

readMAPPER: Interface to MAPPER TFBS database

Alberto Riva and Heike Sichtig (U. of Florida)

October 1, 2012

1 Introduction

The **RMAPPER** package provides an interface to the MAPPER database of predicted Transcription Factor Binding Sites (TFBS). The MAPPER database (<http://genome.ufl.edu/mapper/>) contains predicted binding sites for over 800 Transcription Factors in the promoter regions of all human, mouse and Drosophila genes. This package allows you to query the database according to the desired parameters and to download the resulting set of TFBS into an R object.

2 readMAPPER demonstration

First, the user should load the RMAPPER library.

```
R> library(RMAPPER)
```

Then, examples can be run using the example function or directly using *readMAPPER*. The functions `query` and `hits` are available to display the exact query and details of hits respectively. Note that the value for the "models" parameter should be a comma-separated list of model accession numbers. A table listing all transcription factors known to MAPPER with the corresponding model accession numbers is available at: <http://genome.ufl.edu/mapper/factors-table>. It can also be retrieved in R by evaluating:

```
R> rmapperFactorTable("*")
```

Use the following R command to get the model accession numbers known to MAPPER for a specific transcription factor, i.e. "AbaA".

```
R> rmapperFactorTable("AbaA")
```

Use the following R command to get the model accession numbers known to MAPPER for all transcription factors starting with "Ab...".

```
R> rmapperFactorTable("Ab")
```

2.1 Example 1

Retrieve predicted hits for model "M00027" (abaA) in the promoter of transcript NM_009696 (mouse Apoe). Only hits with a score above the 95th percentile will be returned.

```
R> mh = readMAPPER(gene="NM_009696", perc="95", models="M00027")
R> mh
```

```
http://genome.ufl.edu/mapper/ hits collection on query:
gene=NM_009696&perc=95&models=M00027
```

There were 1 hits.

	GeneID	Gene	Org	Accession	Model	Factor	Strand	Chrom	Start	End
1	11816	Apoe	Mm	NM_009696	M00027	AbaA	+	chr7	20284804	20284819
		Region		Score	Eval	Conserved		AlignModel		
1	Promoter	6.6	11		-	*->tct.cattcct.t.ct<-*				
				Match				Seq		
1	t+t	cattcc+	+ct		TTTcATTCCAaGgCT					

```
R> query(mh)
```

```
[1] "gene=NM_009696&perc=95&models=M00027"
```

```
R> hits(mh)
```

	GeneID	Gene	Org	Accession	Model	Factor	Strand	Chrom	Start	End
1	11816	Apoe	Mm	NM_009696	M00027	AbaA	+	chr7	20284804	20284819
		Region		Score	Eval	Conserved		AlignModel		
1	Promoter	6.6	11		-	*->tct.cattcct.t.ct<-*				
				Match				Seq		
1	t+t	cattcc+	+ct		TTTcATTCCAaGgCT					

2.2 Example 2

Retrieve predicted hits for models "M00027" (abaA) and "M00442" (ABF) in the promoter of transcript NM_009696 (mouse Apoe). Only hits with a score greater than 3 will be returned.

```
R> mh = readMAPPER(gene="NM_009696", score="3", models="M00027,M00442")
R> mh
```

```
http://genome.ufl.edu/mapper/ hits collection on query:
gene=NM_009696&score=3&models=M00027,M00442
```

There were 2 hits.

	GeneID	Gene	Org	Accession	Model	Factor	Strand	Chrom	Start	End
--	--------	------	-----	-----------	-------	--------	--------	-------	-------	-----

```

1 11816 Apoe Mm NM_009696 M00027 AbaA + chr7 20284804 20284819
5 11816 Apoe Mm NM_009696 M00442 ABF - chr7 20286147 20286161
    Region Score Eval Conserved          AlignModel
1 Promoter    6.6 11.0          - *->tct.cattcct.t.ct<-*
5 Promoter    7.1  7.8          - *->g.ccAcGTGgcc.ga<-*
                Match                      Seq
1   t+t cattcc+ +ct      TTTcCATTCCAaGgCT
5   + +cAcGTG c+ g+      CcTCACGTGTCAgGC

```

```
R> query(mh)
```

```
[1] "gene=NM_009696&score=3&models=M00027,M00442"
```

```
R> hits(mh)
```

```

    GeneID Gene Org Accession Model Factor Strand Chrom    Start      End
1 11816 Apoe Mm NM_009696 M00027 AbaA + chr7 20284804 20284819
5 11816 Apoe Mm NM_009696 M00442 ABF - chr7 20286147 20286161
    Region Score Eval Conserved          AlignModel
1 Promoter    6.6 11.0          - *->tct.cattcct.t.ct<-*
5 Promoter    7.1  7.8          - *->g.ccAcGTGgcc.ga<-*
                Match                      Seq
1   t+t cattcc+ +ct      TTTcCATTCCAaGgCT
5   + +cAcGTG c+ g+      CcTCACGTGTCAgGC

```

3 Help

You can issue the following R command for help.

```
R> rmapperHelp()
```