

# Introduction to Bioinformatics

JTMS-19

Marc-Thorsten Hütt

mhuett@constructor.university

Felix Jonas

fjonas@constructor.university

What is this session about?

Basic terminology of clustering trees. UPGMA algorithm is introduced.

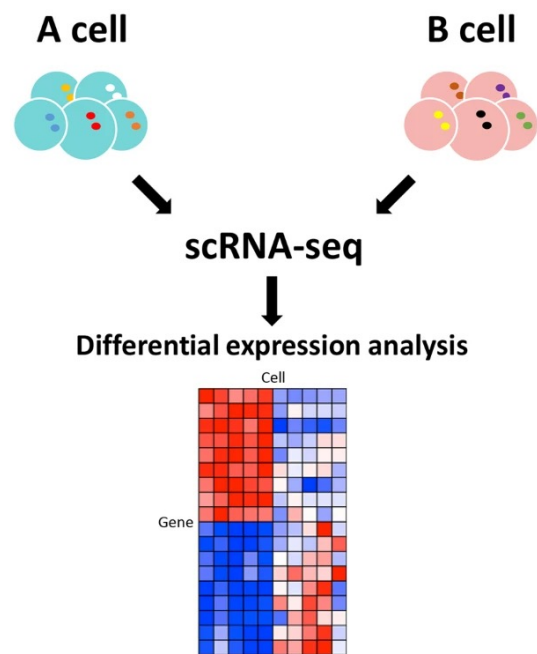
How can you revise the material after the session?

Read Baxevanis/Oullette chapters 14.1, 14.2

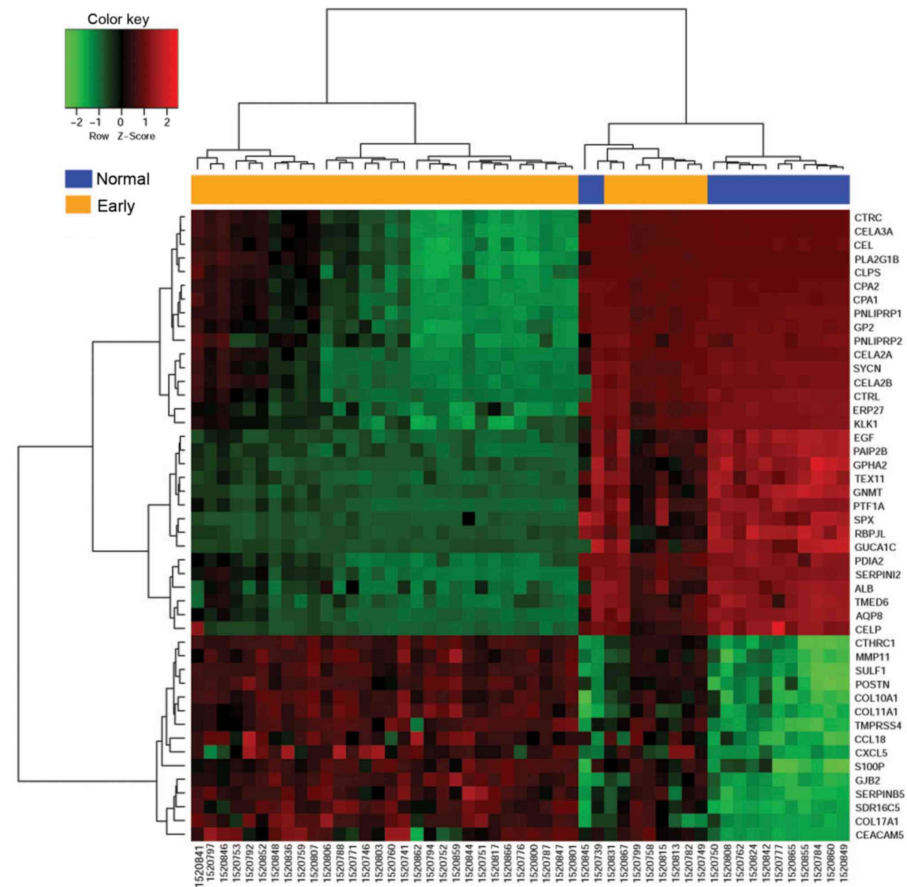
Read Durbin et al. chapter 7.3 (first half)

