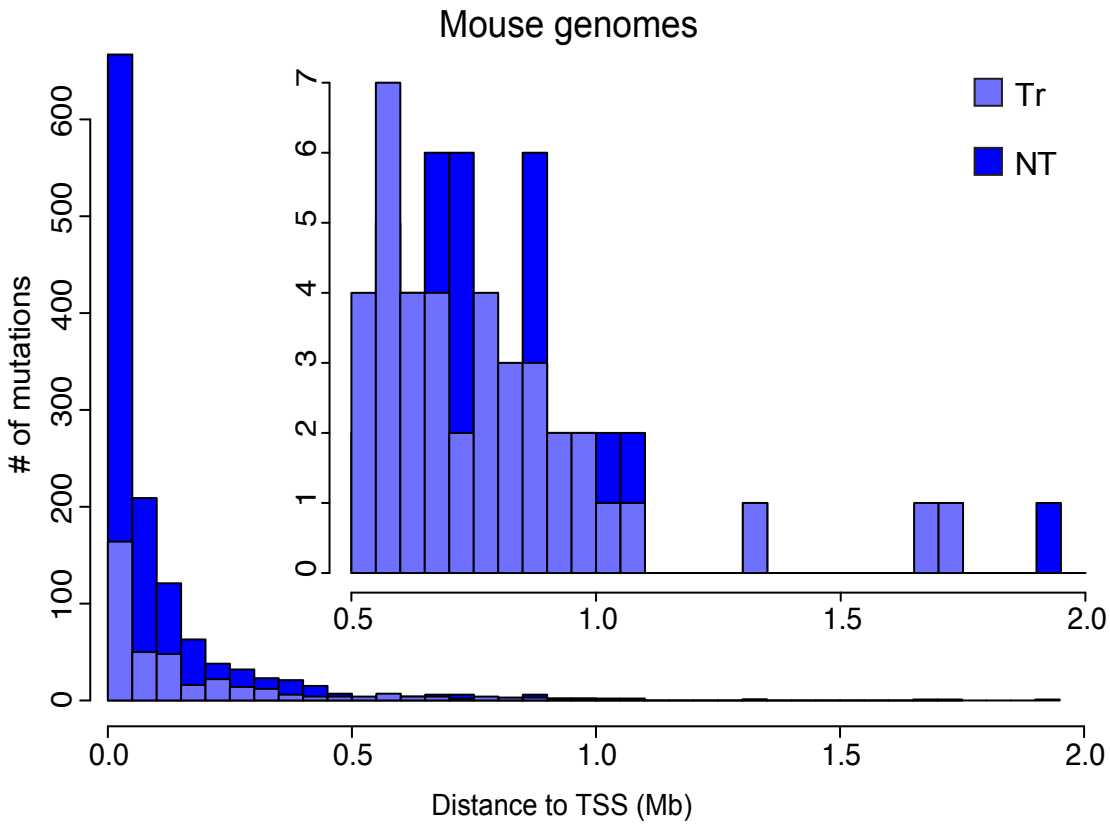




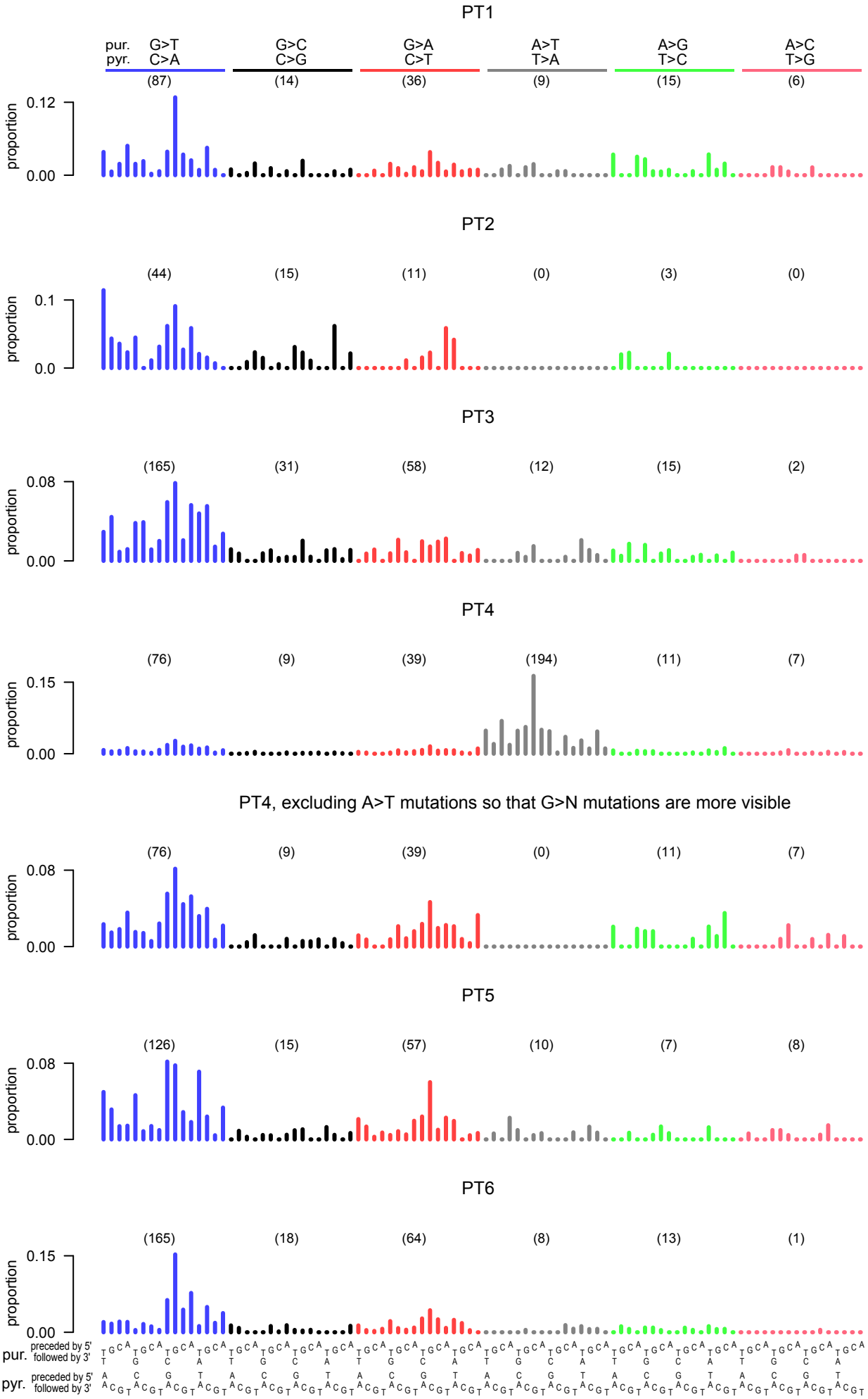
Supplemental Fig. S15 Probability densities of G:C > T:A mutation counts by variant allele frequency (VAF) in different copy number regions of the genome of mouse tumor M6. Probability density histograms at were created using the R function hist() with probability=T; probability density lines were estimated using the R density() function



