

The R package *trioClasses* for definition of the class *FamilyExperiment*, an extension of *SummarizedExperiment*, for use in trio based analyses of genetic data.

Samuel G. Younkin

March 13, 2013

1 Packages & Data

```
> library("trioClasses")
> data("sample")
> data("8q24-european-all.sm")
```

2 SummarizedExperiment

```
> se <- SummarizedExperiment(assays = SimpleList(geno = t(sm)),
  colData = col.DF, rowData = gr)
```

3 Pedigree

```
> ped <- PedClass(ped.DF)
```

4 FamilyExperiment

```
> (ste <- FamilyExperiment(se, pedigree = ped))
```

```
class: FamilyExperiment
dim: 511 968
exptData(0):
assays(1): geno
rownames(511): chr14:54383433 chr14:54383470 ...
  chr14:54444385 chr14:54444750
rowData values names(0):
colnames(968): H_ME-DS10776_2-DS10776_2
  H_ME-DS10776_3-DS10776_3 ... H_ME-DS11313_3-DS11313_3
  H_ME-DS11313_1-DS11313_1
colData names(1): id
pedigree(4139): famid id fid mid sex dx
complete trios(320):
```

5 Methods

5.1 ScanTrio

```
> (ste.rare <- ste[!(MAF(ste) >= 0.01 | is.na(MAF(ste))))])

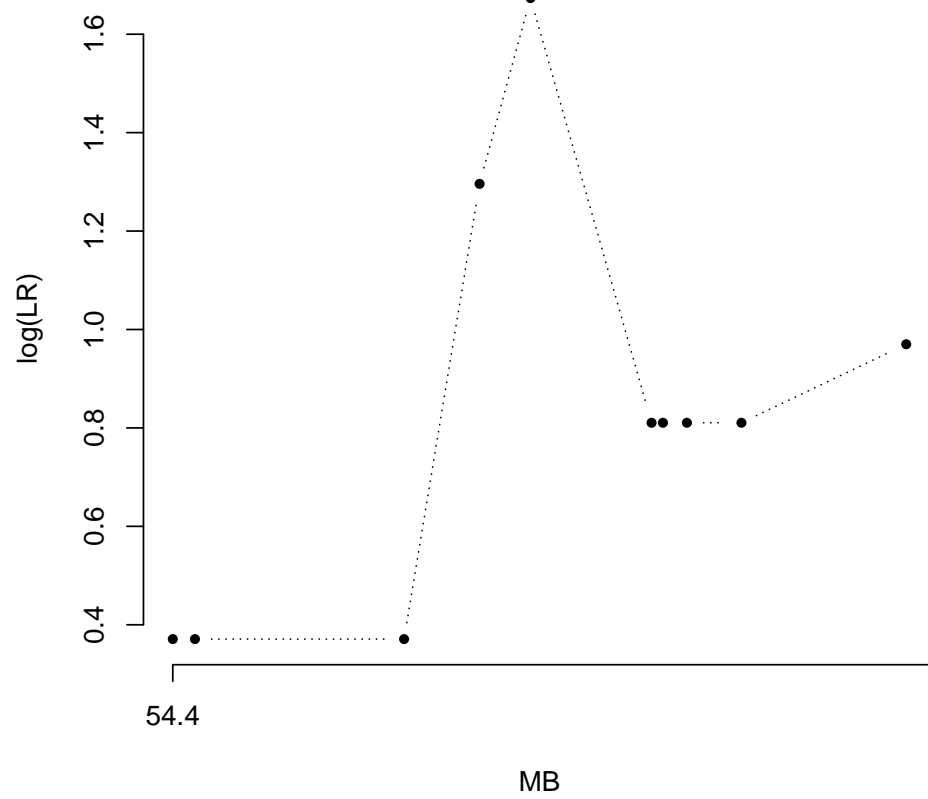
class: FamilyExperiment
dim: 371 968
exptData(0):
assays(1): geno
rownames(371): chr14:54383433 chr14:54383819 ...
               chr14:54444016 chr14:54444385
rowData values names(0):
colnames(968): H_ME-DS10776_2-DS10776_2
               H_ME-DS10776_3-DS10776_3 ... H_ME-DS11313_3-DS11313_3
               H_ME-DS11313_1-DS11313_1
colData names(1): id
pedigree(4139): famid id fid mid sex dx
complete trios(320):

> window <- rowData(ste)[1:10] + 1000
> system.time(scan.trio <- ScanTrio(object = ste.rare,
  window = window, block = range(rowData(ste.rare))))

      user  system elapsed
1.000   0.004   1.005

> scan.trio

DataFrame with 10 rows and 7 columns
      lr  minor.in  major.in  minor.out  major.out  mendel.in
<numeric> <integer> <integer> <integer> <integer> <integer>
1    1.449112      3        6      327      355        0
2    1.449112      3        6      327      355        0
3    1.449112      3        6      327      355        0
4    3.654525      3        9      327      352        0
5    5.330066      3       10      327      351        0
6    2.249029      8       15      322      346        0
7    2.249029      8       15      322      346        0
8    2.249029      8       15      322      346        0
9    2.249029      8       15      322      346        0
10   2.638015     14       24      316      337        0
mendel.out
<integer>
1          0
2          0
3          0
4          0
5          0
6          0
7          0
8          0
```



