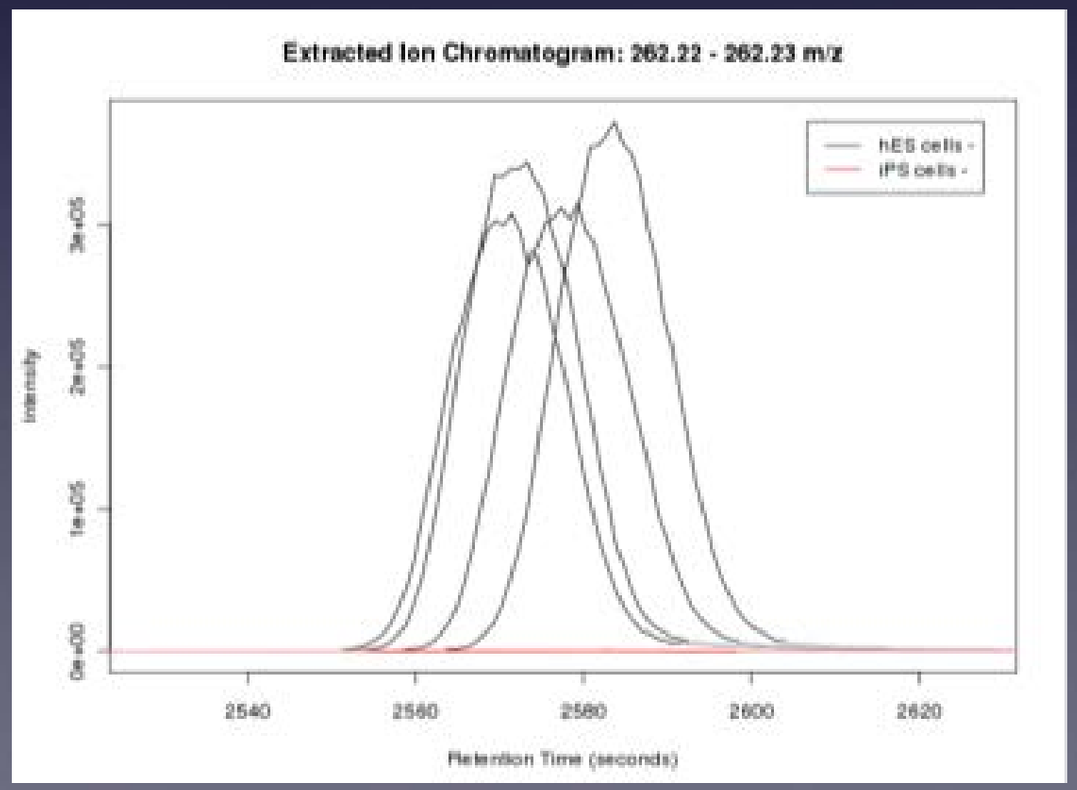
History of XCMS
and
Ideas of LC-MS Processing

H. Paul Benton

- Developed 2004/2006
  - Colin Smith (TSRI)