*alternative reading:* Hütt/Dehnert chapters 3.3.1 – 3.3.3

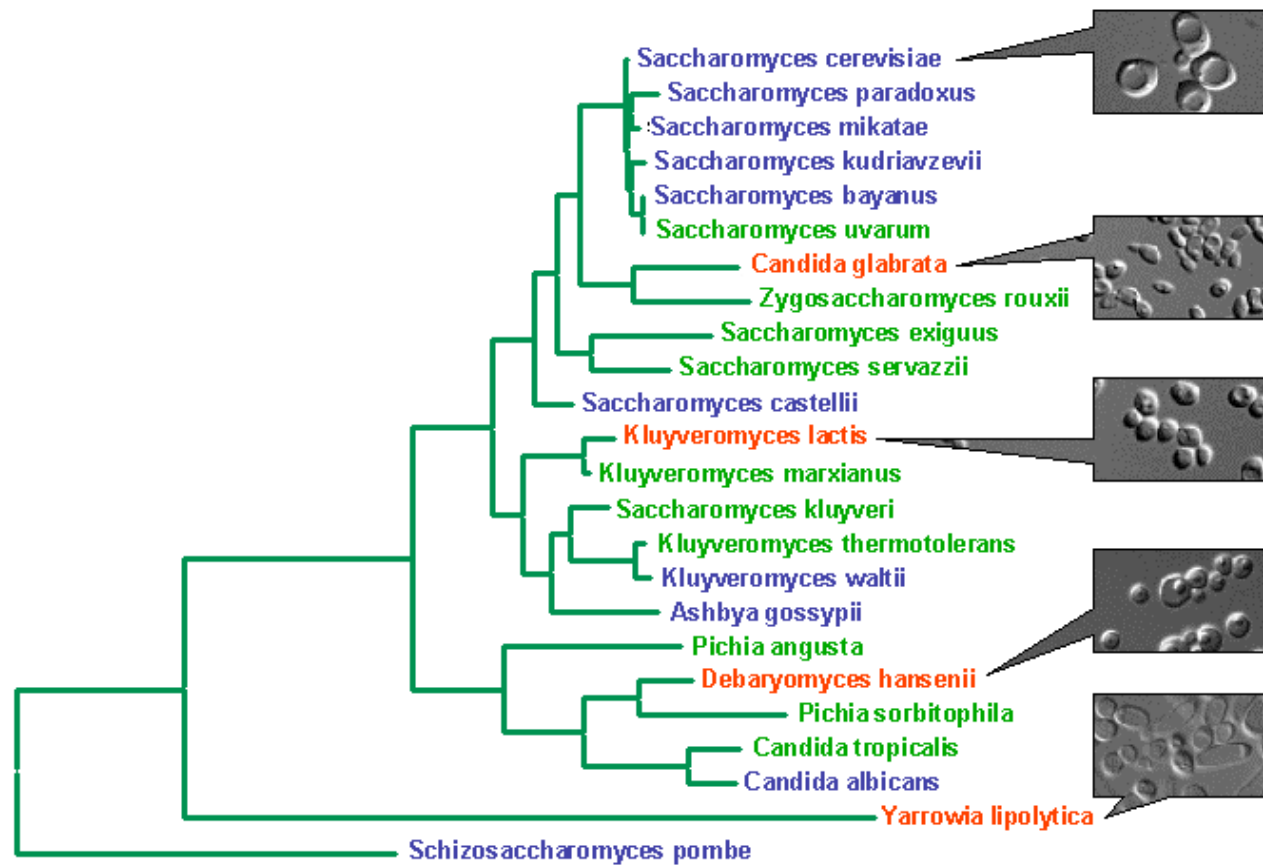
**One more item in the bioinformatics toolbox:**  
**basic principles of clustering algorithms**



