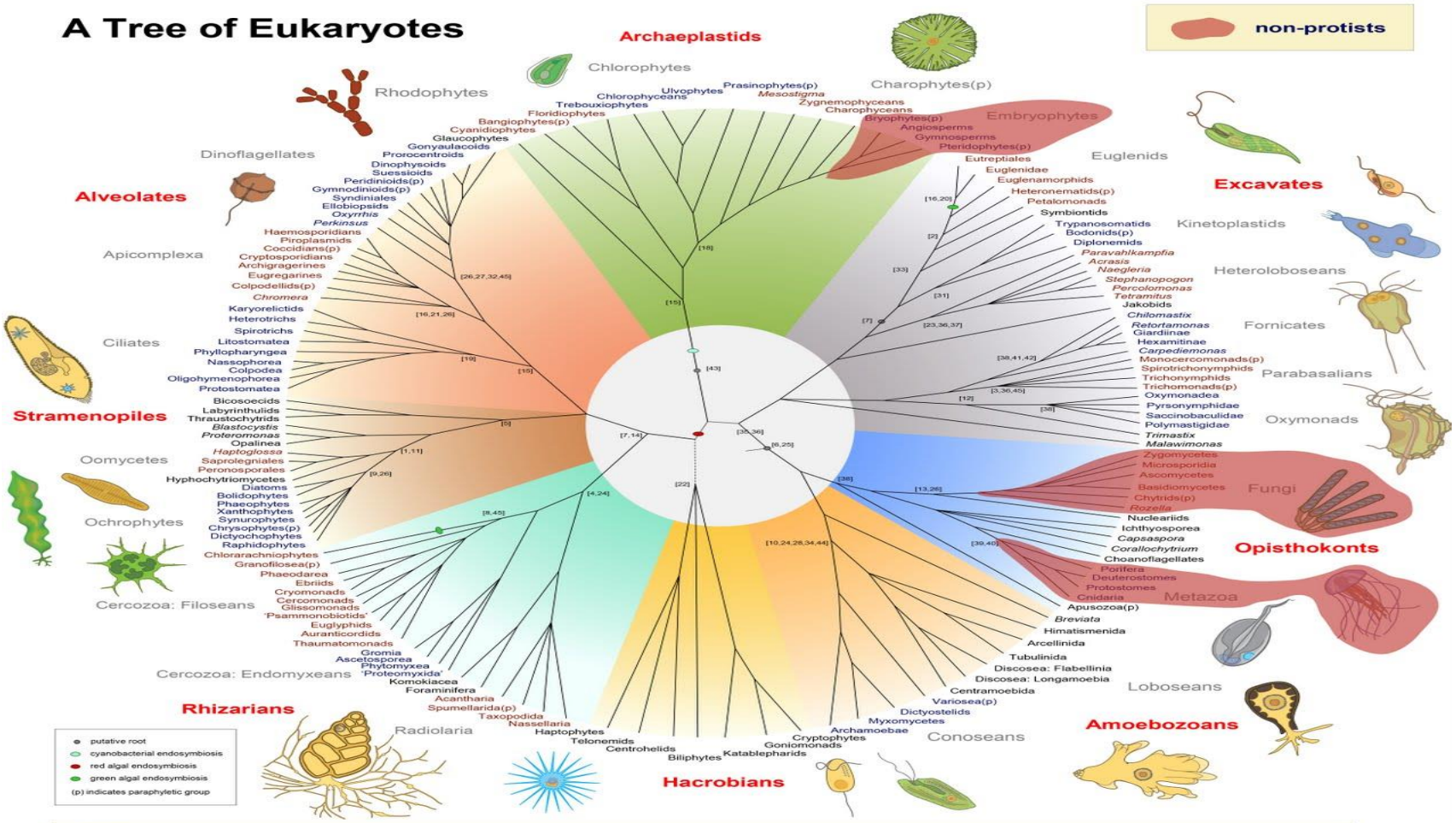
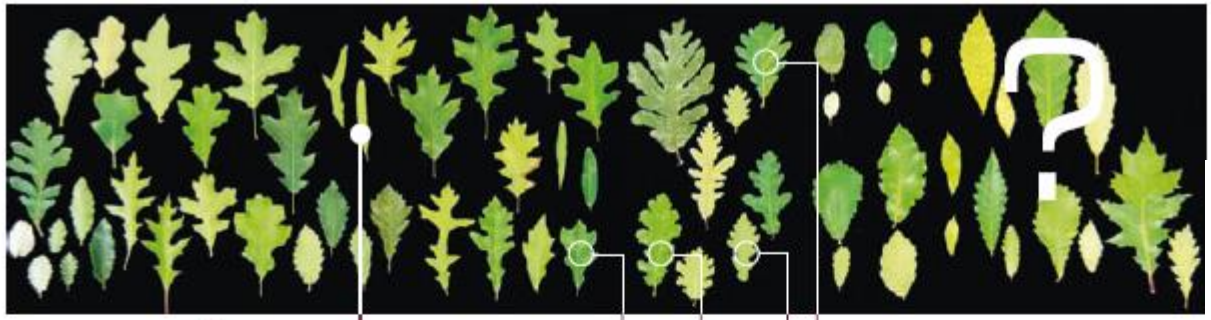
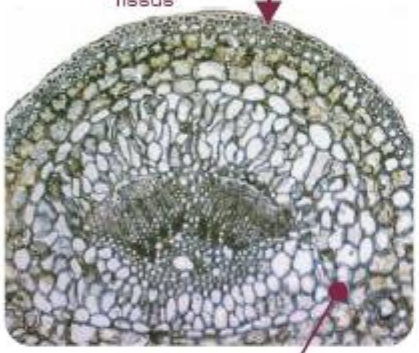


