

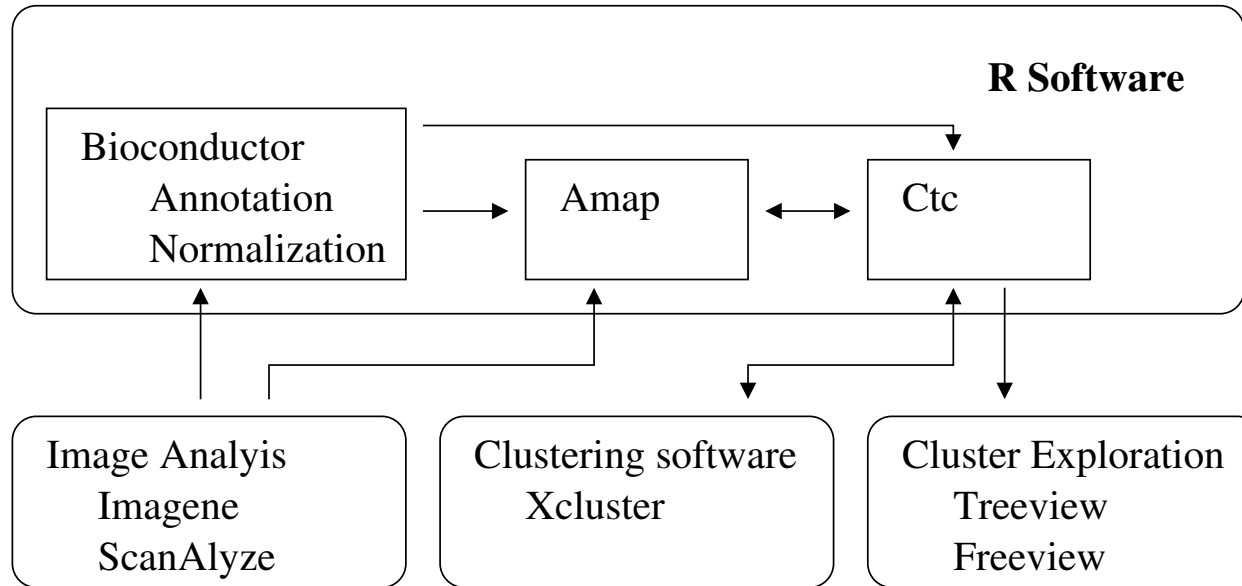
Amap & ctc: R packages for gene clustering

Antoine Lucas and Sylvain Jasson

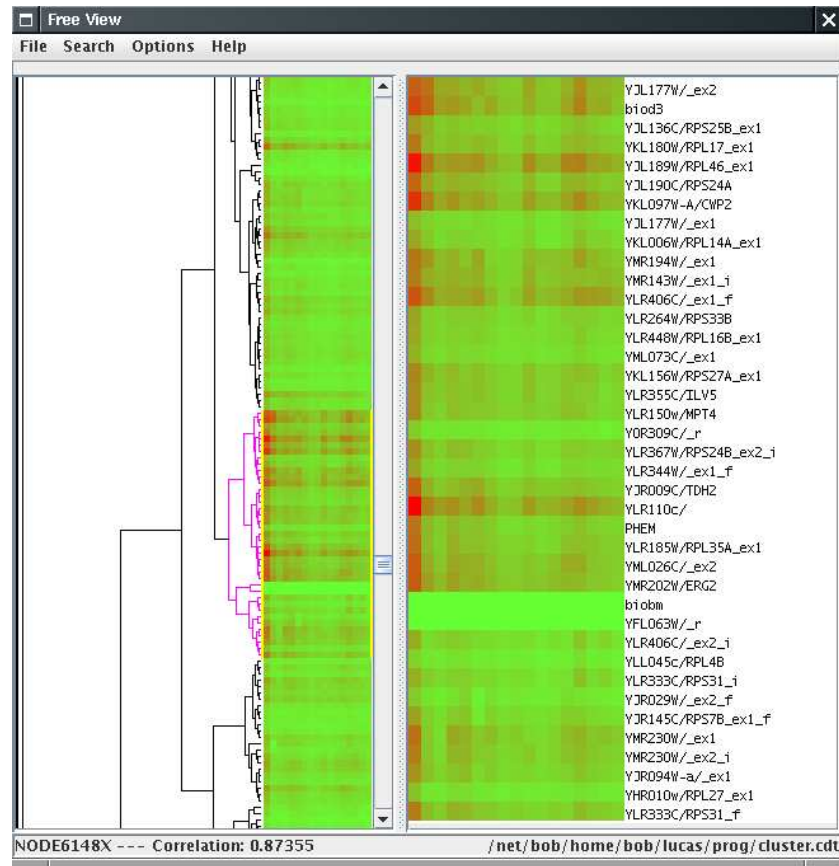
Aim

- Clustering large datasets
 - easy use of Xcluster → *ctc*
 - optimized clustering functions → *amap*
- Graphical and convenient visualization → *ctc*
 - export to Freeview / TreeView softwares
 - export to Newick file format
- Robust statistics tools → *amap*
- Easily used as a web interfaces (examples provided)
- Compatibility with Eisen software suite → *ctc*.

Amap and ctc usage



Visualization



Window of results with Freeview.

Benchmark of R packages

Sample: 6600 genes under 17 conditions, hierarchical clustering.

Package	Function	Time	Memory
ctc	xcluster*	33.9 sec	29 Mo
amap	hcluster	17.6 sec	189 Mo
amap	hclusterpar	11.7 sec	189 Mo
stats	hclust	19 min 56	686 Mo
cluster	agnes	44 min 17	190 Mo

* Xcluster does not compute exhaustive distance while agglomerate clusters.

We used average link and Euclidean distance for clustering. Computer: a dual Xeon processor server, with 4Go Memory RAM. Time and memory are given by system command "time". R version 2.0.1