

PCAWG-Paper E

# Filtering of potential false positive SNV calls

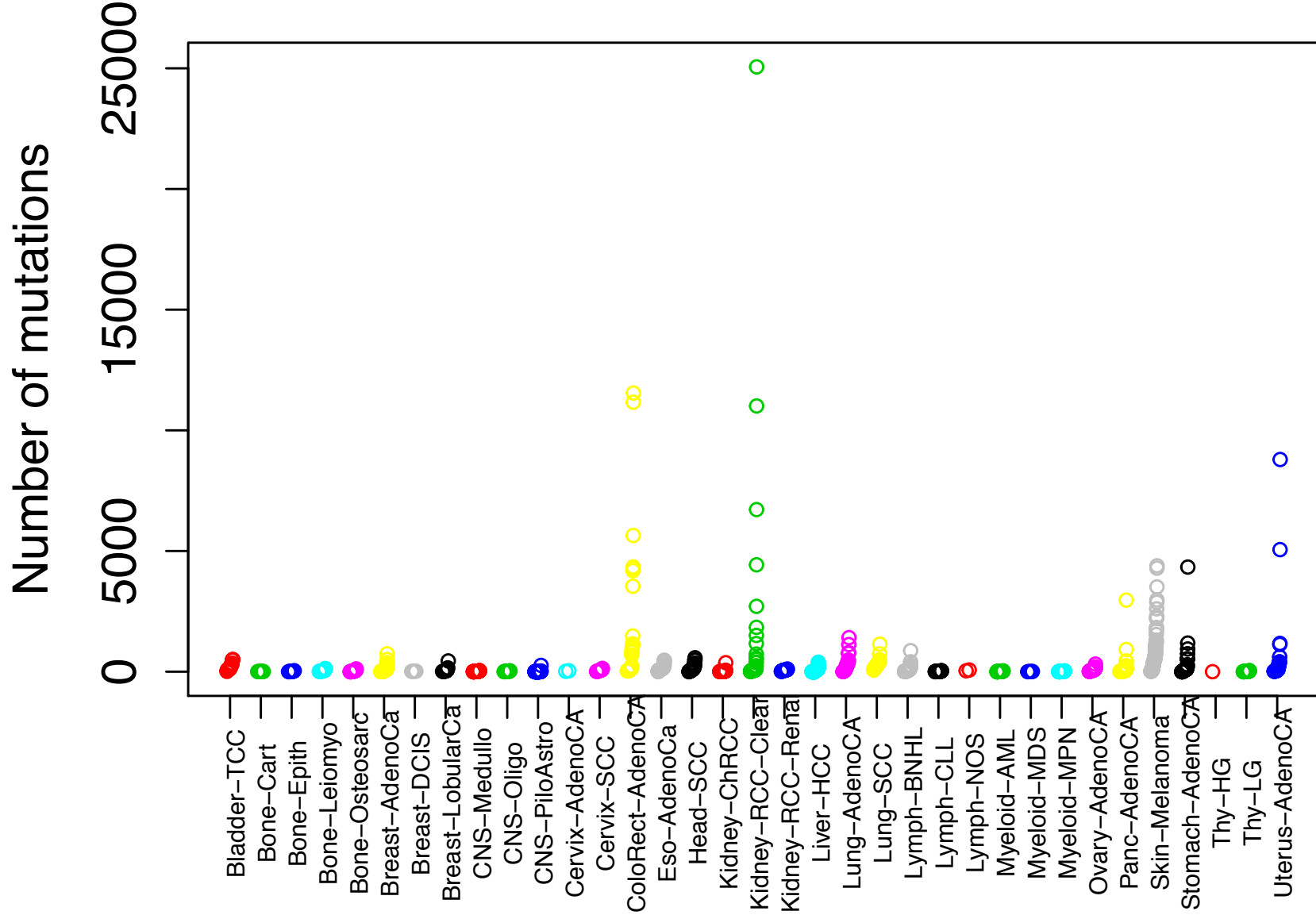
Filter1: remove SNVs marked as Low support in original VCFs

Filter2: remove somatic SNVs observed in the 1KG/DBSNP datasets

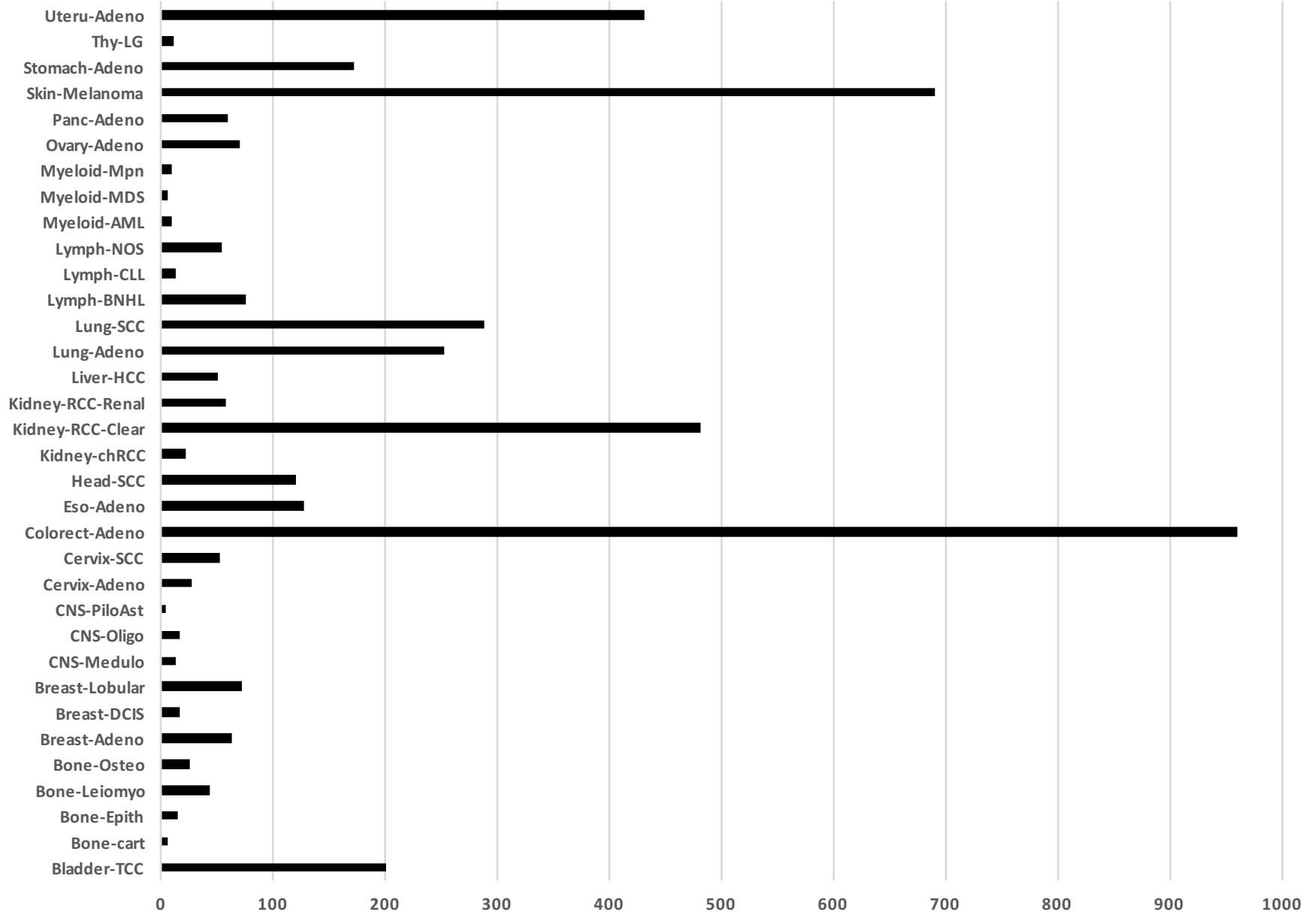
Filter3 : Apply the 1000 Genome Mask (only accounting for SNVs present in highly mapable region of the genome)

Filter4: remove SNVs observed in the low complexity region of the human genome (approx. 2% of the genome)

