

# Supporting Information

## Charting Angiosperm Chemistry – Evolutionary Perspective on Specialized Metabolites Reflected in Chemical Property Space

*Astrid Henz Ryen and Anders Backlund\**

Research group for Pharmacognosy, Department of Medicinal Chemistry, BMC -  
Biomedical Center, Uppsala University, Box 574, 75123 Uppsala, Sweden

## Table of content

**Table S1 and Table S2:** S1+S2\_datasets.xlsx (Separate Excel file)

**High-resolution plots:** High\_resolution\_plots.pdf (Separate PDF file)

## Description

### Table S1:

Complete dataset of specialized metabolites, first sheet in: **S1+S2\_datasets.xlsx**

The dataset used in this study comprised 17,902 unique specialized metabolites from angiosperms, viz. sesquiterpene lactones, tropane alkaloids, betalains and flavonoids. CAS-numbers (Chemical Abstracts Service, Columbus, OH, USA), SMILES strings (Simplified Molecular Input Line Entry Specification), plant family of origin, and principle component scores from the ChemGPS-NP analysis, are all provided in a comprehensive tab separated text file. Note that this list comprises in total more than 20,000 entries since numerous compounds have been reported from several different plant families.

Data were extracted from various open access sources as well as proprietary databases, such as the “Dictionary of Natural Products”,<sup>1</sup> the “Dictionary of Flavonoids”,<sup>2</sup> *SciFinder*,<sup>3</sup> and continuous literature surveys mostly published between the years 2005 - 2015.<sup>4,5</sup> In order to cover the distribution of the analyzed metabolites throughout the angiosperms, systematic searches have been performed up to the taxonomic rank of the genus. Considerable efforts have been made to verify structural accuracy of the compounds in the database by visual inspection. To ensure that a correct taxonomy has

been employed, the taxonomy of proposed plants of origin has been checked using the “International Plant Names Index”<sup>6</sup> and the systematic classifications have been realigned after APG IV.<sup>7</sup>

**Table S2:**

Complete dataset of STLs produced by *Eupatorium*, *Liatris* and *Mikania*:

**Second sheet in: S1+S2\_datasets.xlsx** . This table is organized in the same way as Table S1 and contains in total 353 unique STLs.

**High-resolution plots:**

For better visibility and more detail, every plot presented in this study is in addition provided in the PDF file: **High\_resolution\_plots.pdf**

## REFERENCES

- (1) Dictionary of Natural Products; CRC Press/Taylor & Francis Group: Boca Raton, FL; <http://dnp.chemnetbase.com>.
- (2) Buckingham, J.; Munasinghe, V. R. N. *Dictionary of Flavonoids with CD-ROM*; CRC Press/Taylor & Francis Group: Boca Raton, FL, 2015.
- (3) SciFinder; Chemical Abstracts Service: Columbus, OH; <https://scifinder.cas.org>.
- (4) Larsson, J.; Gottfries, J.; Bohlin, L.; Backlund, A. *J. Nat. Prod.* **2005**, *68*, 985-991.
- (5) Larsson, J.; Gottfries, J.; Muresan, S.; Backlund, A. *J. Nat. Prod.* **2007**, *70*, 789-794.
- (6) The International Plant Names Index; <https://www.ipni.org>.
- (7) Byng, J. W.; Chase, M. W.; Christenhusz, M. J. M.; Fay, M. F.; Judd, W. S.; Mabberley, D. J.; Sennikov, A. N.; Soltis, D. E.; Soltis, P. S.; Stevens, P. F.; Briggs, B.; Brockington, S.; Chautems, A.; Clark, J. C.; Conran, J.; Haston, E.; Moller, M.; Moore, M.; Olmstead, R.; Perret, M.; Skog, L.; Smith, J.; Tank, D.; Vorontsova, M.; Weber, A.; Grp, A. P. *Bot. J. Linn. Soc.* **2016**, *181*, 1-20.