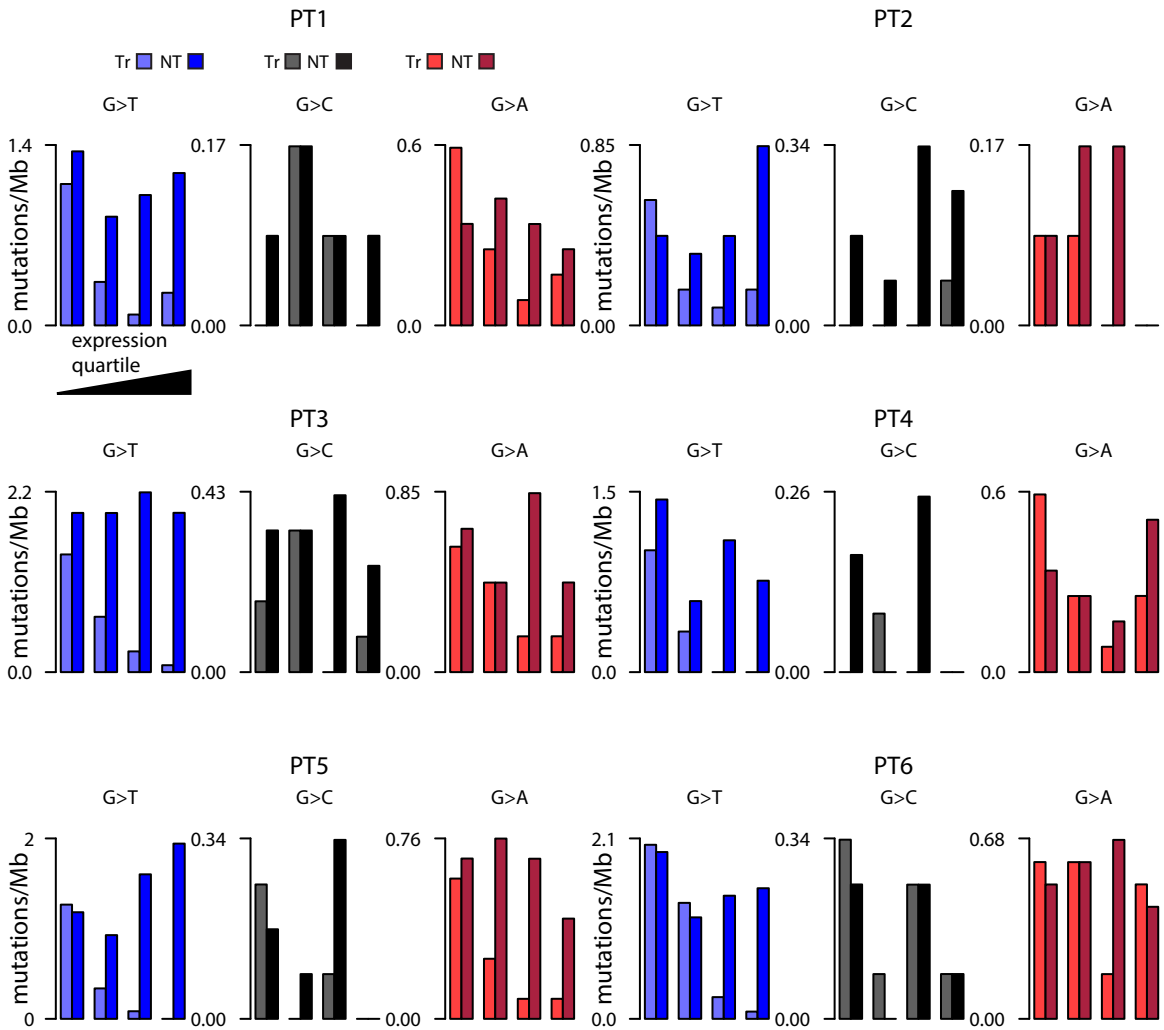
Supplemental Fig. S16 Transcription strand bias by expression quartile for liver tumors from mice treated with AFB1



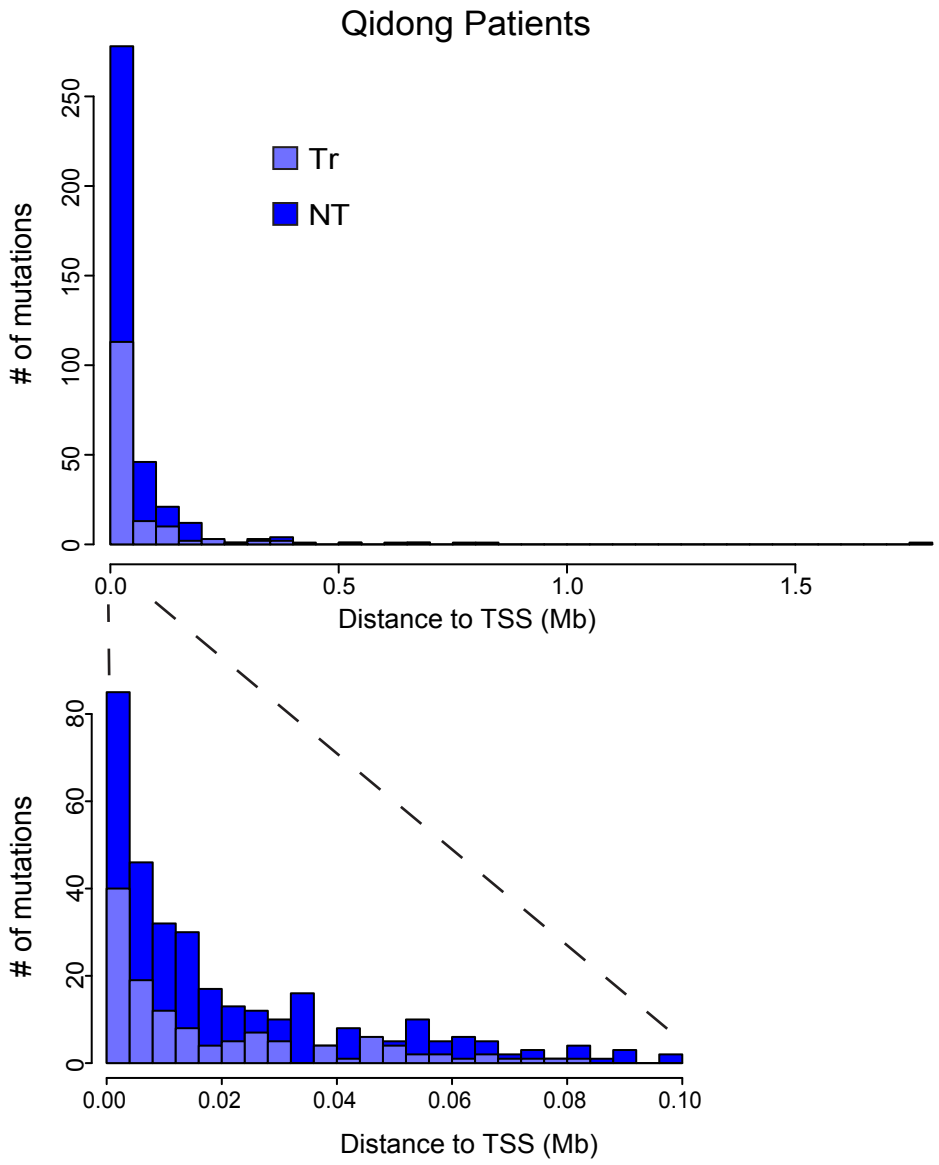
Supplemental Fig. S17 Transcription strand bias along the transcripts for liver tumors from mice treated with AFB1



