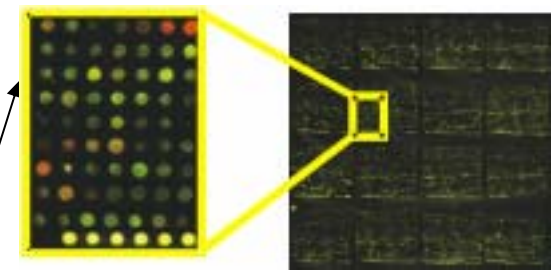
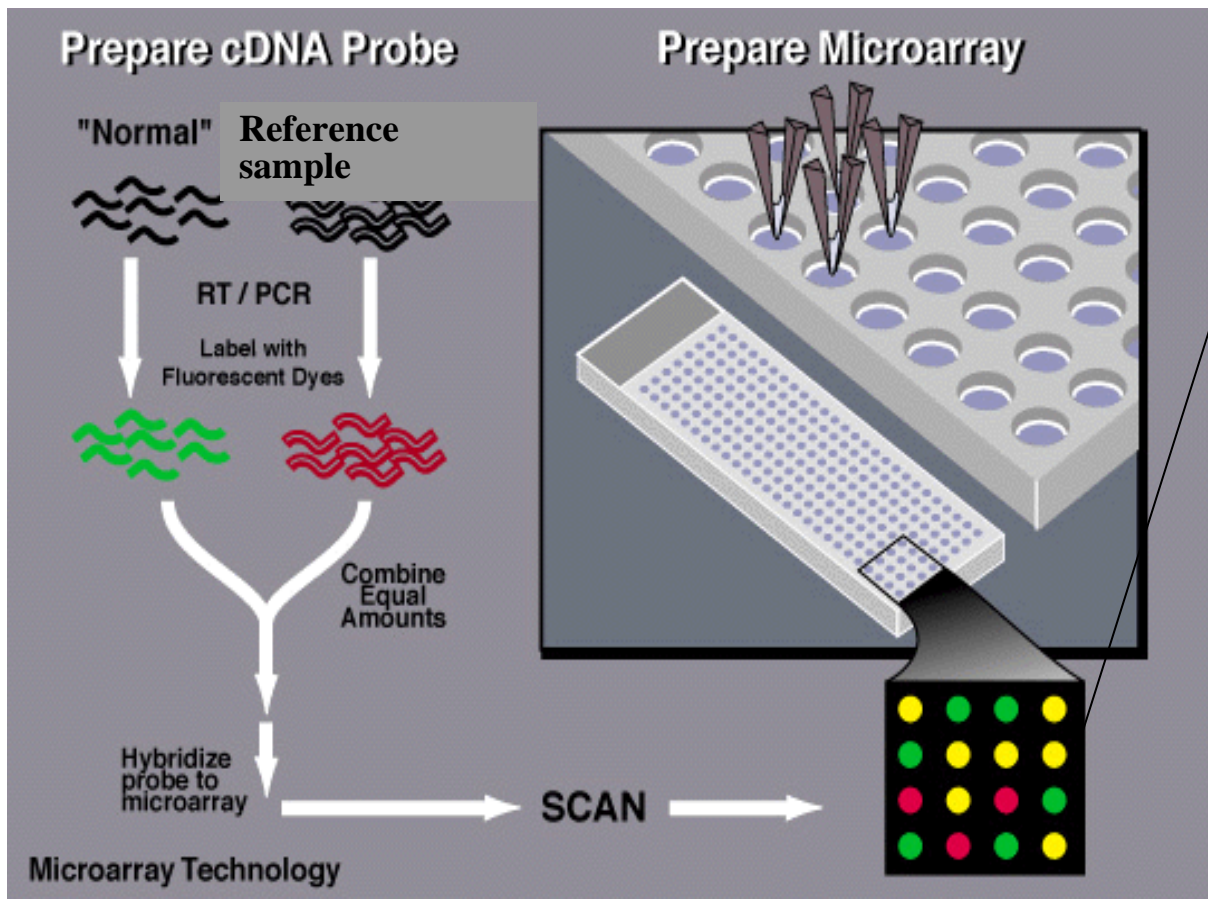


Transcriptomics

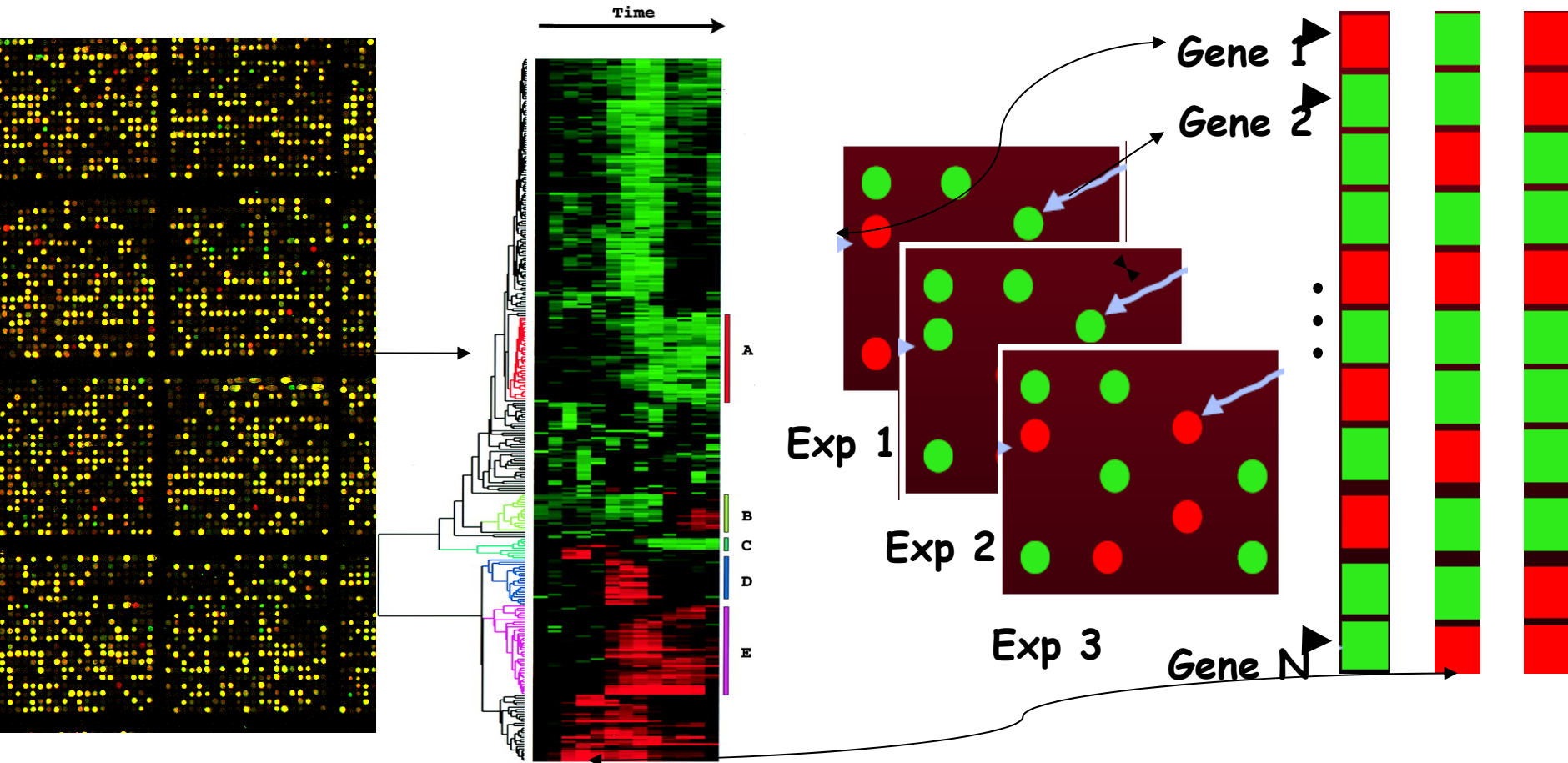
The genome-wide study of gene expression levels, measured in the variation and presence of mRNA

