Balder annotation matching summary

Larson Hogstrom

2024-05-13

Overview

The goal of this document is to report summary statistics for variant annotation matching.

Loading data

```
bDir <- "../../../data/processed/balderResultsDb"
figDir <- "../../../output/actionability_db_curration_20231220"
dbName <- paste0(bDir,"/balder-harmonized-biomarker-data-v20240412.sqlite")
harmonizedDb <- DBI::dbConnect(RSQLite::SQLite(), dbName)

aa.genome.exahustive <- RSQLite::dbGetQuery(harmonizedDb, 'SELECT * FROM exhaustiveClinicalAn
aa.genome.representative <- RSQLite::dbGetQuery(harmonizedDb, 'SELECT * FROM representativeCl
aa.genome.representative.oncokb <- RSQLite::dbGetQuery(harmonizedDb, 'SELECT * FROM represent
```

OncoKB and MOA/CIViC concordance

Summary of columns and proportion of null entries

```
`summarise()` has grouped output by 'annotationMatchGenomicCoord'. You can
override using the `.groups` argument.
```

annotationMatchGenomicCoord	annotationMatchAAChange	counts_and_percent
0	1	6486 (1.1%)
1	0	12511 (2%)
1	1	596217 (96.9%)