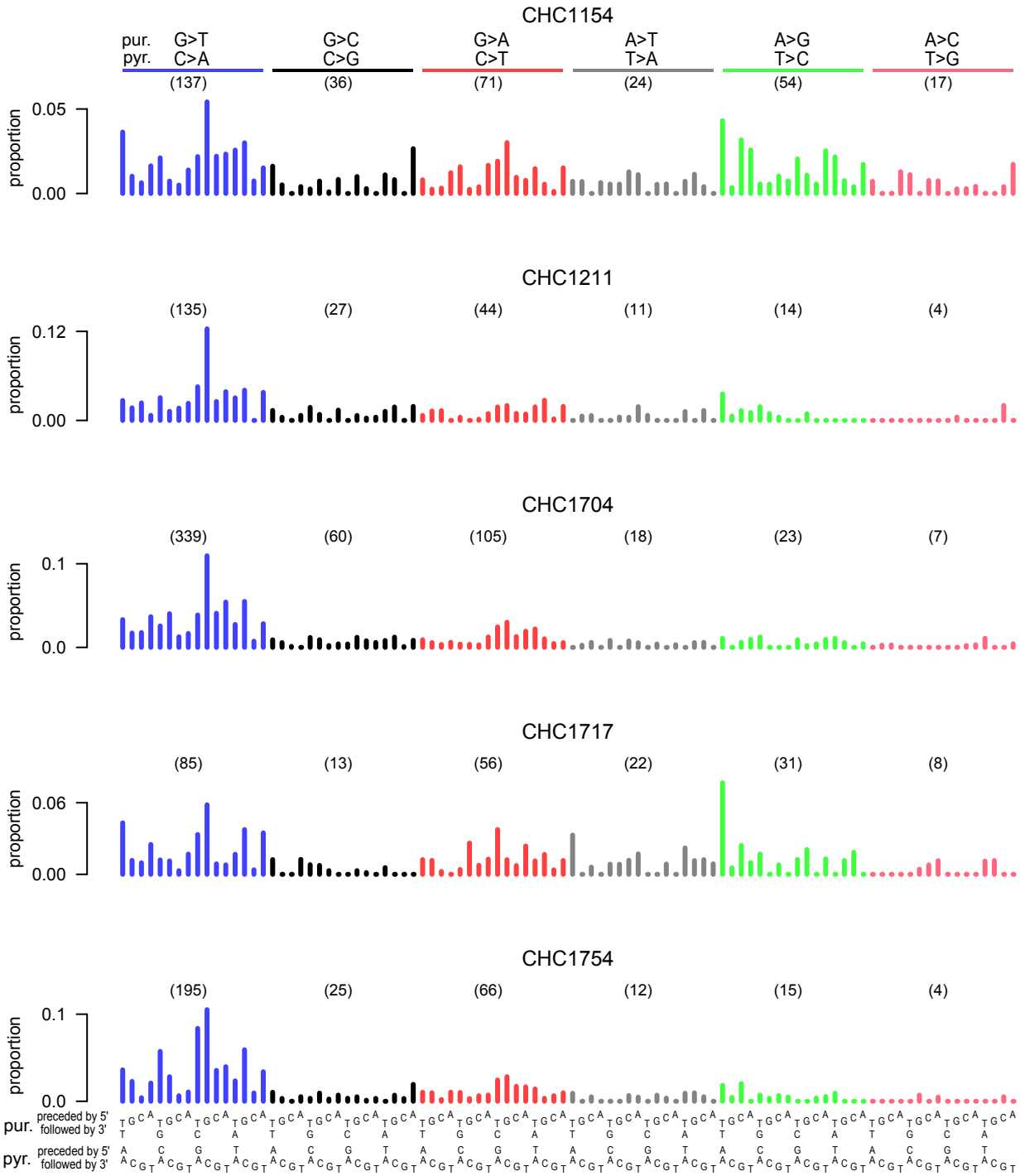
Supplemental Fig. S18 Mutation spectra for Qidong HCCs



