

Vignette for *RedeR*: bridging the gap between hierarchical network representation and functional analysis.

Mauro AA Castro, Xin Wang, Michael NC Fletcher,
Kerstin B Meyer and Florian Markowetz *

<http://genomebiology.com/2012/13/4/R29>

mauro.a.castro@gmail.com

florian.markowetz@cruk.cam.ac.uk

October 17, 2016

Contents

*Cancer Research UK - Cambridge Research Institute, Robinson Way Cambridge, CB2 0RE, UK.

1 Overview

RedeR is an R-based package combined with a Java application for dynamic network visualization and manipulation. It implements a callback engine by using a low-level R-to-Java interface to build and run common plugins. In this sense, *RedeR* takes advantage of *R* to run robust statistics, while the R-to-Java interface bridge the gap between network analysis and visualization.

RedeR is designed to deal with three key challenges in network analysis. Firstly, biological networks are modular and hierarchical, so network visualization needs to take advantage of such structural features. Secondly, network analysis relies on statistical methods, many of which are already available in resources like CRAN or Bioconductor. However, the missing link between advanced visualization and statistical computing makes it hard to take full advantage of R packages for network analysis. Thirdly, in larger networks user input is needed to focus the view of the network on the biologically relevant parts, rather than relying on an automatic layout function. *RedeR* is designed to address these challenges (additional information is available at Castro et al. [?]).

RedeR uses different strategies to link R to Java:

- Data interface: implements the callback engine to make calls from R via xml-rpc protocol. It sets *R* as client and *RedeR* as server.
- Graphic interface: it wraps R graphics into *RedeR* classes.

The design of the software is depicted from Figure ???. Complex graphs with many attributes can be transferred from-and-to *R* using *addGraph* and *getGraph* functions.

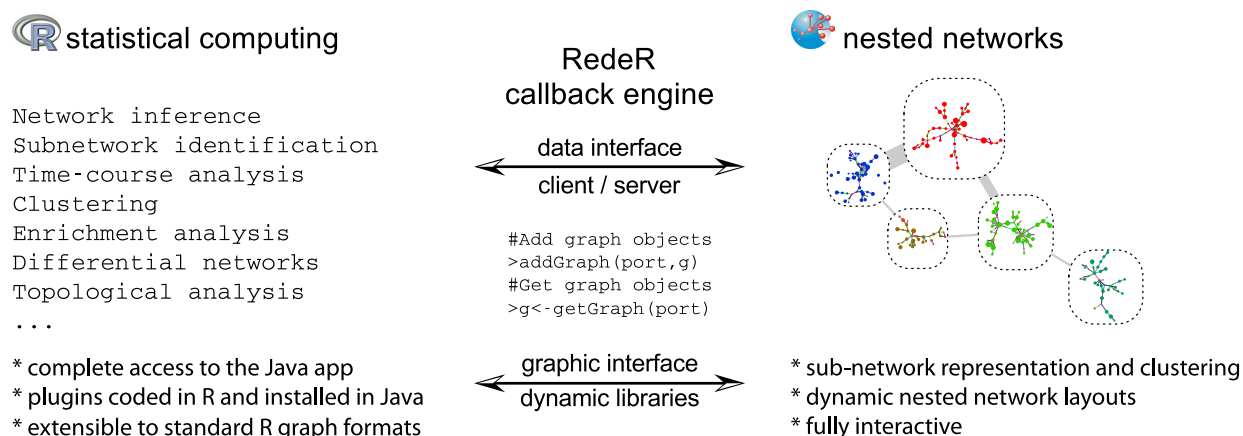


Figure 1: Schematic representation of *RedeR* calls. In the low-level interface, the Apache xmlrpc webserver[?] is used to link R to Java.

