

3 Minimum reporting standards for plant biology context 4 information in metabolomic studies

5 Oliver Fiehn · Lloyd W. Sumner · Seung Y. Rhee ·
6 Jane Ward · Julie Dickerson · Bernd Markus Lange ·
7 Geoff Lane · Ute Roessner · Robert Last ·
8 Basil Nikolau

9 Received: 5 January 2007 / Accepted: 8 June 2007
10 © Springer Science+Business Media, LLC 2007

11 **Abstract** Plant metabolomics has matured over the past
12 8 years. Plant biologists routinely use comprehensive anal-
13 yses of plant metabolites to discover new responses to ge-
14 netic or environmental perturbation, or to validate initial
15 hypotheses on the function and in vivo action of gene
16 products. The wealth of scientific findings has increasingly
17 provoked interest to share and review raw or processed data
18 from plant metabolomics reports. We here suggest a mini-
19 mum of parameters to be reported in order to define details of
20 experimental study designs in plant metabolomics studies.
21

22 **Keywords** Ontology · Semantics · Minimum information ·
23 Metabolite profiling · Metabonomics · Metadata

1 Introduction 24

Four major conferences have been held on the applica- 25
tions and technologies used in plant metabolomics. One 26
of the major observations was that there is no single 27
technology platform that could quantify and identify all 28
plant compounds in a single analysis; instead, technolo- 29
gies are increasingly selected to target specific biological 30
questions which range from evolution of complex traits 31
to adaptation towards abiotic stresses and questions 32
concerning organ- or cell-specific biochemistry. This 33
wealth of scientific findings has increasingly provoked 34
interest to share and review raw or processed data from 35

A1 O. Fiehn (✉)
A2 UC Davis Genome Center, Metabolomics Research Laboratory,
A3 UC Davis, 451 East Health Sci Drive, Davis, CA 95616-8816,
A4 USA
A5 e-mail: ofiehn@ucdavis.edu

A6 L. W. Sumner
A7 Noble Foundation, Ardmore, OK, USA
A8 e-mail: lwsumner@noble.org

A9 S. Y. Rhee
A10 Carnegie Institution, DPB, TAIR, Stanford, CA, USA
A11 e-mail: rhee@coma.Stanford.EDU

A12 J. Ward
A13 Rothamsted Research, Harpenden, Hertfordshire, England
A14 e-mail: jane.ward@bbsrc.ac.uk

A15 J. Dickerson · B. Nikolau
A16 Iowa State University, Ames, IA, USA

A17 J. Dickerson
A18 e-mail: julied@iastate.edu

A19 B. Nikolau
A20 e-mail: dimmas@iastate.edu

A1 B. M. Lange
A2 Washington State University, Pullman, WA, USA
A3
A4 e-mail: lange-m@wsu.edu

A5 G. Lane
A6 AgResearch, Hamilton, New Zealand
A7 e-mail: geoff.lane@agresearch.co.nz

A8 U. Roessner
A9 University of Melbourne, Melbourne, VIC, Australia
A10 e-mail: ute.roessner@acpfg.com.au

A11 R. Last
A12 Michigan State University, East Lansing, MI, USA
A13 e-mail: lastr@msu.edu

36 publications, in order to re-use data to validate preex- 89
 37 isting or generate novel hypotheses. 90

38 What is the baseline of metabolite levels found in model 91
 39 plants such as *Arabidopsis thaliana* or *Oryza sativa*? It has 92
 40 been demonstrated that the metabolomes of higher plants 93
 41 are highly diverse and flexible, spanning an enormous 94
 42 range of complexity and metabolite concentrations. 95
 43 Therefore, answering seemingly simple questions such as 96
 44 ‘how large is the Arabidopsis metabolome?’ is still a dif-
 45 ficult objective because the answer will depend on the
 46 physiological and genotypic settings. For example, which
 47 organs were analyzed, what were the illumination and soil
 48 conditions used for growing plants, how was the nutrient
 49 and watering regime set up, and which genes (if transgenic
 50 plants are studied) were altered?

