

# OpenMS

Development of a Quality Control Pipeline for  
Liquid Chromatography – Mass Spectrometry

**Chris Bielow**

Proteomics

Genomics

Transcriptomics

“OMICS”

Connectomics

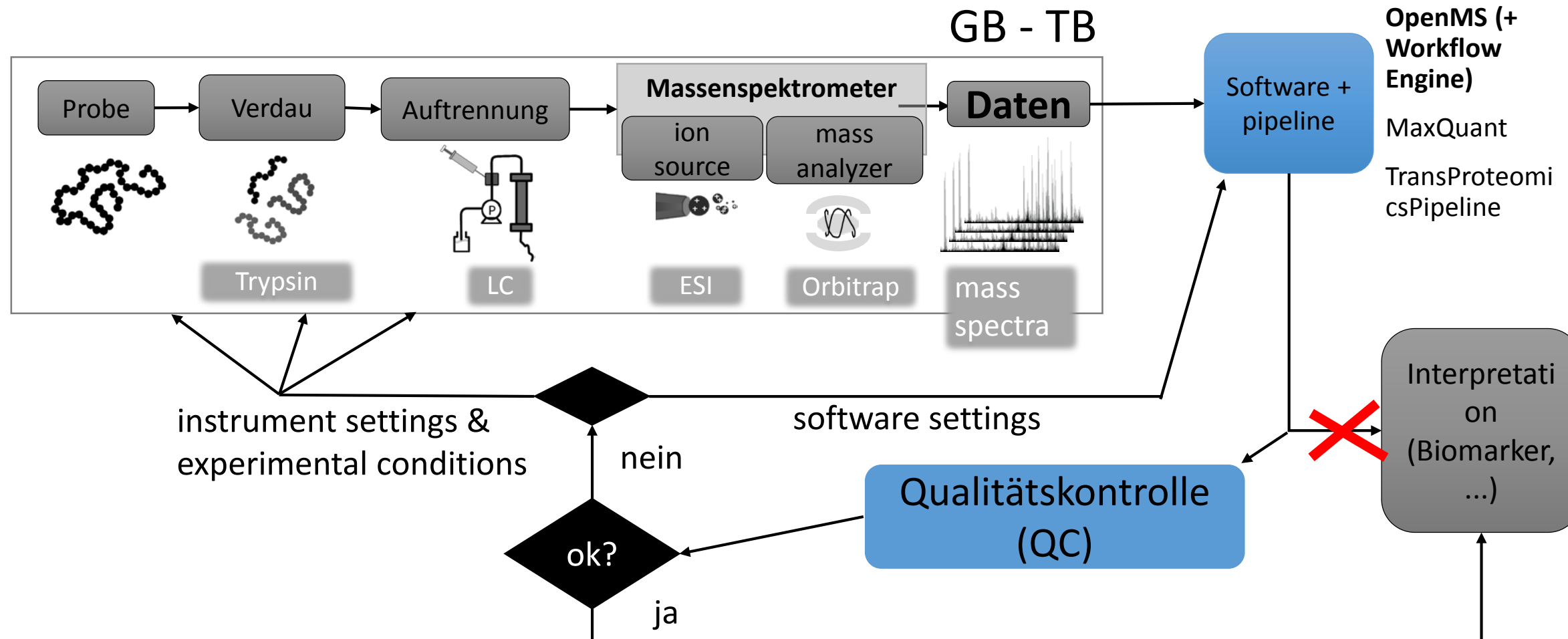
Foodomics

Metabolomics

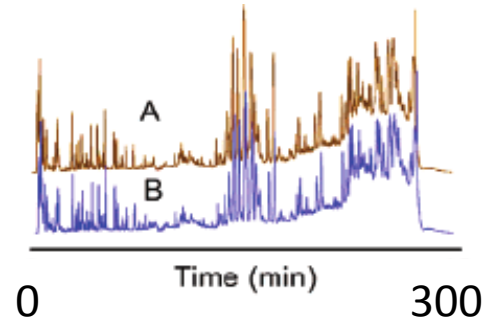
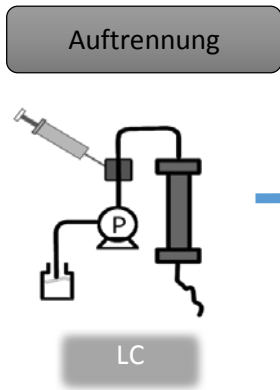
OMICS – Ganzheitliche Charakterisierung und Quantifizierung des entsprechenden Teilgebiets