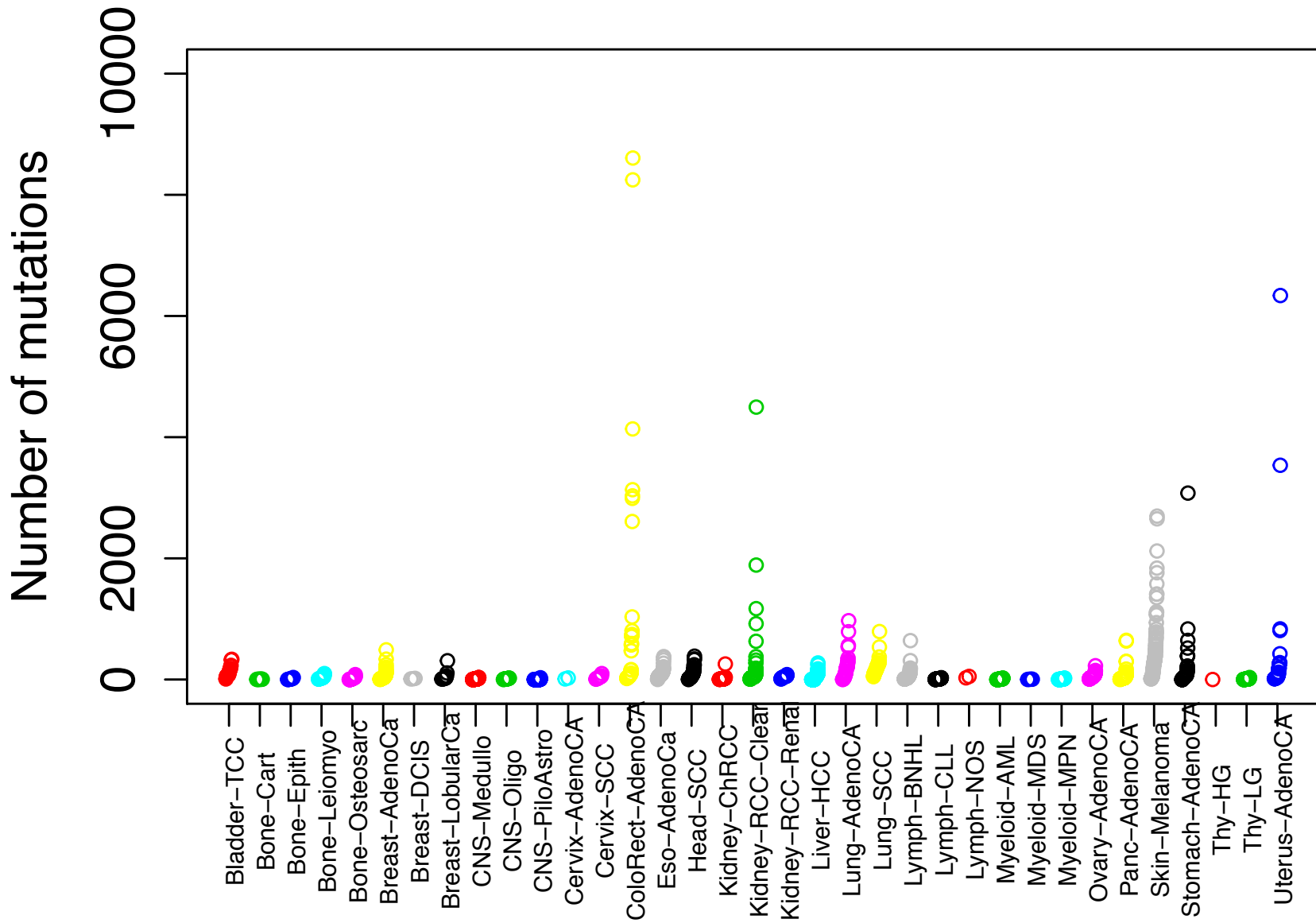
# Total Number of Coding mutations



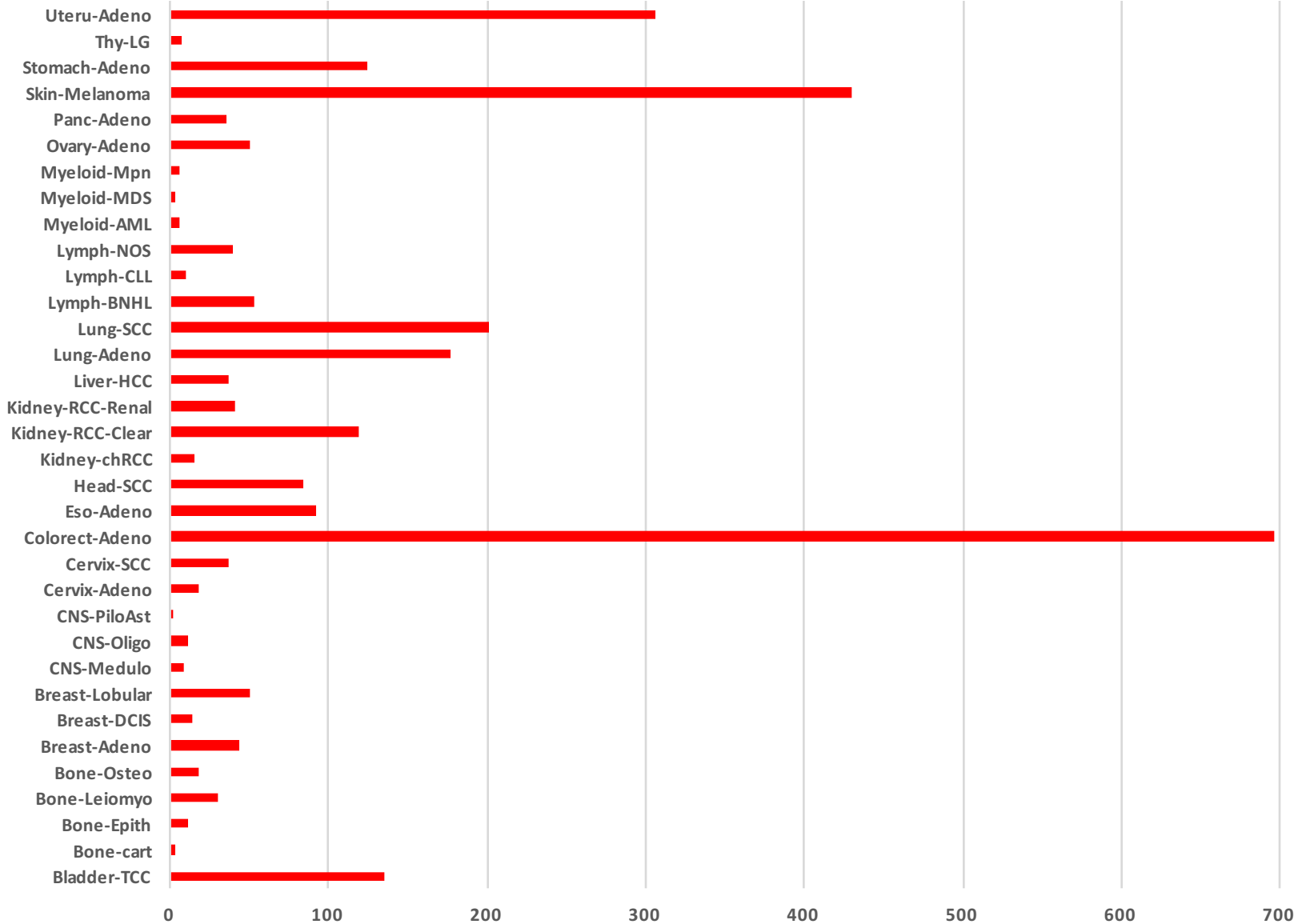
# Average Number of Coding mutations



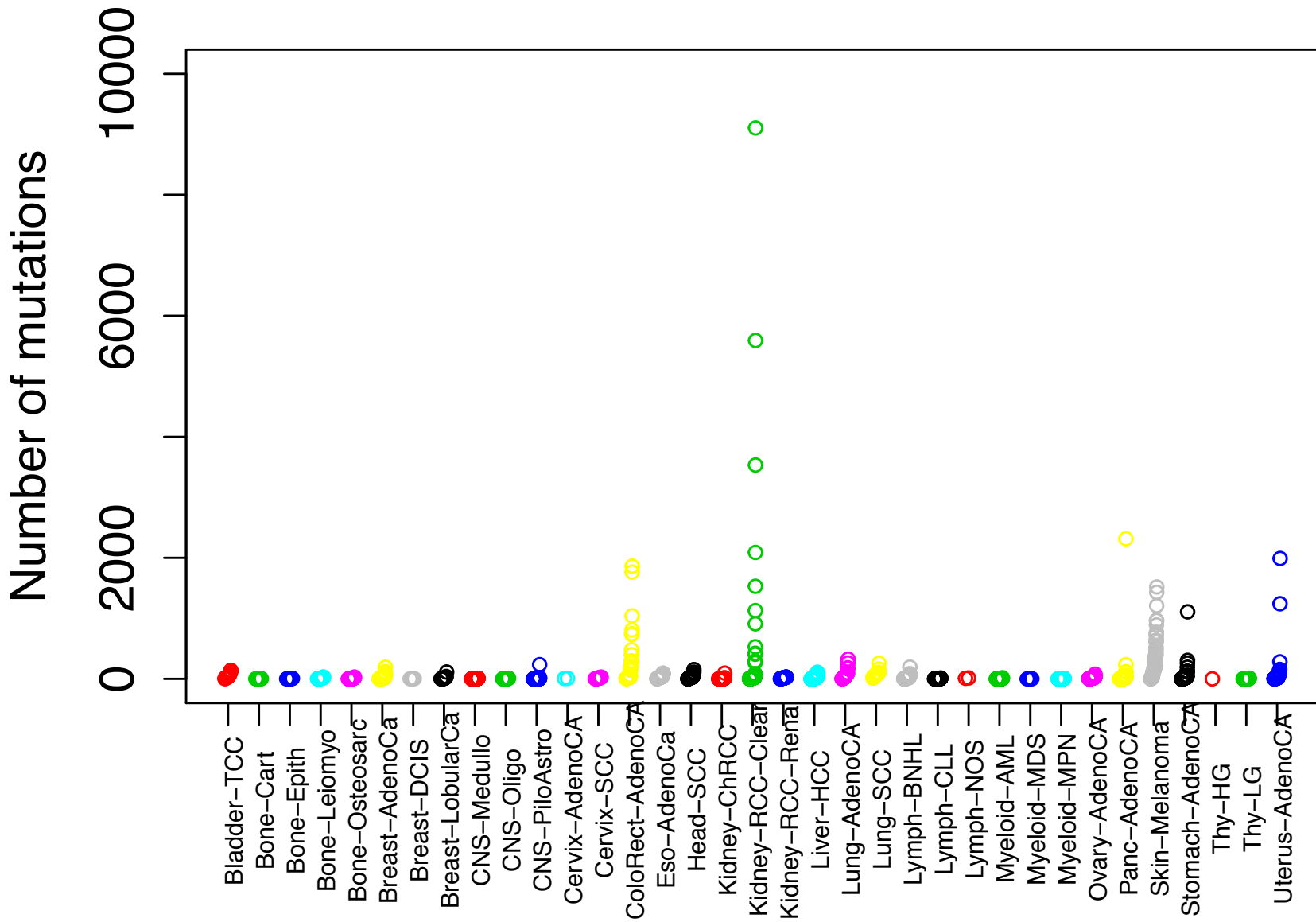
# Total Number of Nonsynonymous Mutations



