```r
# --- 1. LOAD LIBRARIES ---

library(maftools)

library(data.table)



# --- 2. DATA PREPARATION ---

raw_data <- fread("combined_calls.tsv")

colnames(raw_data) <- c("Chromosome", "Start_Position", "Reference_Allele", "Tumor_Se
q_Allele2", "Qual", "Filter")

raw_data[, End_Position := Start_Position + nchar(Reference_Allele) - 1]



# --- 3. UPDATED LABELS ---

raw_data$Tumor_Sample_Barcode = "combined_minion_promethion"

raw_data$Variant_Classification = ifelse(raw_data$Filter == "PASS", "Missense_Mutatio
n", "Silent")

raw_data$Variant_Type = ifelse(nchar(raw_data$Reference_Allele) == 1 & nchar(raw_data
$Tumor_Seq_Allele2) == 1, "SNP",

                            ifelse(nchar(raw_data$Reference_Allele) < nchar(raw_data$Tumo
r_Seq_Allele2), "INS", "DEL"))

raw_data$Hugo_Symbol = raw_data$Chromosome



# --- 4. CREATE MAF OBJECT ---

write.table(raw_data, "combined_platforms.maf", sep="\t", quote=FALSE, row.names=FALS
E)

my_maf = read.maf(maf = "combined_platforms.maf", isTCGA = FALSE, vc_nonSyn = "Missen
se_Mutation")
```
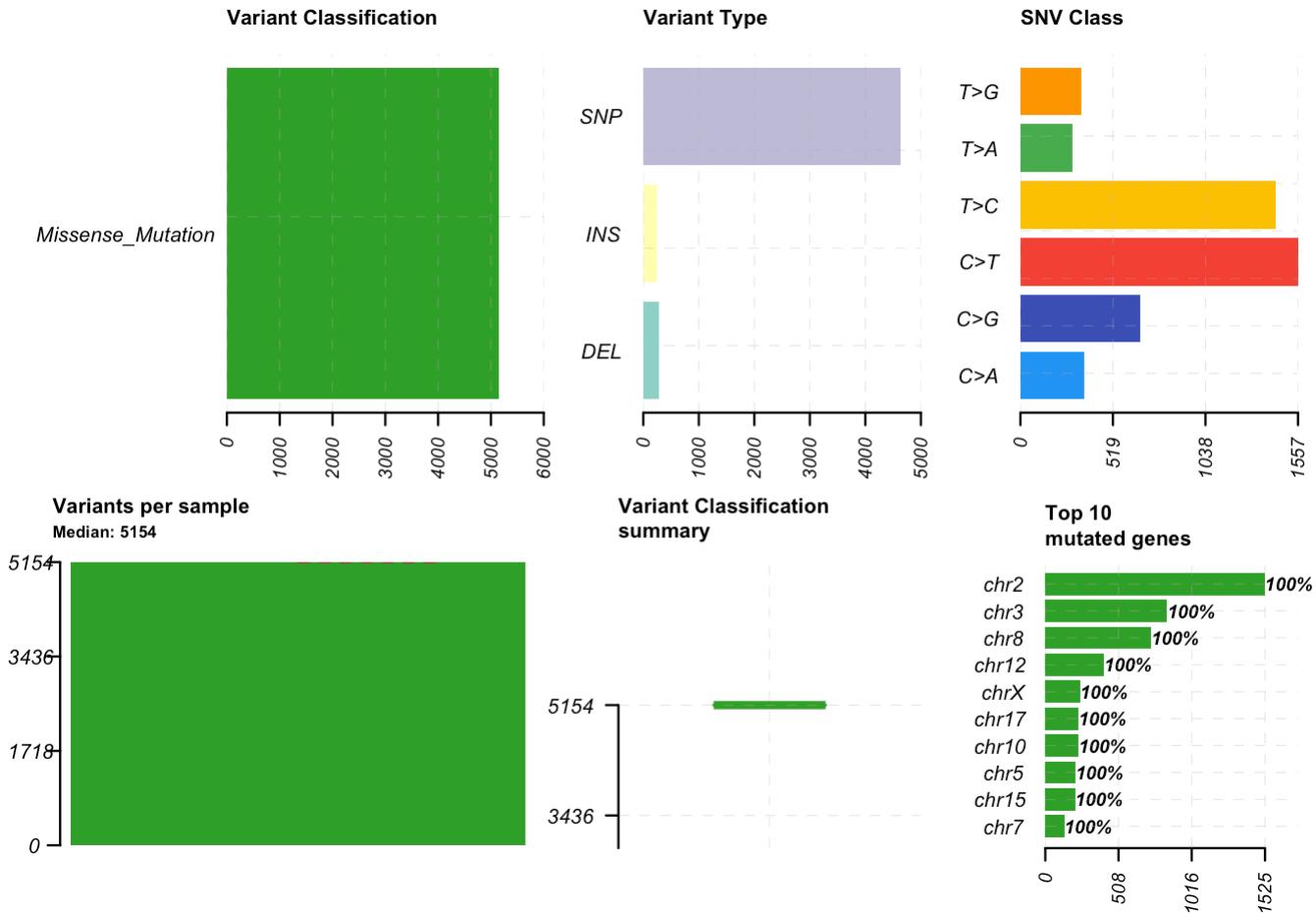
```
## -Reading
## -Validating
## -Silent variants: 638
## -Summarizing
## -Processing clinical data
## --Missing clinical data
## -Finished in 0.045s elapsed (0.038s cpu)
```

```
# --- 5. STABLE VISUALIZATIONS ---



# Plot 1: Summary Dashboard

plotmafSummary(maf = my_maf, addStat = 'median', rmOutlier = TRUE, dashboard = TRUE)
```
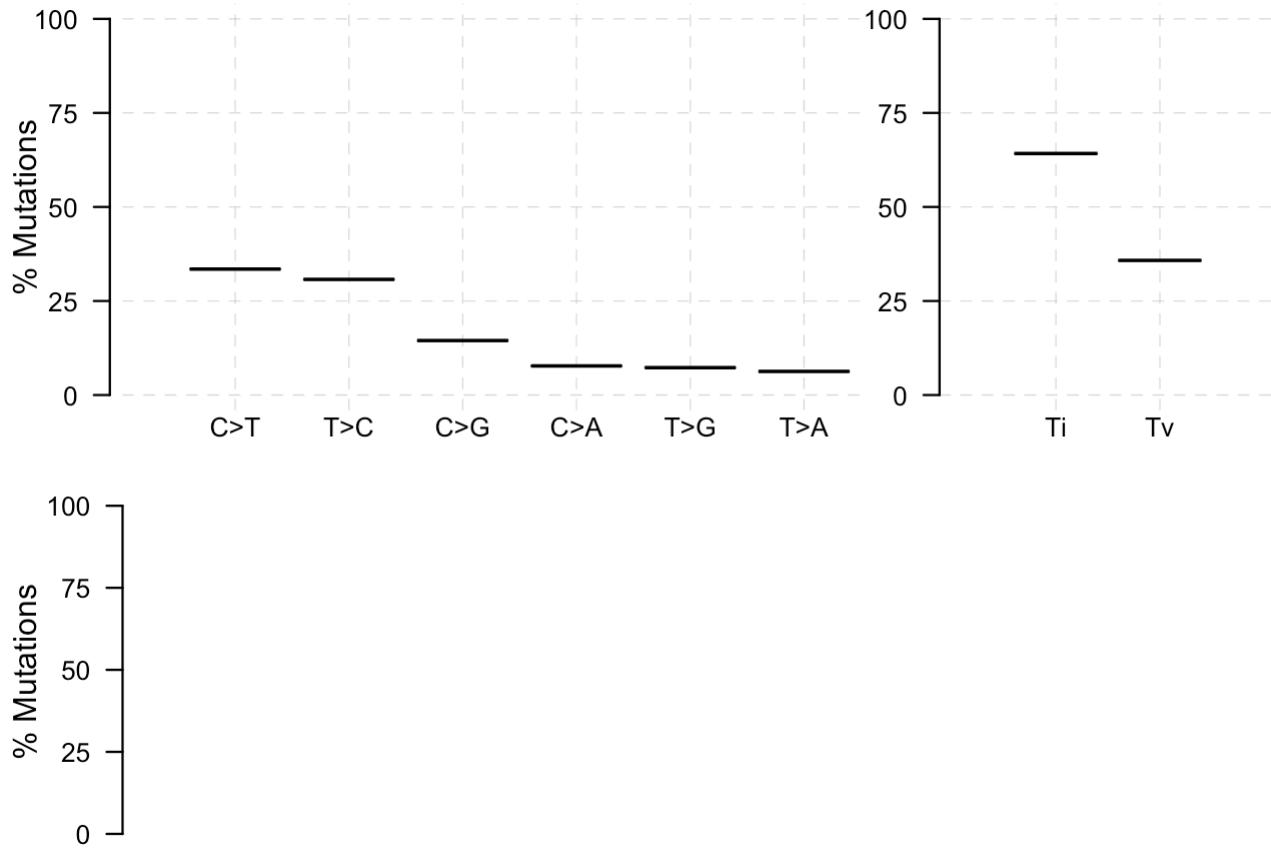


```
# Plot 2: Ti/Tv Analysis

titv_res = titv(maf = my_maf, plot = FALSE, useSyn = TRUE)

plotTiTv(res = titv_res)
```
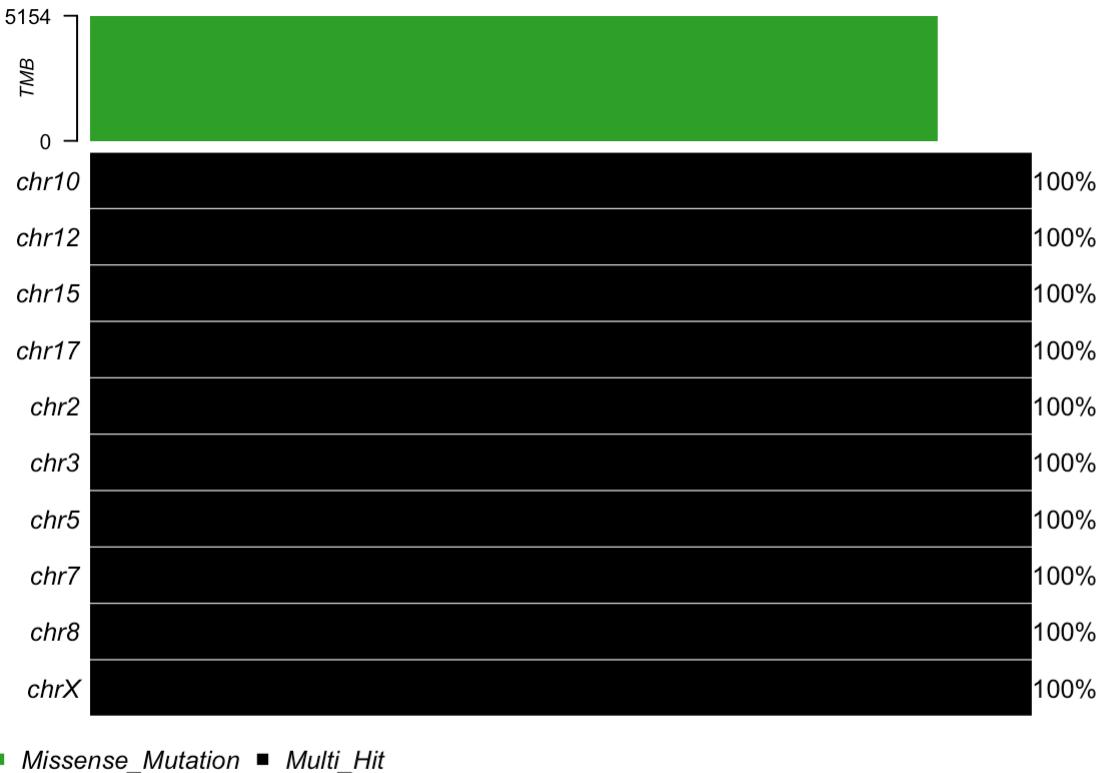
```
# Plot 3: Oncoplot (Top 10 Mutated Chromosomes)

oncoplot(maf = my_maf, top = 10, drawRowBar = FALSE, fontSize = 0.8)
```

```
## Warning in min(x): no non-missing arguments to min; returning Inf
```

```
## Warning in max(x): no non-missing arguments to max; returning -Inf
```

Altered in 1 (100%) of 1 samples.