# Proteomik-Pipeline

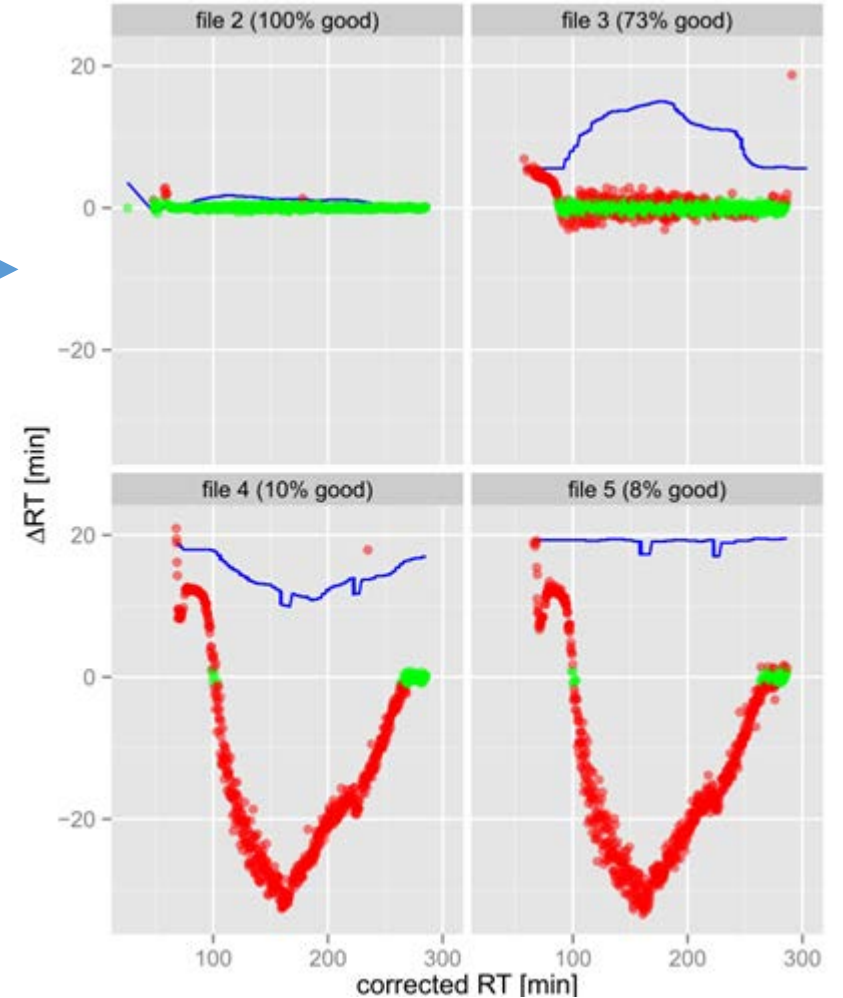
Mittel der Wahl: Liquid Chromatography (LC) + Massenspektrometrie (MS)