2 Quick start

2.1 Main callback methods

The first step is to build the server port, which will be required in all remote procedure calls. By default the constructor *RedPort* should set all details:

```
> library (RedeR)
> rdp <- RedPort ()
```

Next, invoke RedeR using the method *calld*:

```
> calld(rdp)
```

Within an active interface, then the method 'addGraph' can easily send R graphs to the application. For example, the following chunk adds an *igraph*[?] object (Fig.??):

```
> g1 <- graph.lattice(c(5,5,5))
> addGraph( rdp, g1, layout.kamada.kawai(g1) )
```

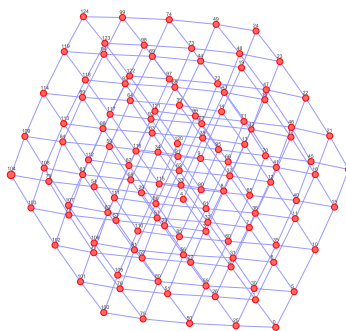


Figure 2: A toy example added to *RedeR* by the *addGraph* function.

Conversely, RedeR graphs can be transferred to R and wrapped in *igraph* objects:

```
> g2 <- getGraph(rdp)
> resetd(rdp)
```

The interface accepts additional graph attributes, as for example edge direction, edge width, edge weight, node shape, node size, node color etc. In *igraph* objects, vertex and edge attributes can be assigned as arbitrary R objects. In order to pass these extensible features to *RedeR* the attributes must be provided in a valid syntax.¹

Another strategy is to wrap graphs into containers and then send it to the Java app. Next, the subgraphs *g3* and *g4* are assigned to different nested structures (Fig.??).

```
> g3 <- barabasi.game(10)
> g4 <- barabasi.game(10)
> V(g3)$name<-paste("sn",1:10,sep="")
```

¹See *getGraph* and *addGraph* specification for additional details.

```

> V(g4)$name<-paste("sm",1:10,sep="")
> addGraph(rdp, g3, isNest =TRUE, gcoord=c(25,25), gscale=50)
> addGraph(rdp, g4, isNest =TRUE, gcoord=c(75,75), gscale=50)

```

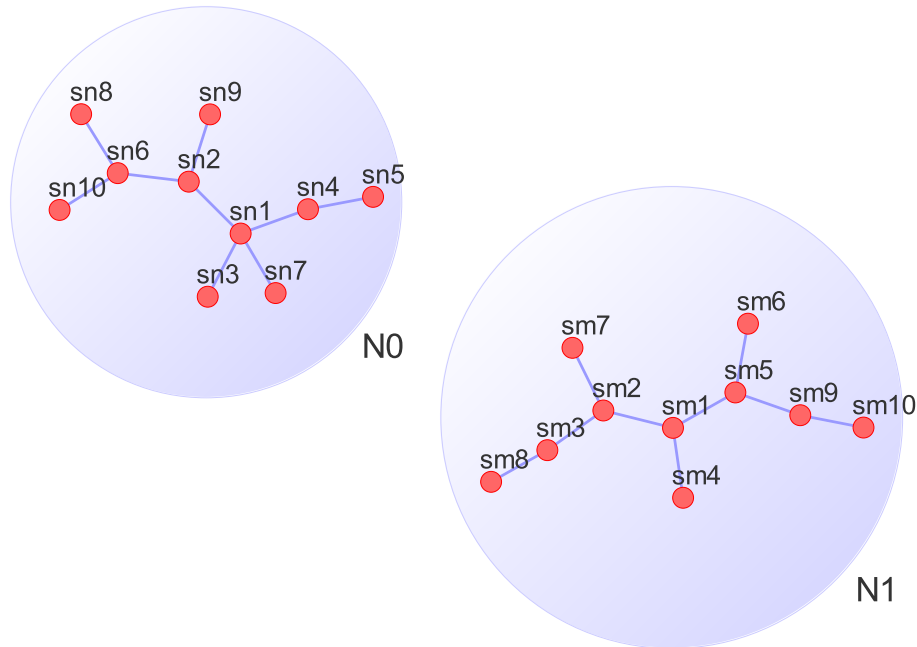


Figure 3: Nested graphs in *RedeR* using the command *addGraph*.