- Missense_Mutation
- Multi_Hit

```
# Plot 4: Rainfall Plot (This will work perfectly for your 55 Kataegis events)

# Use this to visualize the clusters on Chromosome 2

rainfallPlot(maf = my_maf, tsb = "combined_minion_promethion", detectChangePoints = T
RUE)
```

```
## Processing combined_minion_promethion..
```

```
## Kataegis detected at:
```

```
##      Chromosome Start_Position End_Position nMuts Avg_intermutation_dist   Size
##           <num>          <num>        <num> <int>                  <num>  <num>
##   1:          1       17019410     17054002    69                508.7059  34592
##   2:          2       29192926     29920131  1288                565.0389 727205
##   3:          2      113216055    113277617   111                559.6545  61562
##   4:          3       12287696     12434058   209                703.6635 146362
##   5:          3       24123533     24128235     7                783.6667   4702
##   6:          3       24130634     24493919   504                722.2366 363285
##   7:          3       87259403     87275374    19                887.2778  15971
##   8:          3      179186419    179192974     9                819.3750   6555
##   9:          3      179217179    179222517     7                889.6667   5338
##  10:          5      112709435    112716131     8                956.5714   6696
##  11:          5      112718389    112845879   165                777.3780 127490
##  12:          5      177992355    177995875     9                440.0000   3520
##  13:          6      150370221    150405763   101                355.4200  35542
##  14:          7      107661438    107664670     6                646.4000   3232
##  15:          7      107677040    107716314    41                981.8500  39274
##  16:          8       31044583     31055016    12                948.4545  10433
##  17:          8       31069763     31080274    13                875.9167  10511
##  18:          8       31102488     31174361    87                835.7326  71873
##  19:          8      132867967    133131965   543                487.0812 263998
##  20:          9       21976672     21981087     6                883.0000   4415
##  21:          9       21983915     21995144    16                748.6000  11229
##  22:          9       97853378     97856378     8                428.5714   3000
##  23:         10       43077059     43130238   137                391.0221  53179
##  24:         10       87910574     87917350     8                968.0000   6776
##  25:         10       87939302     87949237    11                993.5000   9935
##  26:         10       87957058     87964003    10                771.6667   6945
##  27:         11       64804546     64809675     9                641.1250   5129
##  28:         12       25205716     25250690    78                584.0779  44974
##  29:         12       49020848     49042941    24                960.5652  22093
##  30:         12       49045593     49051609     8                859.4286   6016
##  31:         12       63950137     63954978     7                806.8333   4841
##  32:         12       63982306     63986666     7                726.6667   4360
##  33:         12       63992237     64018932    28                988.7037  26695
##  34:         12       64022548     64035057    24                543.8696  12509
##  35:         12       64058253     64160255   116                886.9739 102002
##  36:         14      104771710    104785733    16                934.8667  14023
##  37:         15       42944414     42948618     8                600.5714   4204
##  38:         15       42967649     42981053    18                788.4706  13404
##  39:         15       42988662     42994605     7                990.5000   5943
##  40:         15       43002020     43052688    54                956.0000  50668
##  41:         15       43074564     43105910    50                639.7143  31346
##  42:         15       45101541     45118049    32                532.5161  16508
##  43:         17        7662815      7686387    31                785.7333  23572
##  44:         17       68422443     68428041     7                933.0000   5598
##  45:         17       68431378     68451163    22                942.1429  19785
##  46:         17       68461865     68493367    36                900.0571  31502
##  47:         17       68498691     68549311    66                778.7692  50620
##  48:         22       28690105     28739901    67                754.4848  49796
##  49:         23       44922028     44930899    12                806.4545   8871
##  50:         23       44963483     44974171    18                628.7059  10688
##  51:         23       74443903     74449700     7                966.1667   5797
##  52:         23       74456451     74463710     9                907.3750   7259
##  53:         23       74490685     74531188    51                810.0600  40503
```

```
## 54:          23       131277902      131282984       7                    847.0000    5082
## 55:          23       131285672      131289385       8                    530.4286    3713
##      Chromosome Start_Position End_Position nMuts Avg_intermutation_dist    Size
##         <num>          <num>         <num> <int>                   <num>  <num>
##          Tumor_Sample_Barcode  C>A  C>G  C>T  T>A  T>C  T>G
##                         <fctr> <int> <int> <int> <int> <int> <int>
##  1: combined_minion_promethion    4    5   29    4   21    6
##  2: combined_minion_promethion  106  175  413   77  418   99
##  3: combined_minion_promethion   10   11   49    2   31    8
##  4: combined_minion_promethion   23   29   65   13   67   12
##  5: combined_minion_promethion    1    1    3    1   NA    1
##  6: combined_minion_promethion   43   74  171   43  140   33
##  7: combined_minion_promethion   NA    3    4    2    9    1
##  8: combined_minion_promethion   NA    2    5   NA    2   NA
##  9: combined_minion_promethion    1    3    2    1   NA   NA
## 10: combined_minion_promethion    1    1    2   NA    3    1
## 11: combined_minion_promethion   15   30   54   12   43   11
## 12: combined_minion_promethion   NA    2    4   NA    3   NA
## 13: combined_minion_promethion   12    8   29    5   38    9
## 14: combined_minion_promethion   NA   NA    4    1    1   NA
## 15: combined_minion_promethion    2    9    9   10    4    7
## 16: combined_minion_promethion    1    1    5    1    4   NA
## 17: combined_minion_promethion    2    2    3    1    3    2
## 18: combined_minion_promethion    7    6   29    5   30   10
## 19: combined_minion_promethion   40   68  214   33  155   33
## 20: combined_minion_promethion   NA    1    3   NA    2   NA
## 21: combined_minion_promethion    2    3    4    1    5    1
## 22: combined_minion_promethion   NA    2    2   NA    3    1
## 23: combined_minion_promethion   10   17   37    8   54   11
## 24: combined_minion_promethion   NA    1    3    1    3   NA
## 25: combined_minion_promethion    1    4    3   NA    2    1
## 26: combined_minion_promethion    2    1    3    2    1    1
## 27: combined_minion_promethion    1   NA    5   NA    3   NA
## 28: combined_minion_promethion    4    9   30    7   17   11
## 29: combined_minion_promethion    3    4    9    1    6    1
## 30: combined_minion_promethion    1   NA    3   NA    4   NA
## 31: combined_minion_promethion   NA    1    1    1    3    1
## 32: combined_minion_promethion    1    1    3   NA    1    1
## 33: combined_minion_promethion   NA    9   11    4    2    2
## 34: combined_minion_promethion   NA    1   12   NA    9    2
## 35: combined_minion_promethion    8   18   31    6   42   11
## 36: combined_minion_promethion    1    4    5   NA    5    1
## 37: combined_minion_promethion    1    1    1   NA    4    1
## 38: combined_minion_promethion    3    1    7   NA    7   NA
## 39: combined_minion_promethion   NA    1    4   NA    2   NA
## 40: combined_minion_promethion    5    7   18    1   17    6
## 41: combined_minion_promethion    2   14   18    2   13    1
## 42: combined_minion_promethion    3    6    9    2    5    7
## 43: combined_minion_promethion    1    5   13   NA   11    1
## 44: combined_minion_promethion    1    1    1   NA    4   NA
## 45: combined_minion_promethion    1    5    3    3   10   NA
## 46: combined_minion_promethion    2    4   15    1   10    4
## 47: combined_minion_promethion    9    9   26    3   15    4
## 48: combined_minion_promethion    2   16   17    5   22    5
## 49: combined_minion_promethion   NA    2    7   NA    2    1
## 50: combined_minion_promethion   NA    4    4    1    4    5
```

```
## 51: combined_minion_promethion     NA     NA      1      1      5     NA
## 52: combined_minion_promethion      1      1      2     NA      3      2
## 53: combined_minion_promethion      2      9     16      4     17      3
## 54: combined_minion_promethion     NA      2      4     NA      1     NA
## 55: combined_minion_promethion      2      1      1     NA      4     NA
##          Tumor_Sample_Barcode    C>A    C>G    C>T    T>A    T>C    T>G
##                               <fctr> <int> <int> <int> <int> <int> <int>
```

combined_minion_promethion

● C>T ● C>A ● T>C
● C>G ● T>A ● T>G

```
# --- 6. DATA EXPORT (Instead of Lollipop Plot) ---



# Get a summary of mutations per chromosome

chr_summary <- getGeneSummary(my_maf)

write.csv(chr_summary, "mutation_counts_per_chromosome.csv", row.names = FALSE)



# Extract the 55 Kataegis clusters to a file for your report

# This captures the Start/End positions of the clusters you saw in the console

kataegis_clusters <- rainfallPlot(maf = my_maf, tsb = "combined_minion_promethion")