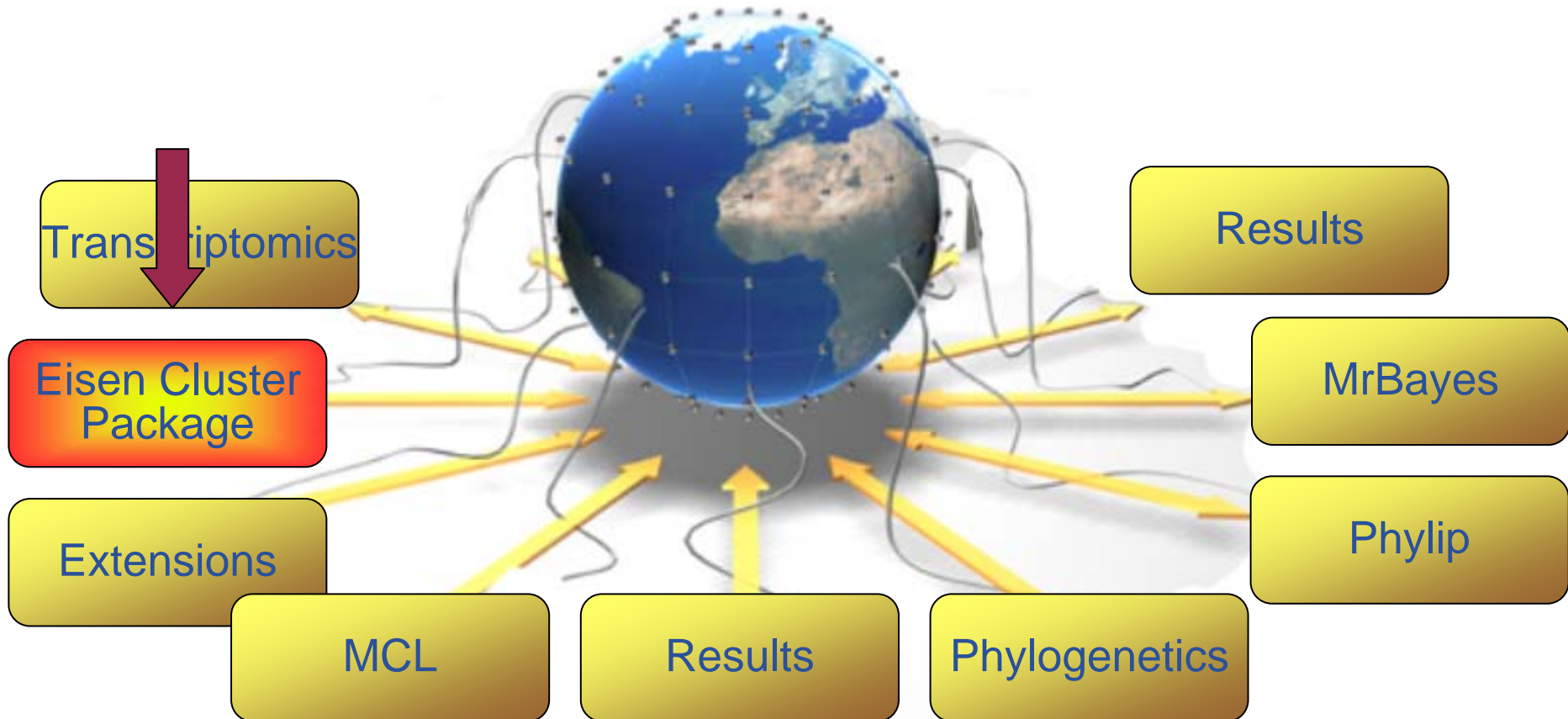


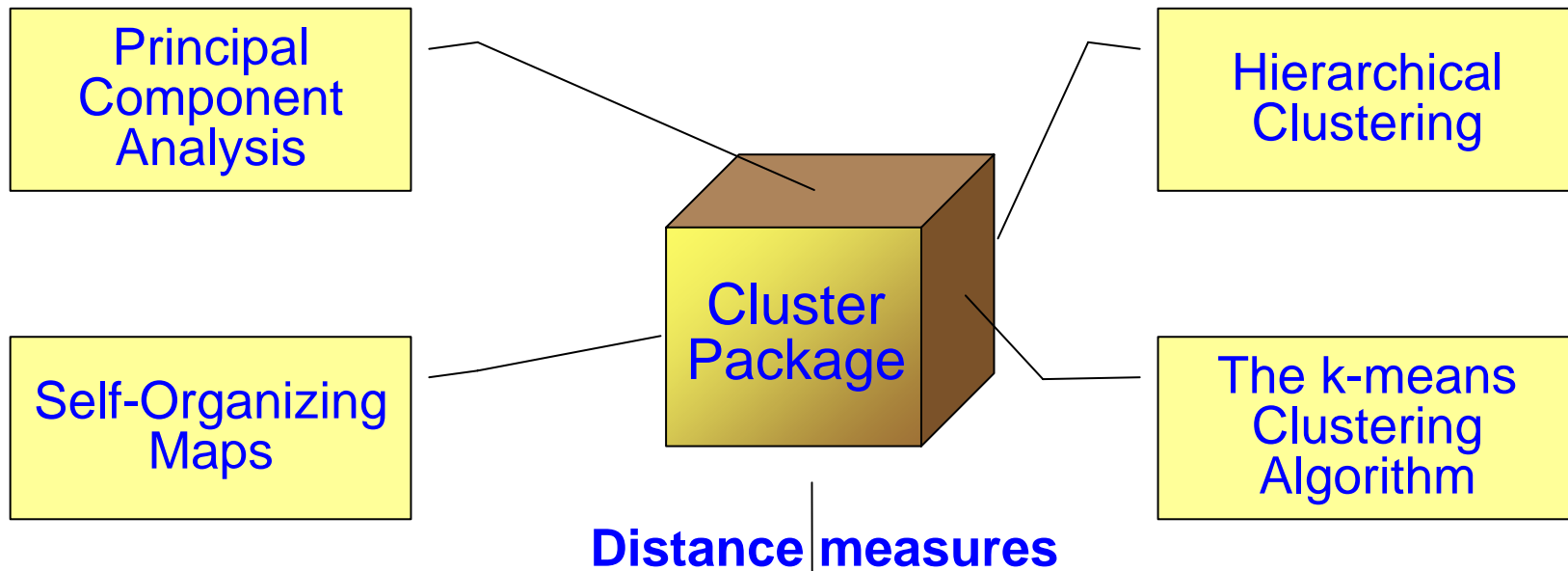


Calculate intensities and ratios

YORF	0 minutes	30 minutes	1 hour
YAL001C	1	1.3	2.4
YAL002W	0.9	0.8	0.7
YAL003W	0.8	2.1	4.2
YAL005C	1.1	1.3	0.8
YAL010C	1.2	1	1.1



