`xmapcore`: The base component for a framework of Genome analysis tools

Tim Yates, Crispin J Miller

December 7, 2010

Important

The `xmapcore` package has similar methods to the deprecated `exonmap` package. We have tried to keep things as close as possible, but some functions, parameters and returned results are crucially different. Please make sure you have read the migration section on page 2 to see how to migrate your code from `exonmap` to `xmapcore`.

Contents

1 Introduction

xmapcore provides access to the genome annotation in the X:Map database (`http://xmap.picr.man.ac.uk`). It is designed to be a base layer for interrogating genome and probeset annotation, and to be extensible, so that later packages for Microarray Expression analysis or SEQ data analysis can be added to extend the functionality of `xmapcore`.

The functions provided by `xmapcore` can be divided into five main groups:

1. Connect/disconnect to an instance of the database

2. Find genome features within a particular set of coordinates (e.g. `genes.in.range()`)

3. Find genome features by name (e.g. `gene.details()`)

4. Map between genome features (e.g. `gene.to.exon()`)

5. Handle Affymetrix Exon array annotation

1

For evaluated examples showing how these tasks are performed, please see the cookbook document, supplied along with this package

2 Migration from `exonmap`

2.1 Connecting to the database

In `exonmap`, when you wished to connect to the database, you called the function `xmapConnect()`. In `xmapcore` this function has been changed to `xmap.connect()`. This is to keep it in line with the formatting of function names across the package.

2.2 Working with features previously in the 'other features' database

In `exonmap` if you wanted to find mappings to EST genes, you were required to pass a `subset='est'` parameter to the method of interest. EST genes are now integrated into the same database with the other features, so now there are seperate EST functions for you to call (`probeset.to.est_gene`, etc)

2.3 Range query parameter order

The parameter order for `in.range` queries has changed. The four parameters required are now in the order:

1. `chr` – Chromosome name as a character vector

2. `start` – Numeric start location

3. `end` – Numeric stop location

4. `strand` – Numeric strand direction (-1 or 1)

2.4 Gene plots

In `exonmap` there exist the functions `gene.graph`, `gene.strip`, and `genePlot`. These do not exist in the `xmapcore` package. The closest comparable method is `xmap.gene.plot`, which can draw the transcripts and exons contained in a gene, and may be passed a function which performs the rendering for each exon. This function is covered in the cookbook pdf supplied with this package.