References:

1. Beakes & Sekimoto 2009 **The evolutionary phylogeny of oomycetes: insights gained from studies of Holocarpic parasites of algae and invertebrates.** in *Oomycete Genetics and Genomics* (eds. Lamour & Kamoun)
2. Breglia *et al.* 2007 **Phylogeny of phagotrophic Euglenids (Euglenozoa) is inferred from Hsp90 gene sequences.** *J Euk Microbiol* **54**:86-92
3. Carpenter & Keeling 2007 **Morphology and Phylogenetic Position of *Eucomonympha imla* (Parabasalia: Hypermastigida)** *J Euk Microbiol* **54**:263-270
4. Cavalier-Smith & Chao 2003 **Phylogeny and classification of phylum cercozoa.** *Protist* **154**:341-358
5. Cavalier-Smith & Chao 2006 **Phylogeny and Megasystematics of Phagotrophic Heterokonts (Kingdom Chromista).** *J Mol Evol* **62**:388–420
6. Cavalier-Smith 2009 **Megaphylogeny, Cell Body Plans, Adaptive Zones: Causes and Timing of Eukaryote Basal Radiations.** *J Euk Microbiol* **56**:26-33
7. Cavalier-Smith 2010 **Kingdoms Protozoa and Chromista and the eozoan root of the eukaryotic tree.** *Biol Lett* AOP
8. Chantangsi *et al.* 2008 **Morphology and molecular phylogeny of a marine interstitial tetraflagellate with putative endosymbionts: Auranticordis quadriverberis n. gen. et sp. (Cercozoa).** *BMC Microbiol* **8**:123
9. Eikrem *et al.* 2004 ***Florenciella parvula* gen. et sp. nov. (Dictyochophyceae, Heterokontophyta), a small flagellate isolated from the English Channel** *Phycologia* **43**:658-668
10. Fiore-Donno *et al.* 2009 **Deep Phylogeny and Evolution of Slime Moulds (Mycetozoa).** *Protist* **161**:55-70
11. Hakariya *et al.* 2007 **A molecular phylogeny of *Haptoglossa* species, terrestrial peronosporomycetes (oomycetes) endoparasitic on nematodes.** *Mycoscience* **48**:169-175
12. Heiss & Keeling 2006 **The Phylogenetic Position of the Oxymonad Saccinobaculus Based on SSU rRNA.** *Protist* **157**:335-344
13. James *et al.* 2006 **Reconstructing the early evolution of Fungi using a six-gene phylogeny.** *Nature* **443**:818-822
14. Keeling 2009 **Chromalveolates and the Evolution of Plastids by Secondary Endosymbiosis.** *J Euk Microbiol* **56**:1-8
15. Keeling et al. 2005 **The tree of eukaryotes.** *Tr Ecol Evol* **20**:670-676
16. Leander 2008 **Marine gregarines: evolutionary prelude to the apicomplexan radiation?** *Tr Parasitol* **24**:60-67
17. Leander *et al.* 2007 **Macroevolution of complex cytoskeletal systems in Euglenids.** *BioEssays* **29**:987-1000
18. Lewis & McCourt 2004 **Green algae and the origin of land plants.** *Am J Bot* **91**:1535-1556
19. Lynn 2003 **Morphology or molecules: How do we identify the major lineages of ciliates (Phylum Ciliophora)** *Eur J Protistol* **39**:356-364
20. Marin *et al.* 2003 **Phylogeny and Taxonomic Revision of Plastid-Containing Euglenophytes based on SSU rDNA Sequence Comparisons and Synapomorphic Signatures in the SSU rRNA Secondary Structure** *Protist* **154**:99-145
21. Morrison 2009 **Evolution of the Apicomplexa: where are we now?** *Tr Parasitol* **25**:375-382
22. Okamoto *et al.* 2009 **Molecular Phylogeny and Description of the Novel Katablepharid *Roombia truncata* gen. et sp. nov., and Establishment of the Hacrobia Taxon nov.** *PLoS ONE* **4**:e7080