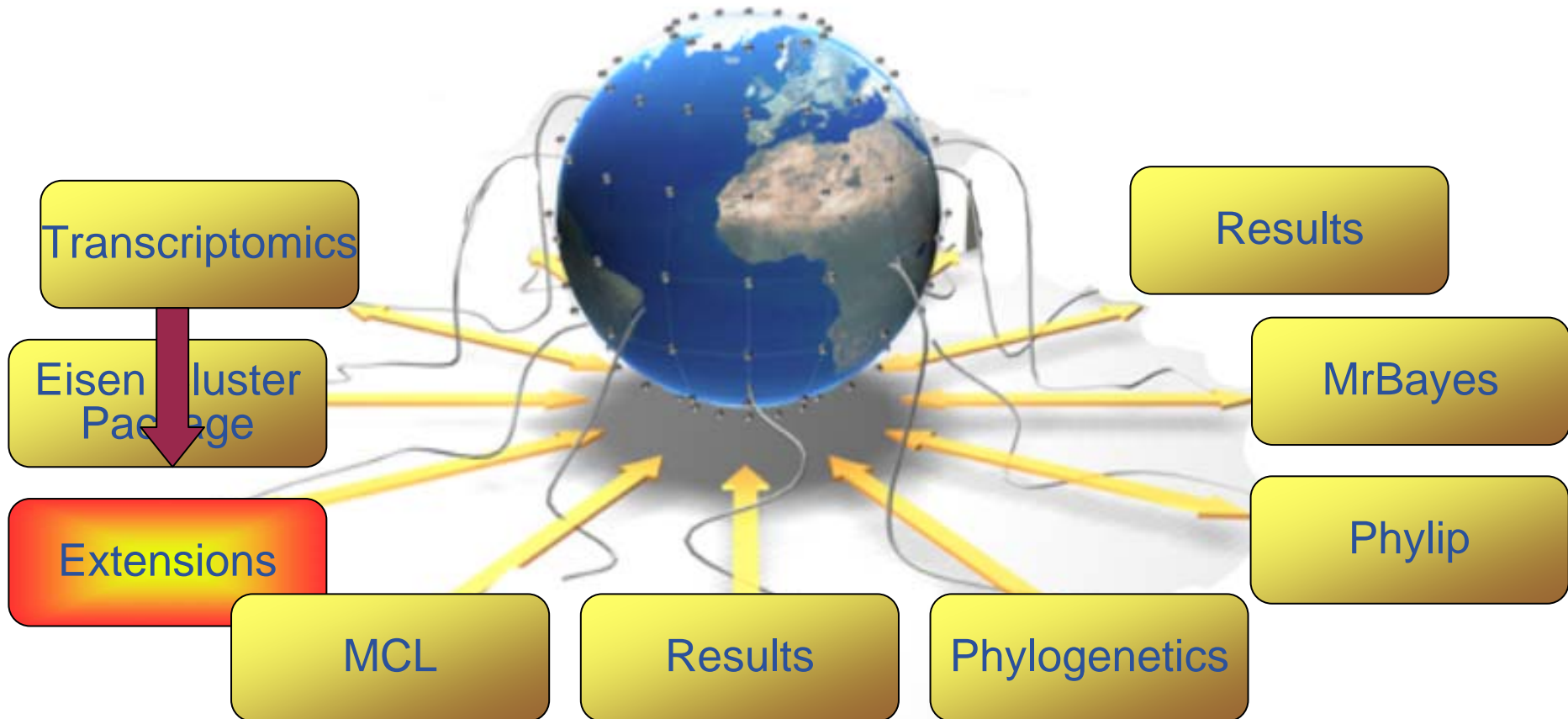


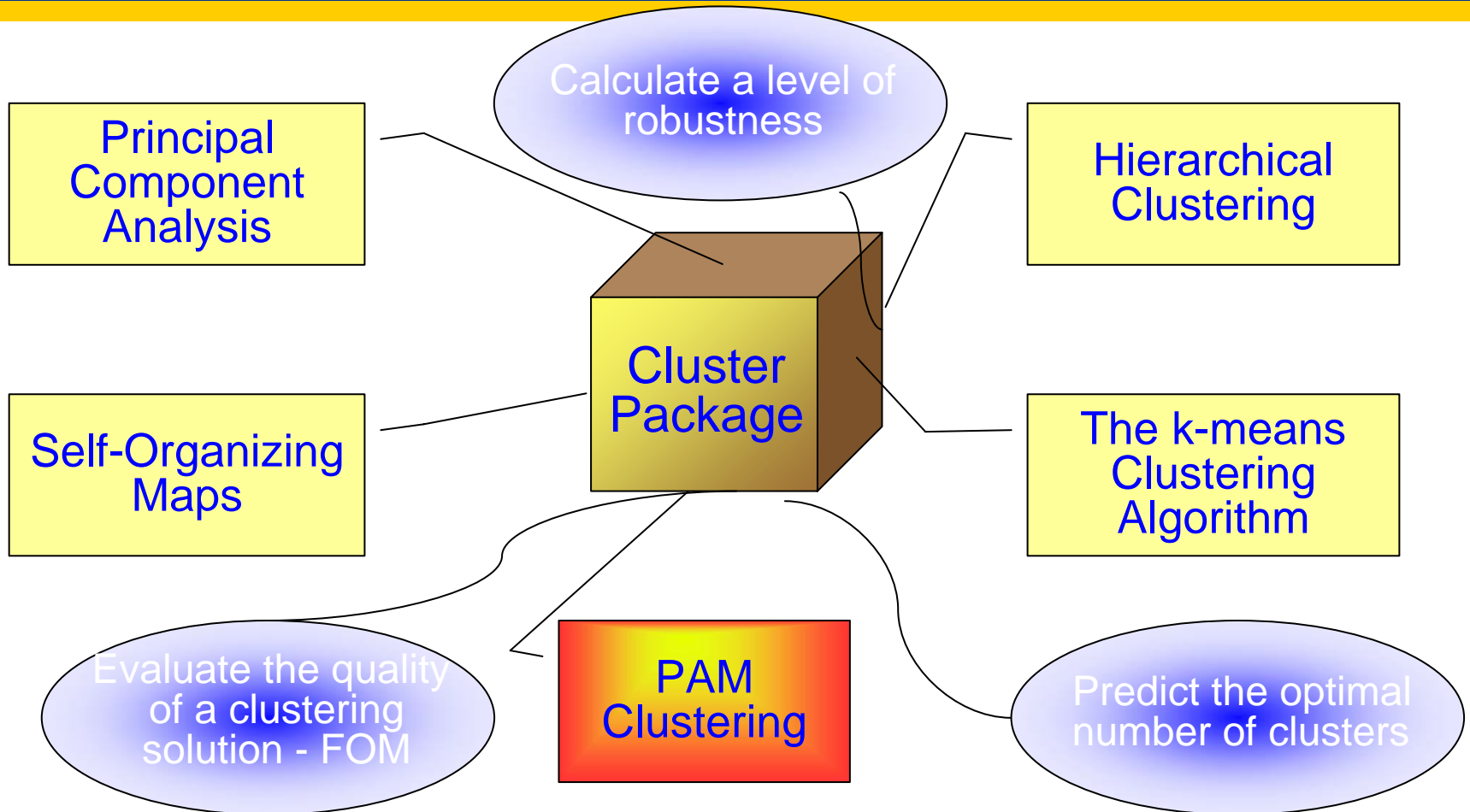


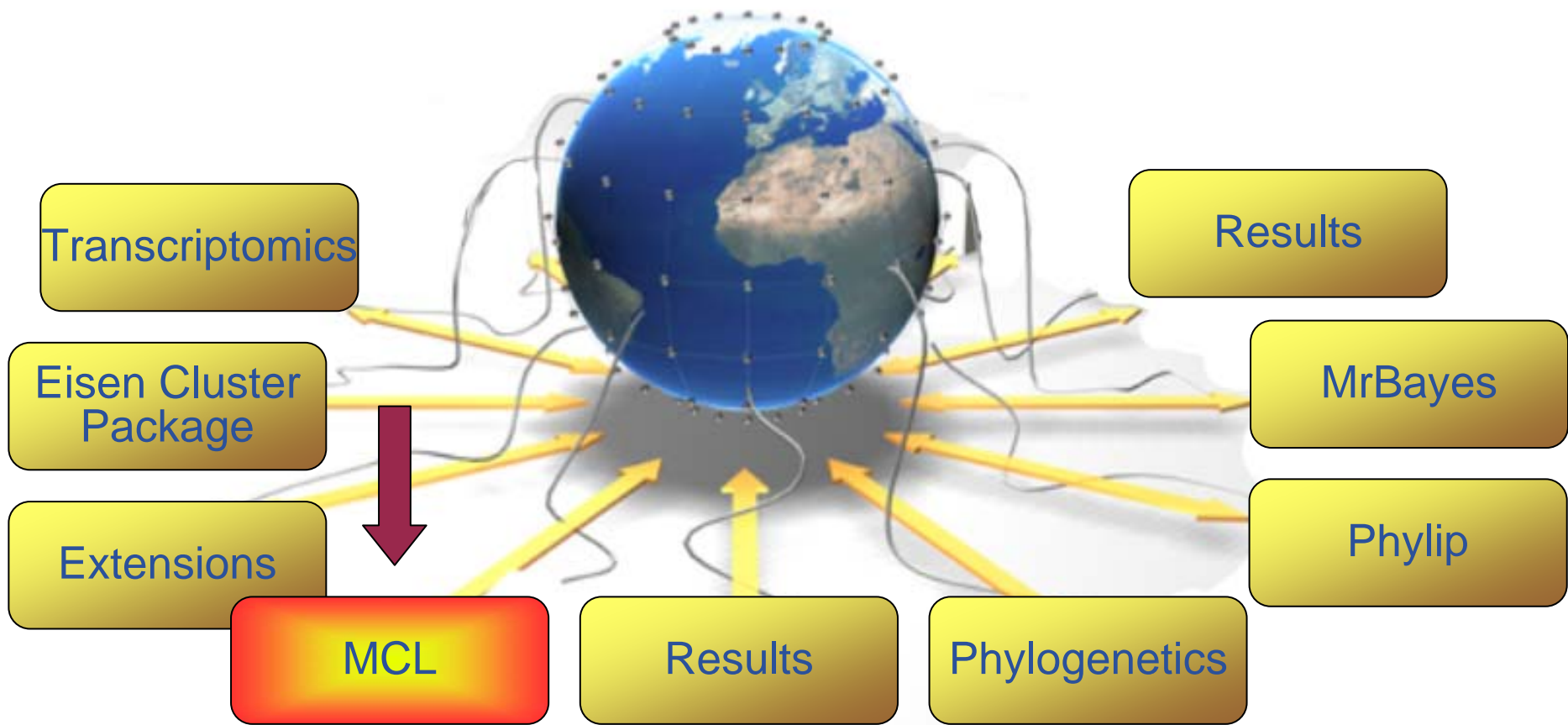
- Pearson's Correlation Coefficient
- Absolute PCC
- Uncentered PCC