```
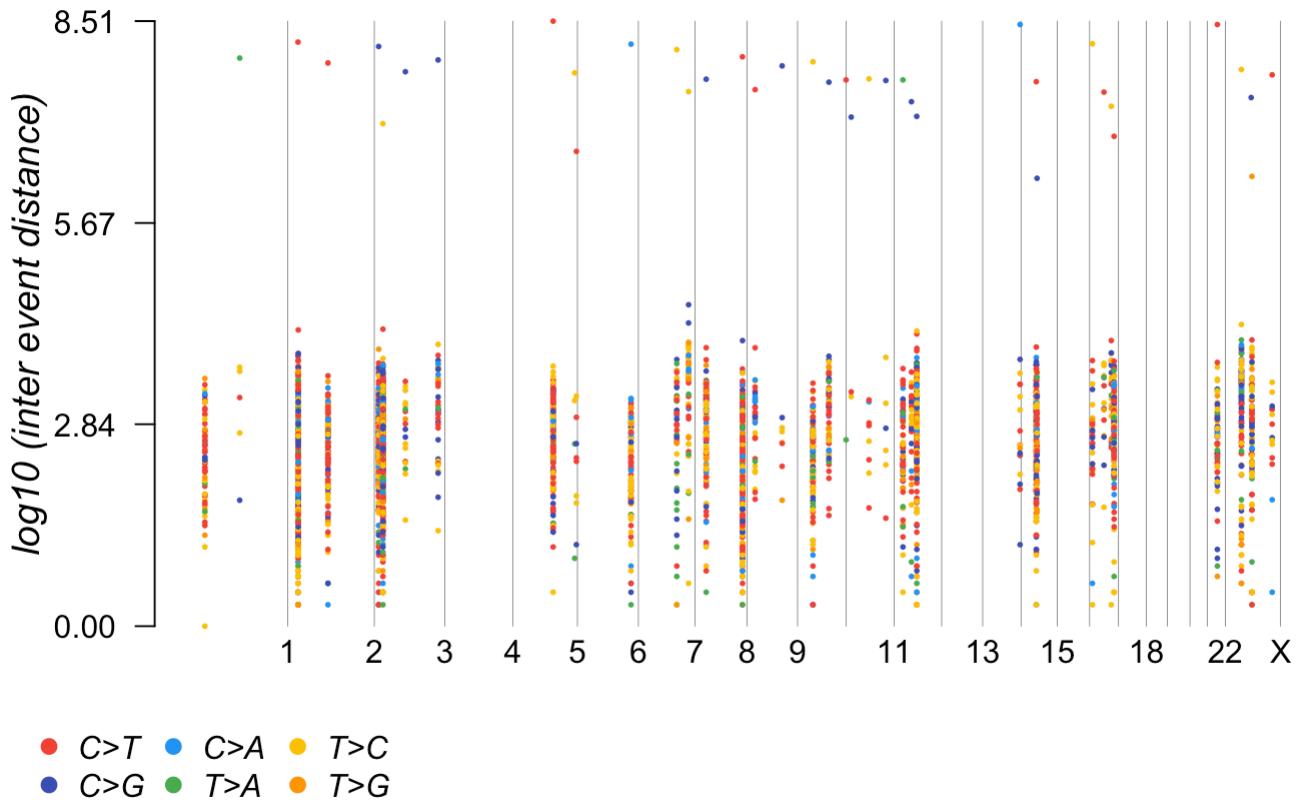
```
## Processing combined_minion_promethion..
```

combined_minion_promethion

Legend:
- 🔴 C>T 🔵 C>A 🟡 T>C
- 🔵 C>G 🟢 T>A 🟠 T>G

```r
if(!is.null(kataegis_clusters)){

  write.csv(kataegis_clusters, "kataegis_hotspots.csv", row.names = FALSE)

}
```

```r
# 1. Load the full panel (Including the missing ones you provided)
panel <- fread("
Chr,Start,End,Gene
chr1,114704464,114716894,NRAS
chr1,17018664,17054151,SDHB
chr2,29192774,29921586,ALK
chr2,113215997,113278921,PAX8
chr3,12287368,12434344,PPARG
chr3,24117153,24495708,THRB
chr3,87259404,87276584,POU1F1
chr3,179148126,179240093,PIK3CA
chr5,112707498,112846239,APC
chr5,173232109,173235206,NKX2-5
chr5,177992235,177996242,PROP1
chr6,150369012,150405969,IYD
chr7,107660828,107717809,SLC26A4
chr7,140713327,140924928,BRAF
chr8,31033810,31176138,WRN
chr8,132866958,133134899,TG
chr9,21967752,21995324,CDKN2A
chr9,22002903,22009313,CDKN2B
chr9,97853226,97856717,FOXE1
chr10,43076840,43130353,RET
chr10,87863625,87971930,PTEN
chr11,532242,537287,HRAS
chr11,14966668,14972351,CALCA
chr11,64803516,64811294,MEN1
chr11,112086873,112095794,SDHD
chr12,25204789,25250936,KRAS
chr12,49018978,49060794,KMT2D
chr12,63844700,64162217,SRGAP1
chr14,104769349,104795748,AKT1
chr15,42942897,43106038,UBR1
chr15,45092650,45114172,DUOX2
chr15,45114326,45118421,DUOXA2
chr17,7661779,7687550,TP53
chr17,40062193,40093867,THRA
chr17,60600193,60666280,PPM1D
chr17,68413623,68551316,PRKAR1A
chr22,28687742,28741834,CHEK2
chrX,44873188,45112779,KDM6A
chrX,72301646,72307157,CITED1
chrX,74421493,74533916,SLC16A2
chrX,131273506,131289457,IGSF1")

# 2. Perform the overlap join
setkey(raw_data, Chromosome, Start_Position, End_Position)
setkey(panel, Chr, Start, End)

annotated_data <- foverlaps(raw_data, panel,
                            by.x=c("Chromosome", "Start_Position", "End_Position"),
                            by.y=c("Chr", "Start", "End"))

# 3. Handle 'Others' and clean up columns
# If it didn't hit a gene in your panel, we'll call it "Intergenic_Region"
```

```
annotated_data[, Hugo_Symbol := ifelse(is.na(Gene), "Intergenic_Region", Gene)]

# Final cleanup to match MAF format
final_maf_df <- annotated_data[, .(Hugo_Symbol, Chromosome, Start_Position, End_Posit
ion,
                                   Reference_Allele, Tumor_Seq_Allele2, Variant_Class
ification,
                                   Variant_Type, Tumor_Sample_Barcode)]

# 4. Re-load into maftools
my_maf_final = read.maf(maf = final_maf_df, isTCGA = FALSE)
```

```
## -Validating
## -Silent variants: 638
## -Summarizing
## -Processing clinical data
## --Missing clinical data
## -Finished in 0.036s elapsed (0.032s cpu)
```

```
# Filter mutations specifically in the ALK cluster
alk_cluster_muts <- final_maf_df[Hugo_Symbol == "ALK"]

# See the mutation types
table(alk_cluster_muts$Reference_Allele, alk_cluster_muts$Tumor_Seq_Allele2)
```

```
##
##                                      A AAAAC AAAACAAAC AAAC AAG AAGAC AAGGC AC ACATG
## A                                    0     1         1    1   1     1     1  1     1
## Aaatg                                1     0         0    0   0     0     0  0     0
## AACC                                 1     0         0    0   0     0     0  0     0
## ACTC                                 1     0         0    0   0     0     0  0     0
## ACTT                                 1     0         0    0   0     0     0  0     0
## Agaggaggagaaggag                     1     0         0    0   0     0     0  0     0
## At                                   6     0         0    0   0     0     0  0     0
## AT                                   2     0         0    0   0     0     0  0     0
## Atg                                  1     0         0    0   0     0     0  0     0
## Atgtgatg                             1     0         0    0   0     0     0  0     0
## C                                   51     0         0    0   0     0     0  0     0
## Ca                                   0     0         0    0   0     0     0  0     0
## Caa                                  0     0         0    0   0     0     0  0     0
## Caaa                                 0     0         0    0   0     0     0  0     0
## Caaaa                                0     0         0    0   0     0     0  0     0
## Caaaaaa                              0     0         0    0   0     0     0  0     0
## Caaaaaaa                             0     0         0    0   0     0     0  0     0
## Cagagagagagagagagagagag             0     0         0    0   0     0     0  0     0
## CAT                                  0     0         0    0   0     0     0  0     0
## Cgt                                  0     0         0    0   0     0     0  0     0
## CGTAT                                0     0         0    0   0     0     0  0     0
## Ct                                   0     0         0    0   0     0     0  0     0
## CT                                   0     0         0    0   0     0     0  0     0
## Ctt                                  0     0         0    0   0     0     0  0     0
## Cttt                                 0     0         0    0   0     0     0  0     0
## Ctttt                                0     0         0    0   0     0     0  0     0
## G                                  188     0         0    0   0     0     0  0     0
## Ga                                   0     0         0    0   0     0     0  0     0
## GA                                   0     0         0    0   0     0     0  0     0
## Gaaagaaaagaaaagaaaagaaaaga           0     0         0    0   0     0     0  0     0
## GGAA                                 0     0         0    0   0     0     0  0     0
## Gggaa                                0     0         0    0   0     0     0  0     0
## Gggtgtgtgtgt                         0     0         0    0   0     0     0  0     0
## Ggtgt                                0     0         0    0   0     0     0  0     0
## Gt                                   0     0         0    0   0     0     0  0     0
## GT                                   0     0         0    0   0     0     0  0     0
## Gtc                                  0     0         0    0   0     0     0  0     0
## GTCTTGAAGGAAA                        0     0         0    0   0     0     0  0     0
## Gtgt                                 0     0         0    0   0     0     0  0     0
## Gtt                                  0     0         0    0   0     0     0  0     0
## GTT                                  0     0         0    0   0     0     0  0     0
## Gttt                                 0     0         0    0   0     0     0  0     0
## Gttttttt                             0     0         0    0   0     0     0  0     0
## T                                   40     0         0    0   0     0     0  0     0
## Ta                                   0     0         0    0   0     0     0  0     0
## TA                                   0     0         0    0   0     0     0  0     0
## Taa                                  0     0         0    0   0     0     0  0     0
## Taaa                                 0     0         0    0   0     0     0  0     0
## Taaaaaaaaaa                          0     0         0    0   0     0     0  0     0
## Taac                                 0     0         0    0   0     0     0  0     0
## Tac                                  0     0         0    0   0     0     0  0     0
## Tacacacacacacacac                    0     0         0    0   0     0     0  0     0
## Tagagag                             0     0         0    0   0     0     0  0     0
```

```
##   Tc                              0    0      0        0   0   0      0      0   0   0
##   TCC                             0    0      0        0   0   0      0      0   0   0
##   Tccaa                           0    0      0        0   0   0      0      0   0   0
##   Tg                              0    0      0        0   0   0      0      0   0   0
##   Ttaaaaaaaaaaaaa                 0    0      0        0   0   0      0      0   0   0
##   Ttctc                           0    0      0        0   0   0      0      0   0   0
##
```

##		AG	AGAGGT	AGCCCTG	AGT	AGTTCAAGGTCACTGAAT	AT
##	A	2	1	1	1	1	4
##	Aaatg	0	0	0	0	0	0
##	AACC	0	0	0	0	0	0
##	ACTC	0	0	0	0	0	0
##	ACTT	0	0	0	0	0	0
##	Agaggaggagaaggag	0	0	0	0	0	0
##	At	0	0	0	0	0	0
##	AT	0	0	0	0	0	0
##	Atg	0	0	0	0	0	0
##	Atgtgatg	0	0	0	0	0	0
##	C	0	0	0	0	0	0
##	Ca	0	0	0	0	0	0
##	Caa	0	0	0	0	0	0
##	Caaa	0	0	0	0	0	0
##	Caaaa	0	0	0	0	0	0
##	Caaaaaa	0	0	0	0	0	0
##	Caaaaaaa	0	0	0	0	0	0
##	Cagagagagagagagagagagag	0	0	0	0	0	0
##	CAT	0	0	0	0	0	0
##	Cgt	0	0	0	0	0	0
##	CGTAT	0	0	0	0	0	0
##	Ct	0	0	0	0	0	0
##	CT	0	0	0	0	0	0
##	Ctt	0	0	0	0	0	0
##	Cttt	0	0	0	0	0	0
##	Ctttt	0	0	0	0	0	0
##	G	0	0	0	0	0	0
##	Ga	0	0	0	0	0	0
##	GA	0	0	0	0	0	0
##	Gaaagaaaagaaaagaaaagaaaaga	0	0	0	0	0	0
##	GGAA	0	0	0	0	0	0
##	Gggaa	0	0	0	0	0	0
##	Gggtgtgtgtgt	0	0	0	0	0	0
##	Ggtgt	0	0	0	0	0	0
##	Gt	0	0	0	0	0	0
##	GT	0	0	0	0	0	0
##	Gtc	0	0	0	0	0	0
##	GTCTTGAAGGAAA	0	0	0	0	0	0
##	Gtgt	0	0	0	0	0	0
##	Gtt	0	0	0	0	0	0
##	GTT	0	0	0	0	0	0
##	Gttt	0	0	0	0	0	0
##	Gttttttt	0	0	0	0	0	0
##	T	0	0	0	0	0	0
##	Ta	0	0	0	0	0	0
##	TA	0	0	0	0	0	0
##	Taa	0	0	0	0	0	0
##	Taaa	0	0	0	0	0	0

##	Taaaaaaaaaa	0	0	0	0		0	0	
##	Taac	0	0	0	0		0	0	
##	Tac	0	0	0	0		0	0	
##	Tacacacacacacac	0	0	0	0		0	0	
##	Tagagag	0	0	0	0		0	0	
##	Tc	0	0	0	0		0	0	
##	TCC	0	0	0	0		0	0	
##	Tccaa	0	0	0	0		0	0	
##	Tg	0	0	0	0		0	0	
##	Ttaaaaaaaaaaaaa	0	0	0	0		0	0	
##	Ttctc	0	0	0	0		0	0	
##									

		ATAGCTAACACATATATGG	ATATATAT	ATATG	ATGTG	ATT	ATTAG
##	A	1	1	1	1	2	1
##	Aaatg	0	0	0	0	0	0
##	AACC	0	0	0	0	0	0
##	ACTC	0	0	0	0	0	0
##	ACTT	0	0	0	0	0	0
##	Agaggaggagaaggag	0	0	0	0	0	0
##	At	0	0	0	0	0	0
##	AT	0	0	0	0	0	0
##	Atg	0	0	0	0	0	0
##	Atgtgatg	0	0	0	0	0	0
##	C	0	0	0	0	0	0
##	Ca	0	0	0	0	0	0
##	Caa	0	0	0	0	0	0
##	Caaa	0	0	0	0	0	0
##	Caaaa	0	0	0	0	0	0
##	Caaaaaa	0	0	0	0	0	0
##	Caaaaaaa	0	0	0	0	0	0
##	Cagagagagagagagagagagag	0	0	0	0	0	0
##	CAT	0	0	0	0	0	0
##	Cgt	0	0	0	0	0	0
##	CGTAT	0	0	0	0	0	0
##	Ct	0	0	0	0	0	0
##	CT	0	0	0	0	0	0
##	Ctt	0	0	0	0	0	0
##	Cttt	0	0	0	0	0	0
##	Ctttt	0	0	0	0	0	0
##	G	0	0	0	0	0	0
##	Ga	0	0	0	0	0	0
##	GA	0	0	0	0	0	0
##	Gaaagaaaagaaaagaaaagaaaaga	0	0	0	0	0	0
##	GGAA	0	0	0	0	0	0
##	Gggaa	0	0	0	0	0	0
##	Gggtgtgtgtgt	0	0	0	0	0	0
##	Ggtgt	0	0	0	0	0	0
##	Gt	0	0	0	0	0	0
##	GT	0	0	0	0	0	0
##	Gtc	0	0	0	0	0	0
##	GTCTTGAAGGAAA	0	0	0	0	0	0
##	Gtgt	0	0	0	0	0	0
##	Gtt	0	0	0	0	0	0
##	GTT	0	0	0	0	0	0
##	Gttt	0	0	0	0	0	0
##	Gttttttt	0	0	0	0	0	0

```