In this case, the subgraphs can be handled apart from each other. For example, the following chunk selects all nodes assigned to the container "N0" and then gets back the subgraph (the selection step can also be done interactively).

```

> selectNodes(rdp, "N0")
> g5 <- getGraph(rdp, status= "selected")
> resetd(rdp)

```

As a suggestion, try some RedeR features in the Java side (e.g. open samples s2 or s3 in the main panel and enjoy the dynamic layout options!).

2.2 Interactive work

The next chunk generates a scale-free graph according to the Barabasi-Albert model[?] and sends the graph to RedeR without any layout information.

```
> g6 <- barabasi.game(500)
> addGraph(rdp, g6, zoom=20)
```

Then using the "relax" options available in the app you can tune the graph layout as presented in Figure ??.

```
> relax(rdp,p2=400,p5=30,ps=T)
```

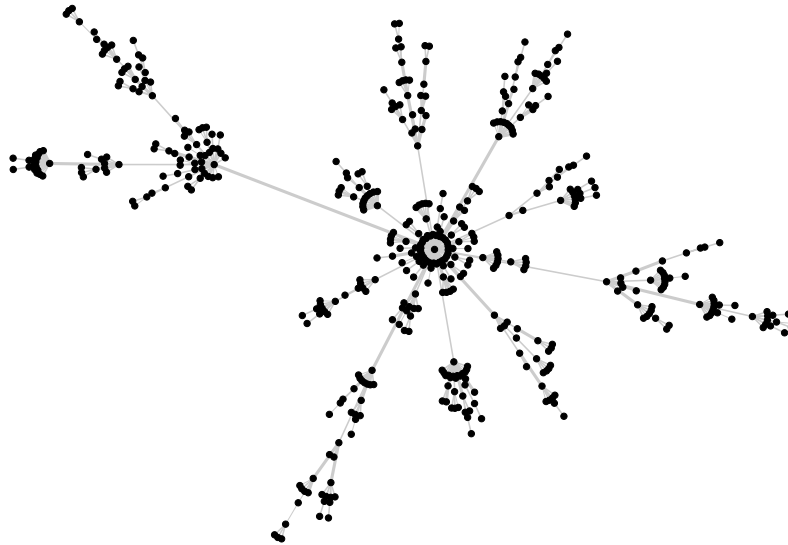
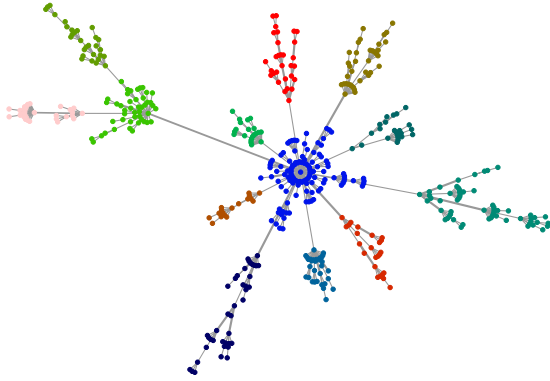


Figure 4: Scale-free graph according to the Barabasi-Albert model[?].

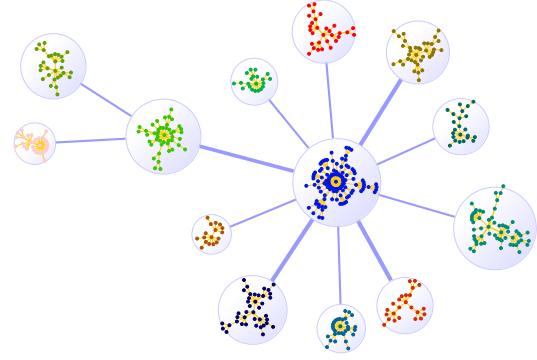
In Figure ?? the same graph is used to exemplify the community structure mapped by the edge-betweenness function available in RedeR. In Figure ?? these communities are nested to containers, which are objects of the same class of the nodes but with additional behaviors (Fig.??). You can build these containers either using *R* or *Java* functions (see options available in the *clustering* main menu and in the shortcuts of the nested objects, i.e., right-click a container).