- Absolute Uncentered PCC

- Spearman's rank
- Kendall's r
- Euclidean distance
- Manhattan Distance







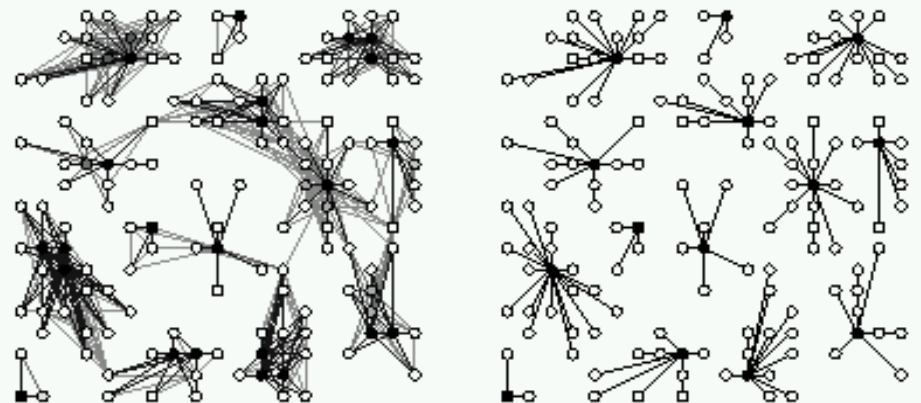
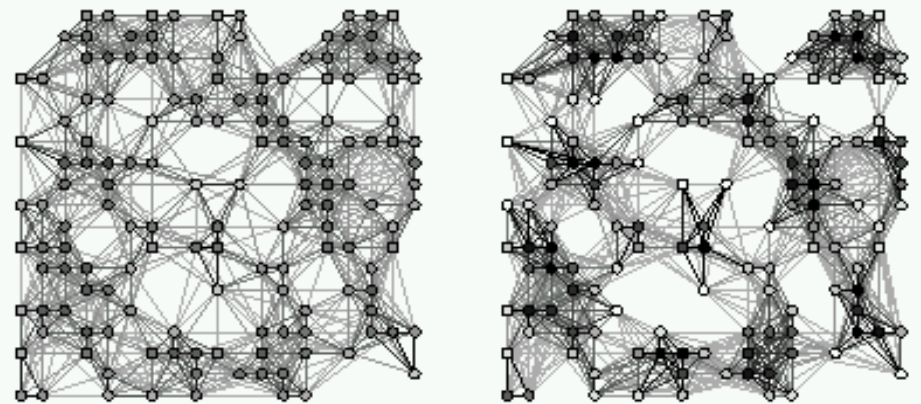