# QC Beispiel – Alignment von Retentionszeiten

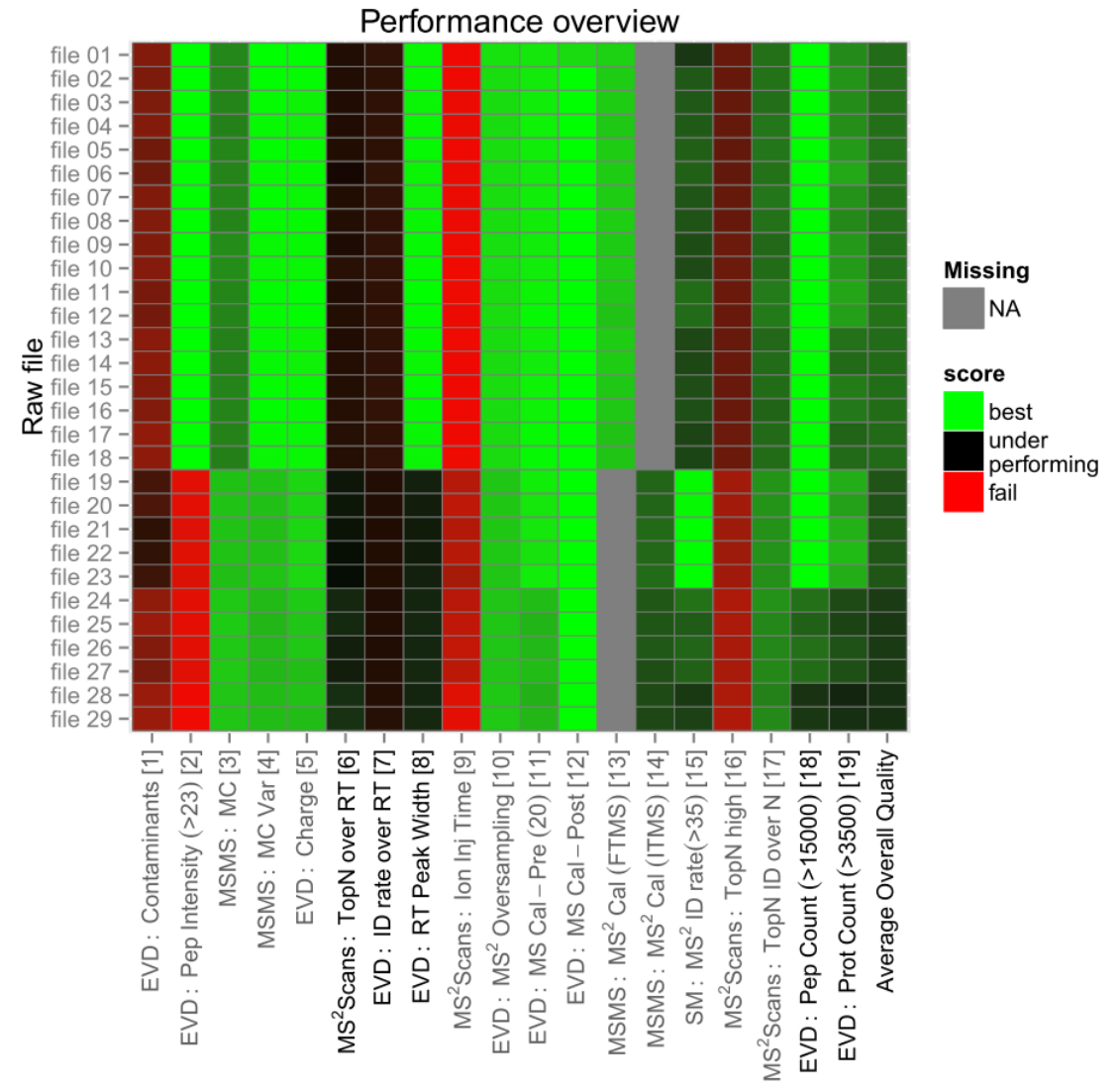


align

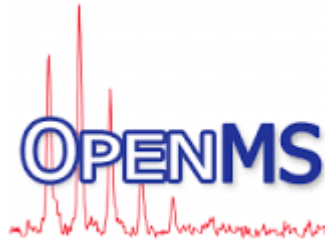


# QC Beispiel - Gesamtübersicht

- Automated scoring of results using adequate



# Software Pipeline (KNIME)



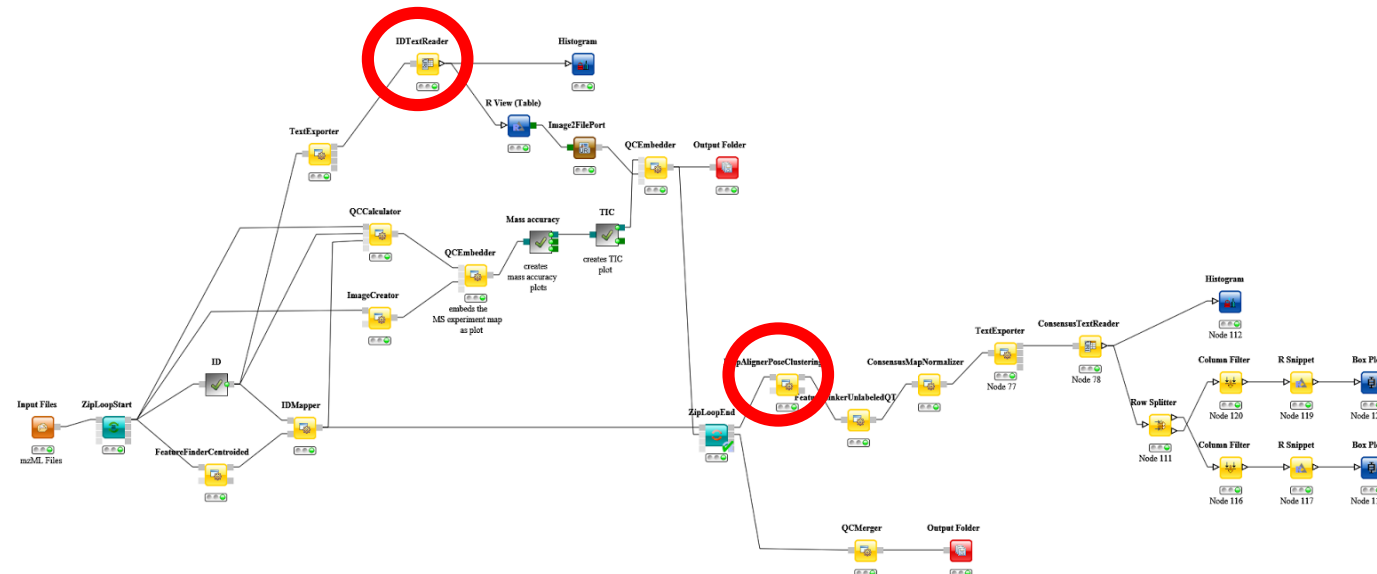
## OpenMS

- C++ Bibliothek für Analyse von LC-MS Daten
- Tools/Executables
  - 150 Tools
  - „Lego“



KNIME (konstanz information miner)

- Pipeline tool
- Visualisierung von Daten



# Zielstellung

- Implementierung von bekannten (oder neuen) Metriken zur Qualitätssicherung von LC-MS Experimenten
  - OpenMS C++
  - ca. 1-4 Metriken pro Team
- Optimierung der Programmlaufzeit (C++ Code Profiling; V-Tune, Xperf,...)
  - “Ausgleich” für zusätzlichen Aufwand durch Erzeugung der Metriken
- Visualisierung der Daten
  - Plot (R)
- Zusammenführung der Metriken in einem automatisierten Report
  - Abstimmung zwischen den Teams
- Erstellung eines Pull-Requests zur Aufnahme der neuen Funktionalität in die OpenMS-Entwicklerversion

# Zeitplan - Seminar

Mi, 01.03., 10:00 Uhr (flexibel +/- 2 Wochen)

- Vorbesprechung

Mo, 19.3. - Fr, 23.3

- Einführung in Flüssigkeitschromatographie & Massenspektrometrie (LC-MS)
- Refresher: C++ und R
- OpenMS (C++ library und Tools) & PTXQC (QC Tool in R)
- Workflow engines: KNIME und TOPPAS
- Git (version control software)
- Quality metrics



