

# Package ‘MiRKATS’

May 30, 2016

**Type** Package

**Title** Microbiome Regression-based Kernel Association Test for Survival  
(MiRKAT-S)

**Version** 1.0

**Date** 2016-04-22

**Author** Anna Plantinga <aplantin@uw.edu>, Michael Wu <mcwu@fhcrc.org>

**Maintainer** Anna Plantinga <aplantin@uw.edu>

**Depends** survival, CompQuadForm, GUniFrac

**Description** Community level test for association between microbiome composition and survival outcomes via user-defined kernel or distance matrices. Recommended distances include the UniFrac and generalized UniFrac distances and the Bray-Curtis dissimilarity metric.

**License** GPL (>= 2)

**NeedsCompilation** no

## R topics documented:

D2K . . . . .	1
MiRKATS . . . . .	2

---

D2K	<i>Construct kernel matrix from distance metric</i>
-----	---

---

### Description

Construct kernel matrix from distance matrix (matrix of pairwise distances) for microbiome data.

### Usage

D2K(D)

**Arguments**

D An n by n matrix giving pairwise distances or dissimilarities, where n is the sample size.

**Details**

The kernel matrix is constructed as  $K = -(I - 11'/n)D^2(I - 11'/n)/2$ , where D is the pairwise distance matrix, I is an identity matrix and 1 is a vector of 1.  $D^2$  represents element wise square. To ensure that  $K$  to be positive semi-definite, a positive semi-definiteness correction is conducted.

**Value**

An n by n kernel or similarity matrix corresponding to the distance matrix.

**Author(s)**

Ni Zhao

**References**

Zhao, Ni, et al. "Testing in microbiome-profiling studies with MiRKAT, the microbiome regression-based kernel association test." *The American Journal of Human Genetics* 96.5 (2015): 797-807.

**Examples**

```
#####
require(GUniFrac)
# Load in data and create a distance metric
data(throat.tree)
data(throat.otu.tab)
unifrac = GUniFrac(throat.otu.tab, throat.tree, alpha = c(1))$unifrac
D1 = unifrac[,,"d_1"]
# Function
K = D2K(D1)
```

---

MiRKATS

*Microbiome Regression-based Kernel Association Test for Survival*

---

**Description**

Community level test for association between microbiome composition and survival outcomes (right-censored time-to-event data) using kernel matrices to compare similarity between microbiome profiles with similarity in survival times.

**Usage**

```
MiRKATS(kd, distance = FALSE, obstime, delta, covar = NULL, beta = NULL,
perm = FALSE, nperm = 1000)
```

**Arguments**

kd	A numeric n by n kernel matrix or matrix of pairwise distances/dissimilarities (where n is the sample size).
distance	Logical, indicating whether kd is a distance matrix (default = FALSE).
obstime	A numeric vector of follow-up (survival/censoring) times.
delta	Event indicator: a vector of 0/1, where 1 indicates that the event was observed for a subject (so "obstime" is survival time), and 0 indicates that the subject was censored.
covar	A vector or matrix of numeric covariates, if applicable (default = NULL).
beta	A vector of coefficients associated with covariates. If beta is NULL and covariates are present, coxph is used to calculate coefficients (default = NULL).
perm	Logical, indicating whether permutation should be used instead of analytic p-value calculation (default=FALSE). Not recommended for sample sizes of 100 or more.
nperm	Integer, number of permutations used to calculate p-value if perm==TRUE (default=1000).

**Details**

obstime, delta, and covar should all have n rows, and the kernel or distance matrix should be a single n by n matrix. If a distance matrix is entered (so distance=TRUE), a kernel matrix will be constructed from the distance matrix.

Missing data is not permitted. Please remove individuals with missing data on y, X or in the kernel or distance matrix prior to using the function.

The Efron approximation is used for tied survival times.

**Value**

P-value obtained using small sample correction

**Author(s)**

Anna Plantinga

**References**

Plantinga, A., Zhan, X., Zhao, N., Chen, J., Jenq, R., and Wu, M.C. MiRKAT-S: a distance-based test of association between microbiome composition and survival times. (In preparation)

Zhao, N., Chen, J., Carroll, I. M., Ringel-Kulka, T., Epstein, M.P., Zhou, H., Zhou, J. J., Ringel, Y., Li, H. and Wu, M.C. (2015). Microbiome Regression-based Kernel Association Test (MiRKAT). *American Journal of Human Genetics*, 96(5):797-807

Chen, J., Chen, W., Zhao, N., Wu, M~C. and Schaid, D~J. (2016) Small Sample Kernel Association Tests for Human Genetic and Microbiome Association Studies. 40:5-19. doi: 10.1002/gepi.21934

Efron, B. (1977) "The efficiency of Cox's likelihood function for censored data." *Journal of the American statistical Association* 72(359):557-565.

Davies R.B. (1980) Algorithm AS 155: The Distribution of a Linear Combination of chi-2 Random Variables, *Journal of the Royal Statistical Society Series C*, 29:323-333.

### Examples

```
#####  
# Generate data  
require(GUniFrac)  
set.seed(1)  
  
# Throat microbiome data  
data(throat.tree)  
data(throat.otu.tab)  
unifrac = GUniFrac(throat.otu.tab, throat.tree, alpha = c(1))$unifrac  
D1 = unifrac[,,"d_1"] # 60 subjects  
  
# Covariates and outcomes  
X <- matrix(rnorm(120), nrow=60)  
S <- rexp(60, 3)  
C <- rexp(60, 1)  
D <- (S<=C) # event indicator  
U <- pmin(S, C) # observed follow-up time  
  
MiRKATS(kd = D1, distance = TRUE, obstime = U, delta = D, covar = X, beta = NULL)
```

# Index

D2K, [1](#)

MiRKATS, [2](#)