

Water, water, every where, but rarely any drop to drink

Royston Goodacre

Published online: 27 December 2013
© Springer Science+Business Media New York 2013

I would like to give all readers a very warm welcome to 2014 and the first issue of the tenth volume of *Metabolomics*. As you may be able to work out: the front cover is a celebration of this achievement, and I thank my colleague Dr Steve O'Hagan for his artistry. I am delighted that the journal is in such good shape and this is due to the excellent papers that are submitted and published, and of course the very valuable reviewing that many of you do. *Metabolomics* has an excellent Editorial board and I am also very grateful to them for their valuable support.

You may be pondering over the title, so let me explain. Whilst I have somewhat moderated the quote from “The Rime of the Ancient Mariner” by Samuel Taylor Coleridge written in 1797–1798, the water does not refer to any liquid substance per se, nor does the drinking to the ‘dryathlon’¹ that I did early last year and will be doing so again to combat any Christmas excesses. Rather the water is an analogy to data—both metabolomics and metadata. Water here is a very apt comparison, as it seems rather ironic that a typical metabolomics experiments generates so much data that it is often referred to in terms of natural disasters—like data floods, data torrents or even data tsunamis. Yet even more ironic that very rarely do we make publicly available the metabolomics data (raw or processed) and the associated metadata with our publications. These metadata are as important as the metabolite data as these refer to the data about the data. We mainly think of these in terms of the important traits or features that we may want to predict, but these also refer to our experimental protocols that are

vital for others to know about so that these experiments may be readily repeated. At this point I refer you to the MSI (Metabolomics Standards Initiative) which in 2007 suggested the best way to report experimental conditions in an objective fashion. If you have not done so I would recommend that you read this collection of papers and they are summarised pictorially within the typical metabolomics workflow in Fig. 1. One of these articles by Sumner et al. (2007) details the various metabolite identification levels and this has been very well received by the metabolomics community. So much so that this journal now requires its authors to follow these guidelines explicitly, and tables with lists of metabolites must now contain reference to which of the four levels of metabolite identification have been made (these are provided in Table 1).

In 2012 the EBI launched MetaboLights (<https://www.ebi.ac.uk/metabolights/>), which is an online repository for metabolomics data along with the associated metadata (Haug et al. 2013), and NIH Metabolomics Workbench (<http://www.metabolomicsworkbench.org>) is establishing similar data storage infrastructures. I was also encouraged to see *Nature Publishing Group* establish a new online resource called *Scientific Data* (<http://www.nature.com/scientificdata/>) which launches in May this year (Editorial 2013). However, rather than this being a new data repository, this resource will aim to have data sets deposited correctly in an appropriate repository, thereby aiding the paper peer-reviewing process.

Metabolomics is already on record in saying that it wishes studies that it publishes and data therein to be as

R. Goodacre (✉)
School of Chemistry and Manchester Institute of Biotechnology,
University of Manchester, 131 Princess Street,
Manchester M1 7DN, UK
e-mail: roy.goodacre@manchester.ac.uk

¹ The term Dryathlon refers to the activity popularised by Cancer Research UK where for the whole of January one is encouraged to become a Dryathlete™ and abstain from alcoholic beverages. <http://www.cancerresearchuk.org/support-us/find-an-event/charity-challenges/dryathlon>.

Fig. 1 Overview of the metabolomics pipeline showing where the MSI reporting standards are aimed (highlighted in blue) (Color figure online)

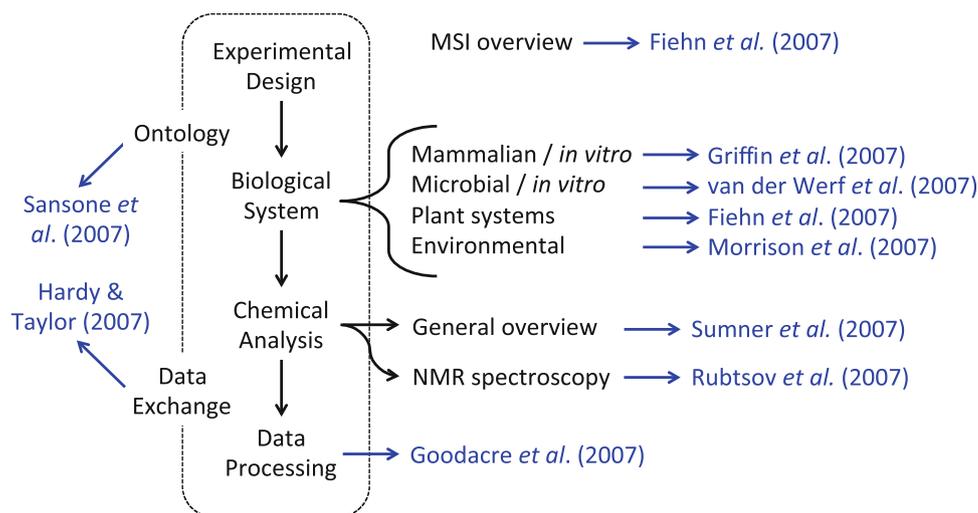


Table 1 A summary of the four metabolite identification levels as defined by MSI (Sumner et al. 2007)