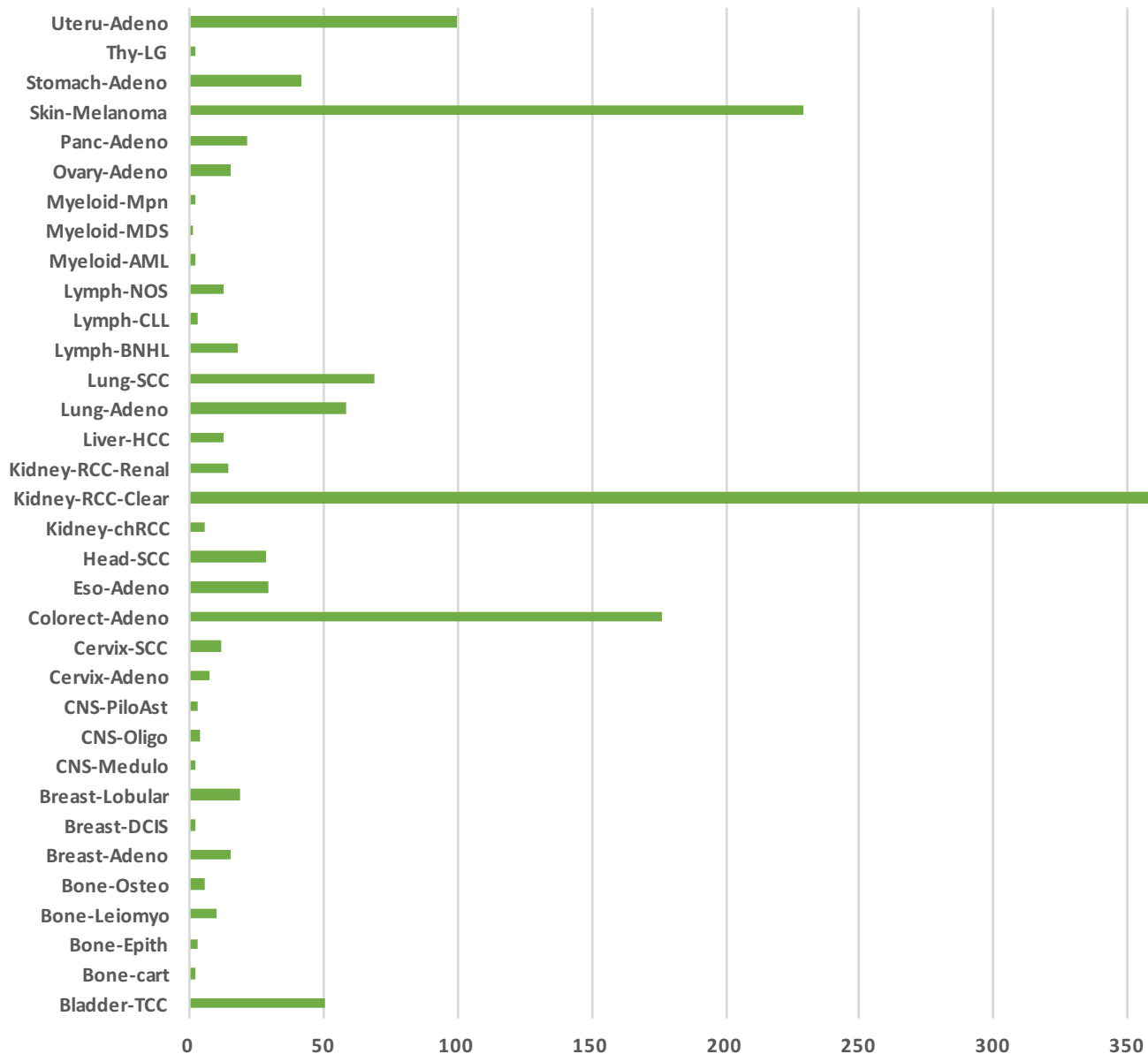
# Average Number of nonsynonymous mutations