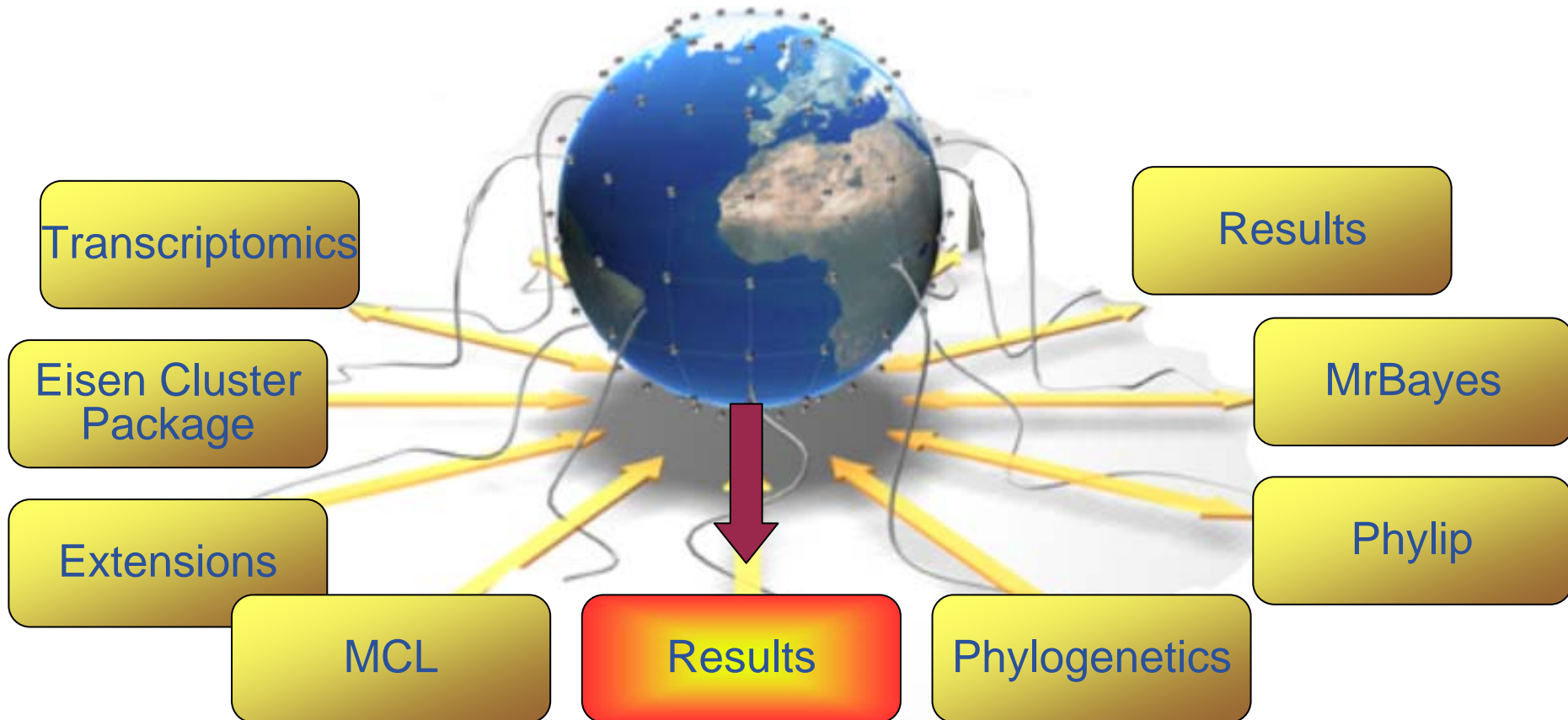


MCL Clustering:

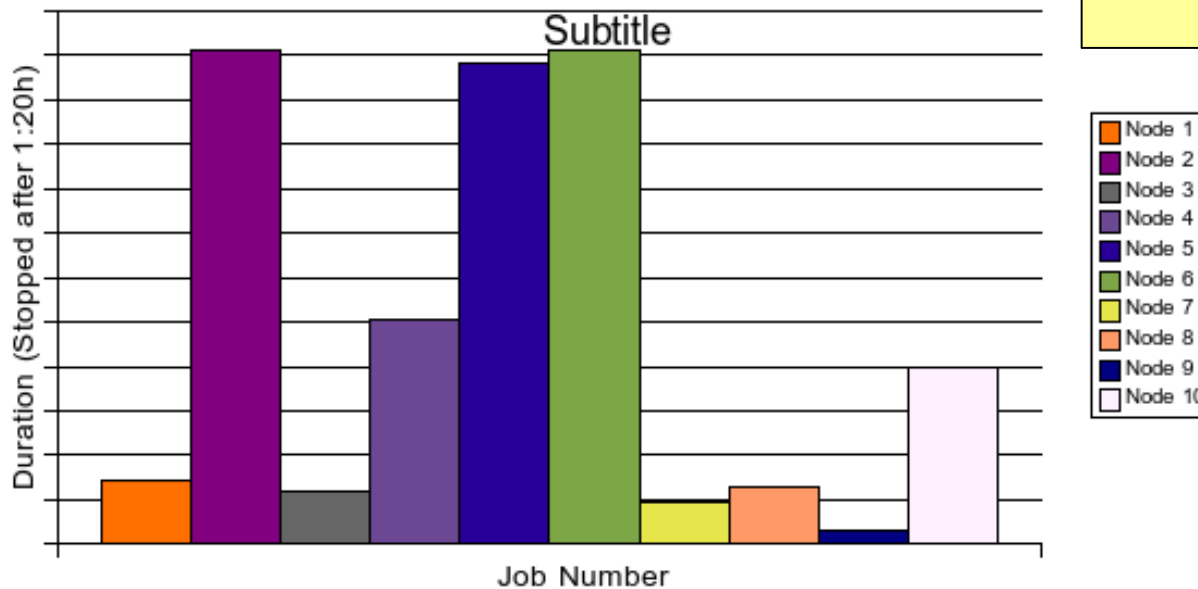
Starts by taking a random walk on the graph described by a similarity matrix

After each step the links are weakened between distant nodes and the links between nearby nodes are strengthened