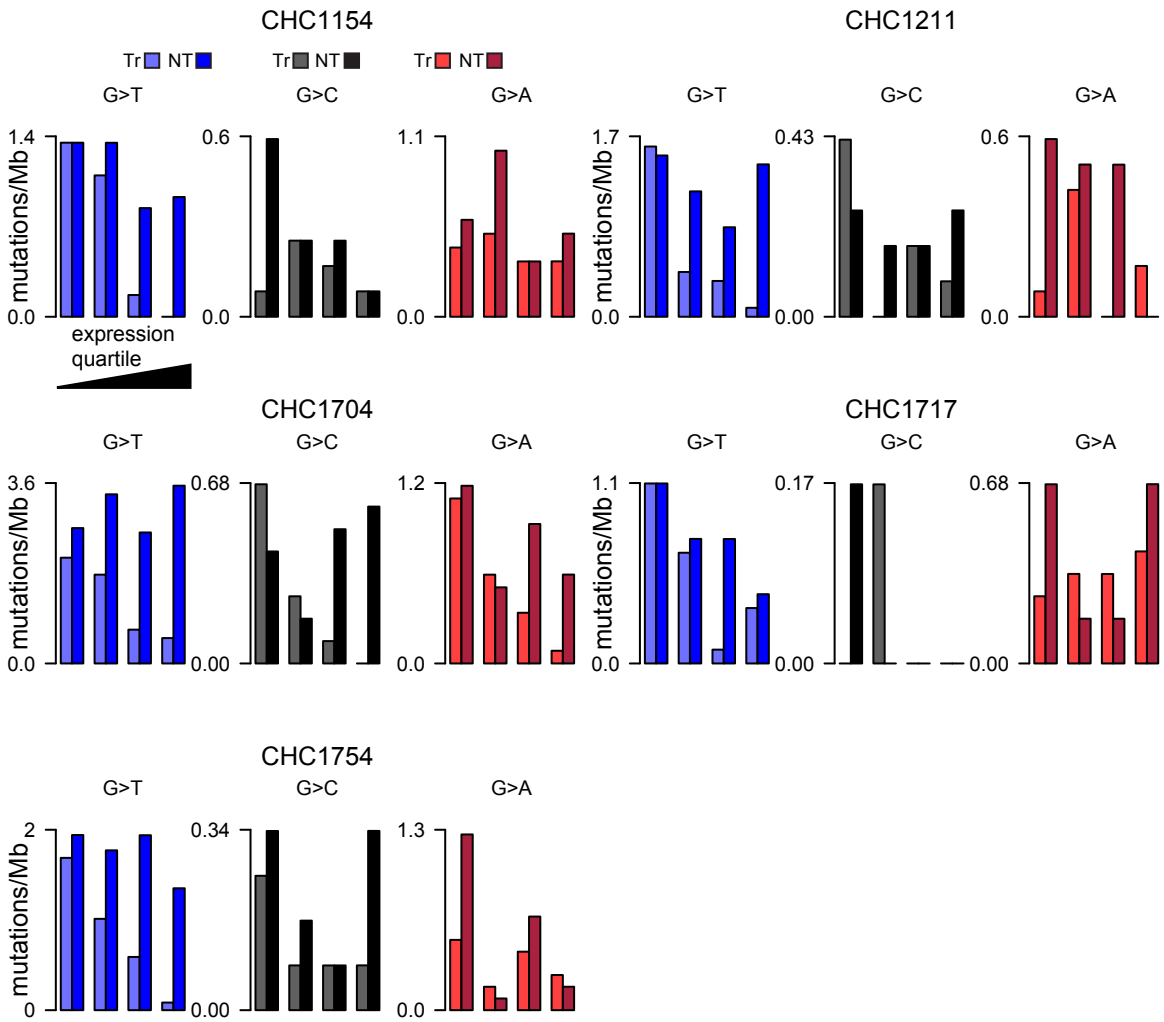
Supplemental Fig. S19 Transcription strand bias by expression quartile for Qidong HCCs



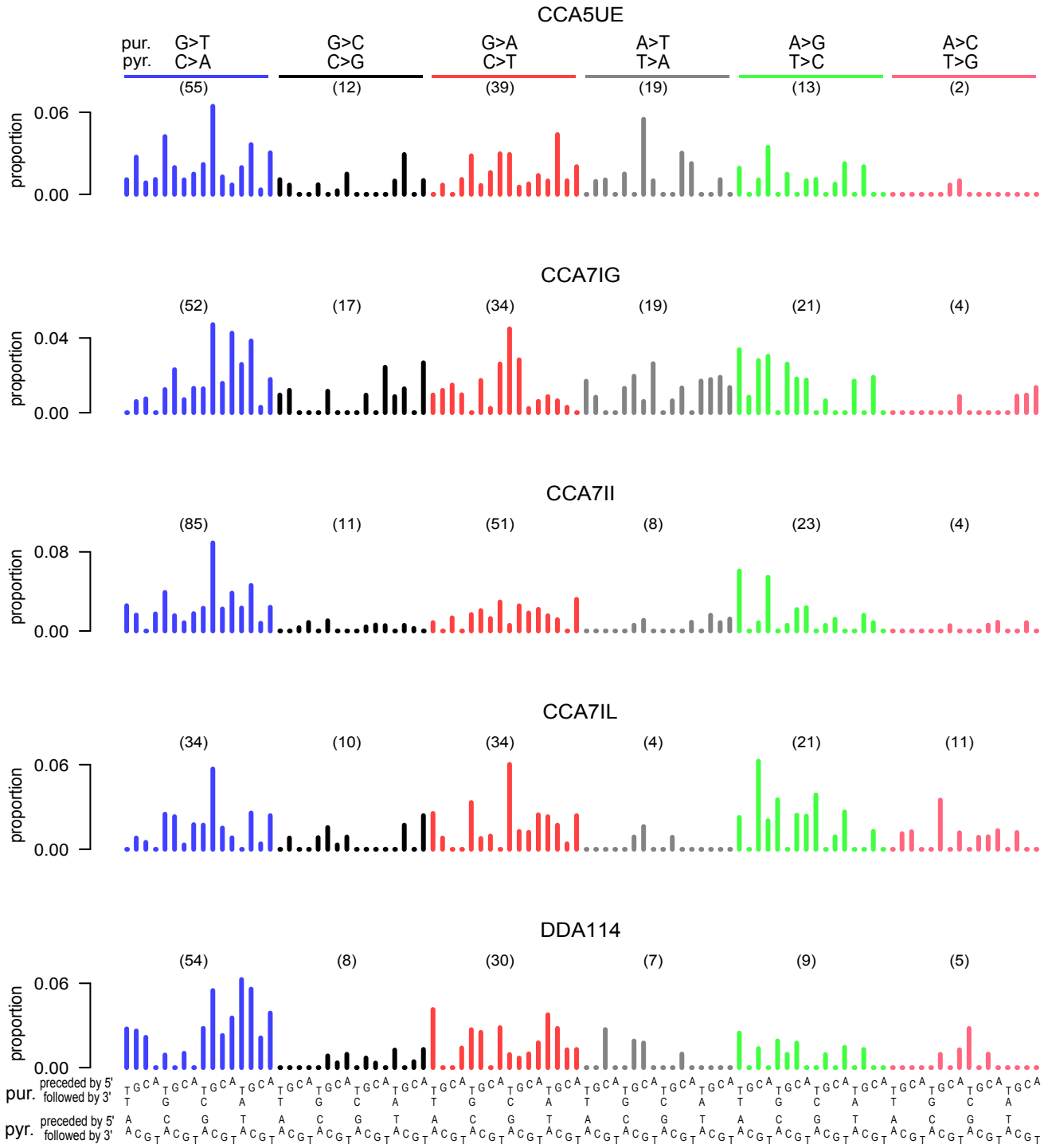
Supplemental Fig. S20 For the Qidong HCCs, the number of mutations is too low to assess whether transcription strand bias declines from the 5' to 3' end of transcripts.