For the next example you will need to reproduce the graph from Figure ?? (or any graph with containers) in the *RedeR* app, then select one of the communities, and run the chunk below: a simple degree distribution should be plotted in the R side.

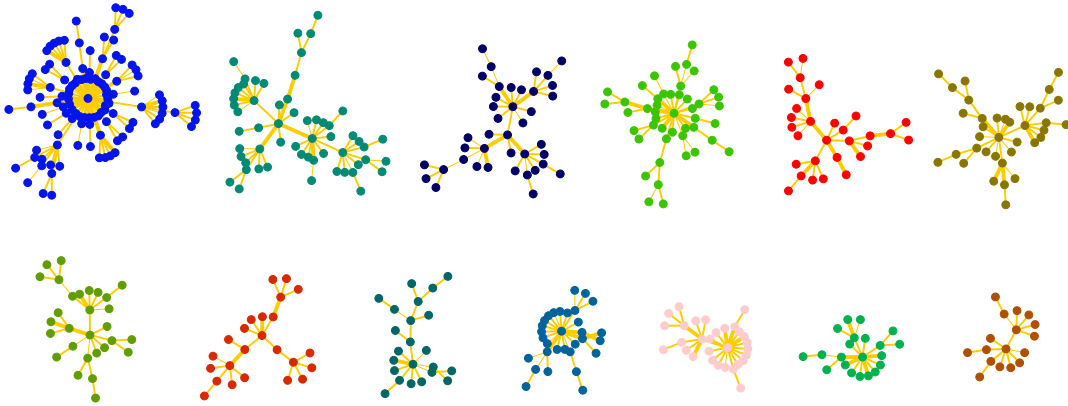
```
> g <- getGraph(rdp, status= "selected")
> if(vcount(g)>0)plot(degree.distribution(g), xlab = "k", ylab = "P(k)", pch=19)
> resetd(rdp)
```



(a) Communities



(b) Graphs into containers



(c) Subnetworks

Figure 5: Community structure: (a) subgraphs detected based on edge betweenness; (b) nested communities into containers; (c) subnetworks in hidden containers.

3 Workflow illustration

This section provides a sequence of steps that illustrates how users might integrate its own pre-processed data in a given network to visualize subgraphs and nested networks. Please refer to Castro et al. [?] for more details about the biological background and experimental design of each example.

3.1 Subgraphs

- 1 - start the app (i.e. 'callD' method).

- 2 - get a dataframe and an interactome:

```
> data(ER.limma)
> data(hs.inter)
> dt <- ER.limma
> gi <- hs.inter
```

- 3 - extract a subgraph and set attributes to RedeR (i.e. logFC from t3-t0 contrast):

```
> gt3 <- subg(g=gi, dat=dt[dt$degenes.t3!=0,], refcol=1)
> gt3 <- att.setv(g=gt3, from="Symbol", to="nodeAlias")
> gt3 <- att.setv(g=gt3, from="logFC.t3", to="nodeColor", breaks=seq(-2,2,0.4), pal=2)
```

ps. some genes will not be found in the interactome!

