The FoodBAII online resources to support discovery of novel dietary biomarkers with metabolomics

Claudine MANACH
Human Nutrition Unit, UMR1019 Clermont-Ferrand
FOODBALL PROJECT - THE FOOD BIOMARKER ALLIANCE

- Janv 2015-Dec 2017. 4,38M€
- 22 partners, PI: E. Feskens, Wageningen, NL

Aim:
Identification and validation of dietary biomarkers using metabolomics

- Intervention studies in 7 centers to identify new biomarkers of food intake
- Inventory of known dietary biomarkers and new validation scoring system
- New tools and resources for research on the food metabolome and biomarker discovery (WP4)
NEW TOOLS: PRIORITY NEEDS IDENTIFIED

1. A compound database with extensive coverage of the food metabolome

2. A food intake biomarkers database

3. A resource to facilitate sharing of standards

4. A resource to facilitate knowledge and data sharing within the FoodBAI community and beyond
- Compound databases are essential to obtain hypotheses of identification for discriminant ions associated with the intake of a given food or dietary pattern.

- Must be as comprehensive as possible, and contain chemical, analytical, and biological data that help to rapidly select the best hypotheses of identification to be further validated.
1- DATABASES TO OBTAIN HYPOTHESES OF IDENTIFICATION: FOODB

- Open access: [www.foodb.ca/](http://www.foodb.ca/)
- Developed by Univ of Alberta
- The most comprehensive database on food constituents
- Covers all types of food constituents
1- DATABASES TO OBTAIN HYPOTHESES OF IDENTIFICATION:

**FooDB**

- **26630 compounds**
- **907 foods**
- **~700,000 concentration values**
1- DATABASES TO OBTAIN HYPOTHESES OF IDENTIFICATION: FooDB

- Chemical data
  - Chemical Names and synonyms
  - Chemical structures
  - Chemical taxonomies
  - Physical properties
  - NMR and MS spectra

- Spectral data

- Biological data
  - Compound description
  - Associated foods
  - Biological effects (Duke’s DB)
  - Flavor
  - Links

- Every metabocards contains > 110 data fields

Showing Compound (±)-Catechin (FDB004453)
Experimental spectra (Quercetin)

- Experimental spectra (Quercetin)
- Predicted spectra (Quercetin 3-glucuronide)

~5000 predicted MS/MS from CFM-ID

Competitive Fragmentation Modeling for Metabolite Identification
http://cfmid.wishartlab.com/
1- Databases to obtain hypotheses of identification: FooDB

Spectra Search Mass Spectrum

- Query Masses (Da):
  - 176.01
  - 238.19
  - 420.16
  - 780.32
  - 956.25
  - 1100.45

- Ionization:
  - Ion Mode: Positive

- Molecular Weight Tolerance ±:
  - 0.05

- Adduct Type:
  - Unknown

Spectra Search Tandem Mass Spectrum

- Parent Ion Mass (Da):
  - 14.05

- Parent Ion Mass Tolerance ±:
  - 0.1

- Ionization:
  - Positive

- CID Energy:
  - Low

- MS/MS Peak list (M/Z RI):
  - 40.048 0.174
  - 56.022 0.424
  - 88.375 0.388
  - 101.508 0.255
  - 112.401 0.775
  - 120.670 100.000
  - 146.966 20.070

- Mass/Charge (m/z) Tolerance ±:
  - 0.5

- Include predicted spectra? [ ]
Many dietary biomarkers are expected to be metabolites of food constituents.
Information is largely missing on metabolites of many non-nutrients.

BioTransformer, an open-access software to predict metabolic by-products of a given compound.
**1- DATABASES TO OBTAIN HYPOTHESES OF IDENTIFICATION: PHYTOHUB**

- Food phytochemicals and their human metabolites
- Designed for metabolomics

**PhytoHub** is a freely available electronic database containing detailed information about dietary phytochemicals and their human metabolites. Around 1,000 dietary phytochemicals are included which represent all the polyphenols, terpenoids, alkaloids and other plant secondary metabolites commonly ingested with the human diet. For each phytochemical, the following will be available: 1) the main dietary sources (extracted from the literature and online databases such as FooDB and Phenol-Explorer) with a direct link to FooDB food cards, 2) known human metabolites (extracted from literature and Phenol-Explorer), 3) predicted in silico metabolites, generated by an in-house tool that uses experimental knowledge of host and microbial metabolism for the various families of phytochemicals, 4) physico-chemical data such as solubility and physiological charge, 5) monoisotopic mass and spectral data (collected from libraries of spectra such as MassBank and ReSpect (RIKEN MSn spectral database) and from our mass spectrometry/metabolomics laboratory and databases.