9	0
10	0

```
> save(scan.trio, file = "../data/scan-trio.RData")
```

5.2 Holger-style Genotype Matrix

Coercion to matrix for *trio*

```
> geno <- as(ste, "matrix")
> aTDT(geno[, 1:5])
```

```
$stat
[1] 24.50000 15.11429 27.12903 11.30769 10.31429
```

```
$pval
[1] 7.430984e-07 1.011946e-04 1.903182e-07 7.718664e-04 1.320044e-03
```

```
$transMinor
[1] 2 6 1 9 8
```

```
$transMajor
[1] 30 29 30 30 27
```

Or apply the aTDT method to the FamilyExperiment directly.

```
> aTDT(ste[1:5])
```

```
$stat
[1] 18.000000 16.780822 3.813559 7.736842 6.391304
```

```
$pval
[1] 2.209050e-05 4.195519e-05 5.083931e-02 5.410534e-03 1.146807e-02
```

```
$transMinor
[1] 18 19 22 18 24
```

```
$transMajor
[1] 54 54 37 39 45
```

A Count of Transmission of Variants (*TransCount()*)

A.1 Window

```
> window <- gr[100] + 10000
> window2 <- gr[100] + 1000
> block <- range(rowData(ste))

> TransCount(ste, window)
```

A.2 Not in window, but in block

```
> TransCount(ste, setdiff(block, window))
```

A.3 In the whole block

```
> TransCount(ste, block)
```

A.4 In both windows

```
> TransCount(ste, GRangesList(window, window2))
```

A.5 Not “In both windows,” but in block

```
> TransCount(ste, GRangesList(setdiff(block, window), setdiff(block,
  window2)))
```

B Count of Transmission of Rare Variants (*TransCount()*)

B.1 In both windows

B.2 Not “In both windows,” but in block

C Classes

C.1 SnpMatrix, DataFrame, GRanges

The four key ingredients are the SNP matrix, the pedigree information as a DataFrame, position of the SNPs given by a GRanges object, and covariate data given as a DataFrame.

```
> sm
> ped.DF
> gr
```

C.2 SummarizedExperiment

We combine three of the key ingredients when we create the SummarizedExperiment object.

```
> se <- SummarizedExperiment(assays = SimpleList(geno = t(sm)),
  colData = col.DF, rowData = gr)
```

C.3 FamilyExperiment

Now, we include the pedigree information as an object of class `PedClass`. We keep `PedClass` independent of `FamilyExperiment` for flexibility.

```
> ped <- PedClass(ped.DF)
> ste <- FamilyExperiment(se, pedigree = ped)
```

Here is the show method.

```
> ste
```

And now we verify that it is indeed an extension of `SummarizedExperiment`.

```
> getClass("FamilyExperiment")
```

C.4 PedClass

Now we investigate the pedigree slot of the `FamilyExperiment` object.

```
> class(pedigree(ste))
> getClass("PedClass")
> pedigree(ste)
```

C.5 geno accessor

```
> class(geno(ste))
> getClass("SnpMatrix")
> geno(ste)
```

C.6 RowData

```
> rowData(ste)
```

C.7 ColData

D Validity