# Total Number of Synonymous Mutations

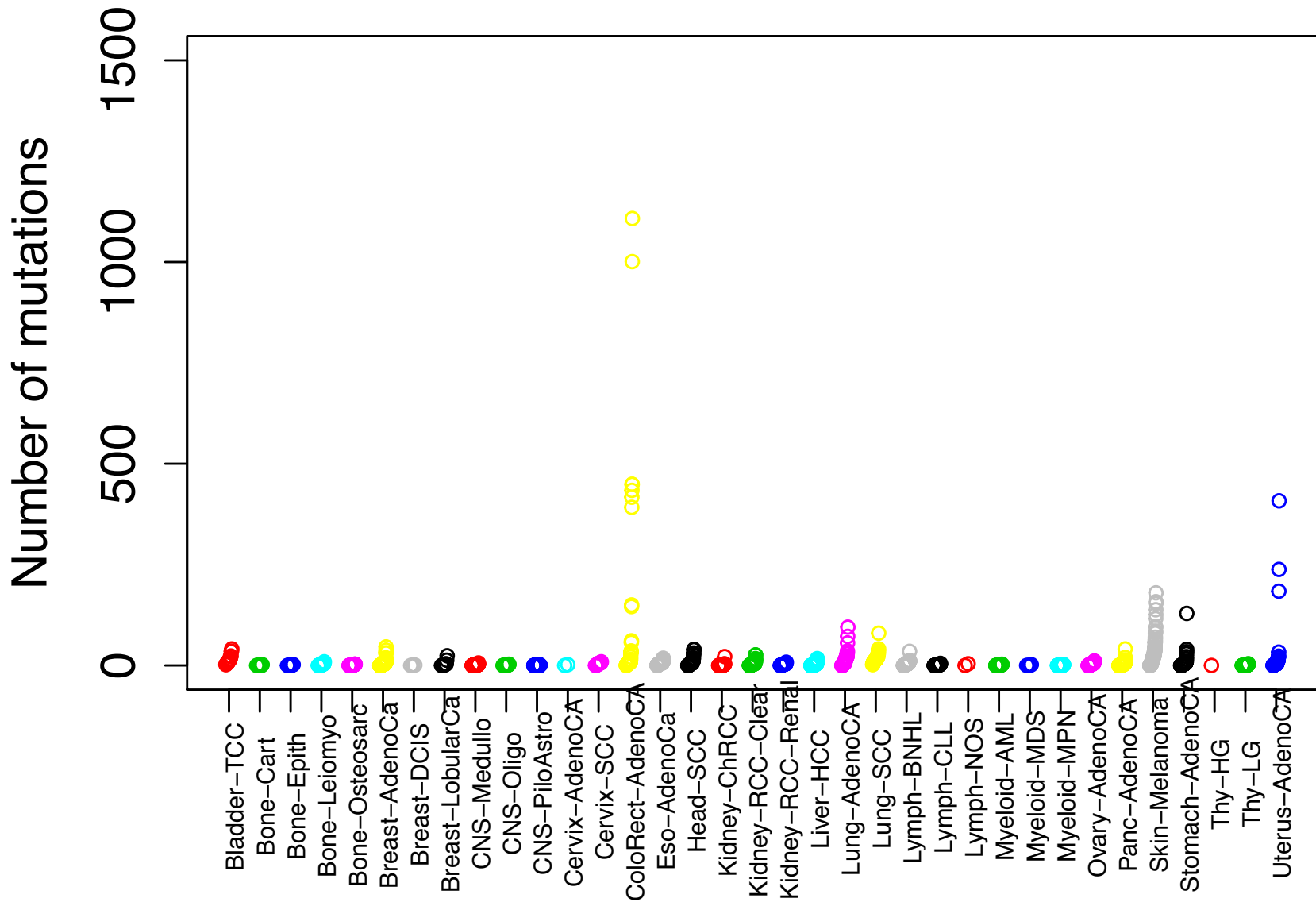


# Average Number of synonymous mutations

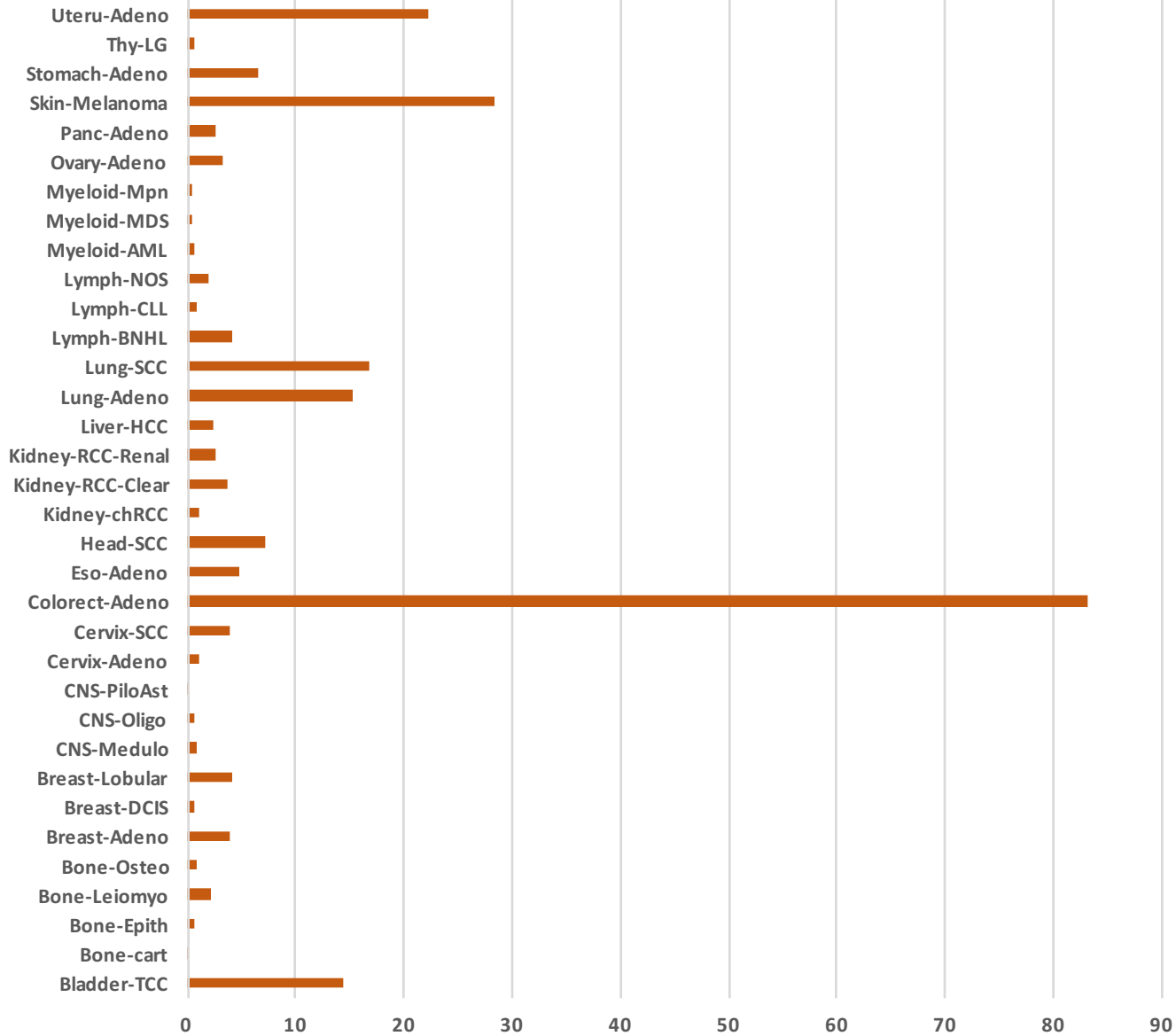




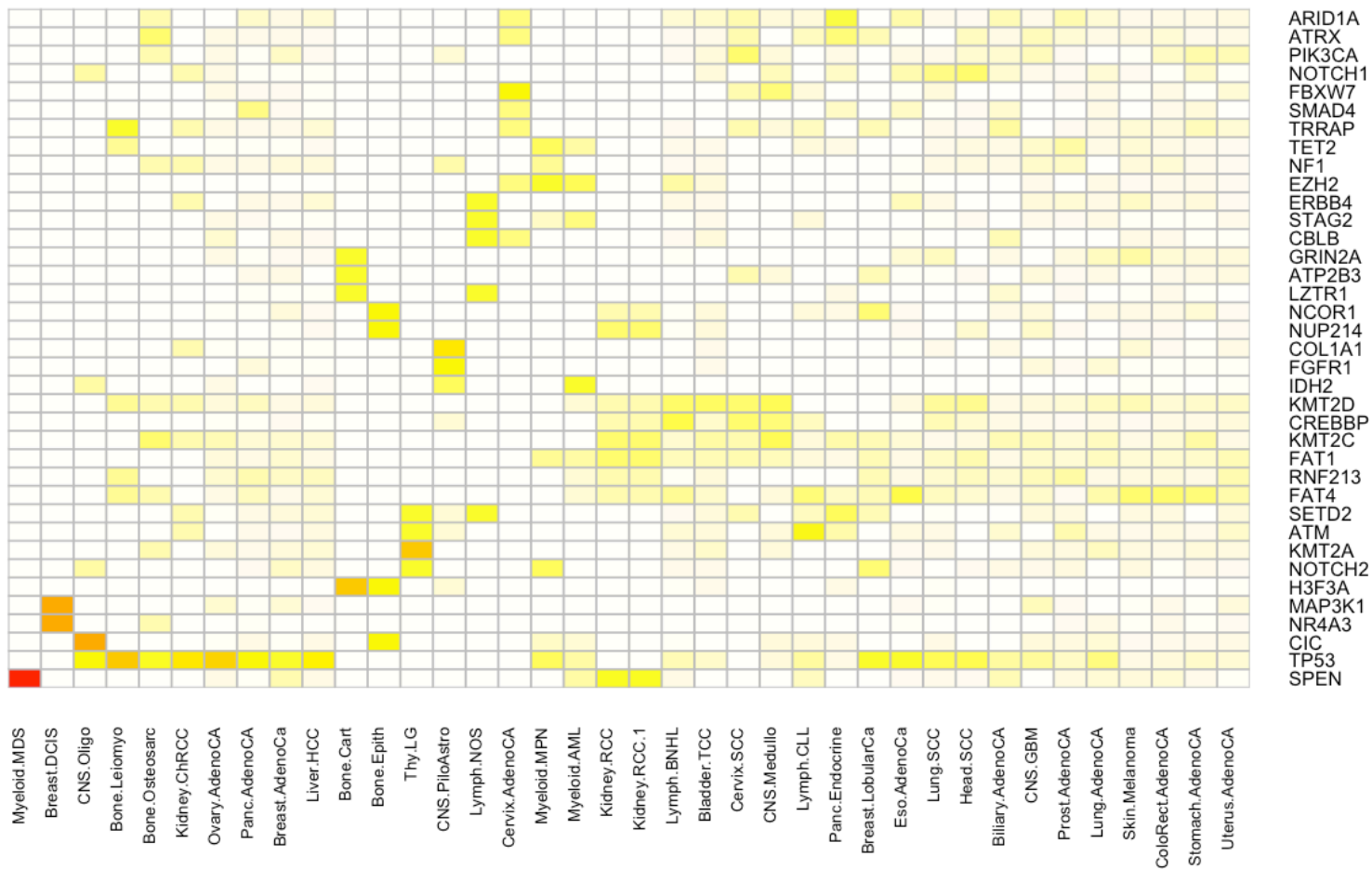
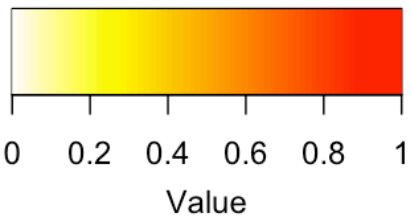
# Total Number of Pre-mature Stop