- Initially developed at **INRA**
- Now a collaborative database

[www.phytohub.eu](http://www.phytohub.eu)
> 1230 dietary phytochemicals

> 660 Polyphenols
- Anthocyanins
- Flavanols
- Flavanones
- Flavones, Flavonols
- Isoflavones
- Other flavonoids
- Hydroxybenzoic acids
- Hydroxycinnamic acids
- Other phenolic acids
- Ellagitannins
- Stilbenes
- Lignans
- Coumarins

> 240 Terpenes
- Monoterpenoids
- Diterpenoids
- Sesquiterpenoids
- Triterpenoids
- Carotenoids
- Phytosterols

> 180 N-containing compounds
- Alkaloids
- Betalaines
- Purine & pyrimidines
- Aminoacids
- Amines
- Glucosinolates
- Thiosulfimates
- Phytoprostanes
- Organic acids
...

> 70 Miscellaneous

PHYTOHUB (www.phytohub.eu)

Chemical structures

In silico predicted metabolites

Physico-chemical properties

Known metabolites

Spectral data

Dietary sources

Compounds
Foods
Search
Advanced Search
About
Glossary
Contact Us
Data curation by experts of the different phytochemical families

<table>
<thead>
<tr>
<th>Name</th>
<th>Monoisotopic mass</th>
<th>Formula</th>
<th>Species</th>
<th>Biofluids</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carvone</td>
<td>150.10446507</td>
<td>C_{13}H_{18}O</td>
<td>rabbit, dog, guinea pig</td>
<td>urine</td>
</tr>
<tr>
<td>Carvone (trans-)</td>
<td>152.120115134</td>
<td>C_{13}H_{18}O</td>
<td>rabbit, rat</td>
<td>urine</td>
</tr>
<tr>
<td>Limonene-1,2-diol</td>
<td>170.13067902</td>
<td>C_{13}H_{18}O_2</td>
<td>human, rat</td>
<td>plasma</td>
</tr>
<tr>
<td>Limonene-6,9-diol</td>
<td>179.13067902</td>
<td>C_{13}H_{18}O_2</td>
<td>rat, rabbit, human, dog</td>
<td>urine, plasma</td>
</tr>
<tr>
<td>Limonene-6,9-diol glucuronide</td>
<td>346.162767008</td>
<td>C_{13}H_{20}O_3</td>
<td>human, rabbit, dog, rat</td>
<td>urine</td>
</tr>
<tr>
<td>Limonene-10-c</td>
<td>152.120115134</td>
<td>C_{13}H_{18}O</td>
<td>human, rabbit, dog, rat</td>
<td>urine</td>
</tr>
<tr>
<td>Limonene-10-c-glucuronide</td>
<td>328.12203122</td>
<td>C_{13}H_{20}O_7</td>
<td>human, rabbit, dog, rat</td>
<td>urine</td>
</tr>
<tr>
<td>Pentyl alcohol</td>
<td>152.120115134</td>
<td>C_{13}H_{18}O</td>
<td>rat, human</td>
<td>urine</td>
</tr>
<tr>
<td>Pentyl acid</td>
<td>166.090379992</td>
<td>C_{13}H_{18}O_2</td>
<td>human, rabbit, dog, rat</td>
<td>urine, plasma</td>
</tr>
<tr>
<td>Pentyl acid glucuronide</td>
<td>342.13146768</td>
<td>C_{13}H_{20}O_3</td>
<td>human, rabbit, rat</td>
<td>urine</td>
</tr>
<tr>
<td>Perillyl glycerine</td>
<td>223.120843415</td>
<td>C_{13}H_{18}O_3</td>
<td>rat, rabbit, dog</td>
<td>urine</td>
</tr>
<tr>
<td>Perillyl acid-8,9-diol</td>
<td>200.104859</td>
<td>C_{13}H_{18}O_4</td>
<td>rat, rabbit, dog</td>
<td>urine</td>
</tr>
<tr>
<td>Dihydropentillic acid (trans-)</td>
<td>168.115029756</td>
<td>C_{13}H_{18}O_2</td>
<td>rat, human</td>
<td>urine, plasma</td>
