**Package ‘Famdenovo’**

**Type** Package
**Title** TP53 mutation carrier estimation
**Version** 0.1.1
**Date** 2020-05-16
**Author** Fan Gao, Wenyi Wang, Xuedong Pan
**Maintainer** Matthew Montierth <mdmontierth@mdanderson.org>, Wenyi Wang <wwang7@mdanderson.org>
**Description** This is a software that predicts the de novo status of deleterious germline mutation(s) in Mendelian diseases based on family history. Currently we apply Famdenovo to de novo TP53 mutations. For new functions for other genes, e.g., BRCA1/2, please check <http://https://bioinformatics.mdanderson.org/public-software/famdenovo/>.

**License** GPL-2

**How to install**

If you have not installed "Famdenovo", download it from "http://bioinformatics.mdanderson.org/main/Famdenovo" and install it from local source in your R console by typing:

install.packages("where\_you\_saved\_the\_file/Famdenovo\_0.1.1.tar.gz", repos = NULL, type = "source")

Or, install from GitHub:

library(devtools)

install\_github("wwylab/Famdenovo\_0.1.1")

**How to use**

**Step 1**. Load package

library(Famdenovo)

**Step 2**. Call "Famdenovo()" function

Famdenovo(family, cancer, mutation, person.id, gene = “TP53”)

**Format of the input files**

Famdenovo requires three data sets as input: family, cancer, and mutation information..

**family: Family Information Data**

The input should be a data frame. The family data should include the following columns with the corresponding column names:

*id*: Index of the person. All individuals should from one family.

*fid*: Index of the person's father. If the individual is the founder of the pedigree, set it as NA.

*mid*: Index of the person's mother. If the individual is the founder of the pedigree, set it as NA.

*gender*: Gender of the person. 0 - female; 1 - male

*age*: Age of the person. If the individual is alive, it is their current age. Otherwise, set it as the age at death.

Example Code:

Data(“TP53.test1.family”)

TP53.test1.family



Figure 1. An example of a pedigree structure



Figure 2. Example of the family data for a pedigree in Figure 1.

**cancer: Cancer Information Data**

The input should be a data frame. The cancer data should include the following columns with the corresponding column names:

*id*: index of the person

*cancer*.type: Type of the cancer. We divided all the cancers into 11 groups according to NCCCN Guidelines Version 1.2012 Li-Fraumeni Syndrome criteria. Check "LFSpro.cancer.type" for details.

*diag.age*: Age when the individual was diagnosed with cancer.

Example Code:

Data(“TP53.test1.cancer”)

TP53.test1.cancer



Figure 3. An example of the cancer file

**mutation: Mutation Information Data**

The input should be a data frame. The mutation data should include the following columns with the corresponding column names:

*id*: Index of the person

*mut.state*: Mutation status of the person. "W" - wild type; "M" - mutated. Individuals who are not sequenced are not included in the mutation information data.

Example Code:

Data(“TP53.test1.mutation”)

TP53.test1.mutation



Figure 4. An example of the mutation file

**person.id**

The input should be either character string(s) or numerical value(s) of the person(s) you want to analyze.

**gene**

The input should be character string(s). The default value is "TP53". We will add other genes in the future. For new functions for other genes, e.g., BRCA1/2, please check https://bioinformatics.mdanderson.org/public-software/famdenovo/

**Format of the output file**

The output is the probability of any deleterious *TP53* mutation being de novo, one mutation carrier per row. Each row contains three elements: "family id", "individual id" and "prob.denovo".

Here is an example:

 [1] "The following ids are not carriers: 1002, 1003"

id prob.denovo

1 2201 0.0001205471

2 2203 0.0001105599

3 3201 0.0126455556