# Average Number of premature-stop mutations

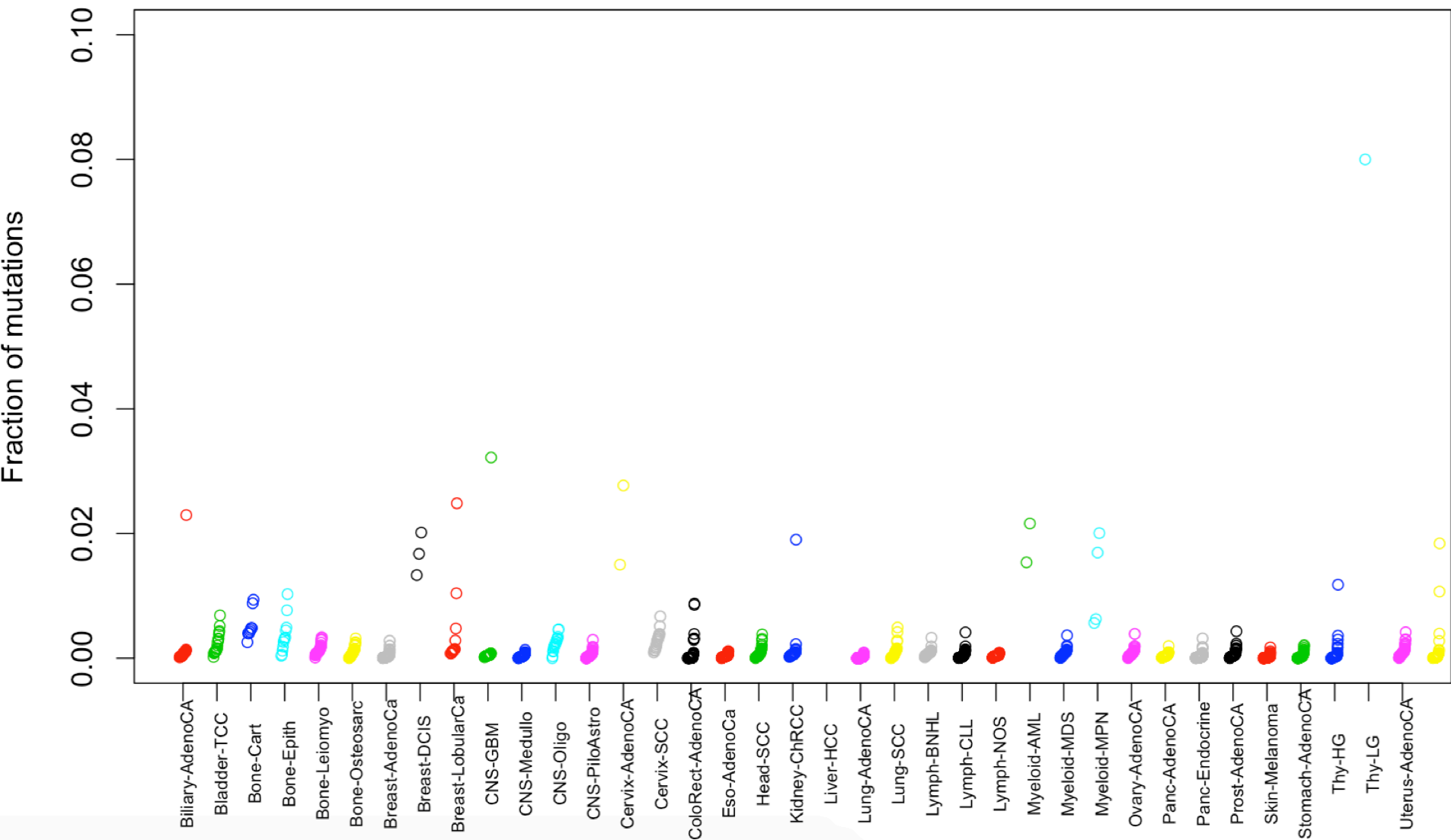


### Color Key

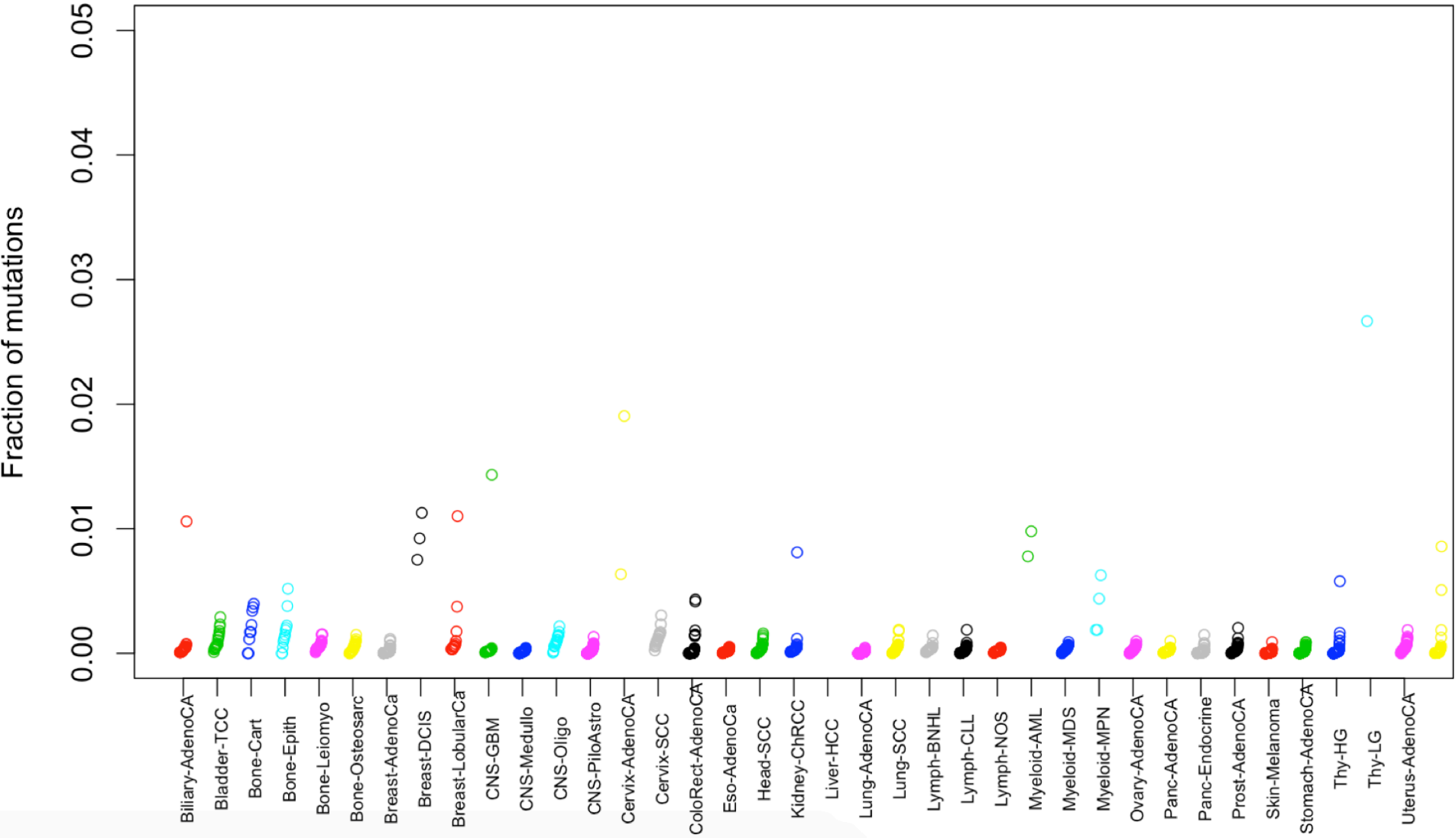




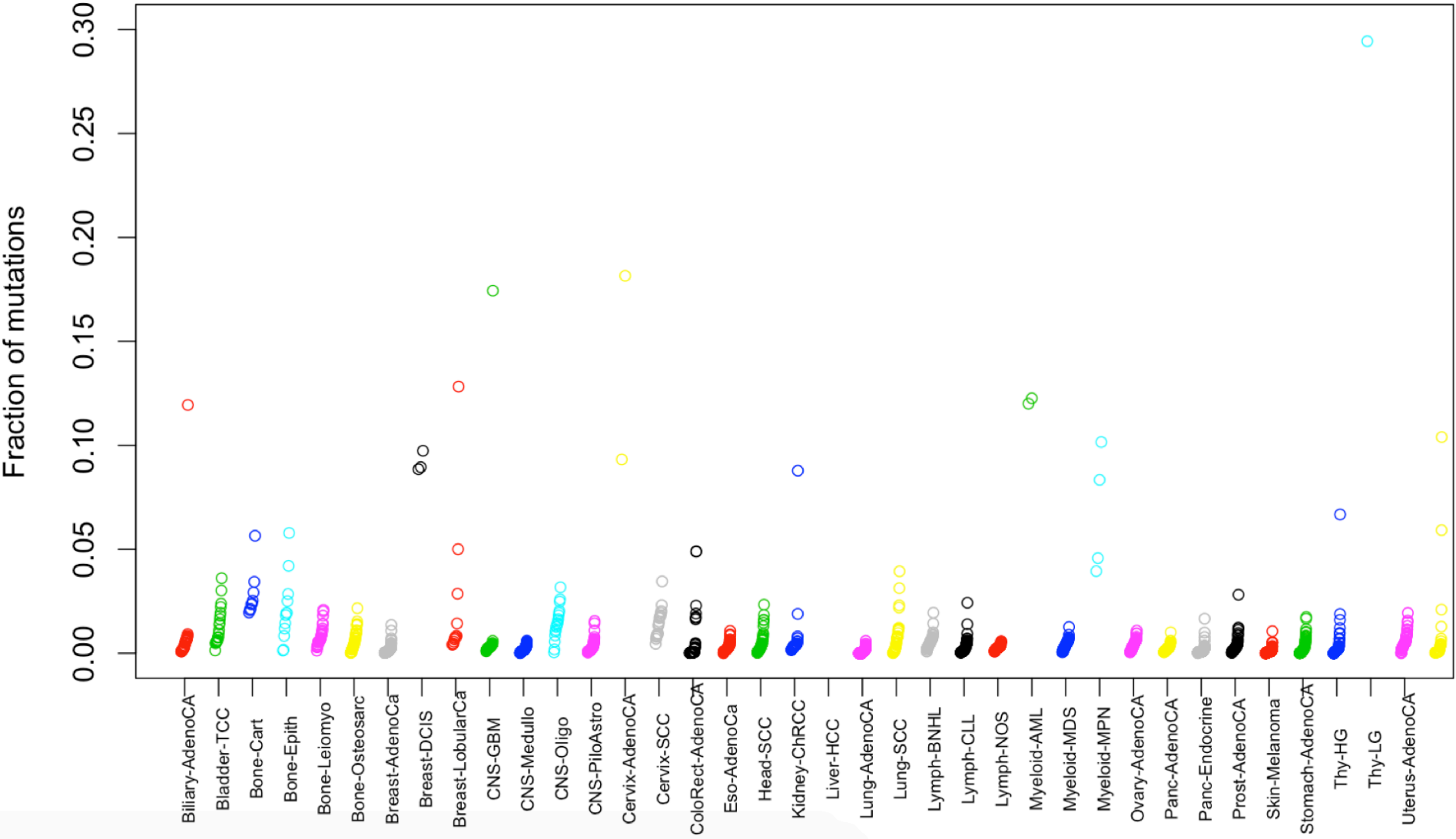
# Fraction of mutations in promoter regions