</tr>
<tr>
<td>Dihydropentillic acid (cis-)</td>
<td>168.115029756</td>
<td>C_{13}H_{18}O_2</td>
<td>rat, human</td>
<td>urine, plasma</td>
</tr>
<tr>
<td>Dihydropentillic acid glucuronide (trans-)</td>
<td>344.147117744</td>
<td>C_{13}H_{20}O_3</td>
<td>human</td>
<td>urine</td>
</tr>
<tr>
<td>Dihydropentillic acid glucuronide (cis-)</td>
<td>344.147117744</td>
<td>C_{13}H_{20}O_3</td>
<td>human</td>
<td>urine</td>
</tr>
</tbody>
</table>
Data curation by experts of the different phytochemical families
<table>
<thead>
<tr>
<th>Compound family</th>
<th>Experts</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alkylresorcinols</td>
<td>Rikard Landberg</td>
<td>Swedish University of Agricultural Sciences, Uppsala, Sweden</td>
</tr>
<tr>
<td>Anthocyanins</td>
<td>Wieslaw Wiczkowski</td>
<td>Institute of Animal Reproduction and Food Research of the Polish Academy of Sciences, Olsztyn, Poland</td>
</tr>
<tr>
<td>Betalaines</td>
<td>Wieslaw Wiczkowski</td>
<td>Institute of Animal Reproduction and Food Research of the Polish Academy of Sciences, Olsztyn, Poland</td>
</tr>
<tr>
<td>Carotenoids</td>
<td>Catherine Caris, Rachel Kopec</td>
<td>INRA Avignon, UMR408, France</td>
</tr>
<tr>
<td>Ellagitannins</td>
<td>Paco Tomas-Barberan, Rocío García-Villalba</td>
<td>CEBAS-CSIC, Murcia, Spain</td>
</tr>
<tr>
<td>Flavonol metabolites</td>
<td>Peter Hollman</td>
<td>RIKILT Wageningen University, The Netherlands</td>
</tr>
<tr>
<td>Flavanones metabolites</td>
<td>Claudine Manach</td>
<td>INRA Clermont-Ferrand, UMR1019, France</td>
</tr>
<tr>
<td>Flavonol metabolites</td>
<td>Claudia Nunes dos Santos</td>
<td>ITQB-UNL/IBET, Oeiras, Portugal</td>
</tr>
<tr>
<td>Hydroxycinnamic acids</td>
<td>Andreia Bento da Silva</td>
<td>Institute of Chemical and Biological Technology (ITQB), Oeiras, Lisbon, Portugal</td>
</tr>
<tr>
<td>Miscellaneous phenolic acids</td>
<td>Andreia Bento da Silva</td>
<td>Institute of Chemical and Biological Technology (ITQB), Oeiras, Lisbon, Portugal</td>
</tr>
<tr>
<td>Phytoprostanes</td>
<td>Cécile Gladine</td>
<td>INRA Clermont-Ferrand, UMR1019, France</td>
</tr>
<tr>
<td>Phytoestrols</td>
<td>Laura Nyström</td>
<td>Institute of Food, Nutrition and Health, ETH Zurich, Switzerland</td>
</tr>
<tr>
<td>Procyanidin metabolites</td>
<td>Begoña Bartolomé Sualdea</td>
<td>Instituto de Investigación en Ciencias de la Alimentación (CIA), CSIC-UAM, Madrid, Spain</td>
</tr>
<tr>
<td>Stilbenes</td>
<td>Mireia Urpi-Sarda</td>
<td>University of Barcelona, Spain</td>
</tr>
<tr>
<td>Terpenoids</td>
<td>Yoann Fillâtre</td>
<td>European Research Institute on Natural Ingredients, Grasse, France</td>
</tr>
</tbody>
</table>

- Curators: min 1-2 experts for every family
  - Check PhytoHub data
  - Import new data from the literature (dietary sources, known metabolites, spectra...)