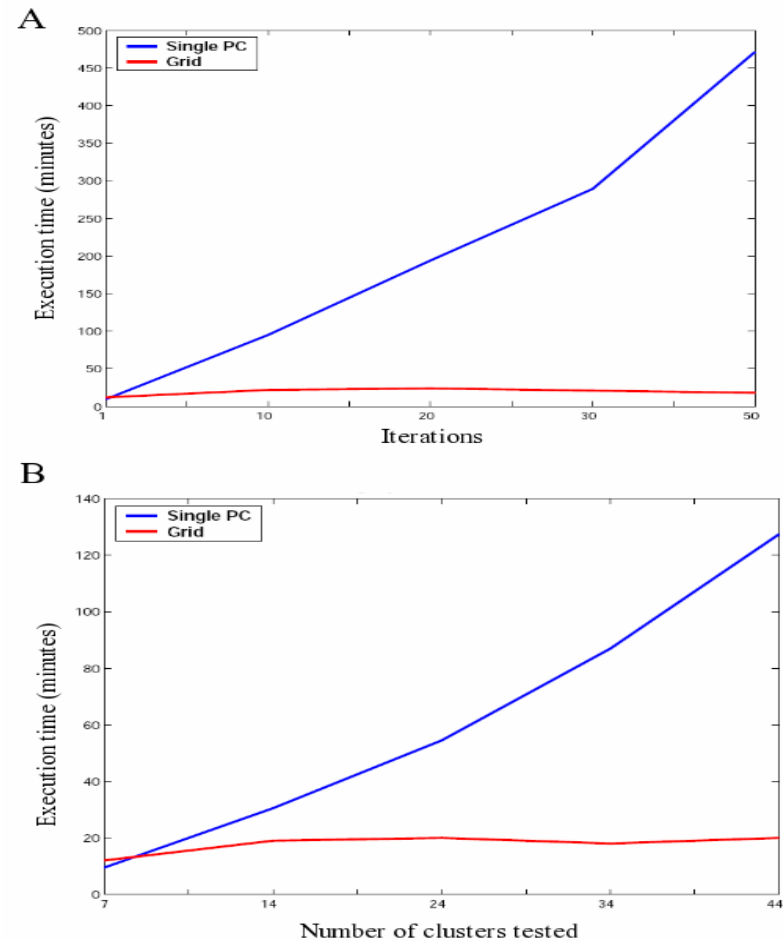
KMeans Grid Execution Times

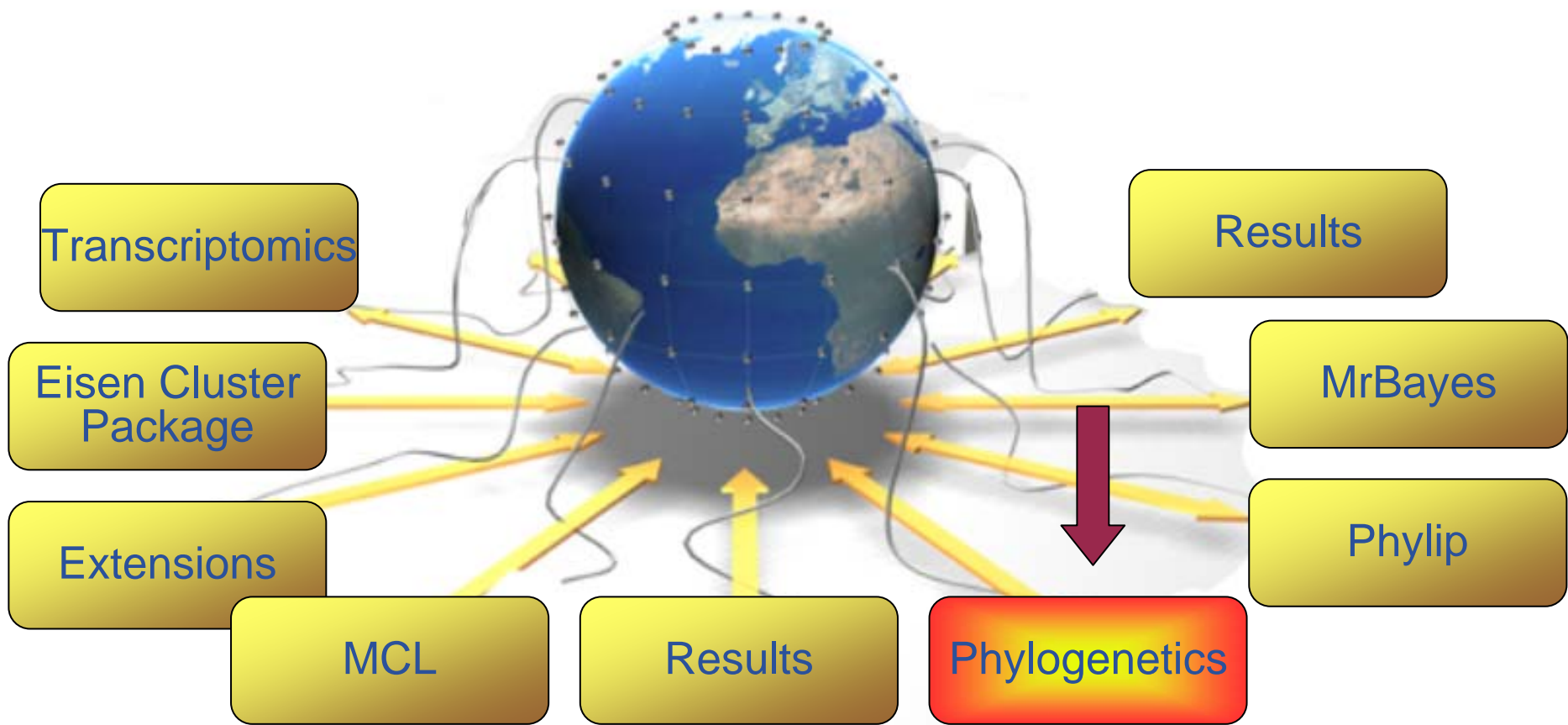


Execution times of the same job in parallel on the grid

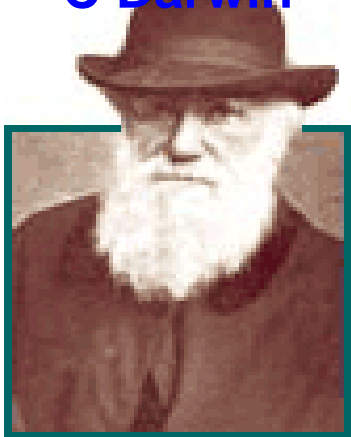
The K-means clustering algorithm was tested on the Grid referenced against a single computer firstly with a fixed number of clusters (7) and a varying amount of iterations and secondly for one iteration and a varying number of clusters (from 7 to 44).

Grid performance is on average very good regardless of the variation in execution times of similar jobs

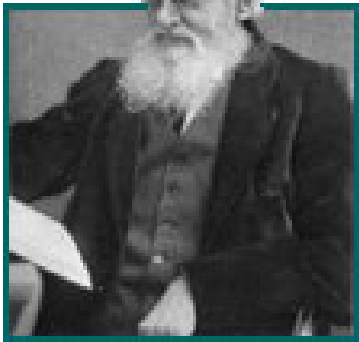




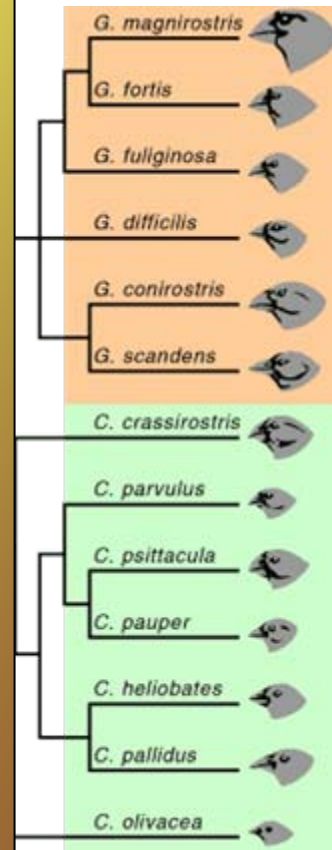
C Darwin



A R Wallace



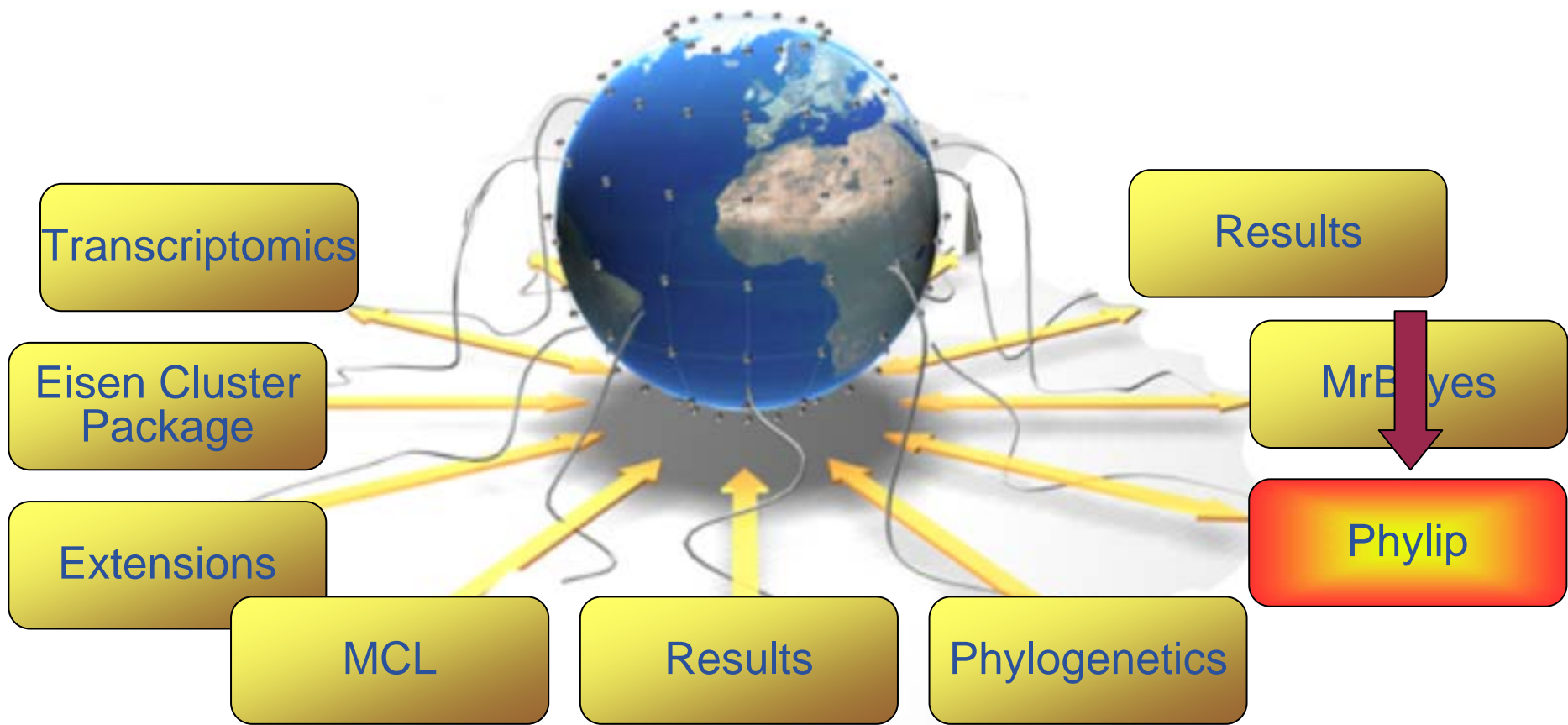
Charles Darwin and the English biologist Alfred Russel Wallace independently conceived of a natural, even observable, way for life to change: a process Darwin called natural selection.

