Discussion on harmonisation potential and needs

Peter Haglund
Umeå University, Sweden

Why harmonize?

- Avoid loss of information
- Comparability of data
- Exchange of data
- Data mining
- Emerging (hazardous) pollutant discovery

Outline

Emerging
Priority
Compound

What need to be harmonized? (What affects the outcome?)

- Nomenclature, quality indicators
- Sampling
- Sample reconstitution
- Chromatography
- Ionization techniques
- MS data collection
- Common suspect list
- LC/GC-MS data processing
- Ranking of candidates
- Data reporting
- Spectra storage
- Raw data storage
- Prioritization of NTS substances
- Possibility for data mining…

What is possible to harmonize?

Area	Possible
Nomenclature, quality indicators	
Sampling	
Sample reconstitution	
Chromatography	
Ionization techniques	
MS data collection	
Common suspect list	
Data processing	
Ranking of candidates	
Data reporting	
Spectra storage	
Raw data storage	
Prioritization of NTS substances	
Data mining tools	

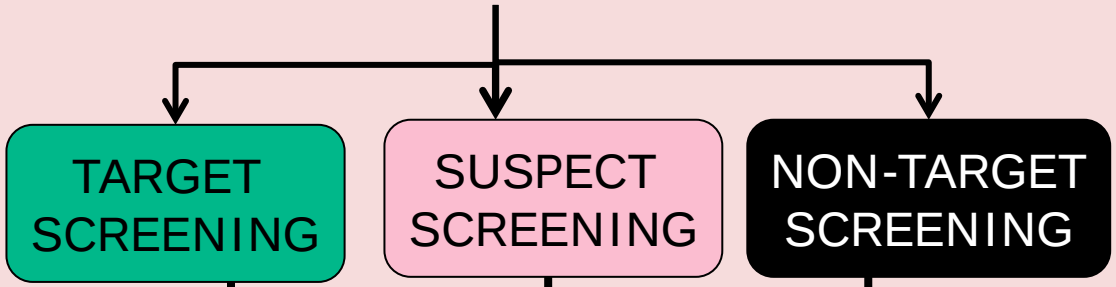
How difficult is it?

Area	Possible?	Easy?
Nomenclature, quality indicators		
Sampling		
Sample reconstitution		
Chromatography		
Ionization techniques		
MS data collection		
Common suspect list		
NTS workflows		
Ranking of candidates		
Data reporting		
Spectra storage		
Raw data storage		
Prioritization of NTS substances		
Data mining tools		

Nomenclature / Quality Indicators

- Suggestion: For all data. Indicate workflow (Target, Suspect, NTS) and confidence (Level, IP)
- Further needs: Refine definitions and quality indicators
- Lead: EAWAG?

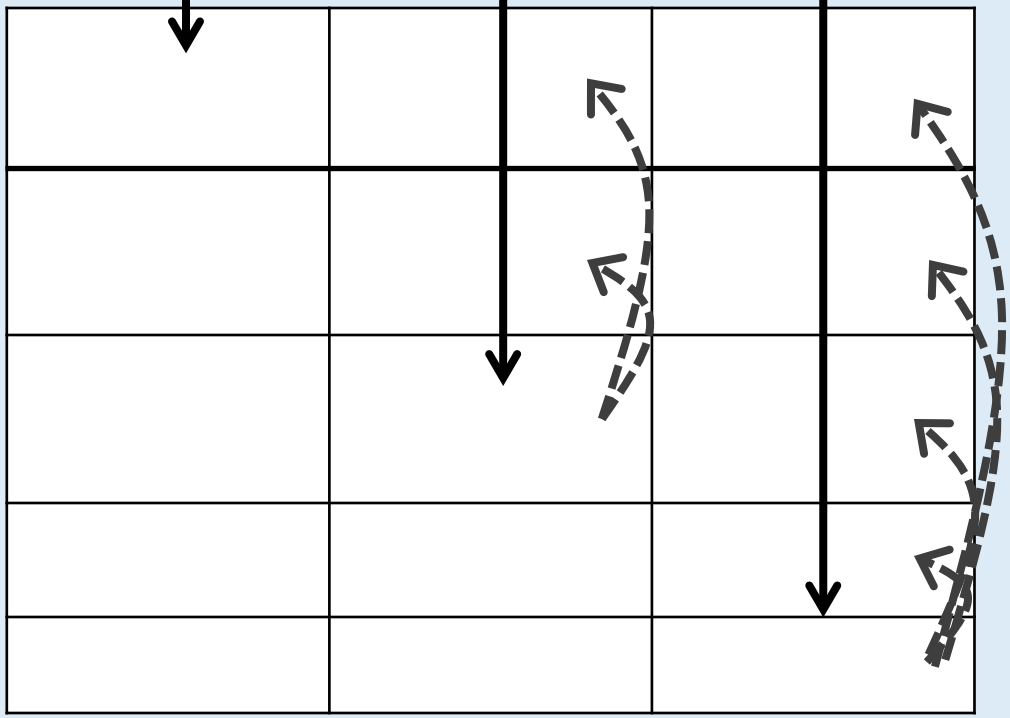
Workflow / Identification confidence (points)