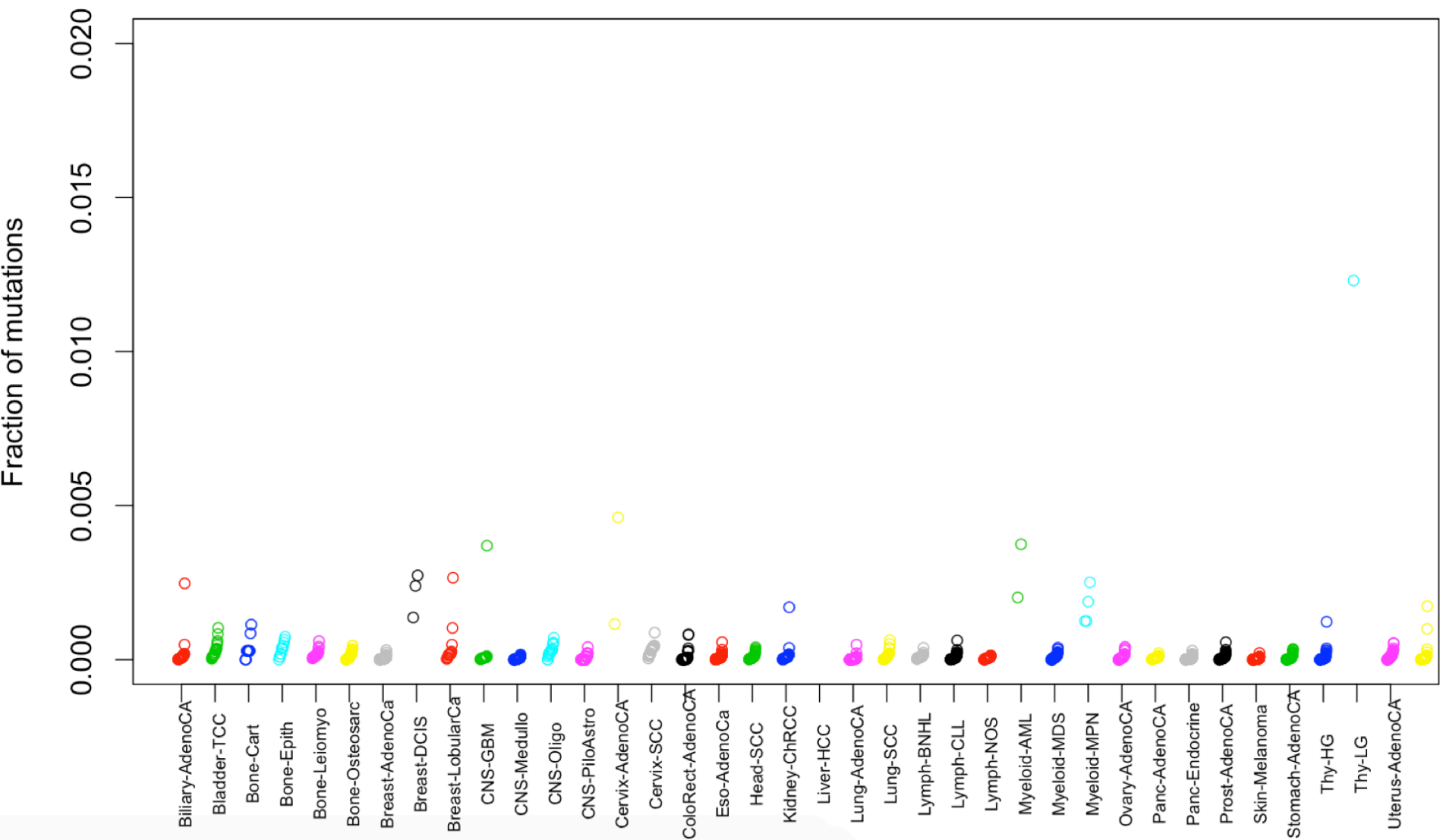
# Fraction of mutations in UTR regions



# Fraction of mutations in DHS regions

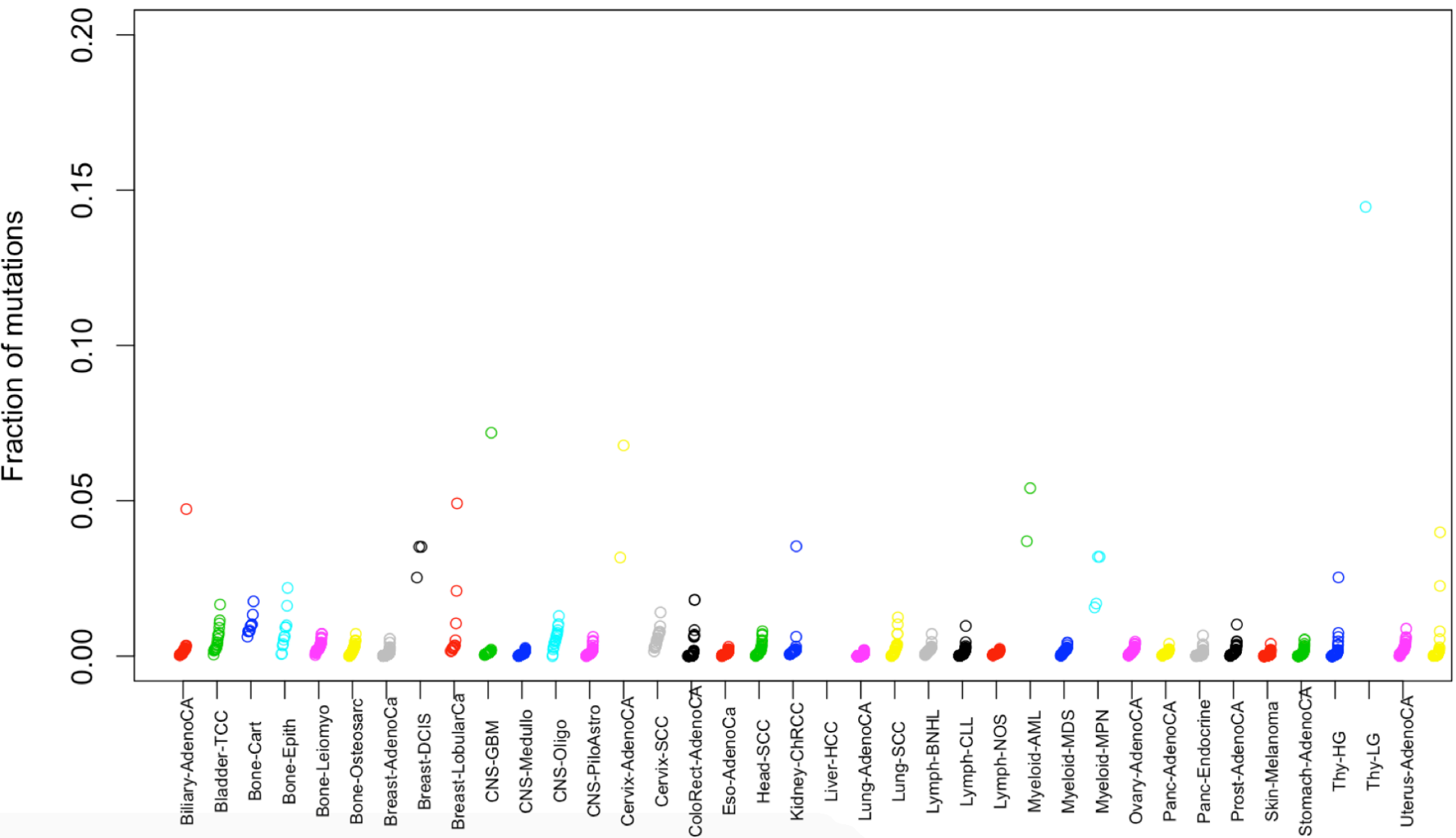


# Fraction of mutations in TFM regions

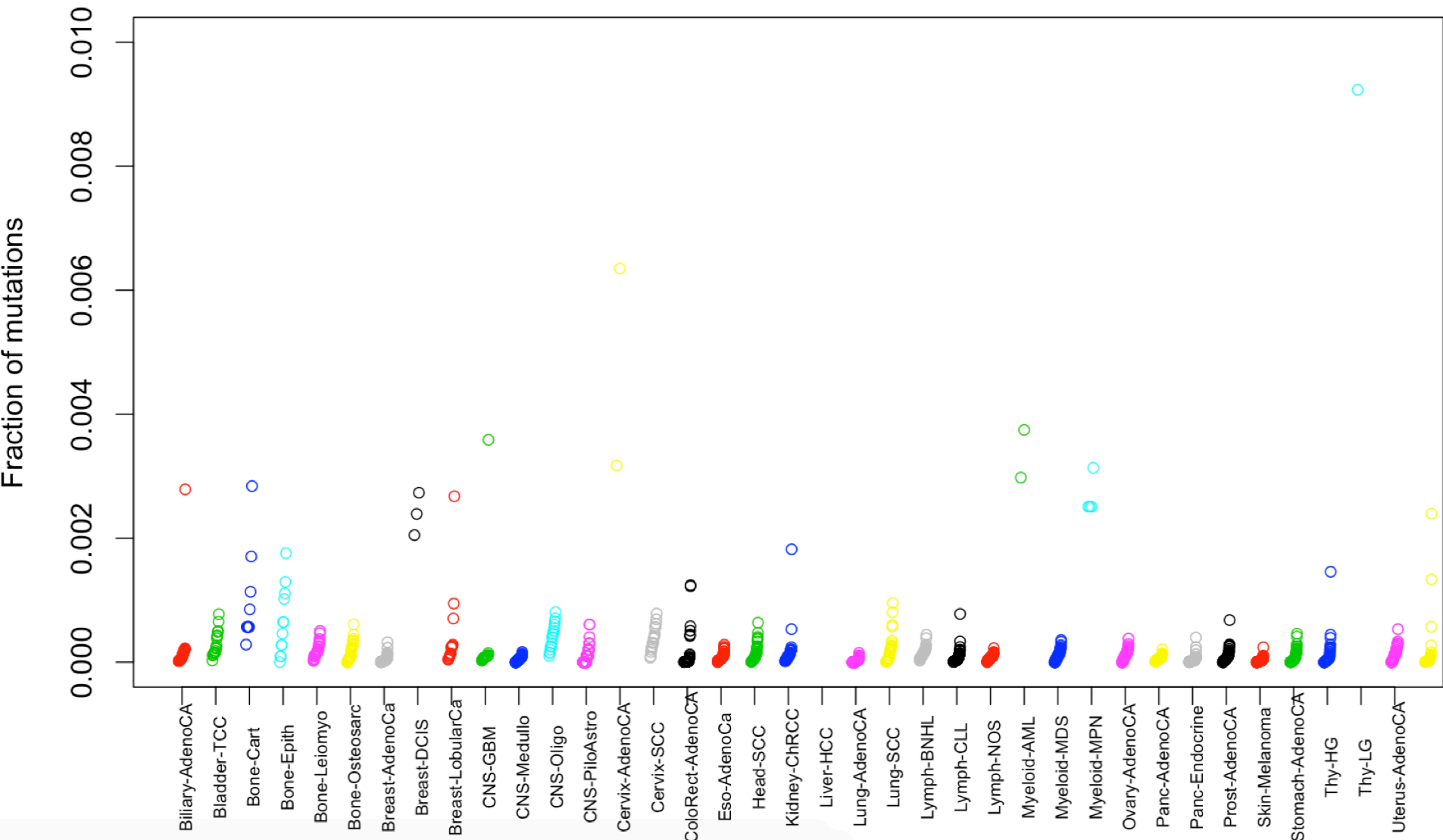