23. Park *et al.* 2009 **Ultrastructure and Molecular Phylogeny of two Heterolobosean *Amoebae*, *Euplaesiobystra hypersalinica* gen. et sp. nov. and *Tulamoeba peronaphora* gen. et sp. nov., Isolated from an Extremely Hypersaline Habitat** Protist **160**:265-283
24. Pawlowski & Burki 2009 **Untangling the phylogeny of amoeboid protists.** *J Euk Microbiol* **56**:16-25
25. Richards & Cavalier-Smith 2005 **Myosin domain evolution and the primary divergence of eukaryotes.** *Nature* **436**:1113-1118
26. Saldarriaga JF (personal comm., 2010)
27. Saldarriaga *et al.* 2004 **Molecular data and the evolutionary history of dinoflagellates.** *Eur J Protistol* **40**:85-111
28. Shadwick *et al.* 2009 **Eumycetozoa = Amoebozoa?: SSU rDNA Phylogeny of Protosteloid Slime Molds and Its Significance for the Amoebozoan Supergroup** *PLoS ONE* **4**:e6754
29. Simpson 2003 **Cytoskeletal organization, phylogenetic affinities and systematics in the contentious taxon Excavata (Eukaryota)** *Int J Sys Evol Microbiol* **53**:1759-1777
30. Simpson *et al.* 2006 **Comprehensive Multigene Phylogenies of Excavate Protists Reveal the Evolutionary Positions of "Primitive" Eukaryotes.** *Mol Biol Evol* **23**:615-625
31. Simpson *et al.* 2006 **The evolution and diversity of kinetoplastid flagellates.** *Tr Parasitol* **22**:168-174
32. Taylor *et al.* 2007 **Dinoflagellate diversity and distribution.** *Biodiv Cons* **17**:407–418
33. Yubuki *et al.* 2009 **Ultrastructure and molecular phylogeny of Calkinsia aureus: cellular identity of a novel clade of deep-sea Euglenids with epibiotic bacteria.** *BMC Microbiol* **9**:16
34. Bernay *et al.* 2009 XII Intl Congress of Protistol, Armação dos Búzio, Brazil (poster)
35. Hampl *et al.* 2009 **Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic "supergroups"** *PNAS* **106**:3859-3864
36. Simpson AGB (personal comm. 2010)
37. Yubuki & Leander 2008 **Ultrastructure and molecular phylogeny of Stephanopogon minuta: An enigmatic microeukaryote from marine interstitial environments** *Eur J Protistol* **44**:241-253
38. Heiss AA (personal comm. 2010)
39. Dunn *et al.* 2008 **Broad phylogenomic sampling improves resolution of the animal tree of life** *Nature* **452**:745-749
40. Glenner *et al.* 2004 **Bayesian Inference of the Metazoan Phylogeny: A Combined Molecular and Morphological Approach** *Curr Biol* **14**:1644-1649
41. Cepicka *et al.* 2008 **Non-monophyly of Retortamonadida and high genetic diversity of the genus Chilomastix suggested by analysis of SSU rDNA** *Mol Phylogenet Evol* **48**:770-775
42. Kolisko *et al.* 2008 **Molecular phylogeny of diplomonads and enteromonads based on SSU rRNA, alpha-tubulin and HSP90 genes: Implications for the evolutionary history of the double karyomastigont of diplomonads.** *BMC Evol Biol* **8**:205
43. Rogozin *et al.* 2009 **Analysis of Rare Genomic Changes Does Not Support the Unikont–Bikont Phylogeny and Suggests Cyanobacterial Symbiosis as the Point of Primary Radiation of Eukaryotes** *Genome Biol & Evol* **2009**:99–113
44. Lahr DJG (personal comm. 2011)
45. Keeling PJ (personal comm. 2011)