- Collaboration COST POSITIVE
  - 77 research institutions with expertise on plant food bioactives

New curators are welcome!
1. Databases to Obtain Hypotheses of Identification:

- Back-end interface for online updates by registered experts

A community tool for research on food phytochemicals

Will be linked to FooDB
NEW TOOLS: PRIORITY NEEDS IDENTIFIED

1. A compound database with extensive coverage of the food metabolome
2. A food intake biomarkers database
3. A resource to facilitate sharing of standards
4. A resource to facilitate knowledge and data sharing within the FoodBAll community and beyond
A FOOD INTAKE BIOMARKERS DATABASE

http://exposome-explorer.iarc.fr/

« dedicated to biomarkers of exposure to environmental risk factors for diseases »
A FOOD INTAKE BIOMARKERS DATABASE

http://exposome-explorer.iarc.fr/

Statistics (June 2016)

- All biomarkers
  - 497 biomarkers
  - 10,480 concentration values
  - 480 publications analyzed

- Dietary biomarkers
  - 147 biomarkers
  - 3,138 concentration values
  - 220 publications analyzed

- Biomarkers
- Cohorts where measured
- Biospecimens
- Analytical methods
- Concentrations
- Correlations with exposures
- Reproducibility over time
- Confounding factors
- Linked to other databases
## Correlation values

<table>
<thead>
<tr>
<th>Population</th>
<th>Intake</th>
<th>Biological specimen</th>
<th>Biomarker</th>
<th>Correlation type</th>
<th>Correlation value</th>
<th>Correlation p-value</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants in the validation study of the MoBa cohort</td>
<td>Citrus fruits</td>
<td>Urine, 24-h</td>
<td>Flavonones</td>
<td>Pearson</td>
<td>0.66</td>
<td>&lt; 0.01</td>
<td>Brantsaeter 2007</td>
</tr>
<tr>
<td>Participants in the validation study of the MoBa cohort</td>
<td>Citrus fruits</td>
<td>Urine, 24-h</td>
<td>Hesperetin</td>
<td>Pearson</td>
<td>0.66</td>
<td>&lt; 0.01</td>
<td>Brantsaeter 2007</td>
</tr>
<tr>
<td>Participants in the validation study of the MoBa cohort</td>
<td>Citrus fruits</td>
<td>Urine, 24-h</td>
<td>Hesperetin</td>
<td>Pearson</td>
<td>0.64</td>
<td>&lt; 0.01</td>
<td>Brantsaeter 2007</td>
</tr>
<tr>
<td>Participants in the validation study of the MoBa cohort</td>
<td>Citrus fruits</td>
<td>Urine, 24-h</td>
<td>Flavonones</td>
<td>Pearson</td>
<td>0.63</td>
<td>&lt; 0.01</td>
<td>Brantsaeter 2007</td>
</tr>
<tr>
<td>Mothers in Maastricht with or without an increased fruit and vegetable diet intervention</td>
<td>Citrus fruits</td>
<td>Plasma, fasting</td>
<td>beta-Cryptoxanthin</td>
<td>Spearman</td>
<td>0.57</td>
<td>&lt; 0.01</td>
<td>Bogers 2004</td>
</tr>
<tr>
<td>Adults living in the Parisian area</td>
<td>Citrus fruits</td>
<td>Urine, first morning spot</td>
<td>Naringenin</td>
<td>Spearman</td>
<td>0.56</td>
<td>&lt; 0.0001</td>
<td>Menneen 2006</td>
</tr>
<tr>
<td>Participants in a dietary validation study</td>
<td>Citrus fruits</td>
<td>Plasma, fasting</td>
<td>Vitamin C</td>
<td>Spearman</td>
<td>0.53</td>
<td>&lt; 0.001</td>
<td>Mohammadifard 2011</td>
</tr>
<tr>
<td>Adults living in the Parisian area</td>
<td>Citrus fruits</td>
<td>Urine, first morning spot</td>
<td>Hesperetin</td>
<td>Spearman</td>
<td>0.52</td>
<td>&lt; 0.0001</td>
<td>Menneen 2006</td>
</tr>
<tr>
<td>Adults living in the Parisian area</td>
<td>Citrus fruits</td>
<td>Urine, first morning spot</td>
<td>Naringenin</td>
<td>Spearman</td>
<td>0.48</td>
<td>&lt; 0.0003</td>
<td>Menneen 2006</td>
</tr>
<tr>
<td>Adults living in the Parisian area</td>
<td>Citrus fruits</td>
<td>Urine, 24-h</td>
<td>Hesperetin</td>
<td>Spearman</td>
<td>0.46</td>
<td>&lt; 0.0006</td>
<td>Menneen 2006</td>
</tr>
</tbody>
</table>