##    T                                        0      0      0      0   0      0
##    Ta                                       0      0      0      0   0      0
##    TA                                       0      0      0      0   0      0
##    Taa                                      0      0      0      0   0      0
##    Taaa                                     0      0      0      0   0      0
##    Taaaaaaaaaa                              0      0      0      0   0      0
##    Taac                                     0      0      0      0   0      0
##    Tac                                      0      0      0      0   0      0
##    Tacacacacacacacac                        0      0      0      0   0      0
##    Tagagag                                  0      0      0      0   0      0
##    Tc                                       0      0      0      0   0      0
##    TCC                                      0      0      0      0   0      0
##    Tccaa                                    0      0      0      0   0      0
##    Tg                                       0      0      0      0   0      0
##    Ttaaaaaaaaaaaaaa                         0      0      0      0   0      0
##    Ttctc                                    0      0      0      0   0      0
##
##                             ATTTC ATTTTT   C C,CGTgt C,Ct  CA CAA CAAAAA
##    A                            1      1  47        0     0   0   0      0
##    Aaatg                        0      0   0        0     0   0   0      0
##    AACC                         0      0   0        0     0   0   0      0
##    ACTC                         0      0   0        0     0   0   0      0
##    ACTT                         0      0   0        0     0   0   0      0
##    Agaggaggagaaggag             0      0   0        0     0   0   0      0
##    At                           0      0   0        0     0   0   0      0
##    AT                           0      0   0        0     0   0   0      0
##    Atg                          0      0   0        0     0   0   0      0
##    Atgtgatg                     0      0   0        0     0   0   0      0
##    C                            0      0   0        0     0   8   1      1
##    Ca                           0      0   6        0     0   0   0      0
##    Caa                          0      0   3        0     0   0   0      0
##    Caaa                         0      0   2        0     0   0   0      0
##    Caaaa                        0      0   1        0     0   0   0      0
##    Caaaaaa                      0      0   1        0     0   0   0      0
##    Caaaaaaa                     0      0   1        0     0   0   0      0
##    Cagagagagagagagagagagag      0      0   1        0     0   0   0      0
##    CAT                          0      0   1        0     0   0   0      0
##    Cgt                          0      0   0        1     0   0   0      0
##    CGTAT                        0      0   1        0     0   0   0      0
##    Ct                           0      0   6        0     0   0   0      0
##    CT                           0      0   3        0     0   0   0      0
##    Ctt                          0      0   3        0     0   0   0      0
##    Cttt                         0      0   0        0     1   0   0      0
##    Cttttt                       0      0   1        0     0   0   0      0
##    G                            0      0  77        0     0   0   0      0
##    Ga                           0      0   0        0     0   0   0      0
##    GA                           0      0   0        0     0   0   0      0
##    Gaaagaaaagaaaagaaaagaaaaga   0      0   0        0     0   0   0      0
##    GGAA                         0      0   0        0     0   0   0      0
##    Gggaa                        0      0   0        0     0   0   0      0
##    Gggtgtgtgtgt                 0      0   0        0     0   0   0      0
##    Ggtgt                        0      0   0        0     0   0   0      0
##    Gt                           0      0   0        0     0   0   0      0
##    GT                           0      0   0        0     0   0   0      0
##    Gtc                          0      0   0        0     0   0   0      0
##    GTCTTGAAGGAAA                0      0   0        0     0   0   0      0
```

```
##    Gtgt                        0      0   0       0     0   0             0
##    Gtt                         0      0   0       0     0   0             0
##    GTT                         0      0   0       0     0   0             0
##    Gttt                        0      0   0       0     0   0             0
##    Gttttttt                    0      0   0       0     0   0             0
##    T                           0      0 243       0     0   0             0
##    Ta                          0      0   0       0     0   0             0
##    TA                          0      0   0       0     0   0             0
##    Taa                         0      0   0       0     0   0             0
##    Taaa                        0      0   0       0     0   0             0
##    Taaaaaaaaaa                 0      0   0       0     0   0             0
##    Taac                        0      0   0       0     0   0             0
##    Tac                         0      0   0       0     0   0             0
##    Tacacacacacacacac           0      0   0       0     0   0             0
##    Tagagag                     0      0   0       0     0   0             0
##    Tc                          0      0   0       0     0   0             0
##    TCC                         0      0   0       0     0   0             0
##    Tccaa                       0      0   0       0     0   0             0
##    Tg                          0      0   0       0     0   0             0
##    Ttaaaaaaaaaaaaa             0      0   0       0     0   0             0
##    Ttctc                       0      0   0       0     0   0             0
##
##                           CAAAGT CAG CATTG CCA CGCGT CGT CGTGTGTGTGT   CT
##    A                           0   0     0   0     0   0             0   0
##    Aaatg                       0   0     0   0     0   0             0   0
##    AACC                        0   0     0   0     0   0             0   0
##    ACTC                        0   0     0   0     0   0             0   0
##    ACTT                        0   0     0   0     0   0             0   0
##    Agaggaggagaaggag            0   0     0   0     0   0             0   0
##    At                          0   0     0   0     0   0             0   0
##    AT                          0   0     0   0     0   0             0   0
##    Atg                         0   0     0   0     0   0             0   0
##    Atgtgatg                    0   0     0   0     0   0             0   0
##    C                           1   3     1   1     1   2             1   7
##    Ca                          0   0     0   0     0   0             0   0
##    Caa                         0   0     0   0     0   0             0   0
##    Caaa                        0   0     0   0     0   0             0   0
##    Caaaa                       0   0     0   0     0   0             0   0
##    Caaaaaa                     0   0     0   0     0   0             0   0
##    Caaaaaaa                    0   0     0   0     0   0             0   0
##    Cagagagagagagagagagag       0   0     0   0     0   0             0   0
##    CAT                         0   0     0   0     0   0             0   0
##    Cgt                         0   0     0   0     0   0             0   0
##    CGTAT                       0   0     0   0     0   0             0   0
##    Ct                          0   0     0   0     0   0             0   0
##    CT                          0   0     0   0     0   0             0   0
##    Ctt                         0   0     0   0     0   0             0   0
##    Cttt                        0   0     0   0     0   0             0   0
##    Ctttt                       0   0     0   0     0   0             0   0
##    G                           0   0     0   0     0   0             0   0
##    Ga                          0   0     0   0     0   0             0   0
##    GA                          0   0     0   0     0   0             0   0
##    Gaaagaaaagaaaagaaaagaaaaga  0   0     0   0     0   0             0   0
##    GGAA                        0   0     0   0     0   0             0   0
##    Gggaa                       0   0     0   0     0   0             0   0
##    Gggtgtgtgtgt                0   0     0   0     0   0             0   0
```

```
##   Ggtgt                         0   0     0       0         0   0      0   0
##   Gt                            0   0     0       0         0   0      0   0
##   GT                            0   0     0       0         0   0      0   0
##   Gtc                           0   0     0       0         0   0      0   0
##   GTCTTGAAGGAAA                 0   0     0       0         0   0      0   0
##   Gtgt                          0   0     0       0         0   0      0   0
##   Gtt                           0   0     0       0         0   0      0   0
##   GTT                           0   0     0       0         0   0      0   0
##   Gttt                          0   0     0       0         0   0      0   0
##   Gttttttt                      0   0     0       0         0   0      0   0
##   T                             0   0     0       0         0   0      0   0
##   Ta                            0   0     0       0         0   0      0   0
##   TA                            0   0     0       0         0   0      0   0
##   Taa                           0   0     0       0         0   0      0   0
##   Taaa                          0   0     0       0         0   0      0   0
##   Taaaaaaaaaa                   0   0     0       0         0   0      0   0
##   Taac                          0   0     0       0         0   0      0   0
##   Tac                           0   0     0       0         0   0      0   0
##   Tacacacacacacacac             0   0     0       0         0   0      0   0
##   Tagagag                       0   0     0       0         0   0      0   0
##   Tc                            0   0     0       0         0   0      0   0
##   TCC                           0   0     0       0         0   0      0   0
##   Tccaa                         0   0     0       