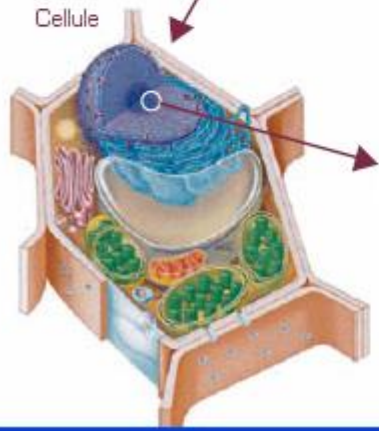


Tissus



ACGTGTGCTAT ▶ *Quercus petraea*
ACGCGTGCTAT ▶ *Quercus robur*
ACGT--GCTAT ▶ *Quercus pubescens*
ACGCAGTCTAT ▶ *Quercus cocinea*

Cellule



Extraction,
Amplification,
Séquençage
ADN

Which tools are available?

ADDED TOOLS

virtual_BiodiversityL@b

- MAFFT alignment (basic options)
- BMGE block selection
- Fasta to Phylip conversion
- Fasta to Nexus conversion
- PHYML wrapper (maximum likelihood phylogenies)
- Draw ascii tree (newick)
- BEAST wrapper

DECLIC

- Disseq
- Distance ape
- Declic Graph
- Declic MDS

DIAGNO-SYST (HIDDEN TOOLS)

- Diagno-Syst char format Formats charfile
- Diagno-Syst
- Disseq reads to taxo
- Taxonomic inventory from

Alignment (multiple)

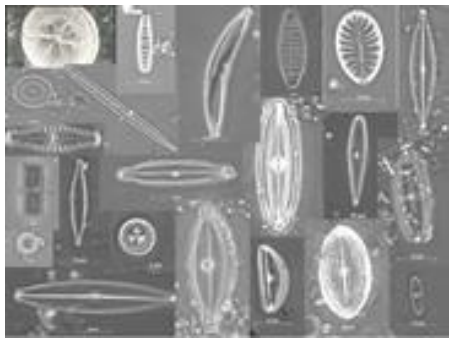
Phylogeny

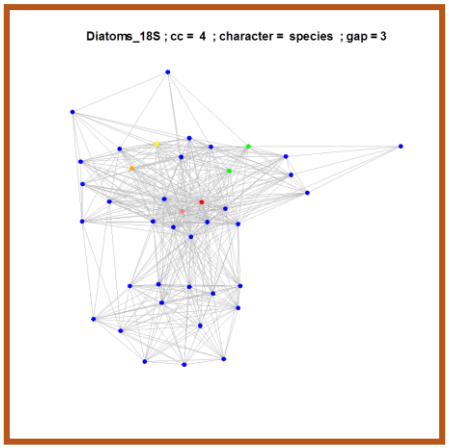
Distance based methods
Smith-Waterman (exact)
Evolution models

Graph clustering

MDS - NLM

Aggregative clustering (fast ,,,,)





Molecular based taxonomy and systematics: An open route for (new) methods

Sequences known by pairwise distances → **Distance geometry**
pattern recognition
machine learning

Clustering

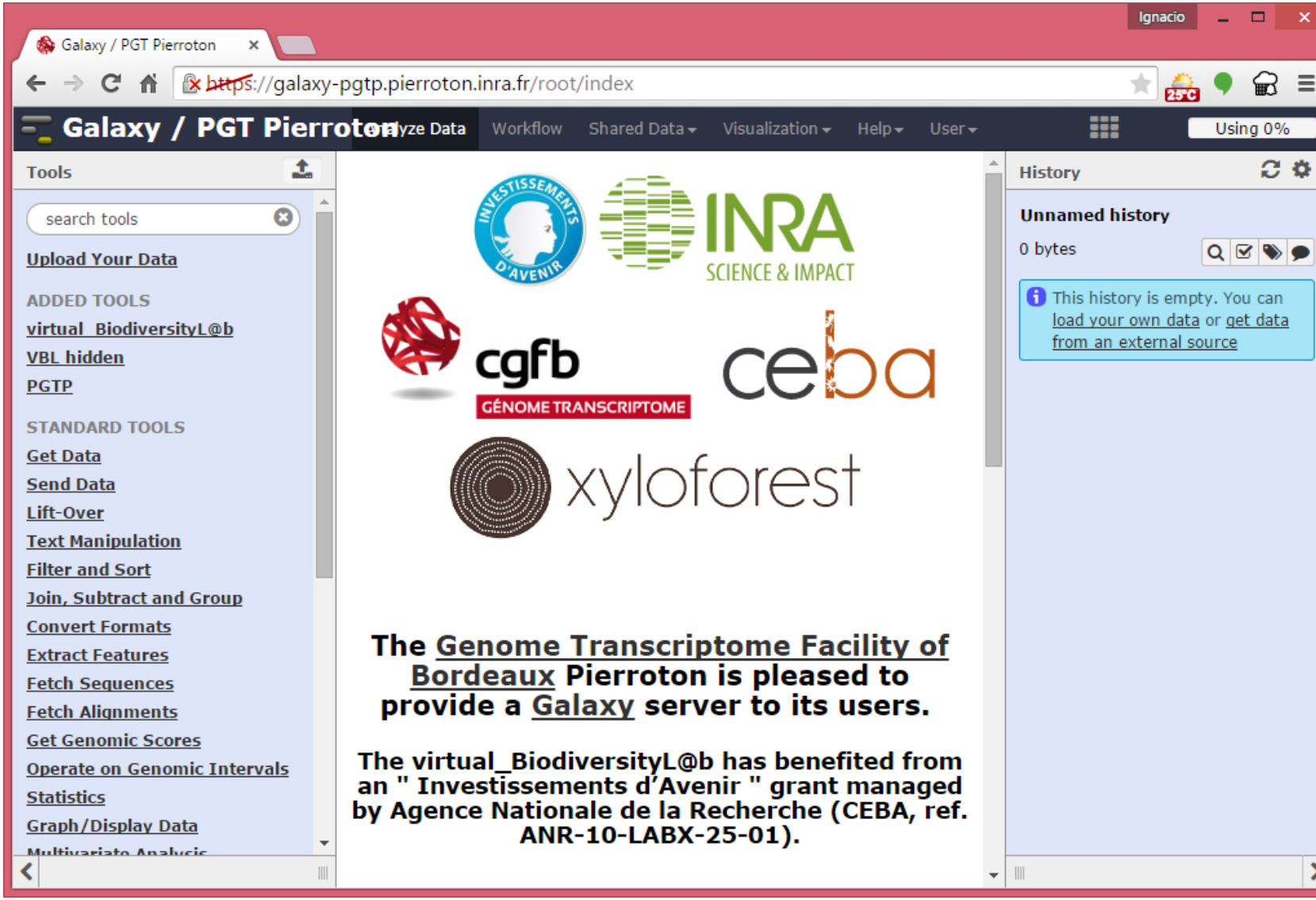
Multidimensional Scaling
linear and nonlinear (e.g. Sammon, 1969)