Available data before analysis:

- **List of targets**
- **List of suspect**

Available data after analysis:



Identification confidence

Level 1: ***Confirmed structure*** by reference standard

Level 2: ***Probable structure*** by library/*diagnostic evidence*

Level 3: ***Tentative candidate(s)*** structure, substituent, class

Level 4: ***Unequivocal molecular formula***

Level 5: ***Exact mass*** of interest

Sampling

- Suggestion: For GC-MS include particulate matter (whole water).

Sample reconstitution

- Suggestion: Reconstitute in a strong solvent (mixture)

- Suggestions?
  - LC
  - GC

Chromatography

- Suggestion:
  - Inject as much as needed to get enough sensitivity
  - Include retention index compounds
  - Common internal standard(s) for semi-quan
  - GC: Use high% methyl-polysiloxane
  - LC: Use C18 for LC
    Use slow gradients (normal HPLC or long-run UHPLC)
    Methanol probably preferrable

- Further developments:
  - HILIC / Mixed mode phases for very polar compounds
  - Evaluate micro/nano-flow
  - Evaluate two-dimensional separations

- Lead: Labs with research interests in the areas

Ionization techniques

- Suggestion: Use ESI +/- and EI
- Further needs:
  - Molecular ion information in GC
  - Evaluation of new (universal) soft ionization techniques
- Lead: UJI and UmU...?