0         0   0      0   0
##   Tg                            0   0     0       0         0   0      0   0
##   Ttaaaaaaaaaaaaa               0   0     0       0         0   0      0   0
##   Ttctc                         0   0     0       0         0   0      0   0
##
##                       CTATGTGTG CTCT CTCTTT CTGTGTGTGTG CTT CTTT CTTTCT
##   A                           0    0      0           0   0    0      0
##   Aaatg                       0    0      0           0   0    0      0
##   AACC                        0    0      0           0   0    0      0
##   ACTC                        0    0      0           0   0    0      0
##   ACTT                        0    0      0           0   0    0      0
##   Agaggaggagaaggag            0    0      0           0   0    0      0
##   At                          0    0      0           0   0    0      0
##   AT                          0    0      0           0   0    0      0
##   Atg                         0    0      0           0   0    0      0
##   Atgtgatg                    0    0      0           0   0    0      0
##   C                           1    1      1           1   1    1      1
##   Ca                          0    0      0           0   0    0      0
##   Caa                         0    0      0           0   0    0      0
##   Caaa                        0    0      0           0   0    0      0
##   Caaaa                       0    0      0           0   0    0      0
##   Caaaaaa                     0    0      0           0   0    0      0
##   Caaaaaaa                    0    0      0           0   0    0      0
##   Cagagagagagagagagagagag     0    0      0           0   0    0      0
##   CAT                         0    0      0           0   0    0      0
##   Cgt                         0    0      0           0   0    0      0
##   CGTAT                       0    0      0           0   0    0      0
##   Ct                          0    0      0           0   0    0      0
##   CT                          0    0      0           0   0    0      0
##   Ctt                         0    0      0           0   0    0      0
##   Cttt                        0    0      0           0   0    0      0
##   Ctttt                       0    0      0           0   0    0      0
##   G                           0    0      0           0   0    0      0
##   Ga                          0    0      0           0   0    0      0
```

```
##    GA                                0    0    0          0    0    0    0
##    Gaaagaaaagaaaagaaaagaaaaga        0    0    0          0    0    0    0
##    GGAA                              0    0    0          0    0    0    0
##    Gggaa                             0    0    0          0    0    0    0
##    Gggtgtgtgtgt                      0    0    0          0    0    0    0
##    Ggtgt                             0    0    0          0    0    0    0
##    Gt                                0    0    0          0    0    0    0
##    GT                                0    0    0          0    0    0    0
##    Gtc                               0    0    0          0    0    0    0
##    GTCTTGAAGGAAA                     0    0    0          0    0    0    0
##    Gtgt                              0    0    0          0    0    0    0
##    Gtt                               0    0    0          0    0    0    0
##    GTT                               0    0    0          0    0    0    0
##    Gttt                              0    0    0          0    0    0    0
##    Gttttttt                          0    0    0          0    0    0    0
##    T                                 0    0    0          0    0    0    0
##    Ta                                0    0    0          0    0    0    0
##    TA                                0    0    0          0    0    0    0
##    Taa                               0    0    0          0    0    0    0
##    Taaa                              0    0    0          0    0    0    0
##    Taaaaaaaaaa                       0    0    0          0    0    0    0
##    Taac                              0    0    0          0    0    0    0
##    Tac                               0    0    0          0    0    0    0
##    Tacacacacacacacac                 0    0    0          0    0    0    0
##    Tagagag                           0    0    0          0    0    0    0
##    Tc                                0    0    0          0    0    0    0
##    TCC                               0    0    0          0    0    0    0
##    Tccaa                             0    0    0          0    0    0    0
##    Tg                                0    0    0          0    0    0    0
##    Ttaaaaaaaaaaaaa                   0    0    0          0    0    0    0
##    Ttctc                             0    0    0          0    0    0    0
##
##                             G  G,GGGAAggaa   GA GAGC GATC GCA GCAA
##    A                      175            0    0    0    0    0    0
##    Aaatg                    0            0    0    0    0    0    0
##    AACC                     0            0    0    0    0    0    0
##    ACTC                     0            0    0    0    0    0    0
##    ACTT                     0            0    0    0    0    0    0
##    Agaggaggagaaggag         0            0    0    0    0    0    0
##    At                       0            0    0    0    0    0    0
##    AT                       0            0    0    0    0    0    0
##    Atg                      0            0    0    0    0    0    0
##    Atgtgatg                 0            0    0    0    0    0    0
##    C                       98            0    0    0    0    0    0
##    Ca                       0            0    0    0    0    0    0
##    Caa                      0            0    0    0    0    0    0
##    Caaa                     0            0    0    0    0    0    0
##    Caaaa                    0            0    0    0    0    0    0
##    Caaaaaa                  0            0    0    0    0    0    0
##    Caaaaaaa                 0            0    0    0    0    0    0
##    Cagagagagagagagagagagag  0            0    0    0    0    0    0
##    CAT                      0            0    0    0    0    0    0
##    Cgt                      0            0    0    0    0    0    0
##    CGTAT                    0            0    0    0    0    0    0
##    Ct                       0            0    0    0    0    0    0
##    CT                       0            0    0    0    0    0    0
```

```
##    Ctt                            0        0    0    0    0    0    0
##    Cttt                           0        0    0    0    0    0    0
##    Cttttt                         0        0    0    0    0    0    0
##    G                              0        0   10    1    1    1    1
##    Ga                             3        0    0    0    0    0    0
##    GA                             3        0    0    0    0    0    0
##    Gaaagaaaagaaaagaaaagaaaaga     1        0    0    0    0    0    0
##    GGAA                           1        0    0    0    0    0    0
##    Gggaa                          0        1    0    0    0    0    0
##    Gggtgtgtgtgt                   1        0    0    0    0    0    0
##    Ggtgt                          2        0    0    0    0    0    0
##    Gt                             4        0    0    0    0    0    0
##    GT                             2        0    0    0    0    0    0
##    Gtc                            1        0    0    0    0    0    0
##    GTCTTGAAGGAAA                  1        0    0    0    0    0    0
##    Gtgt                           1        0    0    0    0    0    0
##    Gtt                            1        0    0    0    0    0    0
##    GTT                            2        0    0    0    0    0    0
##    Gttt                           1        0    0    0    0    0    0
##    Gttttttt                       1        0    0    0    0    0    0
##    T                             52        0    0    0    0    0    0
##    Ta                             0        0    0    0    0    0    0
##    TA                             0        0    0    0    0    0    0
##    Taa                            0        0    0    0    0    0    0
##    Taaa                           0        0    0    0    0    0    0
##    Taaaaaaaaaa                    0        0    0    0    0    0    0
##    Taac                           0        0    0    0    0    0    0
##    Tac                            0        0    0    0    0    0    0
##    Tacacacacacacacac              0        0    0    0    0    0    0
##    Tagagag                        0        0    0    0    0    0    0
##    Tc                             0        0    0    0    0    0    0
##    TCC                            0        0    0    0    0    0    0
##    Tccaa                          0        0    0    0    0    0    0
##    Tg                             0        0    0    0    0    0    0
##    Ttaaaaaaaaaaaaa                0        0    0    0    0    0    0
##    Ttctc                          0        0    0    0    0    0    0
## 
##                   GCAAGTTGGTAAAACC GCACA GGA GGGTGAGTCAGTAAGGAGT    GT
##    A                             0     0   0                     0    0
##    Aaatg                         0     0   0                     0    0
##    AACC                          0     0   0                     0    0
##    ACTC                          0     0   0                     0    0
##    ACTT                          0     0   0                     0    0
##    Agaggaggagaaggag              0     0   0                     0    0
##    At                            0     0   0                     0    0
##    AT                            0     0   0                     0    0
##    Atg                           0     0   0                     0    0
##    Atgtgatg                      0     0   0                     0    0
##    C                             0     0   0                     0    0
##    Ca                            0     0   0                     0    0
##    Caa                           0     0   0                     0    0
##    Caaa                          0     0   0                     0    0
##    Caaaa                         0     0   0                     0    0
##    Caaaaaa                       0     0   0                     0    0
##    Caaaaaaa                      0     0   0                     0    0
##    Cagagagagagagagagagagag       0     0   0                     0    0
```