Manifold learning
IsoMap, EigenMap, etc …

Graph based methods
spectral clustering

Machine learning
SVM, …

Galaxy portal



Where is it possible to compute?

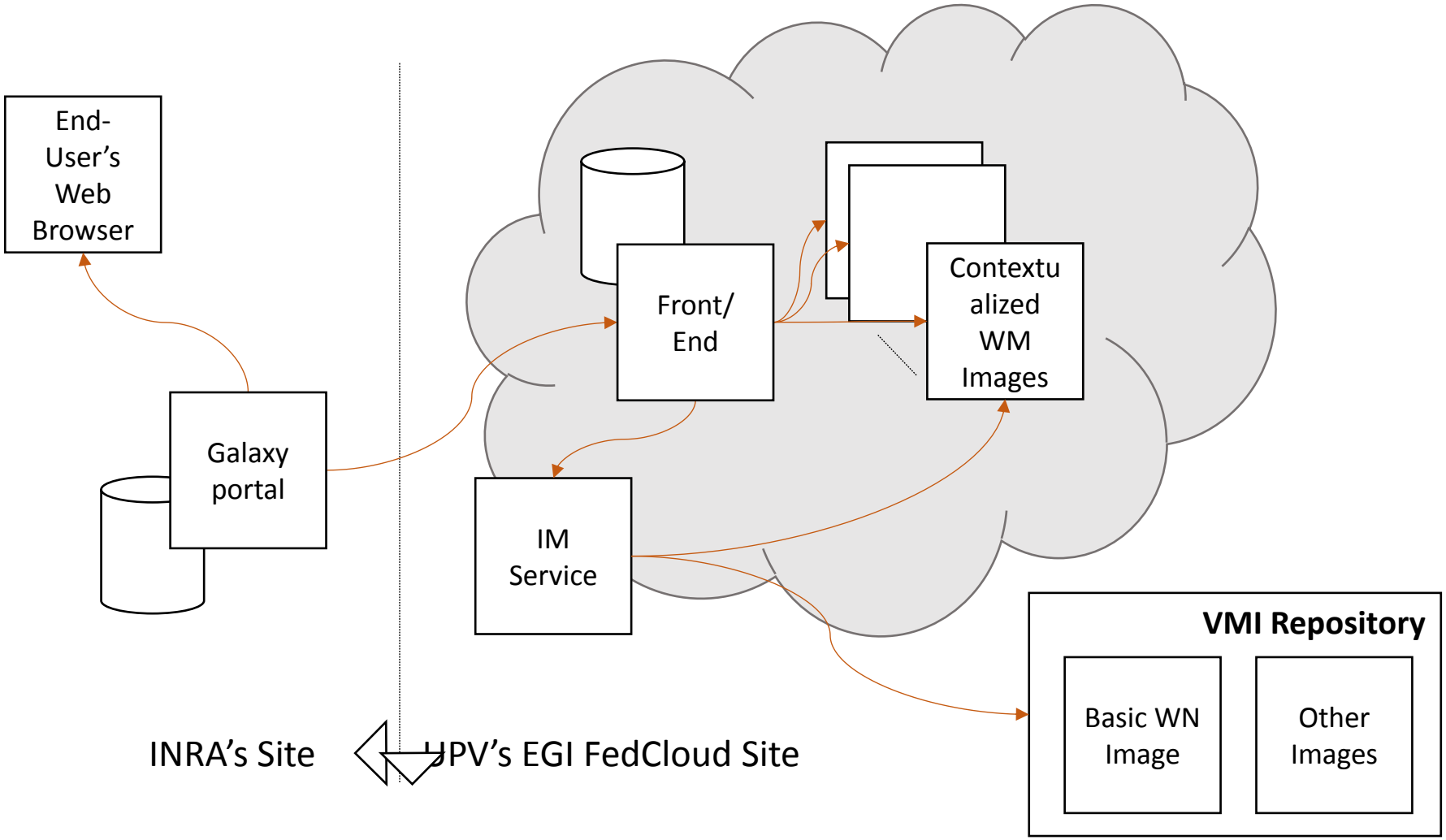
From a unique portal
the Galaxy instance

- Local Galaxy server
- Mesocentre (Tier 2)
  Avakas
  1000 cores
- Tier 1 (IDRIS, one pipeline, not via Galaxy)
- EGI GRID
  France-Grille
- Cloud (next slides …)

Where from?