- 4 - extract another subgraph and set attributes to RedeR (i.e. logFC from t6-t0 contrast):

```
> gt6 <- subg(g=gi, dat=dt[dt$degenes.t6!=0,], refcol=1)
> gt6 <- att.setv(g=gt6, from="Symbol", to="nodeAlias")
> gt6 <- att.setv(g=gt6, from="logFC.t6", to="nodeColor", breaks=seq(-2,2,0.4), pal=2)
```

- 5 - extract another subgraph and set attributes to RedeR (i.e. logFC from t12-t0 contrast):

```
> gt12 <- subg(g=gi, dat=dt[dt$degenes.t12!=0,], refcol=1)
> gt12 <- att.setv(g=gt12, from="Symbol", to="nodeAlias")
> gt12 <- att.setv(g=gt12, from="logFC.t12", to="nodeColor", breaks=seq(-2,2,0.4), pal=2)
```

- 6 - add subgraphs to the app (Fig.??):

```
> addGraph(rdp, gt3, gcoord=c(10,25), gscale=20, isNest=TRUE, theme='tm1', zoom=30)
> addGraph(rdp, gt6, gcoord=c(20,70), gscale=50, isNest=TRUE, theme='tm1', zoom=30)
> addGraph(rdp, gt12, gcoord=c(70,55), gscale=80, isNest=TRUE, theme='tm1', zoom=30)
```

- 7 - nest subgraphs (i.e. overlap time-series):

```
> nestNodes(rdp, nodes=V(gt3)$name, parent="N1", theme='tm2')
> nestNodes(rdp, nodes=V(gt6)$name, parent="N2", theme='tm2')
> nestNodes(rdp, nodes=V(gt3)$name, parent="N4", theme='tm3')
```

- 8 - assign edges to containers

```
> mergeOutEdges(rdp)
```

- 9 - relax the network

```
> relax(rdp, 50, 400)
```

- 10 - add a color legend (other types are available):²

```
> scl <- gt3$legNodeColor$scale
> leg <- gt3$legNodeColor$legend
> addLegend.color(rdp, colvec=scl, labvec=leg, title="node color (logFC)")
```