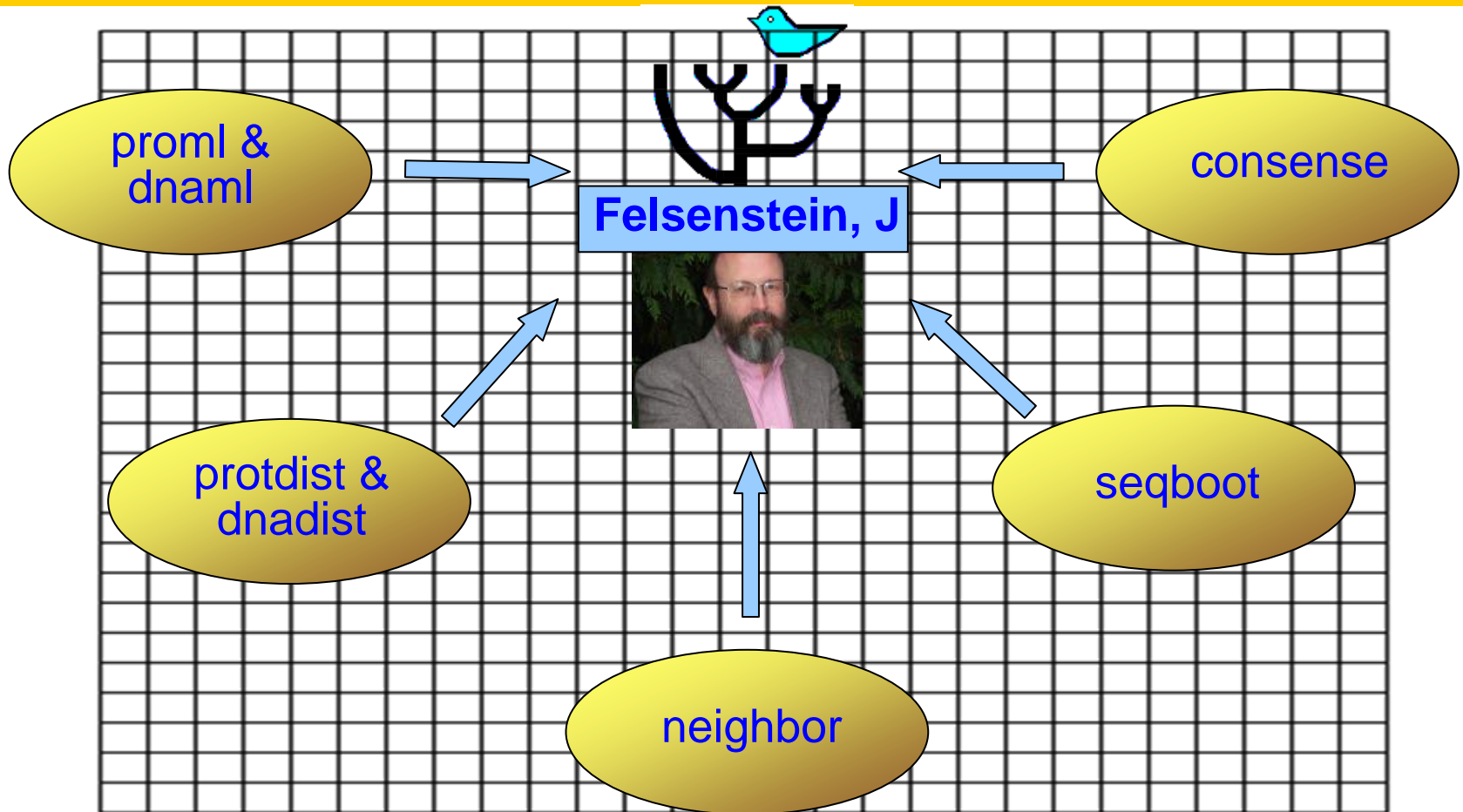


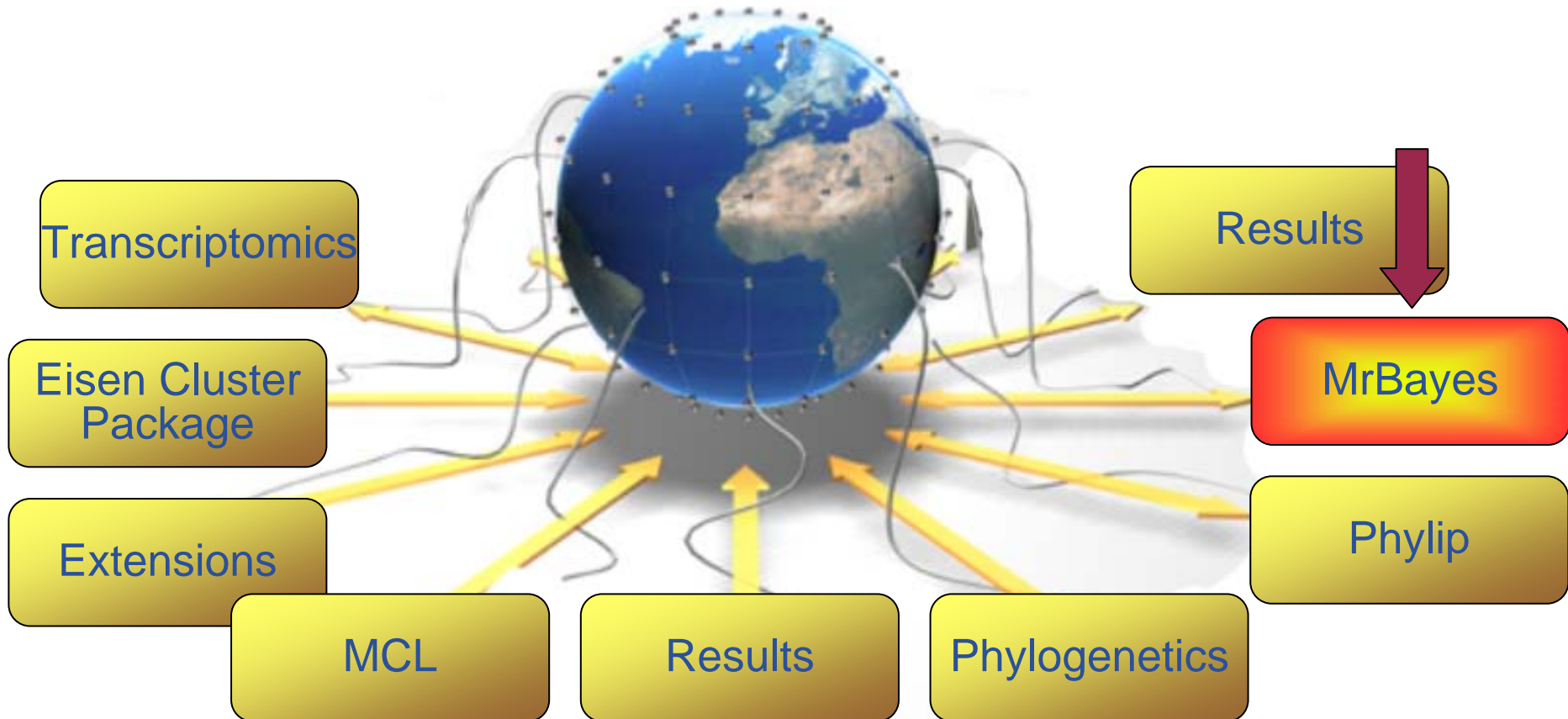
A visit to the Galapagos Islands in 1835 helped Darwin formulate his ideas on natural selection. He found several species of finch adapted to different environmental niches. The finches also differed in beak shape, food source, and how food was captured.