51 For any single study, these data concerning the gener- 100
 52 ation of the biological materials (i.e. the biological 101
 53 context metadata) are not always presented in detail, 102
 54 according to the requirements for method sections of 103
 55 various plant journals. Initiated by the Metabolomics 104
 56 Society, a number of plant researchers have therefore 105
 57 collaborated to refine the minimum set of required 106
 58 reporting parameters that are essential to describe an 107
 59 experiment. It is important to understand that this effort 108
 60 does not attempt to prescribe experimental designs or 109
 61 distinguish scientifically adequate from inadequate 110
 62 designs. Instead, the proposed standards aim to promote 111
 63 “good” plant biology practices with special provisions to 112
 64 enable comparisons of experimental data and designs 113
 65 electronically and between publications in peer-reviewed 114
 66 journals. In a second but separate step, current best 115
 67 practice standards may be developed, which will go 116
 68 beyond the minimum set of core metadata to be reported 117
 69 and will potentially better reflect the ever changing view 118
 70 of the complement of factors that need to be considered 119
 71 for understanding the metabolome of plants. 120

72 The reporting standards proposed here have been 121
 73 reviewed and improved by an in-depth discussion of the 122
 74 participants of the 4th International Conference on Plant 123
 75 Metabolomics, held in Reading, UK, in April 2006. 124
 76 However, despite our best efforts, we may have over- 125
 77 looked important criteria or parameters. In addition, the 126
 78 notion of ‘minimum reporting standards’ cannot refer to 127
 79 an impartial concept but is the result of prolonged 128
 80 discussions to reach consensus. The notion is that 129
 81 ‘minimum reporting standards’ will be endorsed and 130
 82 supported by the plant biology community at large, in 131
 83 order to legitimate mandatory reporting requirements 132
 84 adopted by funding agencies, foundations, scientific 133
 85 organizations and journals. In this respect, the standards 134
 86 presented here do not represent an end point but rather 135
 87 an initial milestone for ongoing discussions. The authors 136
 88 therefore appreciate feedback and constructive criticism 137

which would be incorporated into refined versions of the
 ‘reporting standards’ documents, that will be available
 from the Metabolomics Standards Initiative (MSI) web-
 site (<http://msi-workgroups.sourceforge.net/bio-metadata/>).
 Comments can be also sent to an open list (Msi-work-
 groups-feedback@lists.sourceforge.net) without the need
 to subscribe to one of the specific MSI workgroups
 mailing lists.

2 Materials and methods 97

2.1 The Standards generation process 98

In 2005, the Metabolomics Standards Initiative was formed 99
 as result of a workshop organized by the U.S. National 100
 Institutes of Health. This Initiative was supported and en- 101
 dorsed by the Metabolomics Society. According to the 102
 general metabolomics workflow, one of the key parameters 103
 of standardization for reporting metabolomics data was 104
 identified as biological context information. It was recog- 105
 nized that study designs and emphases of different fields of 106
 biology call for distinct (but small) working groups whose 107
 tasks were to compile initial lists of required standards which 108
 later should be refined by the larger biology context com- 109
 munities. The active participation of governmental agencies 110
 and industrial corporations was actively sought; however, 111
 most collaborators were affiliated with public research 112
 organizations. The biology subgroups comprised plant 113
 biology, mammalian and clinical biology, microbiology, and 114
 environmental biology. Group chairs held contact via 115
 exchange of documents and teleconferences, organized and 116
 chaired by Don Robertson (Pfizer Global Research & 117
 Development, Ann Arbor, MI, USA). The chairs initially 118
 outlined work plans and exchanged information with other 119
 MSI working groups, namely those working on issues of 120
 Chemical Analysis, Data Processing & Statistics, Data 121
 Exchange, and Ontologies. These exchanges occurred via 122
 workshops, conference reports, teleconferences and the MSI 123
 website (see also reports of these working groups in this same 124
 issue of Metabolomics). The metabolomics society further 125
 formed an oversight committee to coordinate activities, led 126
 by Oliver Fiehn (UC Davis, USA). 127