# Zeitplan – praktischer Teil

Do, 29.03.

Vorstellung der Pläne

Mo, 09.04 - Fr, 11.05.

## Woche 1(+2):

- OpenMS-develop installieren
- Projekt auf <https://git.imp.fu-berlin.de> clonen
- Einarbeitung in OpenMS
- Relevante Literatur lesen
- (Externe Programme testen)
- Metriken aussuchen, Zeitplan erstellen, Vortrag

## Woche 3+4

- Neue Metriken implementieren (C++/R)
- Visualisieren
- Tests schreiben

## Woche 5

- Workflows finalisieren (C++/R); evtl. Bugs fixen;
- Abschlussbericht schreiben (Resultate, Metriken, Screenshots, Workflows)

Fr, 18.5. Abschlussbericht und Vorstellung der Ergebnisse

## wöchentliche Besprechung

Freitags: 10-12 Uhr,

Weiteres Vorgehen, Anpassung  
des Zeitplans, Problemlösung

# Vorraussetzungen

- Erfahrung in Objektorientierter Programmierung (Java, C++, ...)
  - C++ Kenntnisse empfehlenswert (wenig Templates)
  - z.B. ALDaBi Praktikum / C++-Kurs

# Empfehlenswert

- R Grundwissen (ggplot2)
- Erfahrung mit Profilern

**Schwierigkeitsgrad (Acht Sterne verteilt auf drei Bereiche):**

A Programmieren: \*\*\*\*

B Biologie/Chemie: \*\*

C Projektmanagement: \*\*

# Fragen?

- mehr Fragen
  - Raum K21,T9
  - [chris.bielow@fu-berlin.de](mailto:chris.bielow@fu-berlin.de)

## Mehr Informationen

- OpenMS
  - [www.OpenMS.de](http://www.OpenMS.de)
- KNIME (konstanz information miner)
  - [www.knime.org](http://www.knime.org)

B.Sc. Arbeit möglich