```
##    CAT                                   0      0    0              0    0
##    Cgt                                   0      0    0              0    0
##    CGTAT                                 0      0    0              0    0
##    Ct                                    0      0    0              0    0
##    CT                                    0      0    0              0    0
##    Ctt                                   0      0    0              0    0
##    Cttt                                  0      0    0              0    0
##    Ctttt                                 0      0    0              0    0
##    G                                     1      1    1              1    3
##    Ga                                    0      0    0              0    0
##    GA                                    0      0    0              0    0
##    Gaaagaaaagaaaagaaaagaaaaga            0      0    0              0    0
##    GGAA                                  0      0    0              0    0
##    Gggaa                                 0      0    0              0    0
##    Gggtgtgtgtgt                          0      0    0              0    0
##    Ggtgt                                 0      0    0              0    0
##    Gt                                    0      0    0              0    0
##    GT                                    0      0    0              0    0
##    Gtc                                   0      0    0              0    0
##    GTCTTGAAGGAAA                         0      0    0              0    0
##    Gtgt                                  0      0    0              0    0
##    Gtt                                   0      0    0              0    0
##    GTT                                   0      0    0              0    0
##    Gttt                                  0      0    0              0    0
##    Gttttttt                              0      0    0              0    0
##    T                                     0      0    0              0    0
##    Ta                                    0      0    0              0    0
##    TA                                    0      0    0              0    0
##    Taa                                   0      0    0              0    0
##    Taaa                                  0      0    0              0    0
##    Taaaaaaaaaa                           0      0    0              0    0
##    Taac                                  0      0    0              0    0
##    Tac                                   0      0    0              0    0
##    Tacacacacacacacac                     0      0    0              0    0
##    Tagagag                               0      0    0              0    0
##    Tc                                    0      0    0              0    0
##    TCC                                   0      0    0              0    0
##    Tccaa                                 0      0    0              0    0
##    Tg                                    0      0    0              0    0
##    Ttaaaaaaaaaaaaa                       0      0    0              0    0
##    Ttctc                                 0      0    0              0    0
##
##              GTA GTATTATAACTTT GTC GTTTTT   T  TA
##    A           0             0   0      0  37   0
##    Aaatg       0             0   0      0   0   0
##    AACC        0             0   0      0   0   0
##    ACTC        0             0   0      0   0   0
##    ACTT        0             0   0      0   0   0
##    Agaggaggagaaggag 0        0   0      0   0   0
##    At          0             0   0      0   0   0
##    AT          0             0   0      0   0   0
##    Atg         0             0   0      0   0   0
##    Atgtgatg    0             0   0      0   0   0
##    C           0             0   0      0 225   0
##    Ca          0             0   0      0   0   0
##    Caa         0             0   0      0   0   0
```

##							
##	Caaa	0	0	0	0	0	0
##	Caaaa	0	0	0	0	0	0
##	Caaaaaa	0	0	0	0	0	0
##	Caaaaaaa	0	0	0	0	0	0
##	Cagagagagagagagagagagag	0	0	0	0	0	0
##	CAT	0	0	0	0	0	0
##	Cgt	0	0	0	0	0	0
##	CGTAT	0	0	0	0	0	0
##	Ct	0	0	0	0	0	0
##	CT	0	0	0	0	0	0
##	Ctt	0	0	0	0	0	0
##	Cttt	0	0	0	0	0	0
##	Ctttt	0	0	0	0	0	0
##	G	1	1	1	1	55	0
##	Ga	0	0	0	0	0	0
##	GA	0	0	0	0	0	0
##	Gaaagaaaagaaaagaaaagaaaaga	0	0	0	0	0	0
##	GGAA	0	0	0	0	0	0
##	Gggaa	0	0	0	0	0	0
##	Gggtgtgtgtgt	0	0	0	0	0	0
##	Ggtgt	0	0	0	0	0	0
##	Gt	0	0	0	0	0	0
##	GT	0	0	0	0	0	0
##	Gtc	0	0	0	0	0	0
##	GTCTTGAAGGAAA	0	0	0	0	0	0
##	Gtgt	0	0	0	0	0	0
##	Gtt	0	0	0	0	0	0
##	GTT	0	0	0	0	0	0
##	Gttt	0	0	0	0	0	0
##	Gttttttt	0	0	0	0	0	0
##	T	0	0	0	0	0	6
##	Ta	0	0	0	0	5	0
##	TA	0	0	0	0	4	0
##	Taa	0	0	0	0	4	0
##	Taaa	0	0	0	0	2	0
##	Taaaaaaaaaa	0	0	0	0	1	0
##	Taac	0	0	0	0	1	0
##	Tac	0	0	0	0	1	0
##	Tacacacacacacacac	0	0	0	0	1	0
##	Tagagag	0	0	0	0	1	0
##	Tc	0	0	0	0	2	0
##	TCC	0	0	0	0	1	0
##	Tccaa	0	0	0	0	1	0
##	Tg	0	0	0	0	1	0
##	Ttaaaaaaaaaaaaa	0	0	0	0	1	0
##	Ttctc	0	0	0	0	1	0
##							

##		TACACACAC,TACACAC	TAGG	TATATTTCATGTAACAGCCAGGG
##	A	0	0	0
##	Aaatg	0	0	0
##	AACC	0	0	0
##	ACTC	0	0	0
##	ACTT	0	0	0
##	Agaggaggagaaggag	0	0	0
##	At	0	0	0
##	AT	0	0	0

```
##    Atg                                   0     0                     0
##    Atgtgatg                              0     0                     0
##    C                                     0     0                     0
##    Ca                                    0     0                     0
##    Caa                                   0     0                     0
##    Caaa                                  0     0                     0
##    Caaaa                                 0     0                     0
##    Caaaaaa                               0     0                     0
##    Caaaaaaa                              0     0                     0
##    Cagagagagagagagagagagag              0     0                     0
##    CAT                                   0     0                     0
##    Cgt                                   0     0                     0
##    CGTAT                                 0     0                     0
##    Ct                                    0     0                     0
##    CT                                    0     0                     0
##    Ctt                                   0     0                     0
##    Cttt                                  0     0                     0
##    Ctttt                                 0     0                     0
##    G                                     0     0                     0
##    Ga                                    0     0                     0
##    GA                                    0     0                     0
##    Gaaagaaaagaaaagaaaagaaaaga           0     0                     0
##    GGAA                                  0     0                     0
##    Gggaa                                 0     0                     0
##    Gggtgtgtgtgt                          0     0                     0
##    Ggtgt                                 0     0                     0
##    Gt                                    0     0                     0
##    GT                                    0     0                     0
##    Gtc                                   0     0                     0
##    GTCTTGAAGGAAA                         0     0                     0
##    Gtgt                                  0     0                     0
##    Gtt                                   0     0                     0
##    GTT                                   0     0                     0
##    Gttt                                  0     0                     0
##    Gttttttt                              0     0                     0
##    T                                     1     1                     1
##    Ta                                    0     0                     0
##    TA                                    0     0                     0
##    Taa                                   0     0                     0
##    Taaa                                  0     0                     0
##    Taaaaaaaaaa                           0     0                     0
##    Taac                                  0     0                     0
##    Tac                                   0     0                     0
##    Tacacacacacacacac                     0     0                     0
##    Tagagag                               0     0                     0
##    Tc                                    0     0                     0
##    TCC                                   0     0                     0
##    Tccaa                                 0     0                     0
##    Tg                                    0     0                     0
##    Ttaaaaaaaaaaaaa                       0     0                     0
##    Ttctc                                 0     0                     0
##
##                          TCCA TCTC  TG TGAGAA TGTGG TTA TTG TTGTG
##    A                        0    0   0      0     0   0   0     0
##    Aaatg                    0    0   0      0     0   0   0     0
##    AACC                     0    0   0      0     0   0   0     0
```

```
