TARGET Osteosarcoma small variant analysis

Sean Davis

9/24/2018

fnames = list.files('~/Downloads/target\_variant\_calling/', pattern='\*maftools.maf', recursive=TRUE, full.names = TRUE)  
names(fnames) = c("discovery", "valid\_exome", "valid\_lcc", "valid\_roche")  
library(readr)  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

vars = bind\_rows(lapply(fnames, function(x) read\_tsv(x) %>% select(-PUBMED)), .id="dataset")

## Parsed with column specification:  
## cols(  
## .default = col\_character(),  
## Entrez\_Gene\_Id = col\_integer(),  
## Start\_Position = col\_integer(),  
## End\_Position = col\_integer(),  
## t\_depth = col\_integer(),  
## t\_ref\_count = col\_integer(),  
## t\_alt\_count = col\_integer(),  
## n\_depth = col\_integer(),  
## n\_ref\_count = col\_integer(),  
## n\_alt\_count = col\_integer(),  
## ALLELE\_NUM = col\_integer(),  
## STRAND\_VEP = col\_integer(),  
## HGNC\_ID = col\_integer(),  
## PUBMED = col\_integer(),  
## PICK = col\_integer(),  
## MINIMISED = col\_integer(),  
## ExAC\_AF = col\_double(),  
## ExAC\_AF\_AFR = col\_double(),  
## ExAC\_AF\_AMR = col\_double(),  
## ExAC\_AF\_EAS = col\_double(),  
## ExAC\_AF\_FIN = col\_double()  
## # ... with 13 more columns  
## )

## See spec(...) for full column specifications.

## Warning in rbind(names(probs), probs\_f): number of columns of result is not  
## a multiple of vector length (arg 1)

## Warning: 3 parsing failures.  
## row # A tibble: 3 x 5 col row col expected actual file expected <int> <chr> <chr> <chr> <chr> actual 1 1085 PUBMED no trailing c… ,25404506,… '/Users/sdavis2/Downloads/targe… file 2 1154 PUBMED no trailing c… ,25164601,… '/Users/sdavis2/Downloads/targe… row 3 1157 PUBMED no trailing c… ,23690342 '/Users/sdavis2/Downloads/targe…

## Parsed with column specification:  
## cols(  
## .default = col\_character(),  
## Entrez\_Gene\_Id = col\_integer(),  
## Start\_Position = col\_integer(),  
## End\_Position = col\_integer(),  
## t\_depth = col\_integer(),  
## t\_ref\_count = col\_integer(),  
## t\_alt\_count = col\_integer(),  
## n\_depth = col\_integer(),  
## n\_ref\_count = col\_integer(),  
## n\_alt\_count = col\_integer(),  
## ALLELE\_NUM = col\_integer(),  
## STRAND\_VEP = col\_integer(),  
## HGNC\_ID = col\_integer(),  
## PICK = col\_integer(),  
## MINIMISED = col\_integer(),  
## ExAC\_AF = col\_double(),  
## ExAC\_AF\_AFR = col\_double(),  
## ExAC\_AF\_AMR = col\_double(),  
## ExAC\_AF\_EAS = col\_double(),  
## ExAC\_AF\_FIN = col\_double(),  
## ExAC\_AF\_NFE = col\_double()  
## # ... with 11 more columns  
## )  
## See spec(...) for full column specifications.

## Parsed with column specification:  
## cols(  
## .default = col\_character(),  
## Entrez\_Gene\_Id = col\_integer(),  
## Start\_Position = col\_integer(),  
## End\_Position = col\_integer(),  
## t\_depth = col\_integer(),  
## t\_ref\_count = col\_integer(),  
## t\_alt\_count = col\_integer(),  
## n\_depth = col\_integer(),  
## n\_ref\_count = col\_integer(),  
## n\_alt\_count = col\_integer(),  
## ALLELE\_NUM = col\_integer(),  
## STRAND\_VEP = col\_integer(),  
## HGNC\_ID = col\_integer(),  
## PUBMED = col\_number(),  
## PICK = col\_integer(),  
## MINIMISED = col\_integer(),  
## ExAC\_AF = col\_double(),  
## ExAC\_AF\_AFR = col\_integer(),  
## ExAC\_AF\_AMR = col\_integer(),  
## ExAC\_AF\_EAS = col\_integer(),  
## ExAC\_AF\_FIN = col\_integer()  
## # ... with 10 more columns  
## )

## See spec(...) for full column specifications.

## Parsed with column specification:  
## cols(  
## .default = col\_character(),  
## Entrez\_Gene\_Id = col\_integer(),  
## Start\_Position = col\_integer(),  
## End\_Position = col\_integer(),  
## t\_depth = col\_integer(),  
## t\_ref\_count = col\_integer(),  
## t\_alt\_count = col\_integer(),  
## n\_depth = col\_integer(),  
## n\_ref\_count = col\_integer(),  
## n\_alt\_count = col\_integer(),  
## ALLELE\_NUM = col\_integer(),  
## STRAND\_VEP = col\_integer(),  
## HGNC\_ID = col\_integer(),  
## PUBMED = col\_number(),  
## PICK = col\_integer(),  
## MINIMISED = col\_integer(),  
## ExAC\_AF = col\_double(),  
## ExAC\_AF\_AFR = col\_double(),  
## ExAC\_AF\_AMR = col\_double(),  
## ExAC\_AF\_EAS = col\_double(),  
## ExAC\_AF\_FIN = col\_double()  
## # ... with 10 more columns  
## )

## See spec(...) for full column specifications.

library(DT)  
datatable(vars)

## Warning in instance$preRenderHook(instance): It seems your data is too  
## big for client-side DataTables. You may consider server-side processing:  
## https://rstudio.github.io/DT/server.html