Reference: http://evolution.berkeley.edu/evolibrary/article/_0/history_14









MrBayes is a program for the Bayesian estimation of phylogeny. Bayesian inference of phylogeny is based upon the posterior probability distribution of trees. Trees are labelled T1, T2, ..., Tn, where n is the number of possible trees. The posterior probability of the i-th tree is calculated using Bayes's formula as

$$\Pr \theta | D = \frac{\Pr D | \theta \Pr \theta}{\Pr D}$$



Bayes, T. 1763. An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London* 53:370-418. Reprinted, E. S. Pearson and M. G. Kendall (eds.). 1970. Pages 131-153 in *Studies in the History of Statistics and Probability*. Charles Griffin, London.

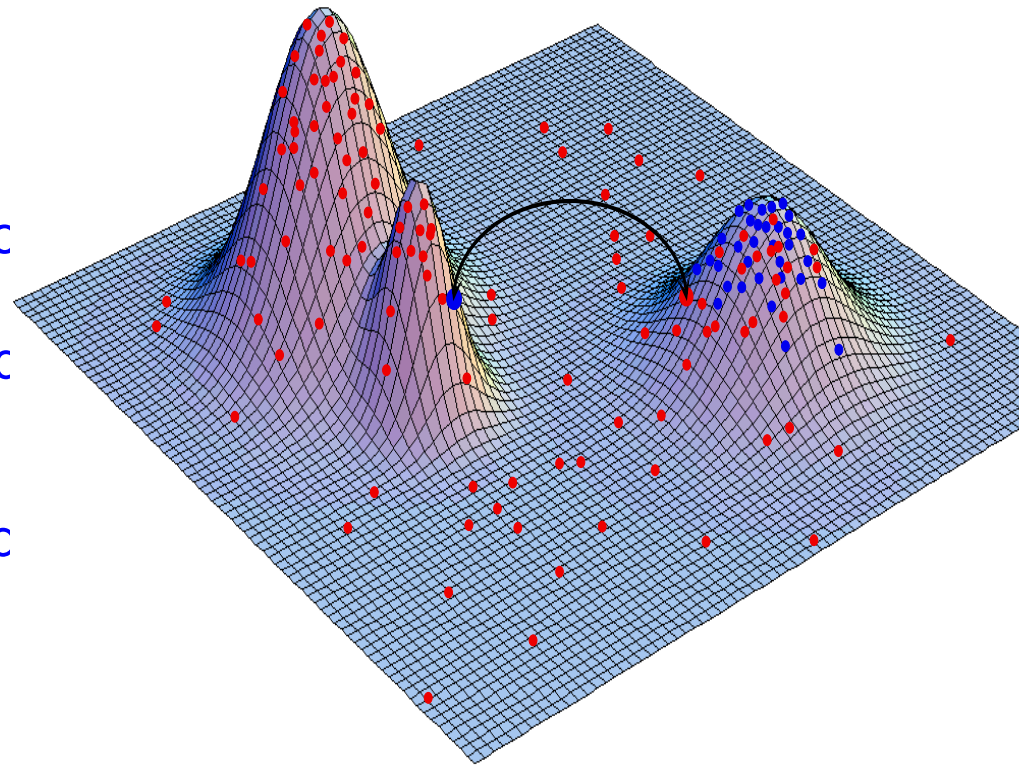


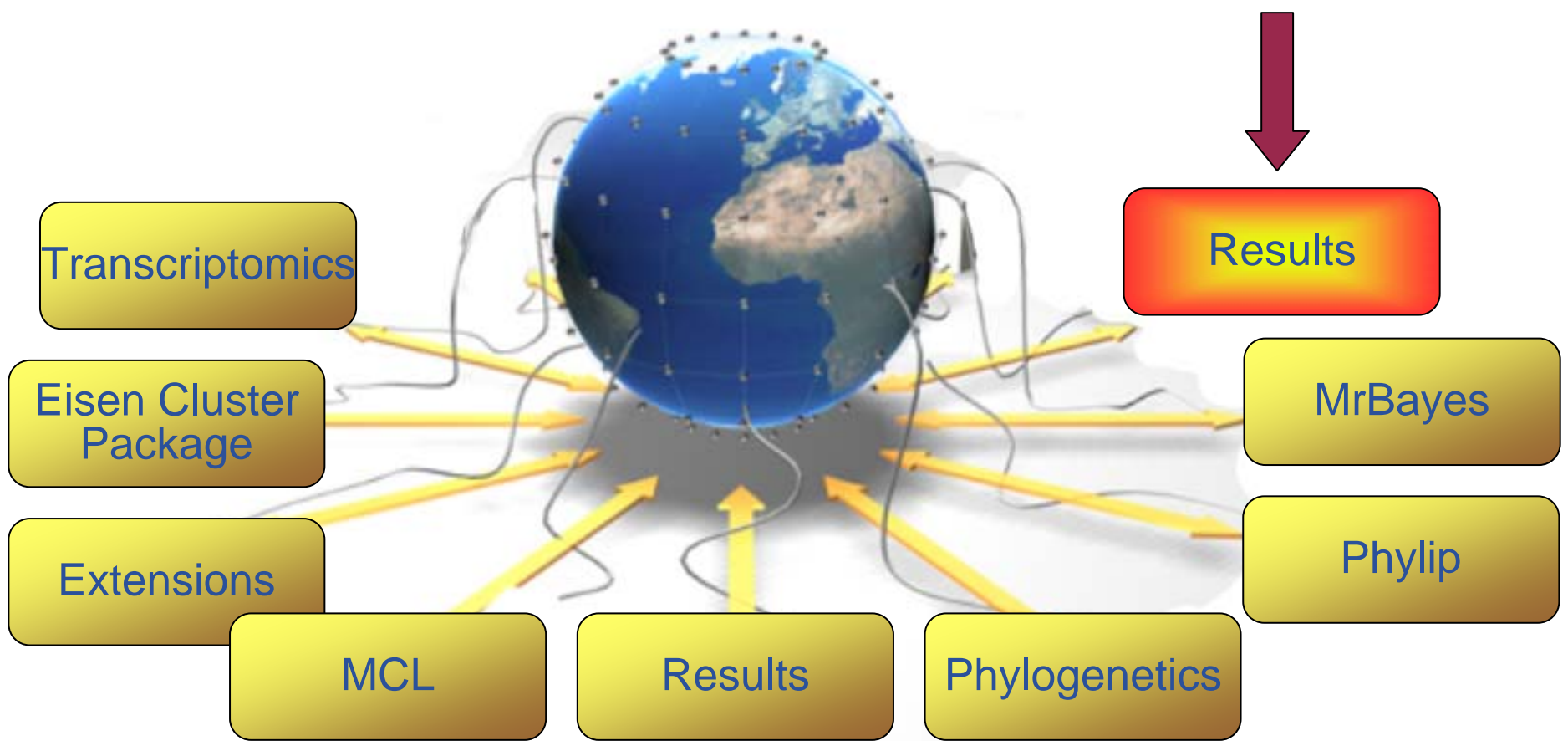
dkfz.



Metropolis-coupled Markov chain Monte Carlo (MCMCMC)

- Run n chains, $n - 1$ of which are heated
- Each generation, every chain proposes and accepts/rejects mc independently
- A swap of the states is attempted between two randomly chosen chains
- States are sampled from the colc chain





There are four steps to a typical Bayesian phylogenetic analysis using MrBayes:

1. Read the Nexus data file
2. Set the evolutionary model
3. Run the analysis
4. Summarize the samples

1. Arthropod nexus file alignment - 86t 3006p

2. 18 jobs, 200000 iterations, burnin 10000.

3. Less than 24h
3,600,000 samples





Constant Technical Support

1. Ivan Merelli
2. Giacinto Donvito
3. Giuseppe La Rocca

WP3 Management

1. Pietro Lio