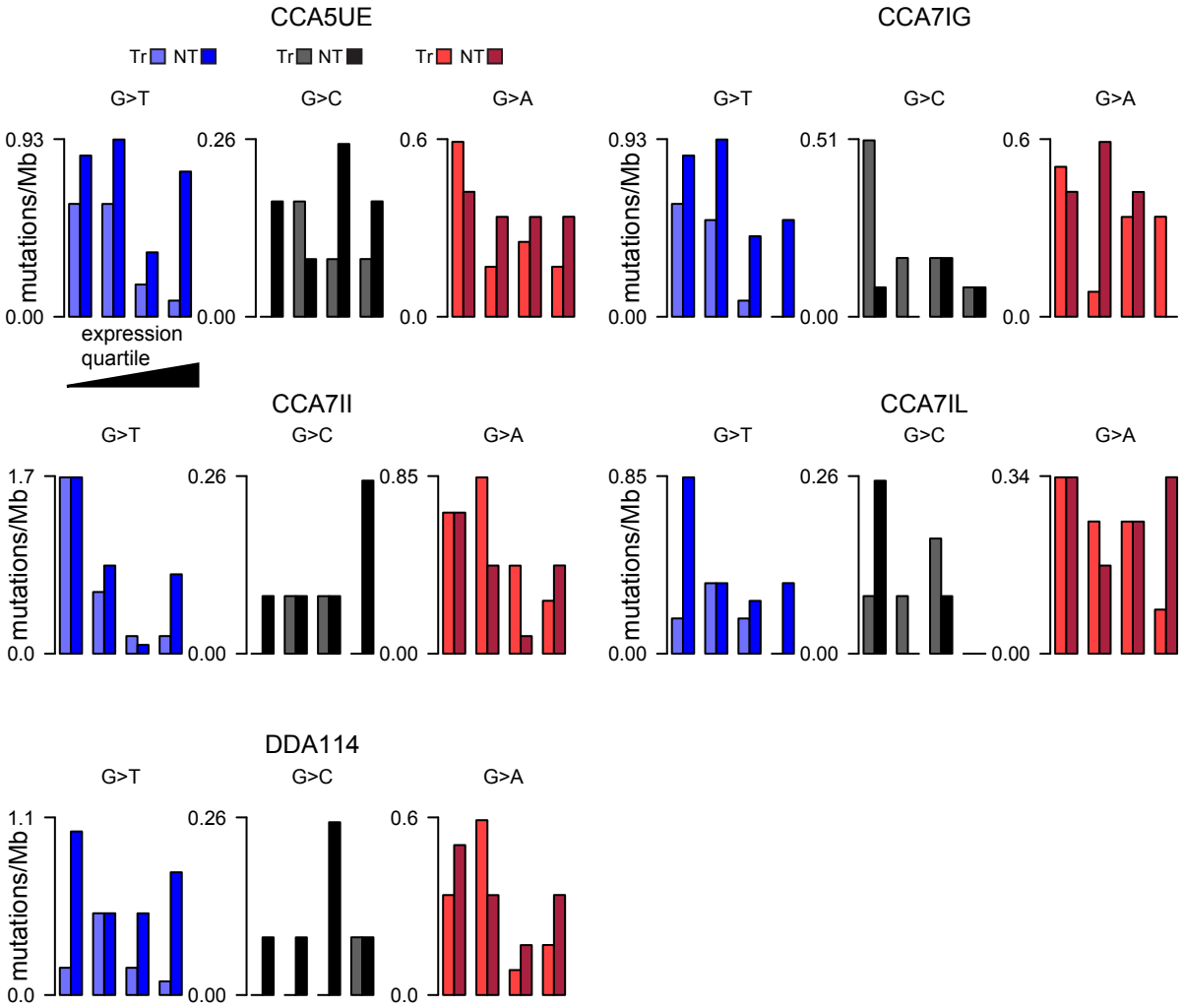
Supplemental Fig. S21 Mutation spectra for 5 HCCs from recent African immigrants from Schulze et al., 2015.



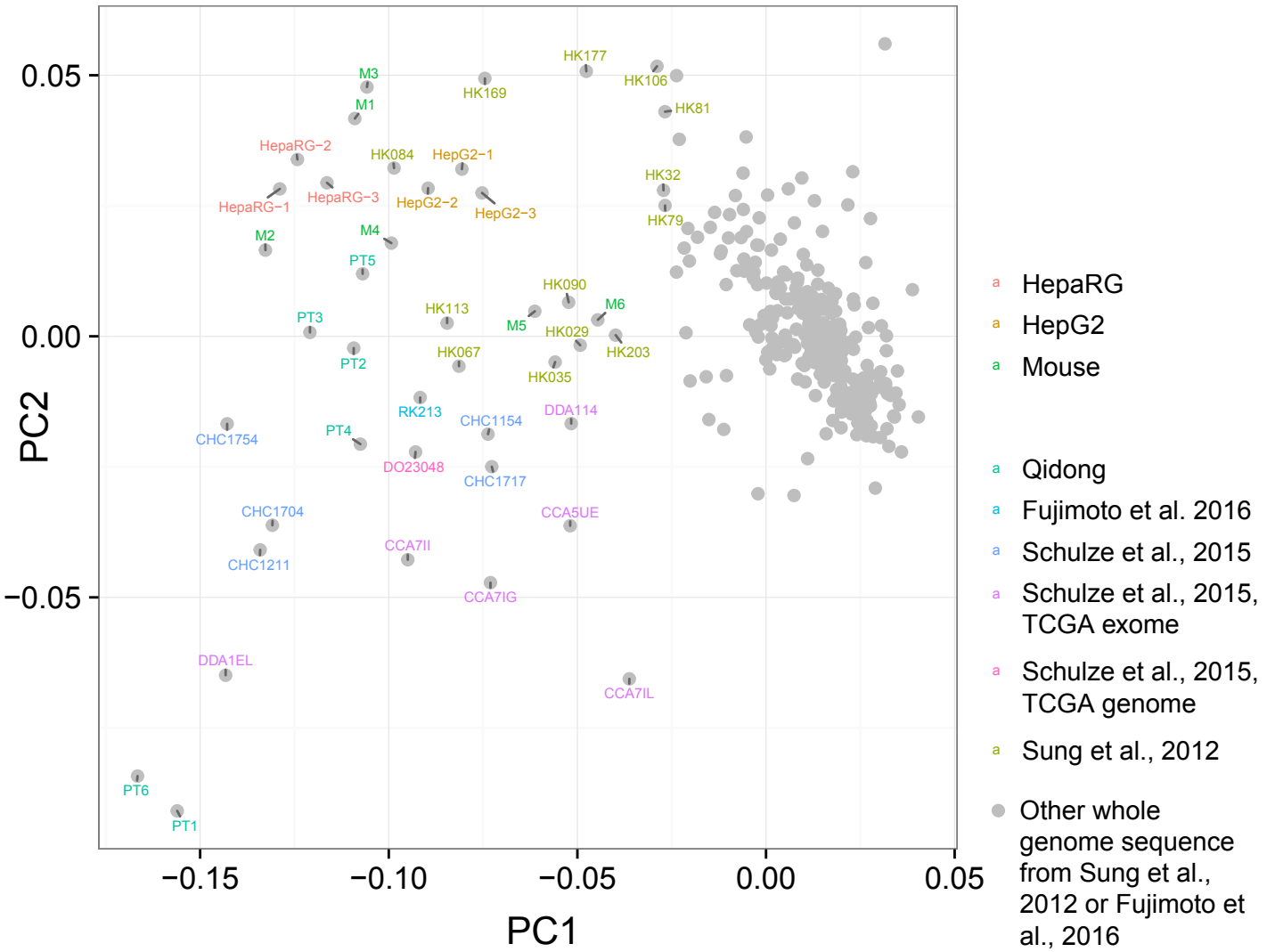
Supplemental Fig. S22 Transcription strand bias by expression quartile for 5 HCCs from recent African immigrants reported in Schulze et al., 2015.