From any computer connected to internet
Currenty available from French Guiana
(IP Cayenne works with it)

Software Architecture

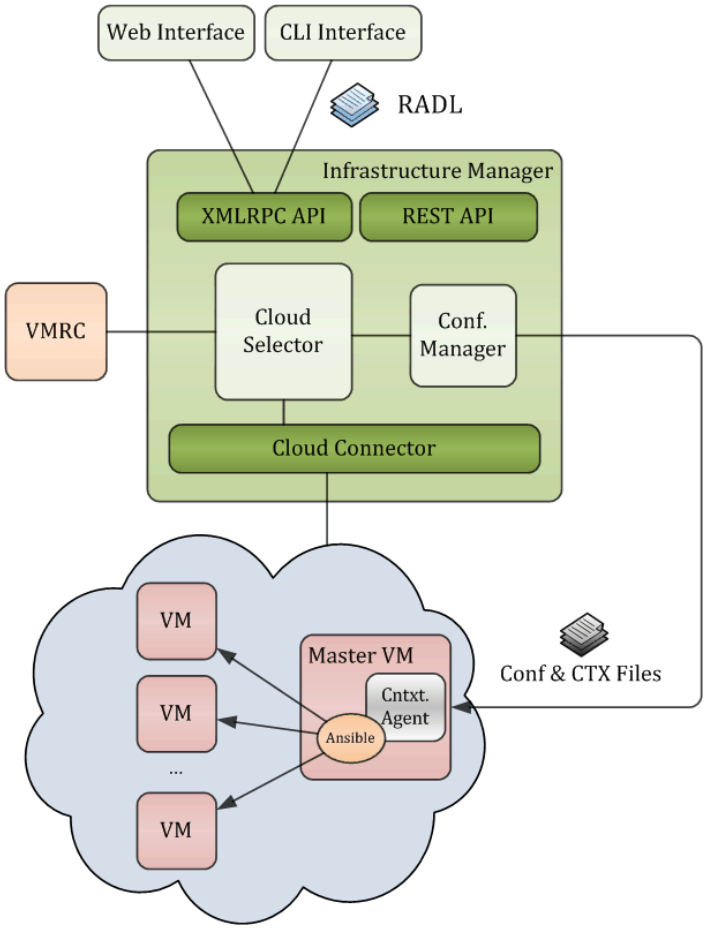


Technologies Used

- Galaxy (galaxyproject.org)
  - Bioinformatics portal that transforms operations into Jobs submitted to a back-end batch queue.
- Infrastructure Manager (IM - www.grycap.upv.es/im)
  - Platform-agnostic, dynamic configuration service that enables re-configuring a cluster.
- Elastic Compute Cloud Clusters (EC3 - www.grycap.upv.es/ec3)
  - Automatic elasticity service based on batch queue length.