- 100+ HPLC-MS
  - Single Quad data
- Retention time shifts

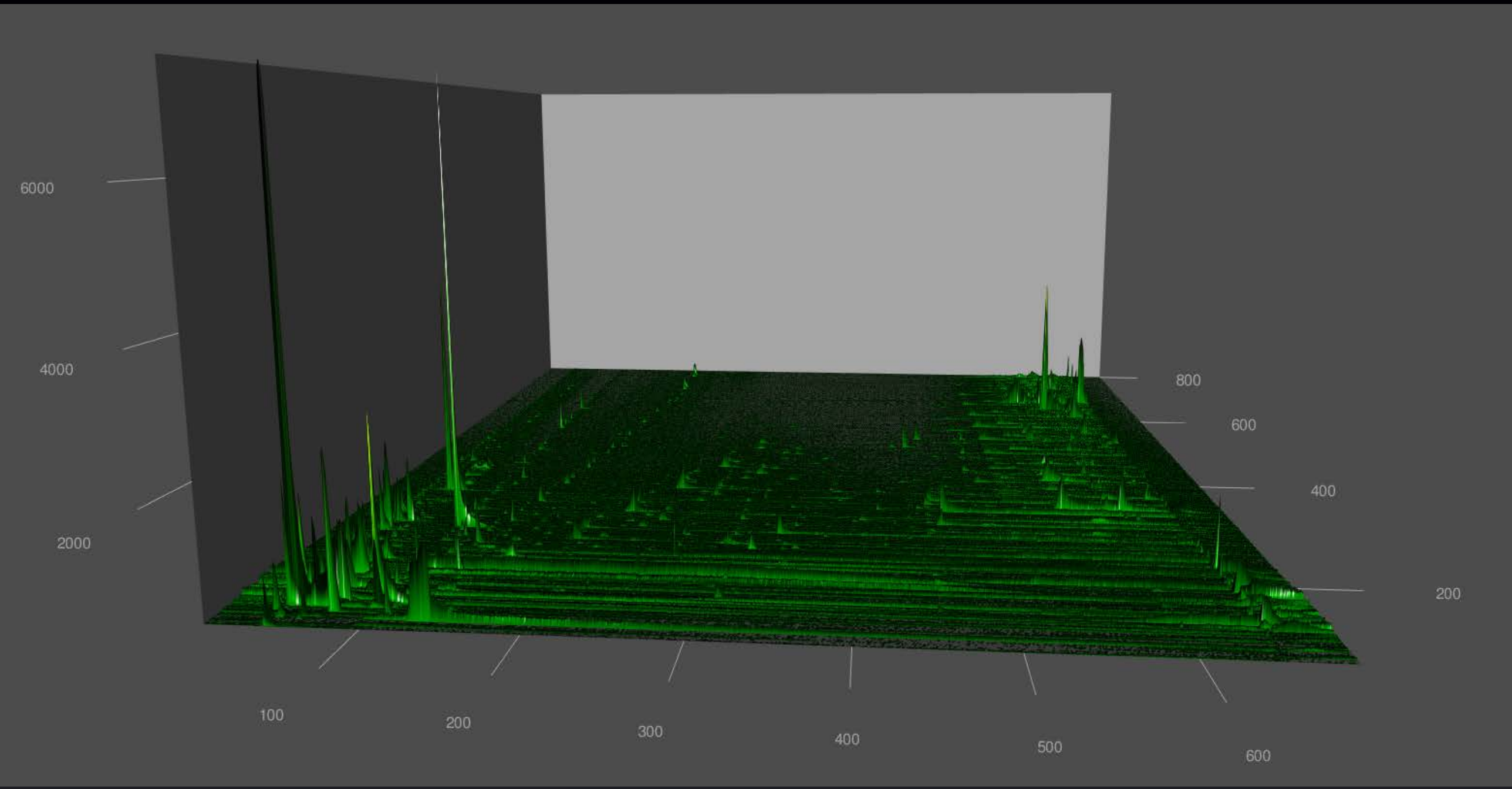


History

- Development of Kernel density peak grouping and LOESS Retention time alignment - 2004 spring
- Checked into Bioconductor - Mid 2005
- CentWave peak detection added - Mid 2007
- Tandem mass data - Mid/late 2007
- Parallel processing - Mid 2008
- OBI-Warp Retention time alignment added - Early 2009
- XCMS Online Work started - Mid 2009

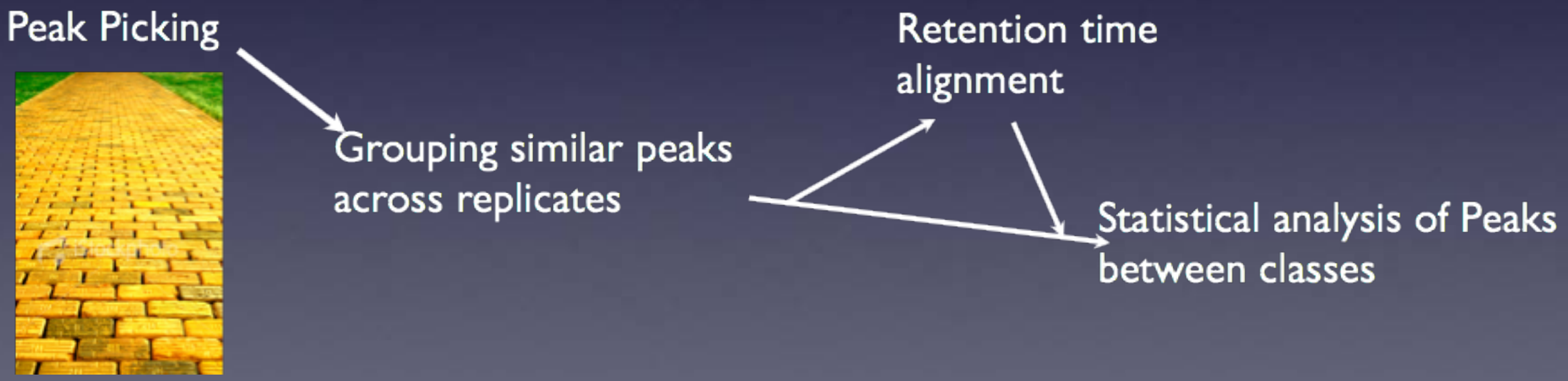
Why do you need XCMS?