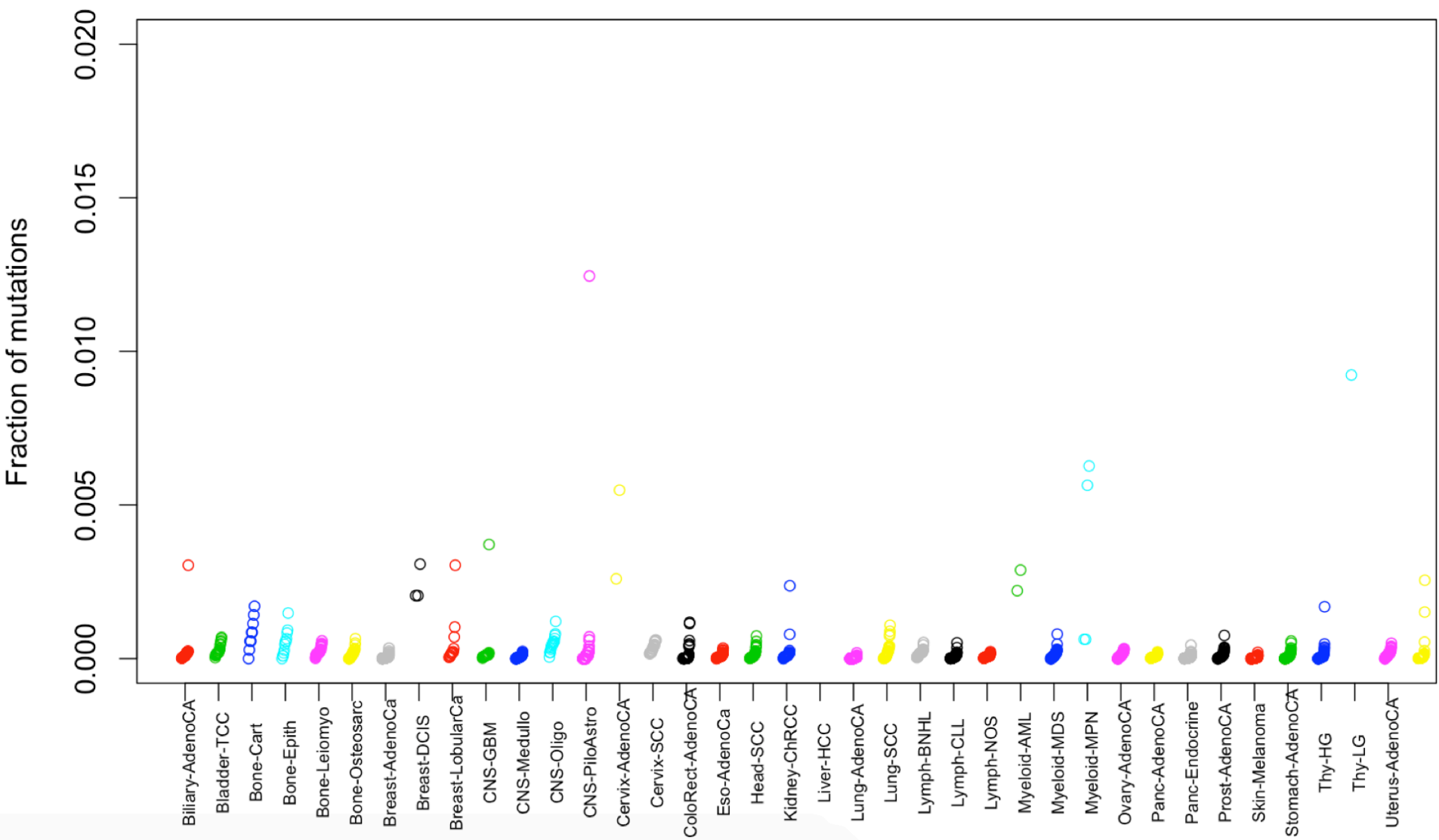
# Fraction of mutations in enhancer regions



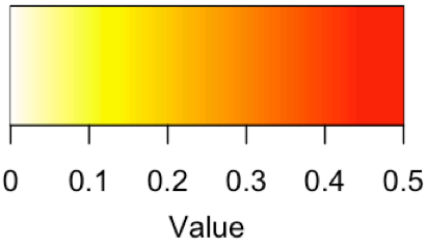
# Fraction of mutations in ncRNA regions



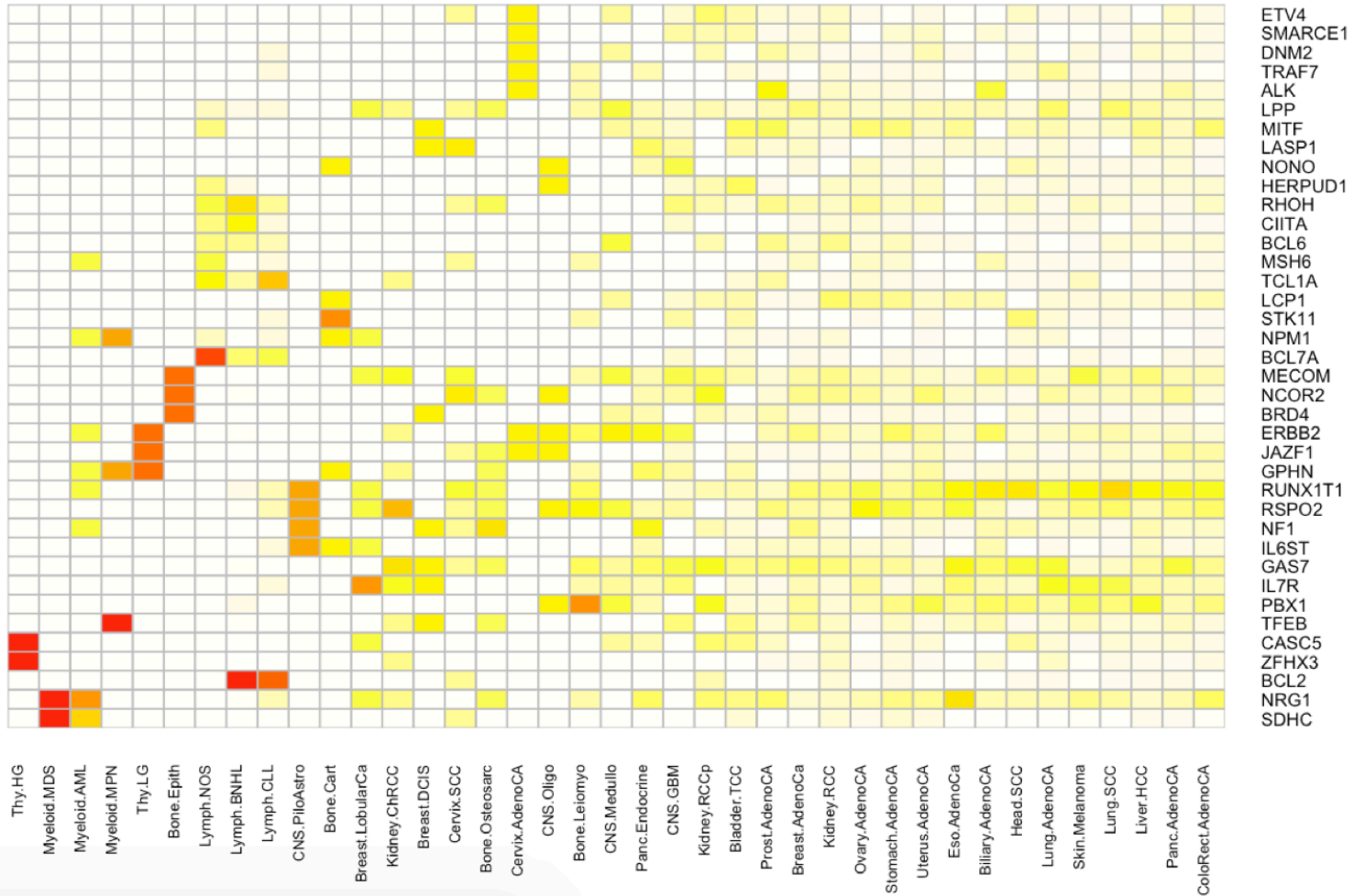
# Fraction of mutations in pseudogene regions