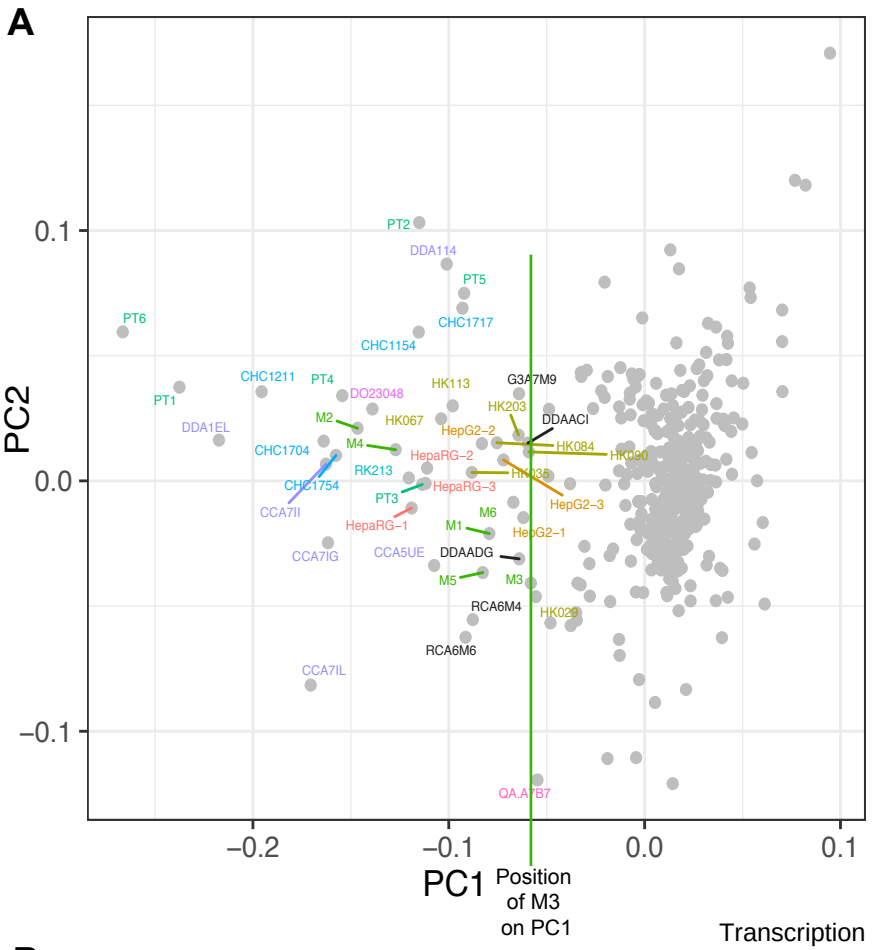
Supplemental Fig. S23 Mutation spectra for TCGA whole exome data reported in Schulze et al., 2015 as likely reflecting aflatoxin exposure.



Supplemental Fig. S24 Transcription strand bias by expression quartile for TCGA WESdatareported as likely reflecting aflatoxin exposure in Schulze et al., 2015.



Supplemental Fig. S25 Principal components analysis on proportions of G > N mutations in trinucleotide context in AFB1 treated cell lines, tumors from AFB1-treated mice, whole-genome-sequenced human HCCs, and selected whole-exome-sequenced human HCCs.