Level	Definition
1	<p><i>Identified compounds</i></p> <p>This definitive identification level requires a minimum of at least two independent and orthogonal data relative to an authentic compound analyzed under identical experimental conditions in the same laboratory on the same analytical platform</p> <p>These orthogonal data must provide different physicochemical properties of the metabolite and for examples may be: retention time/index and mass spectrum; retention time and $^1\text{H}/^{13}\text{C}$ NMR spectrum</p>
2	<p><i>Putatively annotated compounds</i></p> <p>These are very similar to level 1 BUT are identifications that are made without chemical reference standards. The above orthogonal characteristics are still used but are typically matched against public or commercially available spectral libraries</p>
3	<p><i>Putatively characterized compound classes</i></p> <p>This level defines compounds that are based upon characteristic physicochemical properties of a chemical class of compounds, or by spectral similarity to known compounds of a chemical class</p> <p>For example, a fatty acid like C18:1 fatty acid where the unsaturation point is not known (for oleic acid—(9Z)-octadec-9-enoic acid—this would become C18:1 <i>cis</i>-9 if matched with a standard). Another example is a C6 sugar</p>
4	<p><i>Unknown compounds</i></p> <p>Whilst these are unidentified or unclassified compounds, these small molecules can still be differentiated, recognized again by the analyst in further analysis, and therefore quantified based upon spectral data</p>

MSI compliant as possible (Goodacre 2010). Whilst established procedures are not yet in place for metabolite data upload within a MSI compliant framework, I would encourage you to deposit your data to one of the above repositories. COSMOS is a European Union Framework VII project which has Dr Christoph Steinbeck from the EBI as the project coordinator (Steinbeck et al. 2012; <http://www.cosmos-fp7.eu>). COSMOS aims to set and promote community standards for metabolomics and *Metabolomics* is excited to be involved in this new e-infrastructure with data deposition within the MetaboLights database (Salek et al. 2013). In time this and the other activities discussed above will certainly establish objective methods by which metadata and metabolomics data may be associated with publications seamlessly, thus allowing easy access. These data will be tagged with a unique identifier that is provided by MetaboLights that will be associated with the paper.

This transparency is important for the field and will enable meta-analysis of metabolomics data in order to allow hypotheses generated in one laboratory to be tested on data already available within the public domain.

So whilst currently only data dribbles into the public domain, the sluices will soon be open and we shall be awash with important and useful metabolomics data. And then the fun shall really begin!

References

- Editorial. (2013). Announcement: Launch of an online data journal. *Nature*, 502, 142.
- Fiehn, O., Robertson, D., Griffin, J., van der Werf, M., Nikolau, B., Morrison, N., et al. (2007a). The metabolomics standards initiative (MSI). *Metabolomics*, 3, 175–178.

- Fiehn, O., Sumner, L. W., Rhee, S. Y., Ward, J., Dickerson, J., Lange, B. M., et al. (2007b). Minimum reporting standards for plant biology context information in metabolomic studies. *Metabolomics*, 3, 195–201.
- Goodacre, R. (2010). An overflow of... what else but metabolism! *Metabolomics*, 6, 1–2.
- Goodacre, R., Broadhurst, D., Smilde, A., Kristal, B. S., Baker, J. D., Beger, R., et al. (2007). Proposed minimum reporting standards for data analysis in metabolomics. *Metabolomics*, 3, 231–241.
- Griffin, J. L., Nicholls, A. W., Daykin, C., Heald, S., Keun, H., Schuppe-Koistinen, I., et al. (2007). Standard reporting requirements for biological samples in metabolomics experiments: Mammalian/in vivo experiments. *Metabolomics*, 3, 179–188.
- Hardy, N. W., & Taylor, C. F. (2007). A roadmap for the establishment of standard data exchange structures for metabolomics. *Metabolomics*, 3, 243–248.
- Haug, K., Salek, R. M., Conesa, P., Hastings, J., de Matos, P., Rijnbeek, M., et al. (2013). MetaboLights: An open-access general-purpose repository for metabolomics studies and associated meta-data. *Nucleic Acids Research*, 41, 781–786.
- Morrison, N., Bearden, D., Bundy, J. G., Collette, T., Currie, F., Davey, M. P., et al. (2007). Standard reporting requirements for biological samples in metabolomics experiments: Environmental context. *Metabolomics*, 3, 203–210.
- Rubtsov, D. V., Jenkins, H., Ludwig, C., Easton, J., Viant, M. R., Günther, U., et al. (2007). Proposed reporting requirements for the description of NMR-based metabolomics experiments. *Metabolomics*, 3, 223–229.
- Salek, R. M., Steinbeck, C., Goodacre, R., Viant, M. R., & Dunn, W. B. (2013). The role of reporting standards for metabolite annotation and identification in metabolomic studies. *Giga-Science*, 2, 13.
- Sansone, S.-A., Schober, D., Atherton, H. J., Fiehn, O., Jenkins, H., Rocca-Serra, P., et al. (2007). Metabolomics standards initiative—ontology working group: Work in progress. *Metabolomics*, 3, 249–256.
- Steinbeck, C., Conesa, P., Haug, K., Mahendrakar, T., Williams, M., Maguire, E., et al. (2012). MetaboLights: towards a new COSMOS of metabolomics data management. *Metabolomics*, 8, 757–760.
- Sumner, L. W., Amberg, A., Barrett, D., Beger, R., Beale, M. H., Daykin, C., et al. (2007). Proposed minimum reporting standards for chemical analysis. *Metabolomics*, 3, 211–221.
- van der Werf, M. J., Takors, R., Smedsgaard, J., Nielsen, J., Ferenci, T., Portais, J. C., et al. (2007). Standard reporting requirements for biological samples in metabolomics experiments: Microbial and in vitro biology experiments. *Metabolomics*, 3, 189–194.