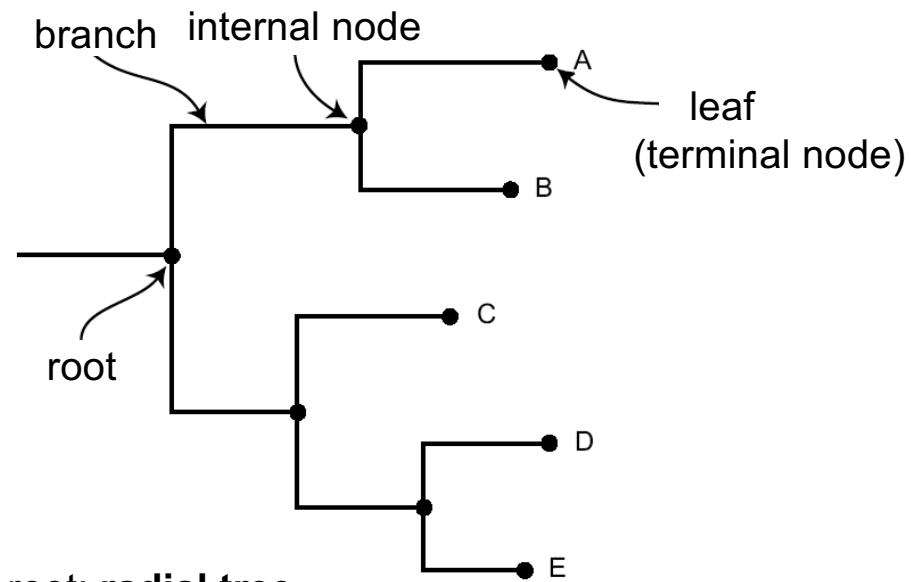
Miao, Z., Zhang, X. Differential expression analyses for single-cell RNA-Seq: old questions on new data. *Quant Biol* **4**, 243–260 (2016).



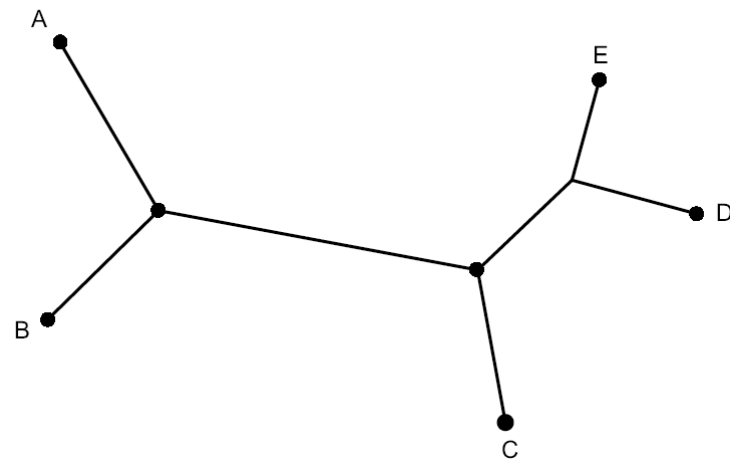
Wu, J., Li, Z., Zeng, K., Wu, K., Xu, D., Zhou, J., & Xu, L. (2019). Key genes associated with pancreatic cancer and their association with outcomes: A bioinformatics analysis. *Molecular Medicine Reports*, **20**, 1343–1352.