##   ACTC                           0    0    0     0    0    0    0    0
##   ACTT                           0    0    0     0    0    0    0    0
##   Agaggaggagaaggag               0    0    0     0    0    0    0    0
##   At                             0    0    0     0    0    0    0    0
##   AT                             0    0    0     0    0    0    0    0
##   Atg                            0    0    0     0    0    0    0    0
##   Atgtgatg                       0    0    0     0    0    0    0    0
##   C                              0    0    0     0    0    0    0    0
##   Ca                             0    0    0     0    0    0    0    0
##   Caa                            0    0    0     0    0    0    0    0
##   Caaa                           0    0    0     0    0    0    0    0
##   Caaaa                          0    0    0     0    0    0    0    0
##   Caaaaaa                        0    0    0     0    0    0    0    0
##   Caaaaaaa                       0    0    0     0    0    0    0    0
##   Cagagagagagagagagagagag        0    0    0     0    0    0    0    0
##   CAT                            0    0    0     0    0    0    0    0
##   Cgt                            0    0    0     0    0    0    0    0
##   CGTAT                          0    0    0     0    0    0    0    0
##   Ct                             0    0    0     0    0    0    0    0
##   CT                             0    0    0     0    0    0    0    0
##   Ctt                            0    0    0     0    0    0    0    0
##   Cttt                           0    0    0     0    0    0    0    0
##   Ctttt                          0    0    0     0    0    0    0    0
##   G                              0    0    0     0    0    0    0    0
##   Ga                             0    0    0     0    0    0    0    0
##   GA                             0    0    0     0    0    0    0    0
##   Gaaagaaaagaaaagaaaagaaaaga     0    0    0     0    0    0    0    0
##   GGAA                           0    0    0     0    0    0    0    0
##   Gggaa                          0    0    0     0    0    0    0    0
##   Gggtgtgtgtgt                   0    0    0     0    0    0    0    0
##   Ggtgt                          0    0    0     0    0    0    0    0
##   Gt                             0    0    0     0    0    0    0    0
##   GT                             0    0    0     0    0    0    0    0
##   Gtc                            0    0    0     0    0    0    0    0
##   GTCTTGAAGGAAA                  0    0    0     0    0    0    0    0
##   Gtgt                           0    0    0     0    0    0    0    0
##   Gtt                            0    0    0     0    0    0    0    0
##   GTT                            0    0    0     0    0    0    0    0
##   Gttt                           0    0    0     0    0    0    0    0
##   Gttttttt                       0    0    0     0    0    0    0    0
##   T                              1    2    2     1    1    1    1    1
##   Ta                             0    0    0     0    0    0    0    0
##   TA                             0    0    0     0    0    0    0    0
##   Taa                            0    0    0     0    0    0    0    0
##   Taaa                           0    0    0     0    0    0    0    0
##   Taaaaaaaaaa                    0    0    0     0    0    0    0    0
##   Taac                           0    0    0     0    0    0    0    0
##   Tac                            0    0    0     0    0    0    0    0
##   Tacacacacacacacac              0    0    0     0    0    0    0    0
##   Tagagag                        0    0    0     0    0    0    0    0
##   Tc                             0    0    0     0    0    0    0    0
##   TCC                            0    0    0     0    0    0    0    0
##   Tccaa                          0    0    0     0    0    0    0    0
##   Tg                             0    0    0     0    0    0    0    0
```

```
##    Ttaaaaaaaaaaaaaa             0    0    0       0     0    0    0     0
##    Ttctc                        0    0    0       0     0    0    0     0
```

```
# Check if these are mostly SNPs, Insertions, or Deletions
table(alk_cluster_muts$Variant_Type)
```

```
##
##  DEL  INS  SNP
##   98  107 1288
```
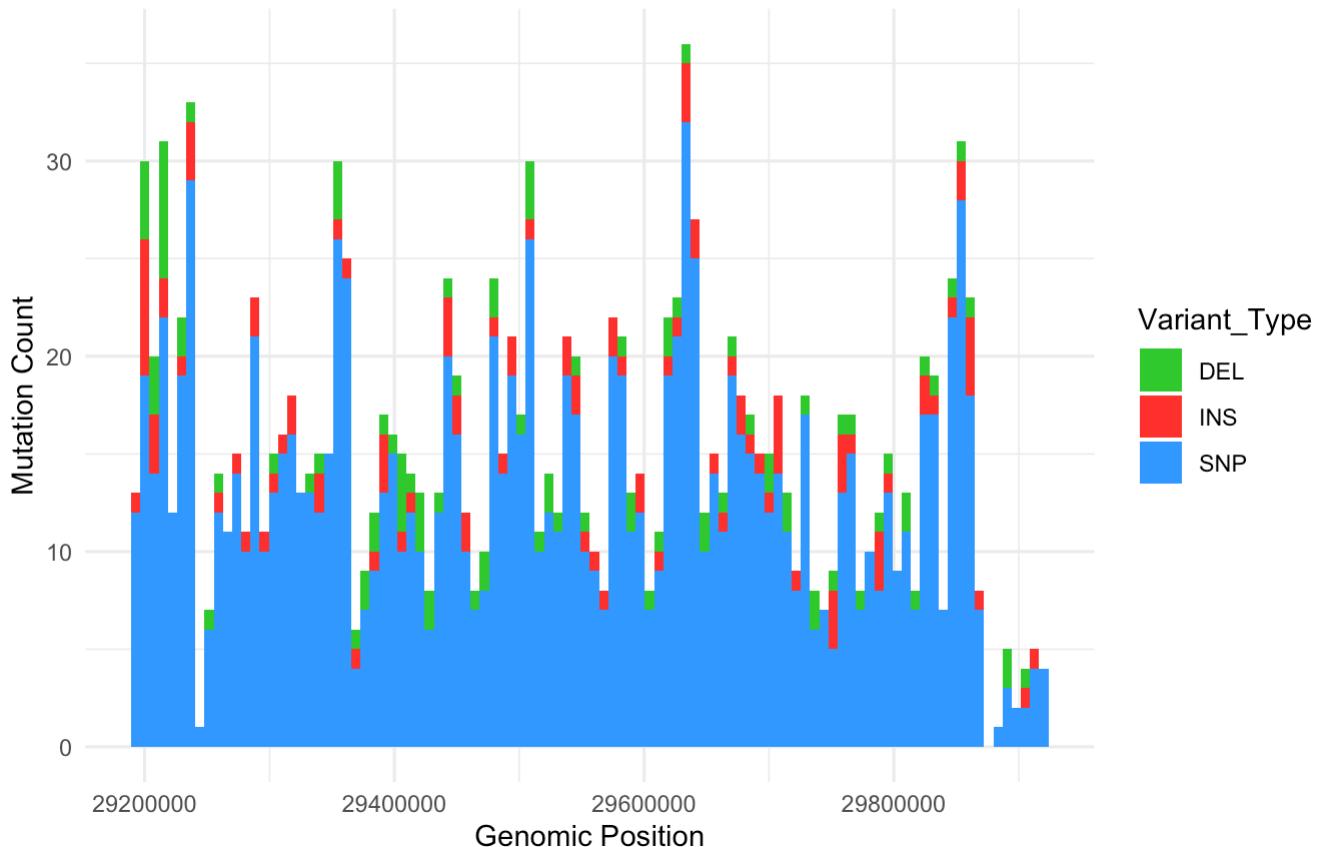
```
# This will likely show a high number of INS (Insertions) and DEL (Deletions)
```

```
# Visualizing where the Indels fall compared to SNPs
library(ggplot2)

ggplot(alk_cluster_muts, aes(x = Start_Position, fill = Variant_Type)) +
  geom_histogram(bins = 100, position = "stack") +
  scale_fill_manual(values = c("SNP" = "#3399FF", "INS" = "#FF3333", "DEL" = "#33CC3
3")) +
  theme_minimal() +
  labs(title = "Mutation Distribution Across ALK Gene",
       subtitle = "Concentration of SNPs and Indels",
       x = "Genomic Position",
       y = "Mutation Count")
```

```r
# Compare SNP/Indel ratios across your top hit genes
panel_hits <- final_maf_df[Hugo_Symbol %in% c("ALK", "RET", "TP53", "TG")]
table(panel_hits$Hugo_Symbol, panel_hits$Variant_Type)
```

```
##
##         DEL  INS  SNP
##   ALK    98  107 1288
##   RET     6   11  137
##   TG     53   30  543
##   TP53    8    5   31
```

```r
# Find the 1kb window in ALK with the highest density of Indels
alk_indels <- alk_cluster_muts[Variant_Type != "SNP"]

# Create 1kb bins and count mutations
alk_indels[, bin := floor(Start_Position / 1000) * 1000]
breakpoint_counts <- alk_indels[, .N, by = bin][order(-N)]

# View the top "hotspots"
head(breakpoint_counts)
```

```
##         bin     N
##       <num> <int>
## 1: 29200000     4
## 2: 29217000     4
## 3: 29216000     3
## 4: 29198000     2
## 5: 29201000     2
## 6: 29203000     2
```

```r
# Calculate Indel-to-SNP ratio for Panel vs Intergenic
instability_summary <- final_maf_df[, .(
  Total_Muts = .N,
  SNP_Count = sum(Variant_Type == "SNP"),
  Indel_Count = sum(Variant_Type %in% c("INS", "DEL")),
  Indel_Ratio = sum(Variant_Type %in% c("INS", "DEL")) / .N
), by = .(In_Panel = Hugo_Symbol != "Intergenic_Region")]

print(instability_summary)
```

```
##    In_Panel Total_Muts SNP_Count Indel_Count Indel_Ratio
##      <lgcl>      <int>     <int>       <int>       <num>
## 1:     TRUE       5753      4617        1136   0.1974622
## 2:    FALSE         39        35           4   0.1025641
```

```r
# Compare the specific types of Indels in ALK vs TG
shattered_comparison <- final_maf_df[Hugo_Symbol %in% c("ALK", "TG")]

# Create a table of Variant Type by Gene
table(shattered_comparison$Hugo_Symbol, shattered_comparison$Variant_Type)
```
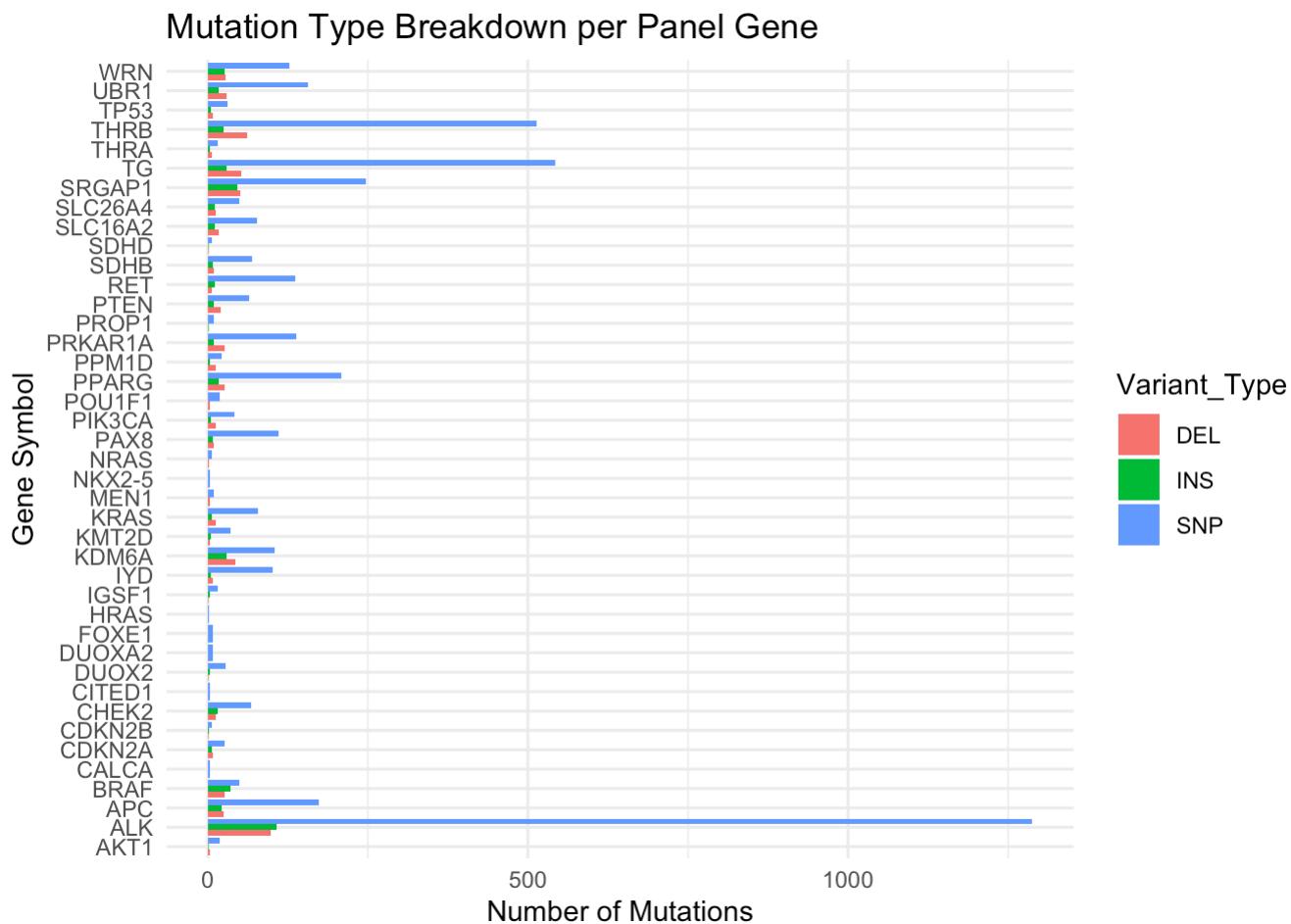
```
##
##        DEL  INS  SNP
##  ALK   98  107 1288
##  TG    53   30  543
```

```
# Plotting the mutation types across the top genes in your panel
library(ggplot2)

# Filter out the intergenic stuff for a cleaner plot
panel_only <- final_maf_df[Hugo_Symbol != "Intergenic_Region"]

ggplot(panel_only, aes(x = Hugo_Symbol, fill = Variant_Type)) +
  geom_bar(position = "dodge") +
  theme_minimal() +
  coord_flip() + # Makes gene names easier to read
  labs(title = "Mutation Type Breakdown per Panel Gene",
       x = "Gene Symbol",
       y = "Number of Mutations")
```


Mutation Type Breakdown per Panel Gene

```
# 1. Isolate ALK mutations from your annotated data
alk_data <- final_maf_df[Hugo_Symbol == "ALK"]

# 2. Identify the densest 1kb windows for Indels (potential structural scars)
alk_indels <- alk_data[Variant_Type != "SNP"]
alk_indels[, bin := floor(Start_Position / 1000) * 1000]

# 3. Summarize and find the peak "Shatter Zone"
alk_hotspots <- alk_indels[, .(
  Indel_Count = .N,
  Unique_Start_Sites = uniqueN(Start_Position)
), by = bin][order(-Indel_Count)]

# 4. Export the top hotspots to CSV
write.csv(head(alk_hotspots, 10), "ALK_shatter_hotspots.csv", row.names = FALSE)

# 5. Print to console for immediate review
print("Top 5 high-density Indel bins in ALK:")
```

```
## [1] "Top 5 high-density Indel bins in ALK:"
```

```
print(head(alk_hotspots, 5))
```

```
##           bin Indel_Count Unique_Start_Sites
##         <num>       <int>             <int>
## 1: 29200000           4                 4
## 2: 29217000           4                 4
## 3: 29216000           3                 3
## 4: 29198000           2                 2
## 5: 29201000           2                 2
```

```
# Load GenomicRanges to handle coordinate overlaps
library(GenomicRanges)
```

```
## Loading required package: stats4
```

```
## Loading required package: BiocGenerics
```

```
## Loading required package: generics
```

```
##
## Attaching package: 'generics'
```

```
## The following objects are masked from 'package:base':
##
##     as.difftime, as.factor, as.ordered, intersect, is.element, setdiff,
##     setequal, union
```

```
## 
## Attaching package: 'BiocGenerics'
```

```
## The following objects are masked from 'package:stats':
## 
##     IQR, mad, sd, var, xtabs
```

```
## The following objects are masked from 'package:base':
## 
##     anyDuplicated, aperm, append, as.data.frame, basename, cbind,
##     colnames, dirname, do.call, duplicated, eval, evalq, Filter, Find,
##     get, grep, grepl, is.unsorted, lapply, Map, mapply, match, mget,
##     order, paste, pmax, pmax.int, pmin, pmin.int, Position, rank,
##     rbind, Reduce, rownames, sapply, saveRDS, table, tapply, unique,
##     unsplit, which.max, which.min
```

```
## Loading required package: S4Vectors
```

```
## 
## Attaching package: 'S4Vectors'
```

```
## The following objects are masked from 'package:data.table':
## 
##     first, second
```

```
## The following object is masked from 'package:utils':
## 
##     findMatches
```

```
## The following objects are masked from 'package:base':
## 
##     expand.grid, I, unname
```

```
## Loading required package: IRanges
```

```
## 
## Attaching package: 'IRanges'
```

```
## The following object is masked from 'package:data.table':
## 
##     shift
```

```
## Loading required package: Seqinfo
```

```
# Define the ALK hotspot regions from your result
hotspot_ranges <- GRanges(seqnames = "chr2",
                          ranges = IRanges(start = alk_hotspots$bin,
                                           end = alk_hotspots$bin + 1000))

# Check if these bins are specifically impacting exons
# (This assumes you have an exon database loaded, or you can check manually)
print("The primary shatter zone is localized between 29,198,000 and 29,217,000.")
```

```
## [1] "The primary shatter zone is localized between 29,198,000 and 29,217,000."
```

```
# --- 7. AUTOMATED RESULTS EXPORT ---

# Define the output directory name
output_dir <- paste0("MAF_Analysis_Results_", Sys.Date())

# Create the directory if it doesn't exist
if (!dir.exists(output_dir)) {
  dir.create(output_dir)
}

message(paste("Saving results to:", output_dir))
```

```
## Saving results to: MAF_Analysis_Results_2026-01-07
```

```r
# --- A. Exporting Data Tables (CSV) ---

# 1. Mutation counts per chromosome
write.csv(chr_summary,
          file = file.path(output_dir, "01_chromosome_mutation_summary.csv"), row.nam
es = FALSE)

# 2. Kataegis clusters (Rainfall data)
if(!exists("kataegis_clusters")) {
  kataegis_clusters <- rainfallPlot(maf = my_maf, tsb = "combined_minion_promethion",
plot = FALSE)
}
write.csv(kataegis_clusters,
          file = file.path(output_dir, "02_kataegis_hotspots.csv"), row.names = FALS
E)

# --- 3. APOBEC Enrichment Scores (Revised) ---

# Check if object exists; if not, try to generate it again
if (!exists("apobec_enrich")) {
  message("APOBEC scores not found. Attempting to calculate...")