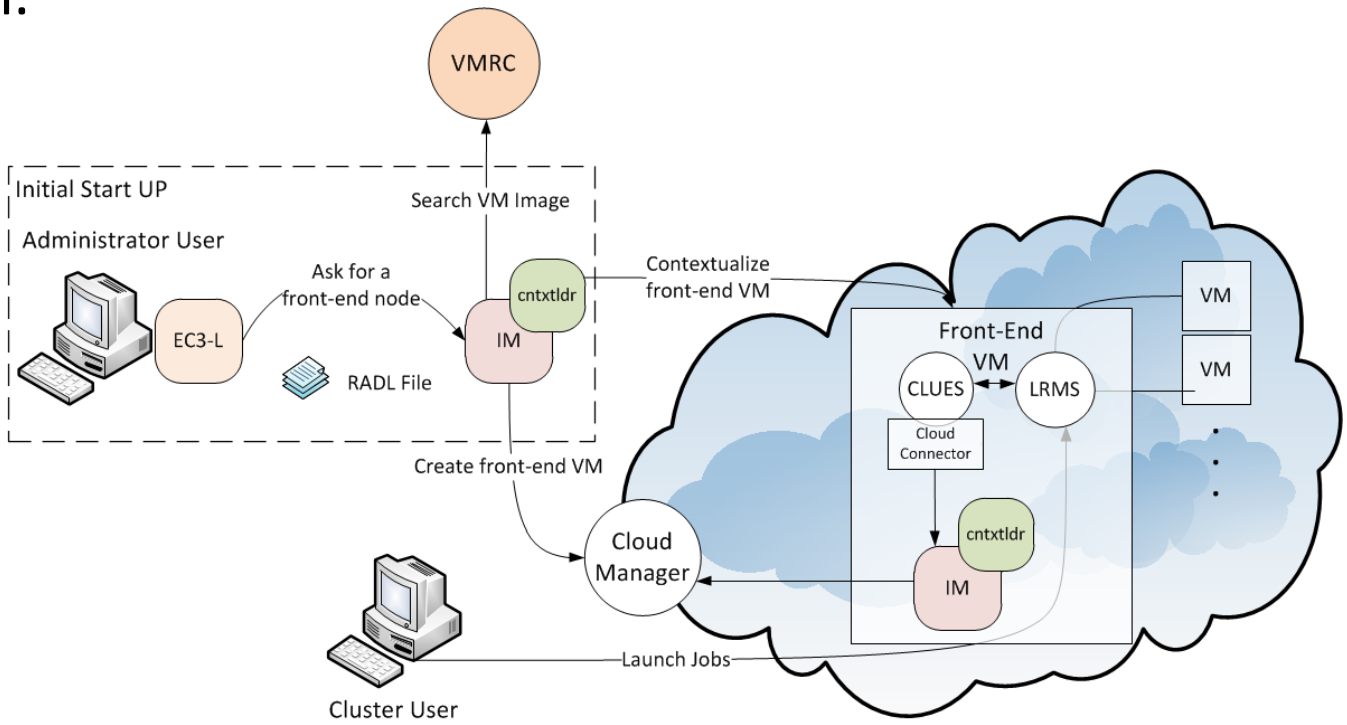
(Re)-Configuration

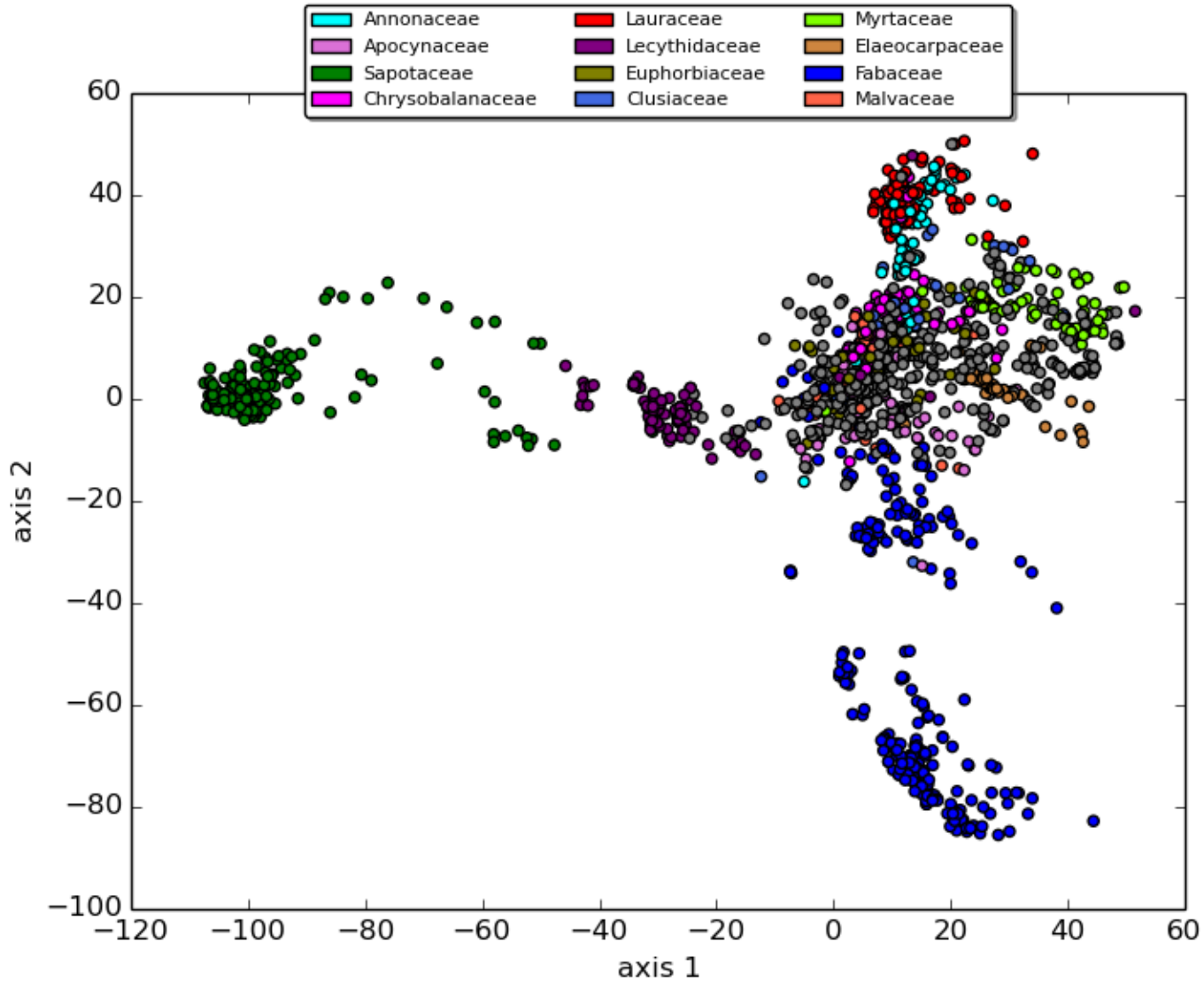
- Infrastructure Manager configures VMs from recipes
  - It can restart a broken configuration or rearrange the changes on an active configuration.
  - It Works with a plethora of CMPs, including OCCI and EGI Federated Cloud.