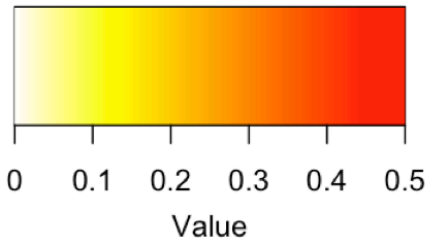
## Color Key



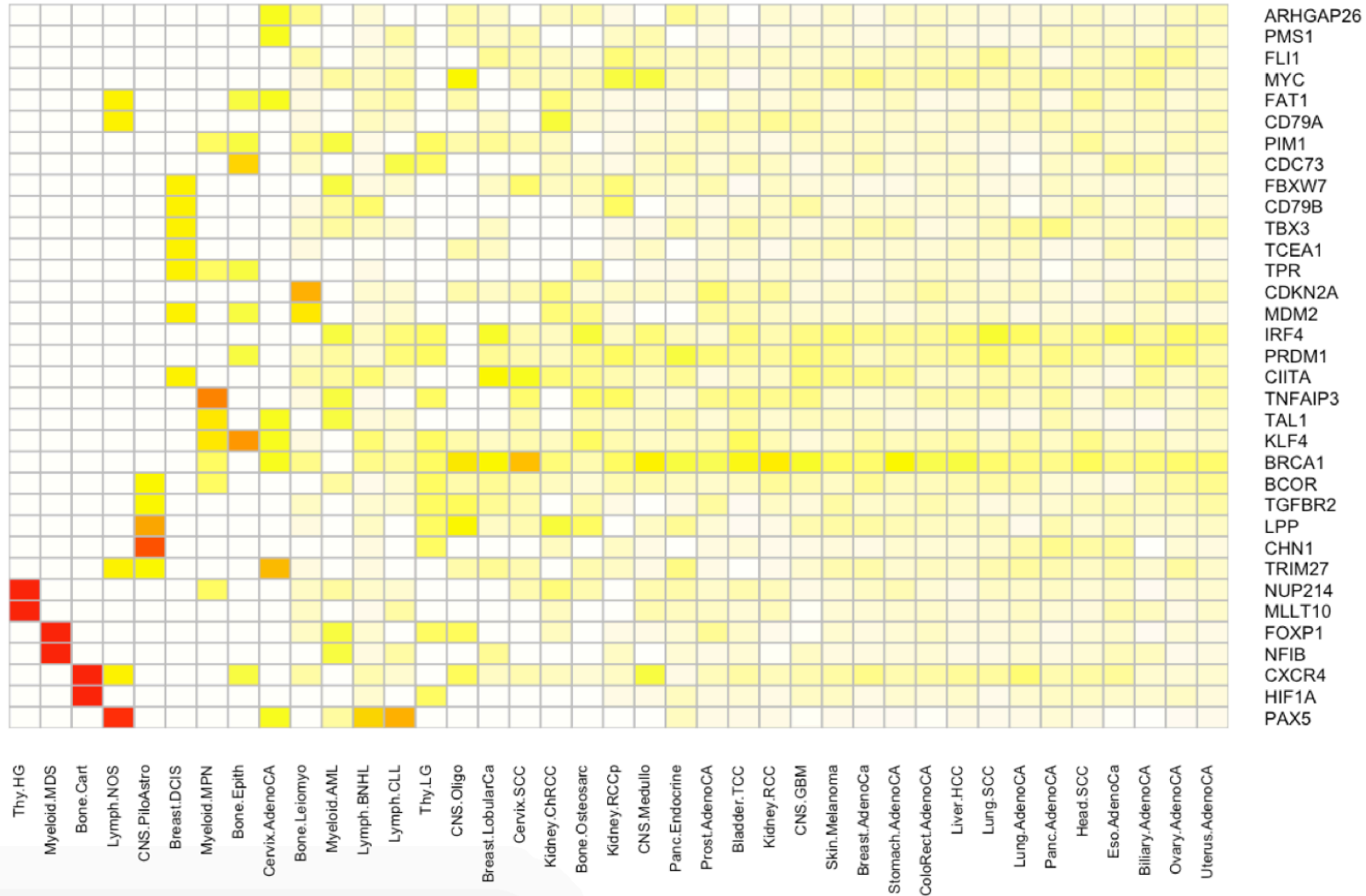
# Promoter regions of CGC genes



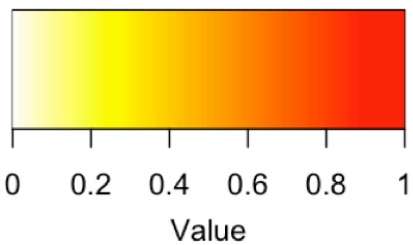
## Color Key



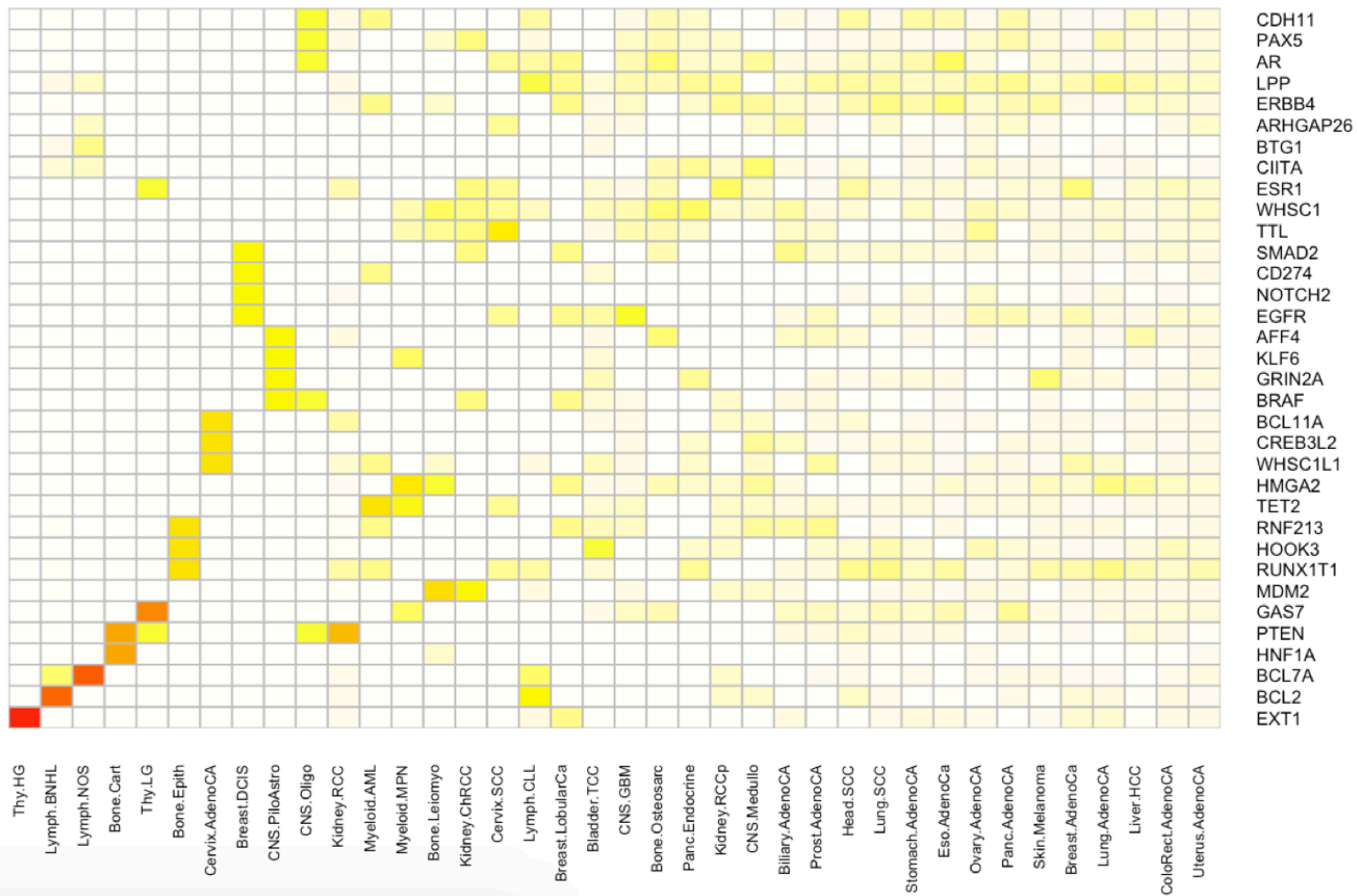
# Enhancer regions of CGC genes



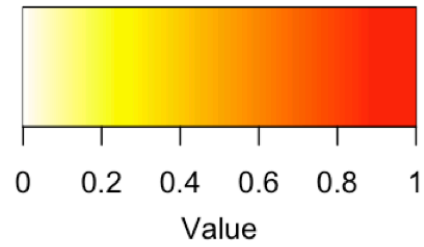
### Color Key



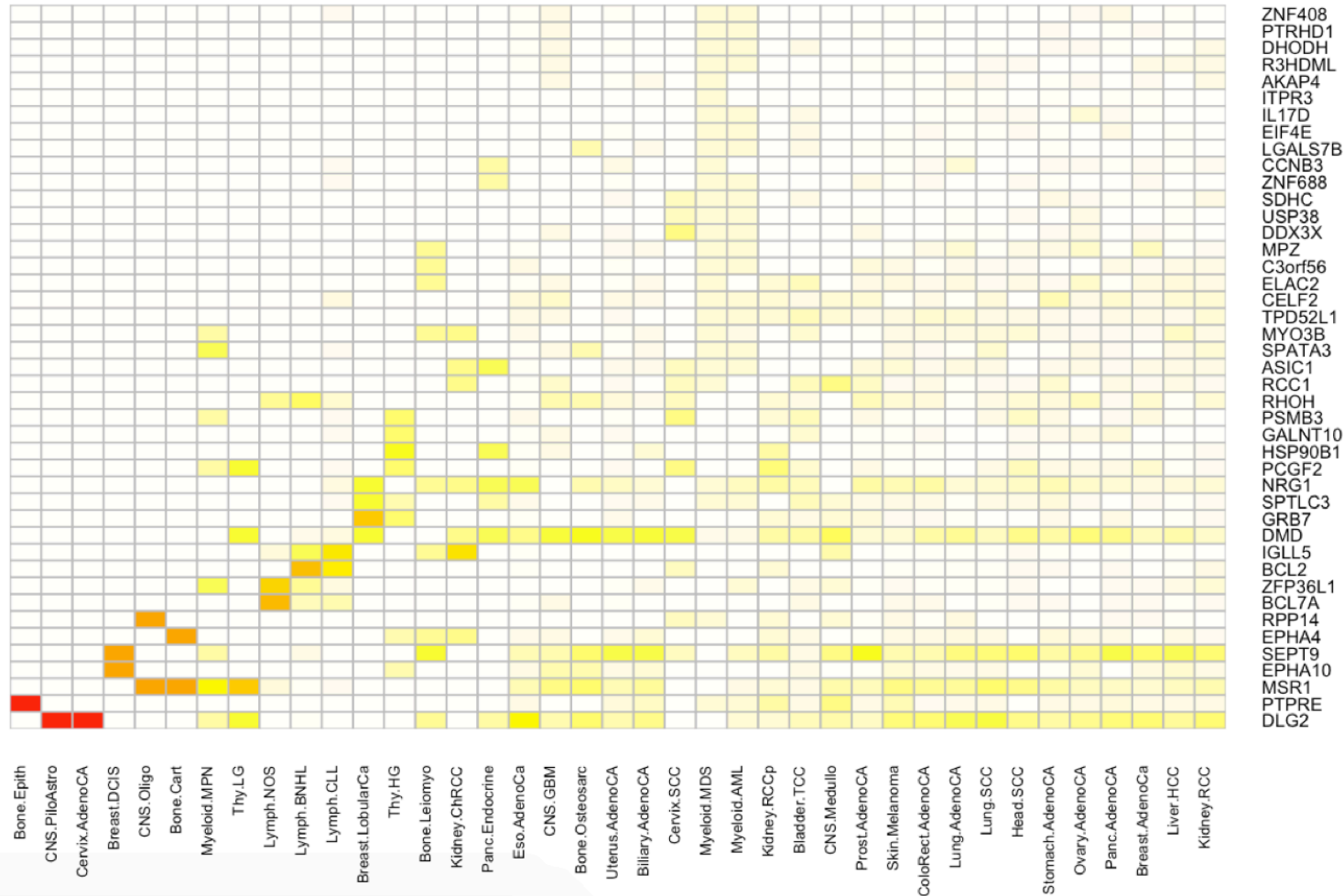
# UTR regions of CGC genes



### Color Key



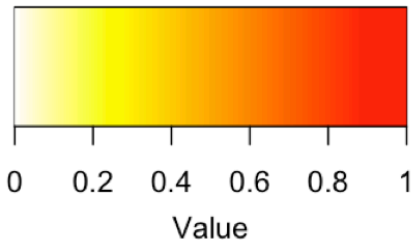
# Promoter regions of all genes







## Color Key



# UTR regions of all genes

- UTRN: Utrophin
  - may serve as a functional substitute for the dystrophin gene
  - Increased expression of utrophin in plasma of ovarian and breast cancer patients