  # Wrapping in try() prevents the script from crashing if the calculation fails
  try({
    apobec_enrich = enrichments(maf = my_maf_final,
                                ref_genome = "BSgenome.Hsapiens.UCSC.hg38")
  }, silent = TRUE)
}
```

```
## APOBEC scores not found. Attempting to calculate...
```

```r
# Only write if the object was successfully created
if (exists("apobec_enrich") && !is.null(apobec_enrich)) {
  write.csv(apobec_enrich,
            file = file.path(output_dir, "03_apobec_enrichment_scores.csv"),
            row.names = FALSE)
  message("APOBEC scores saved successfully.")
} else {
  warning("APOBEC enrichment could not be calculated. Skipping export.")
}
```

```
## Warning: APOBEC enrichment could not be calculated. Skipping export.
```

```r
# 4. Genomic Instability Summary (Panel vs Intergenic)
write.csv(instability_summary,
          file = file.path(output_dir, "04_genomic_instability_comparison.csv"), row.
names = FALSE)

# 5. ALK Shatter Zone Hotspots
write.csv(alk_hotspots,
          file = file.path(output_dir, "05_ALK_shatter_hotspots.csv"), row.names = FA
LSE)



# --- B. Exporting High-Resolution Plots (PDF) ---

pdf(file = file.path(output_dir, "06_genomic_visualizations.pdf"), width = 10, height
= 8)

  # Plot 1: Summary Dashboard
  plotmafSummary(maf = my_maf_final, addStat = 'median', rmOutlier = TRUE, dashboard
= TRUE)

  # Plot 2: Oncoplot of Panel Genes
  oncoplot(maf = my_maf_final, genes = panel$Gene, drawRowBar = TRUE, fontSize = 0.6)
```

```
## Warning in min(x): no non-missing arguments to min; returning Inf
```

```
## Warning in max(x): no non-missing arguments to max; returning -Inf
```

```r
  # Plot 3: Rainfall Plot (Kataegis)
  rainfallPlot(maf = my_maf_final, tsb = "combined_minion_promethion", detectChangePo
ints = TRUE)
```

```
## Processing combined_minion_promethion..
```

```
## Kataegis detected at:
```

```
##      Chromosome Start_Position End_Position nMuts Avg_intermutation_dist    Size
##           <num>         <num>        <num> <int>                  <num>   <num>
##  1:          1      17019410     17054002    69                508.7059   34592
##  2:          2      29192926     29920131  1288                565.0389  727205
##  3:          2     113216055    113277617   111                559.6545   61562
##  4:          3      12287696     12434058   209                703.6635  146362
##  5:          3      24123533     24128235     7                783.6667    4702
##  6:          3      24130634     24493919   504                722.2366  363285
##  7:          3      87259403     87275374    19                887.2778   15971
##  8:          3     179186419    179192974     9                819.3750    6555
##  9:          3     179217179    179222517     7                889.6667    5338
## 10:          5     112709435    112716131     8                956.5714    6696
## 11:          5     112718389    112845879   165                777.3780  127490
## 12:          5     177992355    177995875     9                440.0000    3520
## 13:          6     150370221    150405763   101                355.4200   35542
## 14:          7     107661438    107664670     6                646.4000    3232
## 15:          7     107677040    107716314    41                981.8500   39274
## 16:          8      31044583     31055016    12                948.4545   10433
## 17:          8      31069763     31080274    13                875.9167   10511
## 18:          8      31102488     31174361    87                835.7326   71873
## 19:          8     132867967    133131965   543                487.0812  263998
## 20:          9      21976672     21981087     6                883.0000    4415
## 21:          9      21983915     21995144    16                748.6000   11229
## 22:          9      97853378     97856378     8                428.5714    3000
## 23:         10      43077059     43130238   137                391.0221   53179
## 24:         10      87910574     87917350     8                968.0000    6776
## 25:         10      87939302     87949237    11                993.5000    9935
## 26:         10      87957058     87964003    10                771.6667    6945
## 27:         11      64804546     64809675     9                641.1250    5129
## 28:         12      25205716     25250690    78                584.0779   44974
## 29:         12      49020848     49042941    24                960.5652   22093
## 30:         12      49045593     49051609     8                859.4286    6016
## 31:         12      63950137     63954978     7                806.8333    4841
## 32:         12      63982306     63986666     7                726.6667    4360
## 33:         12      63992237     64018932    28                988.7037   26695
## 34:         12      64022548     64035057    24                543.8696   12509
## 35:         12      64058253     64160255   116                886.9739  102002
## 36:         14     104771710    104785733    16                934.8667   14023
## 37:         15      42944414     42948618     8                600.5714    4204
## 38:         15      42967649     42981053    18                788.4706   13404
## 39:         15      42988662     42994605     7                990.5000    5943
## 40:         15      43002020     43052688    54                956.0000   50668
## 41:         15      43074564     43105910    50                639.7143   31346
## 42:         15      45101541     45118049    32                532.5161   16508
## 43:         17       7662815      7686387    31                785.7333   23572
## 44:         17      68422443     68428041     7                933.0000    5598
## 45:         17      68431378     68451163    22                942.1429   19785
## 46:         17      68461865     68493367    36                900.0571   31502
## 47:         17      68498691     68549311    66                778.7692   50620
## 48:         22      28690105     28739901    67                754.4848   49796
## 49:         23      44922028     44930899    12                806.4545    8871
## 50:         23      44963483     44974171    18                628.7059   10688
## 51:         23      74443903     74449700     7                966.1667    5797
## 52:         23      74456451     74463710     9                907.3750    7259
## 53:         23      74490685     74531188    51                810.0600   40503
```

```
## 54:        23    131277902    131282984     7                    847.0000   5082
## 55:        23    131285672    131289385     8                    530.4286   3713
##      Chromosome Start_Position End_Position nMuts Avg_intermutation_dist   Size
##          <num>          <num>        <num> <int>                   <num> <num>
##      Tumor_Sample_Barcode  C>A   C>G   C>T   T>A   T>C   T>G
##                    <fctr> <int> <int> <int> <int> <int> <int>
##  1: combined_minion_promethion     4     5    29     4    21     6
##  2: combined_minion_promethion   106   175   413    77   418    99
##  3: combined_minion_promethion    10    11    49     2    31     8
##  4: combined_minion_promethion    23    29    65    13    67    12
##  5: combined_minion_promethion     1     1     3     1    NA     1
##  6: combined_minion_promethion    43    74   171    43   140    33
##  7: combined_minion_promethion    NA     3     4     2     9     1
##  8: combined_minion_promethion    NA     2     5    NA     2    NA
##  9: combined_minion_promethion     1     3     2     1    NA    NA
## 10: combined_minion_promethion     1     1     2    NA     3     1
## 11: combined_minion_promethion    15    30    54    12    43    11
## 12: combined_minion_promethion    NA     2     4    NA     3    NA
## 13: combined_minion_promethion    12     8    29     5    38     9
## 14: combined_minion_promethion    NA    NA     4     1     1    NA
## 15: combined_minion_promethion     2     9     9    10     4     7
## 16: combined_minion_promethion     1     1     5     1     4    NA
## 17: combined_minion_promethion     2     2     3     1     3     2
## 18: combined_minion_promethion     7     6    29     5    30    10
## 19: combined_minion_promethion    40    68   214    33   155    33
## 20: combined_minion_promethion    NA     1     3    NA     2    NA
## 21: combined_minion_promethion     2     3     4     1     5     1
## 22: combined_minion_promethion    NA     2     2    NA     3     1
## 23: combined_minion_promethion    10    17    37     8    54    11
## 24: combined_minion_promethion    NA     1     3     1     3    NA
## 25: combined_minion_promethion     1     4     3    NA     2     1
## 26: combined_minion_promethion     2     1     3     2     1     1
## 27: combined_minion_promethion     1    NA     5    NA     3    NA
## 28: combined_minion_promethion     4     9    30     7    17    11
## 29: combined_minion_promethion     3     4     9     1     6     1
## 30: combined_minion_promethion     1    NA     3    NA     4    NA
## 31: combined_minion_promethion    NA     1     1     1     3     1
## 32: combined_minion_promethion     1     1     3    NA     1     1
## 33: combined_minion_promethion    NA     9    11     4     2     2
## 34: combined_minion_promethion    NA     1    12    NA     9     2
## 35: combined_minion_promethion     8    18    31     6    42    11
## 36: combined_minion_promethion     1     4     5    NA     5     1
## 37: combined_minion_promethion     1     1     1    NA     4     1
## 38: combined_minion_promethion     3     1     7    NA     7    NA
## 39: combined_minion_promethion    NA     1     4    NA     2    NA
## 40: combined_minion_promethion     5     7    18     1    17     6
## 41: combined_minion_promethion     2    14    18     2    13     1
## 42: combined_minion_promethion     3     6     9     2     5     7
## 43: combined_minion_promethion     1     5    13    NA    11     1
## 44: combined_minion_promethion     1     1     1    NA     4    NA
## 45: combined_minion_promethion     1     5     3     3    10    NA
## 46: combined_minion_promethion     2     4    15     1    10     4
## 47: combined_minion_promethion     9     9    26     3    15     4
## 48: combined_minion_promethion     2    16    17     5    22     5
## 49: combined_minion_promethion    NA     2     7    NA     2     1
## 50: combined_minion_promethion    NA     4     4     1     4     5
```

```
## 51: combined_minion_promethion     NA     NA     1     1     5     NA
## 52: combined_minion_promethion      1      1     2    NA     3      2
## 53: combined_minion_promethion      2      9    16     4    17      3
## 54: combined_minion_promethion     NA      2     4    NA     1     NA
## 55: combined_minion_promethion      2      1     1    NA     4     NA
##            Tumor_Sample_Barcode    C>A    C>G   C>T   T>A   T>C    T>G
##                              <fctr> <int> <int> <int> <int> <int> <int>
```

```r
  # Plot 4: Mutation Distribution Across ALK (the ggplot)
  p1 <- ggplot(alk_cluster_muts, aes(x = Start_Position, fill = Variant_Type)) +
    geom_histogram(bins = 100, position = "stack") +
    scale_fill_manual(values = c("SNP" = "#3399FF", "INS" = "#FF3333", "DEL" = "#33CC
33")) +
    theme_minimal() +
    labs(title = "Shatter Zone: Mutation Distribution Across ALK", x = "Genomic Posit
ion", y = "Count")
  print(p1)

  # Plot 5: Mutation Type Breakdown
  p2 <- ggplot(panel_only, aes(x = Hugo_Symbol, fill = Variant_Type)) +
    geom_bar(position = "dodge") +
    theme_minimal() + coord_flip() +
    labs(title = "Mutation Type Breakdown per Panel Gene")
  print(p2)

dev.off()
```

```
## quartz_off_screen
##                   2
```

```r
message("All files successfully saved.")
```

```
## All files successfully saved.
```