- 11 - select a gene:

```
> selectNodes(rdp, "RET")
```

- 12 - reset graph:

```
> resetd(rdp)
```

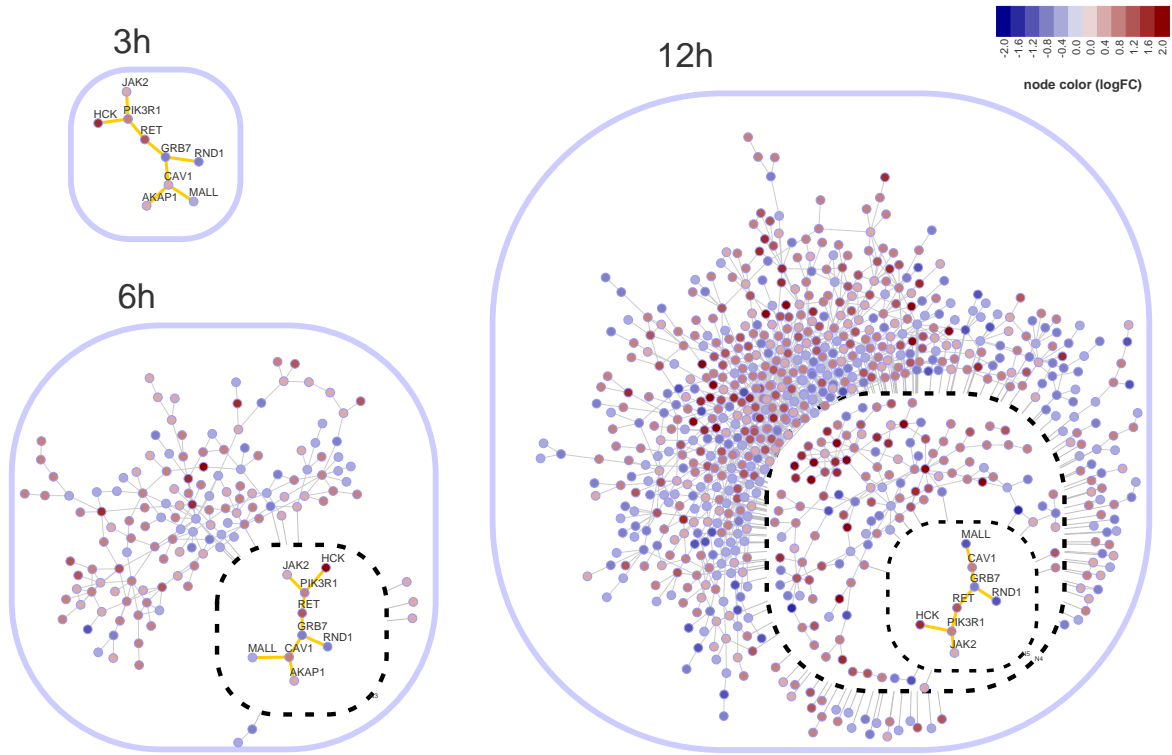


Figure 6: Nested subnetworks. This graph shows genes differentially expressed in estrogen-treated MCF-7 cells at 3, 6 and 12 h (relative to 0 h). The insets correspond to the overlap between consecutive time points (adapted from Castro et al.[?]).

²Legends are set only via command line in the current version.

3.2 Nested networks and clustering

- 1 - get a dataframe and an igraph object:

```
> data(ER.deg)
> dt <- ER.deg$dat
> sg <- ER.deg$ceg
```

- 2 - map the dataframe to the graph:

```
> sg <- att.mapv(sg, dat=dt, refcol=1)
```

- 3 - set attributes to RedeR (i.e. gene symbols and two available numeric data):

```
> sg <- att.setv(sg, from="Symbol", to="nodeAlias")
> sg <- att.setv(sg, from="logFC.t3", to="nodeColor", breaks=seq(-1,1,0.2), pal=2)
> sg <- att.setv(sg, from="ERbdist", to="nodeSize", nquant=10, isrev=TRUE, xlim=c(5,40,1))
```

- 4 - add graph to the app (Fig.??):

```
> addGraph(rdp,sg)
```

- 5 - compute a hierarchical clustering using standard R functions:

```
> hc <- hclust(dist(get.adjacency(sg, attr="weight")))
```

- 6 - map the hclust object onto the network (pvclust objects are also compatible!):

```
> nesthc(rdp,hc, cutlevel=3, nmemb=5, cex=0.3, labels=V(sg)$nodeAlias)
```

...at this point nested objects from the network should also be mapped to a dendrogram! different levels of the nested structure can be set by the nesthc method. Additionally, significant clusters can be computed by the pvclust package, which is already compatible with RedeR interface.

- 7 - assign edges to containers:

```
> mergeOutEdges(rdp,nlev=2)
```

- 8 - relax the network:

```
> relax(rdp)
```

- 9 - add color and size legends:³

```
> scl <- sg$legNodeColor$scale
> leg <- sg$legNodeColor$legend
> addLegend.color(rdp, colvec=scl, labvec=leg, title="diff. gene expression (logFC)")
```

```
> scl <- sg$legNodeSize$scale
> leg <- sg$legNodeSize$legend
> addLegend.size(rdp, sizevec=scl, labvec=leg, title="bd site distance (kb)")
```

- 10 - reset graph:

```
> resetd(rdp)
```

³Legends are set only via command line in the current version.

