Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Ultrahigh resolution metabolomics for S-containing metabolites

Ryo Nakabayashi[1] and Kazuki Saito[1,2]

The advent of the genome-editing era greatly increases the opportunities for synthetic biology research that aims to enhance production of potentially useful bioactive metabolites in heterologous hosts. A wide variety of sulfur (S)-containing metabolites (S-metabolites) are known to possess bioactivities and health-promoting properties, but finding them and their chemical assignment using mass spectrometry-based metabolomics has been difficult. In this review, we highlight recent advances on the targeted metabolomic analysis of S-metabolites (S-omics) in plants using ultrahigh resolution mass spectrometry. The use of exact mass and signal intensity differences between ^{32}S-containing monoisotopic ions and counterpart ^{34}S isotopic ions exploits an entirely new method to characterize S-metabolites. Finally, we discuss the availability of S-omics for synthetic biology.

Addresses
[1] RIKEN Center for Sustainable Resource Science, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan
[2] Graduate School of Pharmaceutical Sciences, Chiba University, 1-8-1 Inohana, Chuo-ku, Chiba 260-8675, Japan

Corresponding authors: Nakabayashi, Ryo (ryo.nakabayashi@riken.jp) and Saito, Kazuki (kazuki.saito@riken.jp)

Current Opinion in Biotechnology 2016, **43**:8–16

This review comes from a themed issue on **Analytical biotechnology**

Edited by **Jurre Kamphorst** and **Ian Lewis**

http://dx.doi.org/10.1016/j.copbio.2016.07.003

Introduction

It is no exaggeration to say that the appearance of next generation DNA sequencers is dramatically changing all types of biological research. Before its appearance, functional genomics could be only performed in a few selected model organisms. Nowadays, the accumulation of big sequence data enables functional genomics in all sequence-readable organisms. In synthetic biology, genes identified in functional genomics are heterologously expressed in host organisms, to create machineries for enhancing production of new target metabolites. Microbe and plant specialized/secondary metabolites, such as bleomycin, lovastatin, artemisinin, morphine, paclitaxel, prostratin, resveratrol, and vincristine, can now be produced using techniques of synthetic biology [1–6].

In the near future, one can expect that almost all known specialized metabolites with potent biological activities will be investigated using functional genomics and synthetic biology. As researches go on, the available pool of known bioactive metabolites will be possibly exhausted. To avoid this, we should continue to develop analytical methods to discover new metabolites with potentially useful bioactivities in metabolome resources. Therefore, it is important how such metabolites can be precisely profiled using high-throughput metabolomics approaches. Among newly approved drugs from 1981 to 2014, substantial number of entities contain heteroatoms such as N, O, S, and halogens [7•]. Some of these drugs are derived from natural products. These suggest that efficient profiling methods need to be developed for discovering seed and lead heteroatom-containing natural products. In this review, we focus on the targeted analysis of S-containing metabolites (S-metabolites) using mass spectrometry (MS) in plants. Furthermore, we discuss the availability to use S-omics in a biotechnological research, for instance, in synthetic biology.

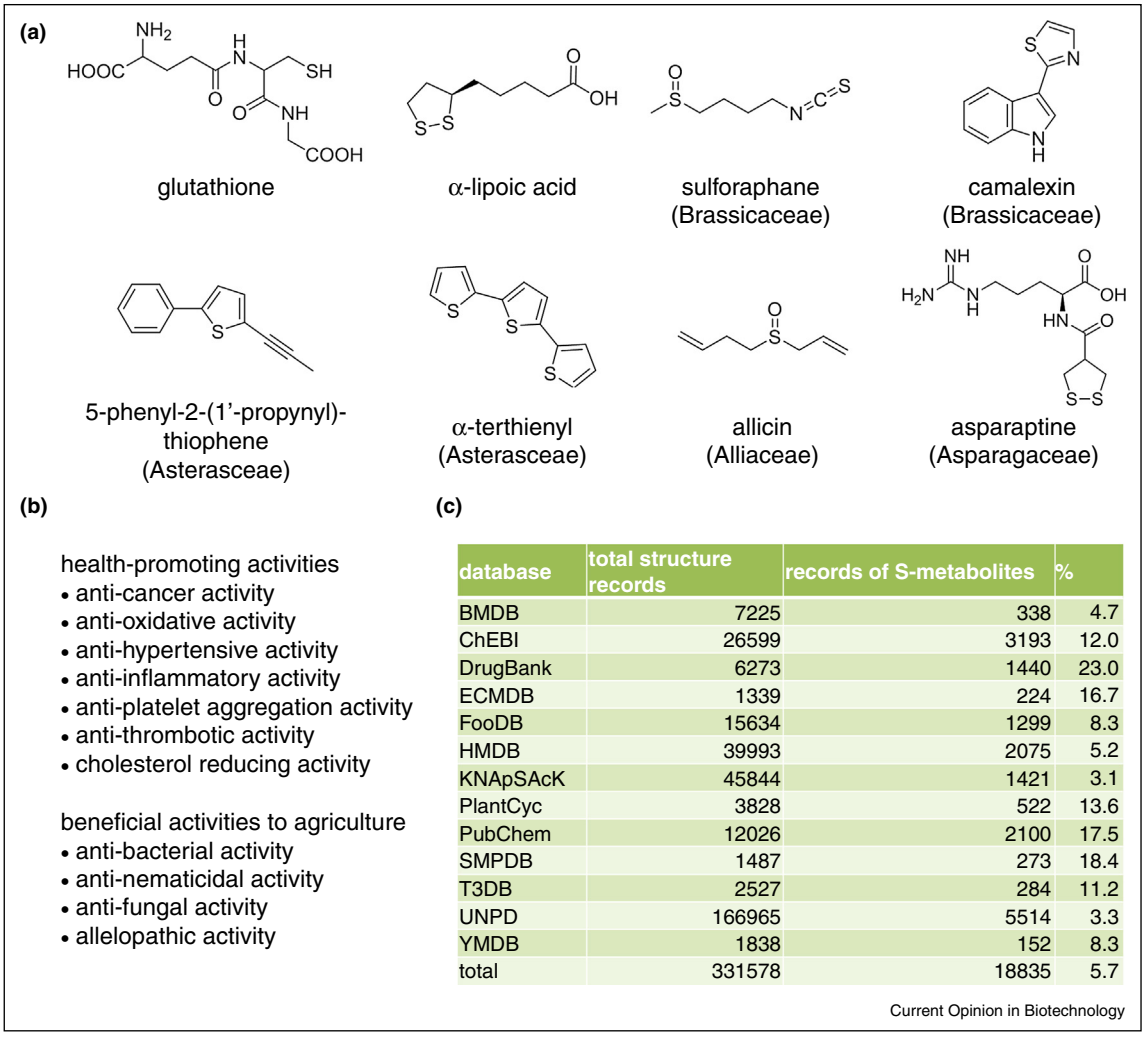
Chemical diversity and usefulness of S-metabolites