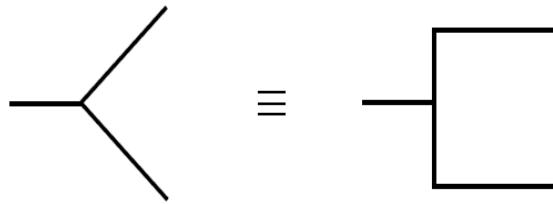
cbi.labri.fr



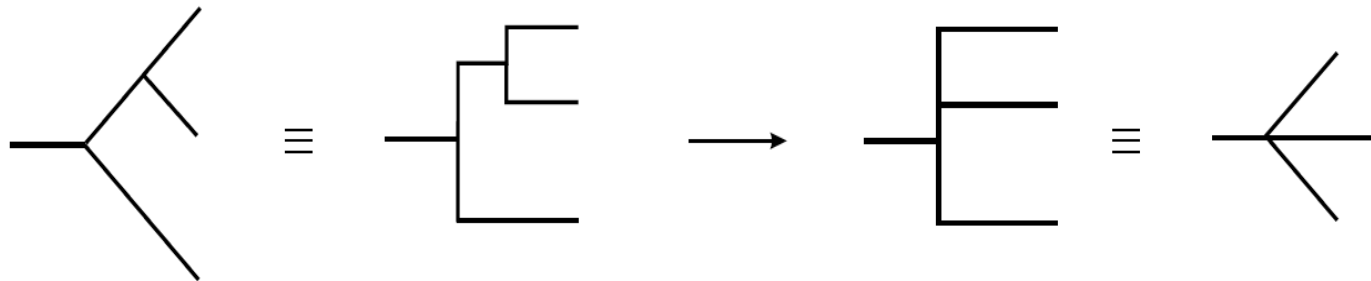
eliminating the root: **radial tree**



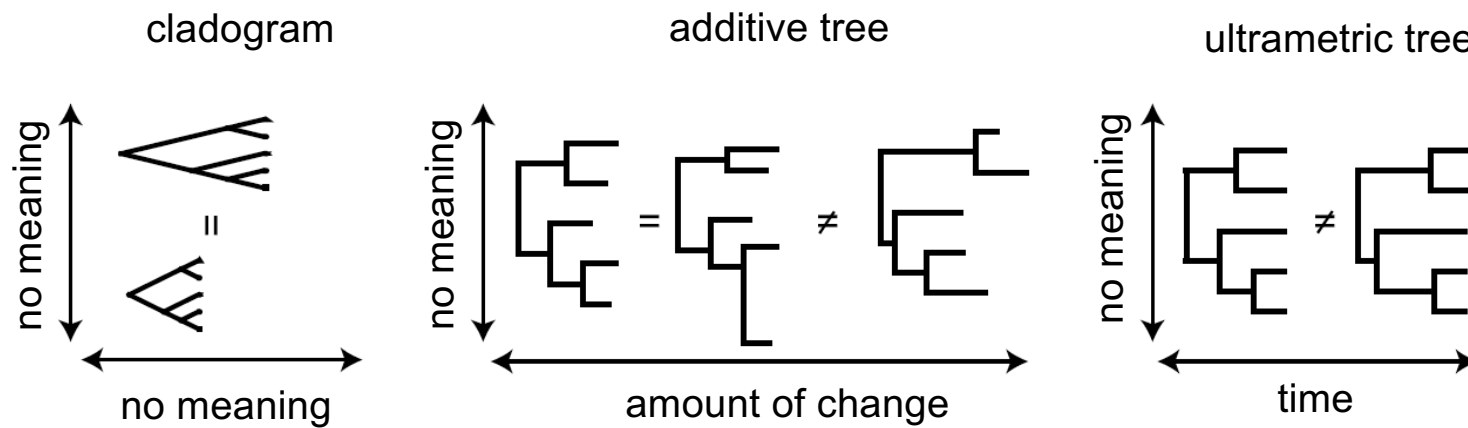
a



b

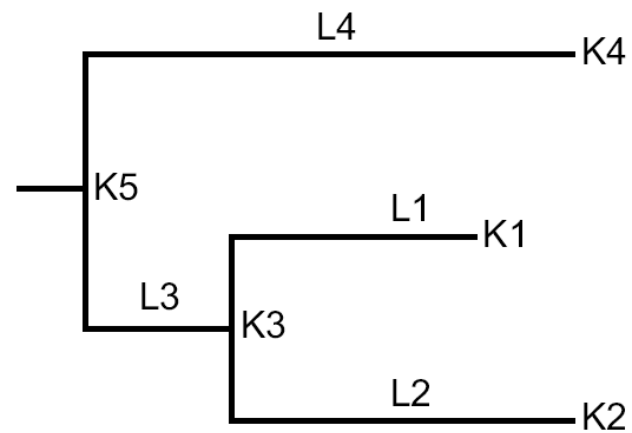


## What do the axes mean in a tree?





## Newick notation for a tree

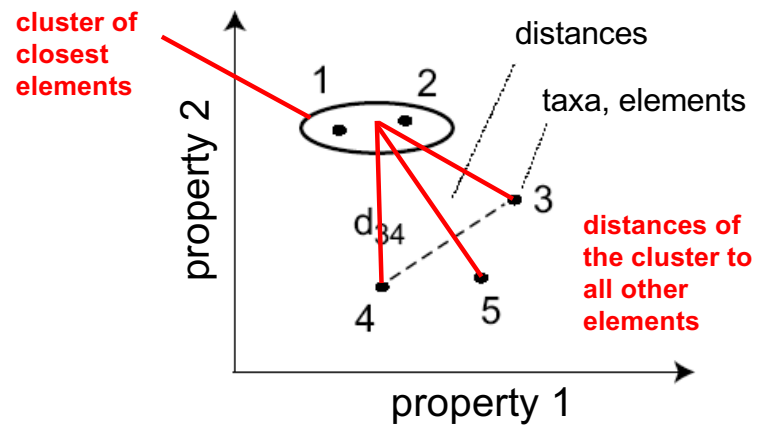