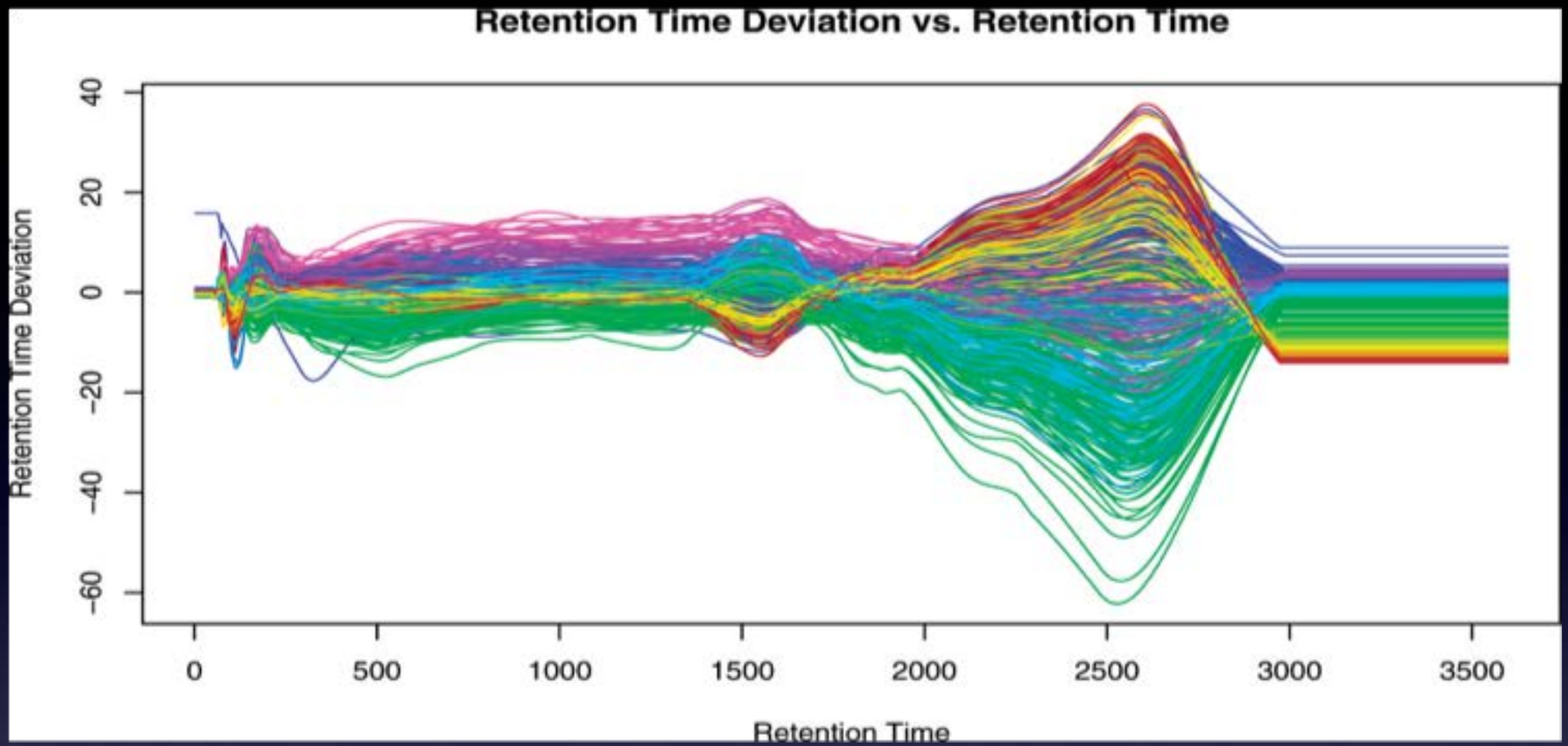
- Data processing
  - Complex data complex files - 100s - 1000s - 10000s
  - Statistical analysis
- Retention time alignment
  - Non-linear alignment


6000
4000
2000
100
200
300
400
500
600
800
600
400
200

General principals

- What is the general workflow?





- Retention time shifts in a nonlinear fashion between sample during the LC-MS run
- A nonlinear model is needed to align these chromatograms
- HPLC shifts can be as big as 30s!

Modular

- XCMS was written in a language called ‘R’
- It is Open Source GNU code
  - This means you can read all the code
- ‘R’ makes writing your own extension *easier*
- It is however command line based
  - You have to know the commands to run it

In action

- Questions?
- Next, Basic use of XCMS - locally