S-Metabolites are biosynthesized using sulfate, cysteine, and/or methionine in all organisms including plants [8]. Biosynthesis produces a wide variety of chemical structures containing highly reactive functional moieties including S (Figure 1a). This chemical variety provides beneficial properties, especially health-promoting activities, to human activities (Figure 1b). The occupancy of known S-metabolites in several databases is 5.7% in average, but interestingly, this goes up to 23% in the DrugBank database [9], highlighting the bioactive roles of these compounds (Figure 1c). These facts indicate that developing targeted analysis methods need to be addressed to find novel seed or lead S-metabolites for future drug developments.

A structural feature of S-metabolites

One of the general targeted analysis methods applied is liquid chromatography–ultra violet–tandem mass spectrometry (LC–UV–MS/MS). Here, a specific metabolite group is detected using a common structural feature such the UV spectrum (e.g., key UV absorption) or MS/MS spectrum (e.g., key product ion). In the case of S-metabolites, such

Figure 1



database	total structure records	records of S-metabolites	%
BMDB	7225	338	4.7
ChEBI	26599	3193	12.0
DrugBank	6273	1440	23.0
ECMDB	1339	224	16.7
FooDB	15634	1299	8.3
HMDB	39993	2075	5.2
KNApSAcK	45844	1421	3.1
PlantCyc	3828	522	13.6
PubChem	12026	2100	17.5
SMPDB	1487	273	18.4
T3DB	2527	284	11.2
UNPD	166965	5514	3.3
YMDB	1838	152	8.3
total	331578	18835	5.7

Current Opinion in Biotechnology

A chemical diversity of S-containing metabolites (S-metabolites). **(a)** Structure of plant S-metabolites. Glutathione and α-lipoic acid are primary metabolite: the others are specialized/secondary metabolites. Family name of plants accumulating these metabolites are shown in parentheses. **(b)** Beneficial properties of plant S-metabolites to human activities. **(c)** Structural records of S-metabolites in databases. The data was downloaded on April 1st in 2015. S-metabolites were searched using the condition of S atom number ($1 \leq n \leq 5$). Percentage means the occupancy of S-metabolites to the total records. BMDB, http://www.cowmetdb.ca/cgi-bin/browse.cgi; ChEBI, https://www.ebi.ac.uk/chebi/; DrugBank, http://www.drugbank.ca/; ECMDB, http://ecmdb.ca/; FooDB, http://foodb.ca/; HMDB, http://www.hmdb.ca/; KNApSAcK, http://kanaya.naist.jp/KNApSAcK/; PlantCyc, http://www.plantcyc.org/; PubChem Classification Browser (Biosystems and Pathways), https://pubchem.ncbi.nlm.nih.gov/classification/#hid=72; SMPDB, http://smpdb.ca/; T3DB, http://www.t3db.ca/; UNPD, http://pkuxxj.pku.edu.cn/UNPD/; and YMDB: http://www.ymdb.ca/.

common features were not available due to the chemical diversity.

The exact mass and natural abundance of the stable S isotopes ^{32}S (31.972072 Da, 95.02%) and ^{34}S (33.967868 Da, 4.21%) are present in the monoisotopic and isotopic ions respectively of all S-metabolites. Theoretically, these differences in exact mass and signal intensity can be detected among the monoisotopic and isotopic ions given sufficient instrument sensitivity and mass resolution. Nevertheless, the information intrinsic to such ions has been overlooked in metabolomics analysis. One reason is that, at the beginning of the metabolomics era, isotopic ions were typically deleted from the data matrix since they were redundant amongst the hundreds of ion peaks in data analysis. At that time, mass accuracy was mainly used for chemical assignment of detected metabolites. Opportunities to use isotopic ions increase with the improving ability of peak resolution and mass accuracy in Fourier transform ion cyclotron resonance-mass spectrometry (FTICR-MS) instruments. As shown in Supplementary Table 1, the specifications of