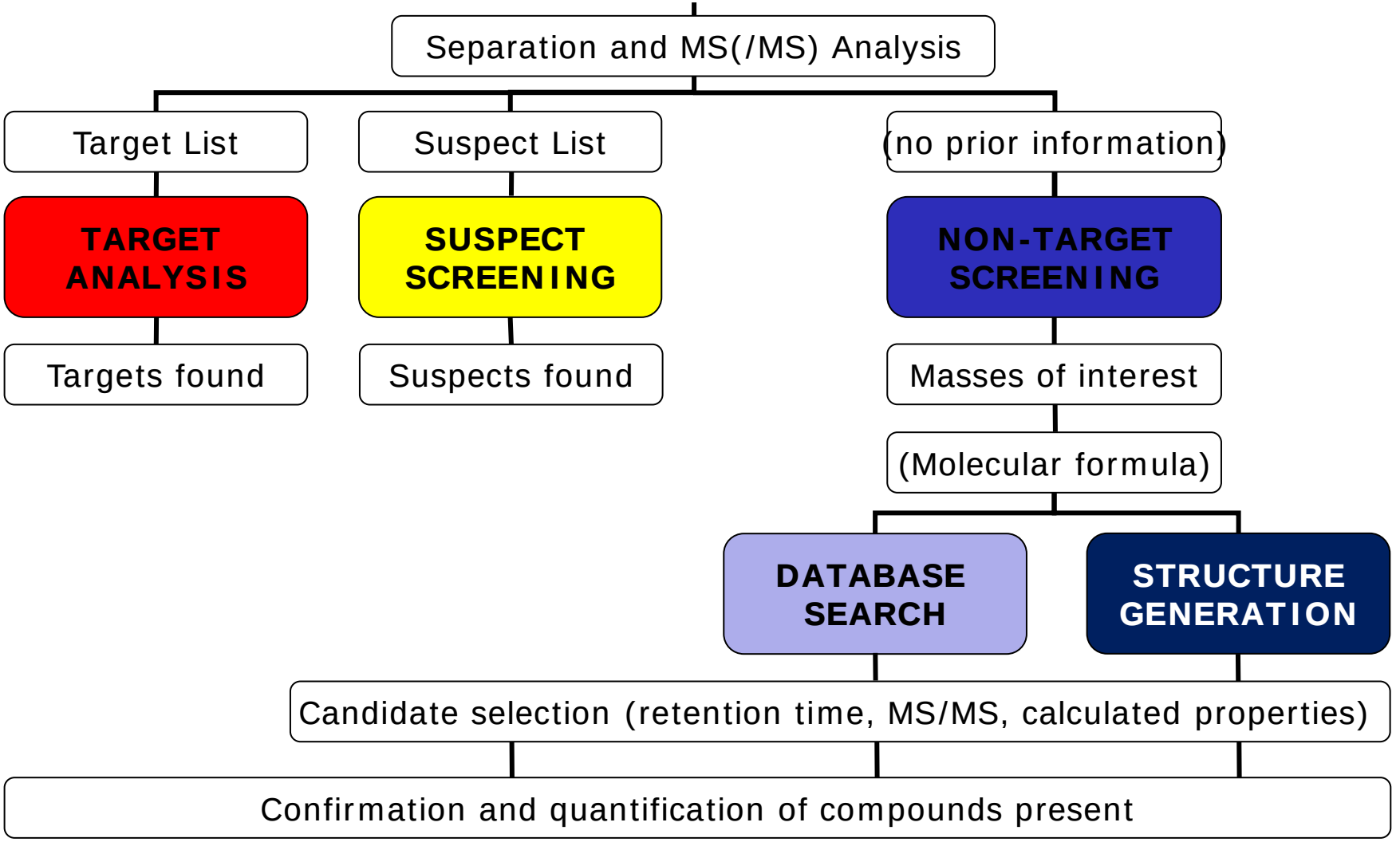
MS data collection

- Suggestion:
  - Use the highest resolution possible (40–60,000 feasible)
  - Collect data over 35 – 2000 (if possible)

- Further needs:
  - For NTS evaluate (segmented) DIA + deconvolution

- Lead: Labs with research interests in the area?

Workflows


Separation and MS(/MS) Analysis
Target List
Suspect List
(no prior information)
TARGET ANALYSIS
SUSPECT SCREENING
NON-TARGET SCREENING
Targets found
Suspects found
Masses of interest
(Molecular formula)
DATABASE SEARCH
STRUCTURE GENERATION
Candidate selection (retention time, MS/MS, calculated properties)
Confirmation and quantification of compounds present

Suspect workflow

- Suggestion:
  - **Screen smart** first
  - **Screen big** if time allows

- Further needs:
  - Common suspect lists
  - Useful retention time information format
  - Easy exchange of retention and MS/MS-MS information (share in-house databases)

- Lead: Retention time UfZ (Martin), Common Suspect list LfU StoffIdent…?

Common suspect list

- Suggestion: Yes?
- Further needs: Ranking/prioritization scheme
  - EU priority (information from European institutions?)
  - Emission potential (information from European institutions?)
  - Occurrence
  - Aquatic effects
  - Newly discovered compounds of concern
- Lead: ?

Ranking of NTS candidates

- Suggestion:
  - Use literature citations
  - Retention time information
  - In-silico fragmentation
  - Toxicity prediction?
  - Exposure: Occurrence, formation during processes

- Further needs:
  - Best model to weigh information including statistical tools
  - Can we benefit from BINGO?

- Lead: IPB, UFZ, EAWAG, Labs with research interests in the area ?

Data reporting (NTS trials)

- Suggestion: Improve reporting template, automate as much as possible
- Further needs: automatic upload should be possible
- Lead: EI with input from bioinformatics (IPB)?

Spectra storage & exchange

- Suggestion: Use MassBank

- Further needs:

- Import routines (RMassBank) are there but not yet fully implemented

  ⇨ MassBank workshop today!

- Willingness to share and invest time for that
- How can we benefit from well curated mzCloud?

- Lead: MassBank UFZ/Eawag/IPB

Raw data storage

- Suggestion: We provide an option to store raw data

- Further needs:
  Selection of appropriate storing place,
  Decision on meta data
  Development of tools to retrieve information

- Lead: UFZ, IPB?

Prioritization of NTS substances

- Suggestion:
  - Develop/use a common Prio scheme
  - Find synergies
  - Exchange information on NTS (exact mass or MS/MS,...)
- Further needs: platform to exchange NTS in a regular period
- Lead: NIVA, Norman?

Data mining tools

- Suggestion:
- Further needs:
  - Tools to interrogate raw data to find, semi-quantify and trend new and emerging pollutants in digital archives
- Lead: IPB Halle?