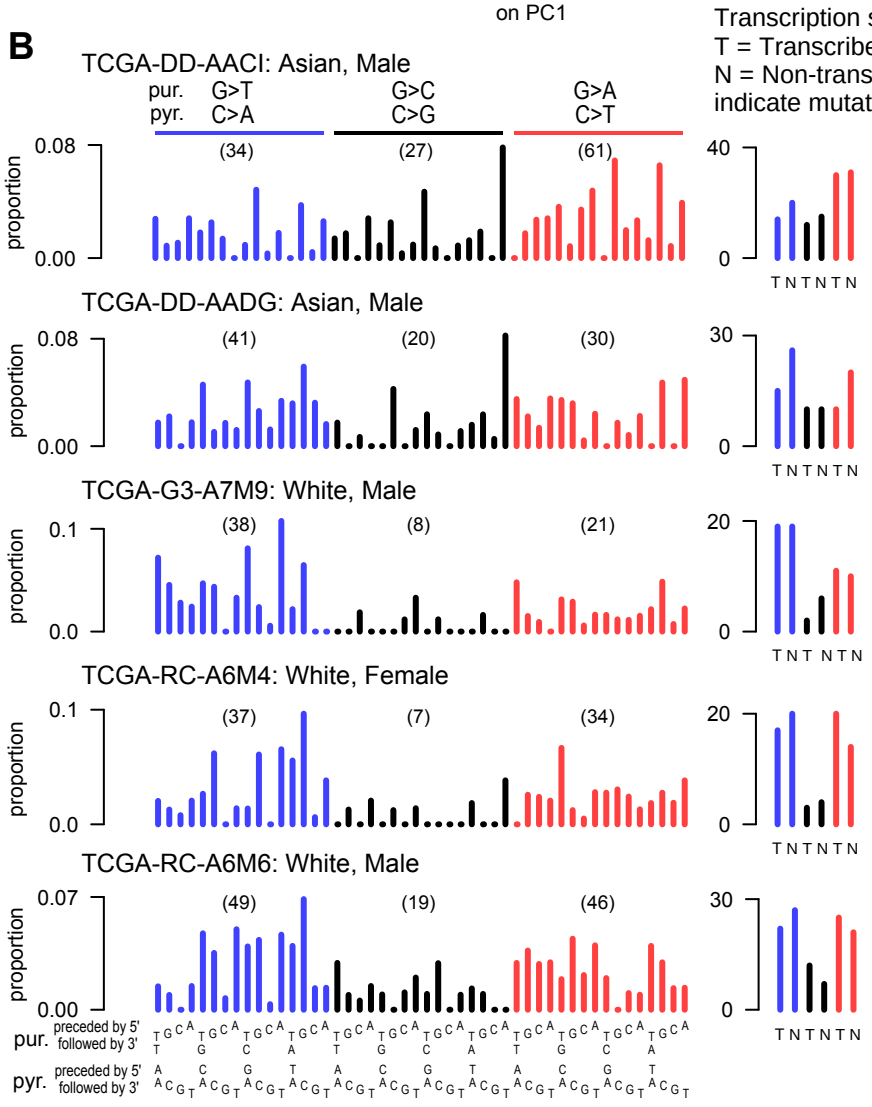


HepaRG
HepG2

Mouse

Qidong
Fujimoto et al., 2016
Shulze et al., 2015
Schulze et al., 2015 TCGA exome
Schulze et al., 2015,TCGA genome
Sung et al., 2012

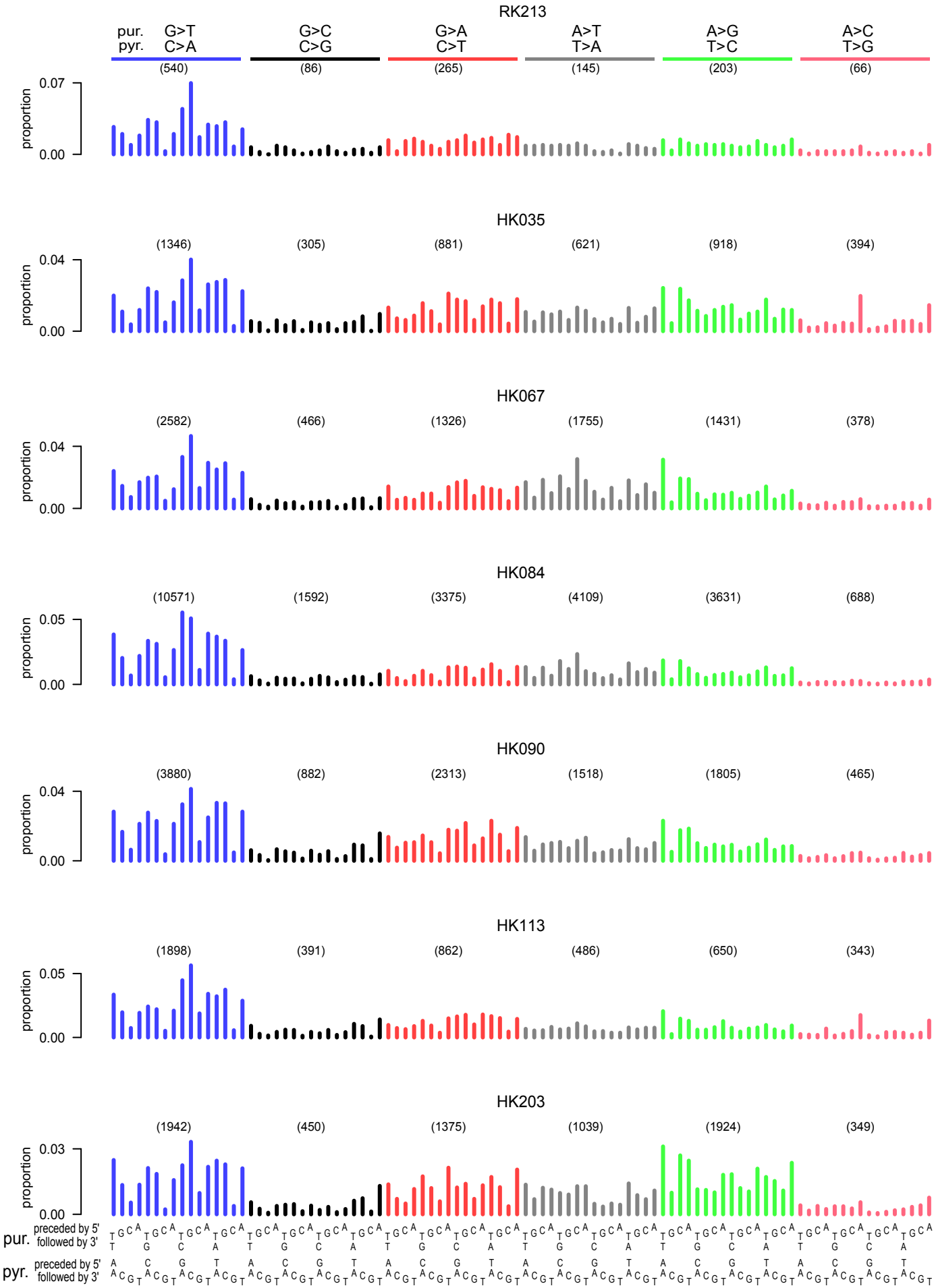
Other whole exome sequence from TCGA, or whole genome sequence from Sung et al., 2012 or Fujimoto et al., 2016