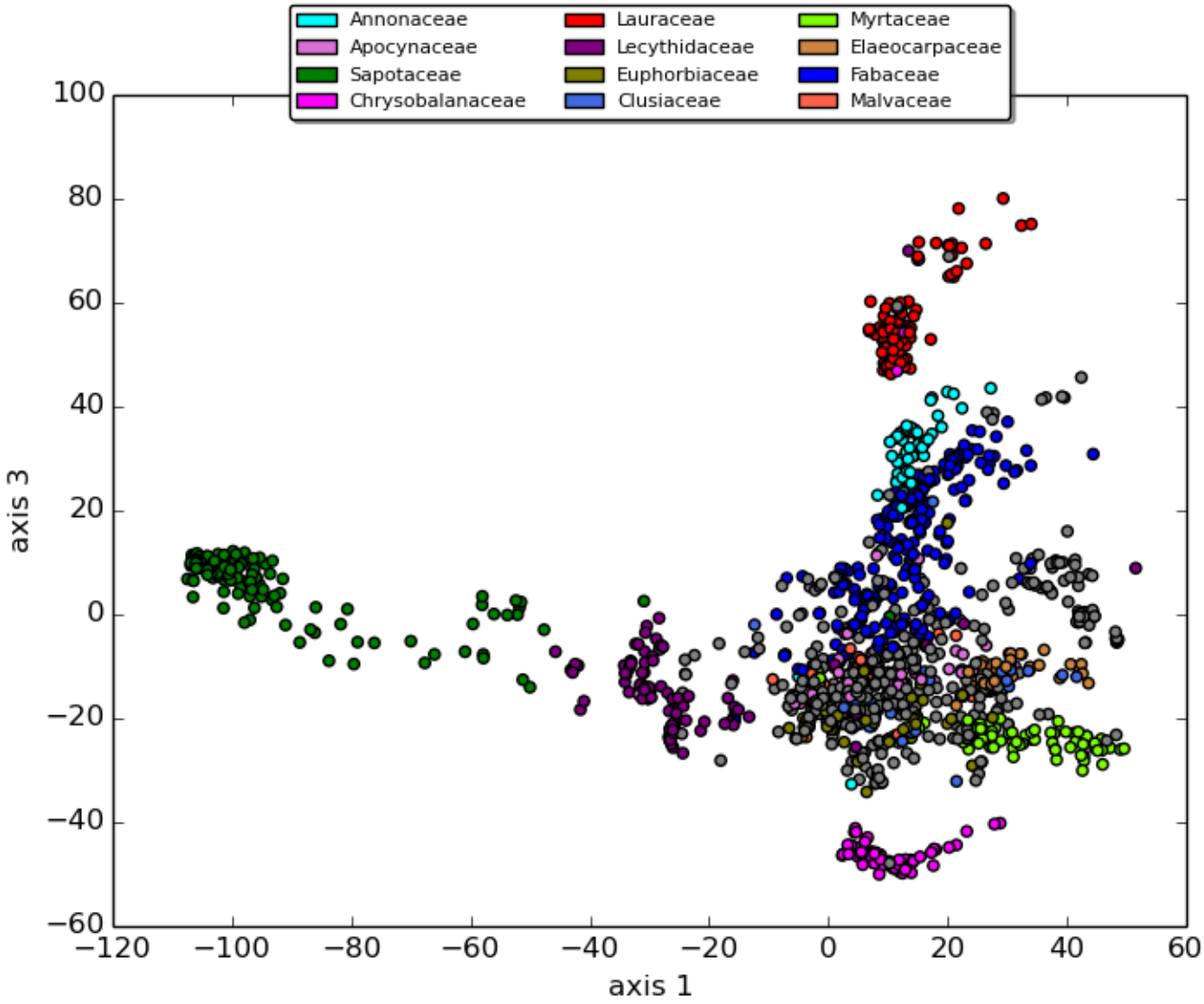
Elastic Compute Cloud Cluster

•EC3 (Elastic Cloud Computing Cluster) leverages CLUES to create elastic virtual clusters in the Cloud.
•It also uses Infrastructure Manager (IM) for the reconfiguration.
•It provides
•automatic
•elasticity
•based on
•the batch
•queue
•workload.





Centre d'Etude de la Biodiversité Amazonienne - © 2012



Centre d'Etude de la Biodiversité Amazonienne - © 2012