$(K4, (K2, K1))$     **or**     $((K1, K2), K4)$     **or**     $(K4, (K1, K2))$

$(K4 : L4, (K1 : L1, K2 : L2) : L3)$

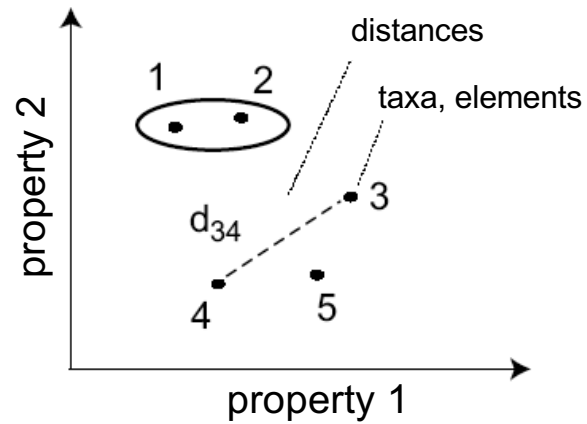
## UPGMA

### observables

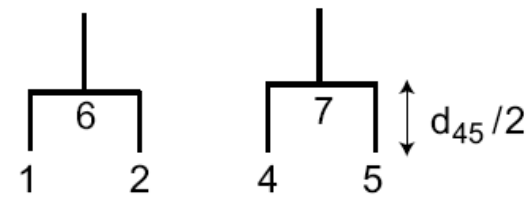
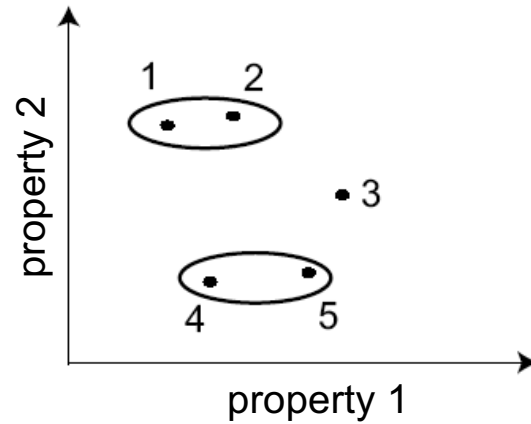
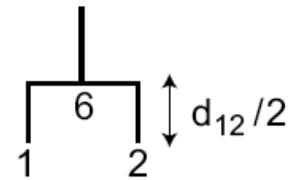


# UPGMA

## observables

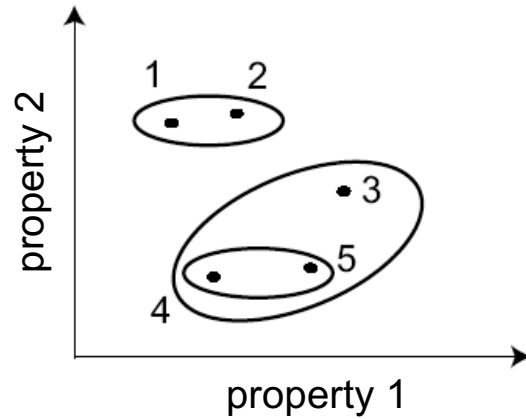


## tree



# UPGMA

## observables



## tree