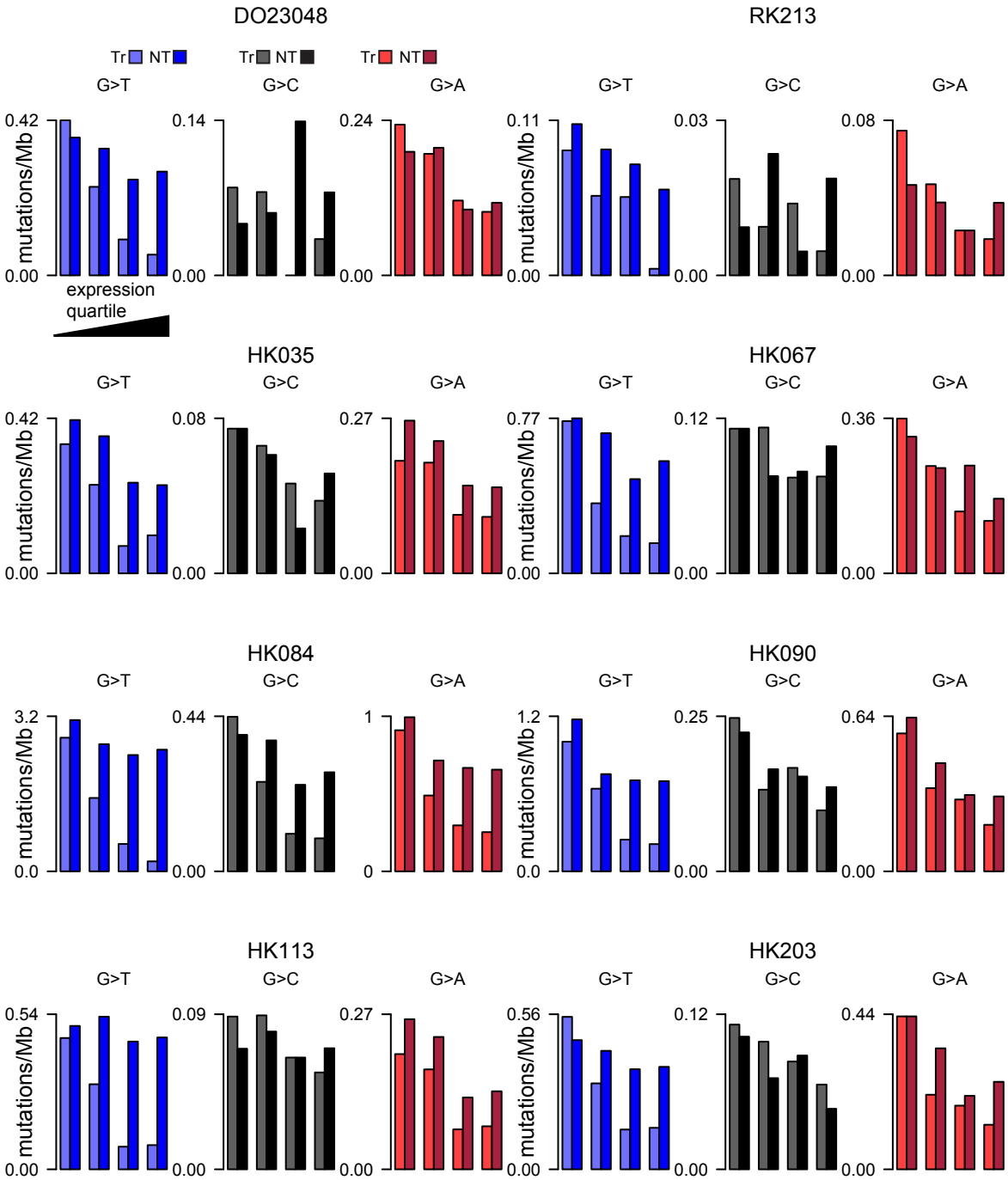
Supplemental Fig. S26 No convincing evidence of aflatoxin exposure in additional TCGA HCCs treated in North America.

(A) PCA of the proportions of G > T mutations in trinucleotide context as in main text Figure 4.

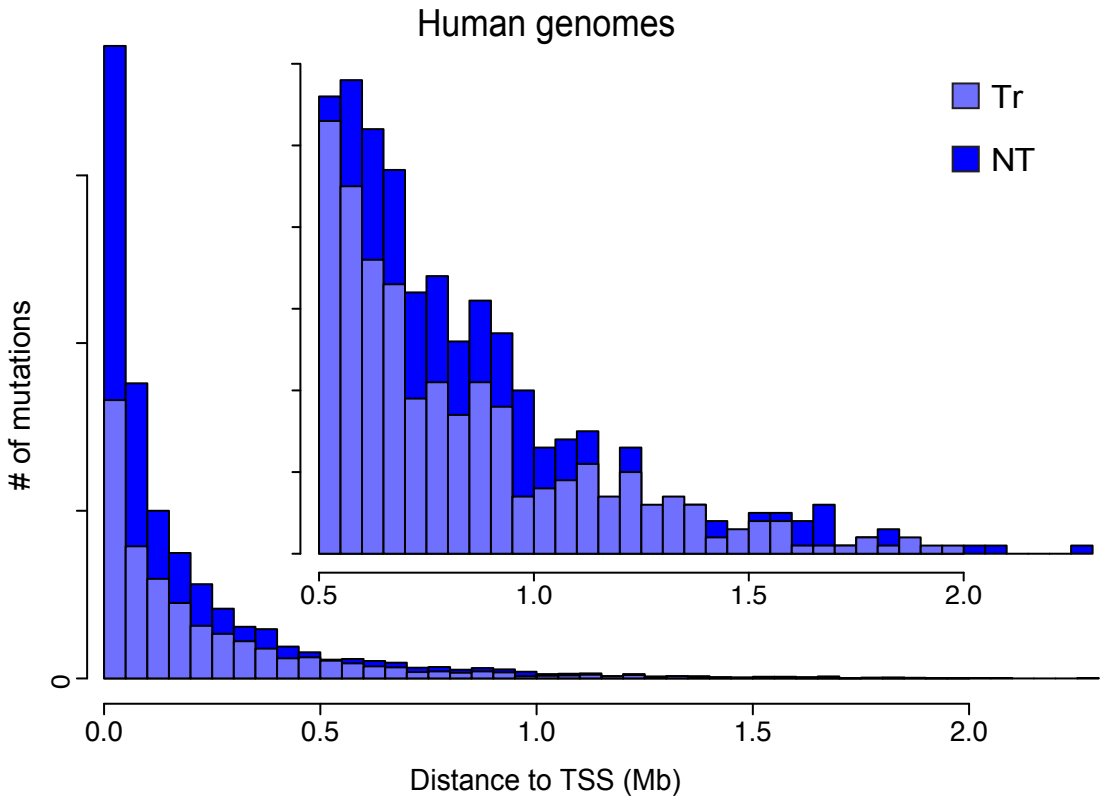
(B) Mutation spectra and transcriptional strand bias of G > N mutations of 5 TCGA HCCs.



Supplemental Fig. S27 Whole genome somatic mutation spectra likely reflecting aflatoxin exposure in human HCCs.



Supplemental Fig. S28 Transcription strand bias by expression quartile for whole genome somatic mutations likely reflecting aflatoxin exposure in human HCCs.



Supplemental Fig. S29 Transcription strand bias along the transcripts for likely aflatoxin exposed human genomes..