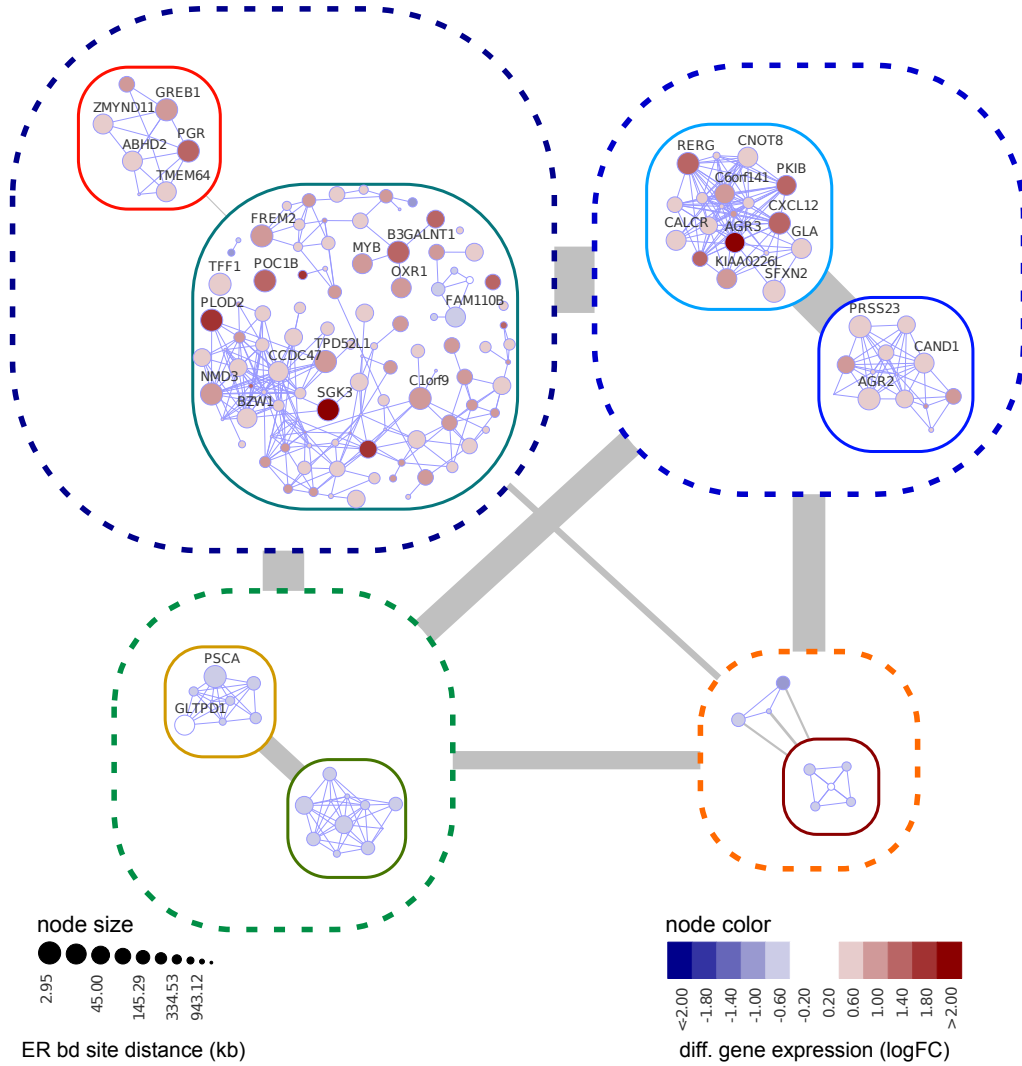


Figure 7: Hierarchical networks. This graph is obtained by superimposing a dendrogram onto the corresponding co-expression gene network (adapted from Castro et al.[?]).

4 Installation

R package

The RedeR package is freely available from Bioconductor at <http://bioconductor.org/packages/release/bioc/html/RedeR.html>.

Java application

The RedeR jar file is already included in the R package and, as usual, to run Java applications your system should have a copy of the JRE (Java Runtime Environment, version ≥ 6).

5 Session information

```
R version 3.3.1 (2016-06-21)
Platform: x86_64-apple-darwin13.4.0 (64-bit)
Running under: OS X 10.9.5 (Mavericks)

attached base packages:
[1] stats      graphics  grDevices  utils      datasets  methods   base

other attached packages:
[1] RedeR_1.22.0 igraph_1.0.1

loaded via a namespace (and not attached):
[1] Matrix_1.2-7.1  magrittr_1.5    tools_3.3.1     RCurl_1.95-4.8
[5] grid_3.3.1      bitops_1.0-6    lattice_0.20-34 pvclust_2.0-0
[9] XML_3.98-1.4
```