Showing 1 to 10 of 79 entries
Next step will be to integrate dietary biomarkers identified in intervention studies *(in link with literature reviews in FoodBAI/ WP3)*
- Authentic standards are needed for validation of identifications in metabolomics studies
- Lack of commercial standards for many food-derived compounds: putative identifications
- Also hampers the development of quantitative methods of analysis for candidate biomarkers
Many compounds have been synthetised or isolated in academic laboratories.

http://foodcomex.org/
Information on available compounds & contact details

Online catalog

- Chemical synthesis
- Extraction-Purification from foods
- Biofluids from animals fed pure compounds or food extracts
- In vitro incubations

*Microsomes, isolated enzymes, microbiota*
### 3- A Resource to Facilitate Sharing of Standards: FoodComEx

**Foodcomex Compounds**

<table>
<thead>
<tr>
<th>FoodComEx Id</th>
<th>Compound Name</th>
<th>Origin</th>
<th>Storage Form</th>
<th>Maximum Quantity</th>
<th>Storage Condition</th>
<th>Contact Name</th>
<th>Contact Address</th>
<th>Contact Email</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDB029964</td>
<td>Vanillic acid 4-sulfate</td>
<td>Chemical synthesis</td>
<td>Powder</td>
<td>Production upon request</td>
<td>Rita Ventura, Claudia Santos</td>
<td>ITQB-UNL/IBET, Oeiras, Portugal</td>
<td><a href="mailto:rventura@itqb.unl.pt">rventura@itqb.unl.pt</a>, <a href="mailto:csantos@itqb.unl.pt">csantos@itqb.unl.pt</a></td>
<td><img src="image" alt="Structure" /></td>
<td></td>
</tr>
<tr>
<td>FDB029941</td>
<td>Quercetin 3'-sulfate</td>
<td>Chemical synthesis</td>
<td>Freeze-dried</td>
<td>Milligram amounts when produced. Production upon request</td>
<td>Celestino Santos Buelga</td>
<td>Grupo de Investigación de Polímeros de Silicato, Facultad de Farmacia, Universidad de Salamanca, Spain</td>
<td><a href="mailto:csb@usal.es">csb@usal.es</a></td>
<td><img src="image" alt="Structure" /></td>
<td></td>
</tr>
<tr>
<td>FDB031313</td>
<td>4-hydroxybenzoic acid 4-O-sulfate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><img src="image" alt="Structure" /></td>
<td></td>
</tr>
</tbody>
</table>

**Bilateral negociation with the provider** *(Chart of good practices online)*
3- A RESOURCE TO FACILITATE SHARING OF STANDARDS

http://foodcomex.org/

- A virtual board to request unavailable compounds
- A providers directory to foster collaborations
- A repository of shared production methods
1. A compound database with extensive coverage of the food metabolome
2. A food intake biomarkers database
3. A resource to facilitate sharing of standards
4. A resource to facilitate knowledge and data sharing within the FoodBAII community and beyond
Welcome to the FOODBALL Project

The Food Biomarkers Alliance (FOODBALL) is an initiative aimed at identifying and quantifying dietary biomarkers in order to improve the capabilities of nutritional assessment and research.

Foodomics

Foodomics is a discipline in the field of food and nutrition that applies and integrates comprehensive high-throughput -omics technologies to improve health and well-being, and nutritional knowledge. Foodomics encompasses the global food domain, including fields such as nutrigenomics. Studies in foodomics often focus on areas such as the mechanisms of different bioactive food components, quantification of dietary biomarkers to identify different health states, the assessment of food quality and safety, or examination of the body's biological response to different nutritional patterns.

Description useful resources + links + tutorials
Twitter account @foodmetabolome
Wiki for FoodBALL partners
RESOURCES DEVELOPED IN FOODBALL


Please use these resources and give us your feedback!
ACKNOWLEDGEMENTS

Rosa Vazquez Fresno
Tanvir Sayed
David Arndt
Allison Pon
Albert remus Rosana
Yannick Djoumbou Feunang
David Wishart

Jarlei Fiamoncini
Franck Giacomoni
Craig Knox
Balthazar Pavot
Estelle Pujos-Guillot
Christine Morand

Joe Rothwell
Vanessa Neveu
Augustin Scalbert