The plant biology context work was founded on previ- 128
 ous publications that laid the groundwork for reporting 129
 standards. Specifically, the architecture for metabolomics 130
 (ArMet) (Jenkins et al. 2004, 2005) and the ‘Standard 131
 Metabolic Reporting Structure’ document (SMRS) (Lindon 132
 et al. 2005) were utilized as starting points. These evolving 133
 standards were complemented by demands for ‘Minimum 134
 Information About a Metabolomics Experiment’ [MIA- 135
 Met] (Bino et al. 2004); which recognized efforts by other 136
 communities. Especially for genomics studies, the need for 137

138 standardized reporting had been recognized, and several
139 initiatives have evolved, such as the Minimum Information
140 About a Microarray Experiment [MIAME] (Brazma et al.
141 2001); the Reporting Structure for Biological Investiga-
142 tions [RSBI] (<http://www.mged.org/workgroups/rsbi/rsbi.html>), the Functional Genomics Ontology [FuGO] (<http://obi.sourceforge.net/>), 'Chemical effects in biological systems—data dictionary' [CEBS-DD] (Fostel et al. 2005);
144 and the Proteomics Standards Initiative [PSI] of the Human
145 Proteome Organization.

146 An initial draft of the document presented here was cir-
147 culated between the members of the working group, chaired
148 by Basil Nikolau (Iowa State University, Ames, IA, USA).
149 This document was then released and discussed at a 90-min
150 workshop with about 150 participants at the 4th Plant Met-
151 abolomics Conference, Reading, UK, in April 2006. Ulti-
152 mately, the refined and updated version presented here was
153 released for discussion at the 2nd Annual Conference of the
154 Metabolomics Society (Boston, 24–29 June 2006).
155
156

157 3 Results

158 The reporting standards for detailing plant metabolomics
159 studies build upon the commonly accepted practice of
160 reporting plant biological, and specifically, plant physio-
161 logical experiments. However, the best practice for such
162 reporting has never been formally laid out and enforced by
163 journals. For example, author guidelines in the *Plant*
164 *Journal* detail that 'Experimental procedures should be
165 sufficiently detailed to enable the experiments to be
166 reproduced'. The level of experimental detail presented in
167 *Plant Journal* is therefore only subject to the peer-review
168 process, which often focuses on the justification and rele-
169 vance of the scientific content, rather than methodological
170 aspects. The journal *Plant Cell* is more specific, detailing
171 the author guidelines for 'Method' descriptions by the
172 subheadings 'Large scale experiments' and 'Quantification
173 of molecules'. Nevertheless, the instructions in many
174 journals are necessarily more general than specific com-
175 ments and metadata that are needed for electronic reposi-
176 tories and wide scale re-use of quantitative and qualitative
177 data are not necessarily reported. We have therefore
178 included such classic descriptors of good practice of plant
179 biological experiments and have consequently structured
180 our considerations for a minimum list of core metadata to
181 the following four major classes:
182

- 183 (i) The description of the physical object under investi-
184 gation (the 'biosource'), which includes genotypic and
185 spatial information
186 (ii) metadata relating to the (average) growth history of
187 plants, excluding treatments

- (iii) specifications of the physiological or biochemical
188 intervention(s) to which plants were subjected as
189 treatment
190
(iv) details of the harvest and post-harvest conditions, in
191 order to assess the conditions at harvest, and likely
192 alterations of metabolic contents due to quenching
193 and post-harvest storage conditions.
194

195 Consideration was also given to how the requirement
196 of such minimal metadata should be enforced when
197 metabolomics data are being submitted. For example,
198 although parameters are left sufficiently vague enough to
199 suit many different studies, certain metadata might still be
200 omitted from submission to journals or databases. In such
201 cases, it would be required to state and justify where
202 these omissions in the data occur. Valid reasons for not
203 reporting plant context metadata might be inaccessibility
204 of some data (e.g., in studies relating to plant products for
205 end consumers, or for certain field trials), or intellectual
206 property or commercial restrictions. However, even if
207 exact details cannot be given, authors of plant meta-
208 bolomics data would be required to give general
209 descriptions. Eventually, data or conclusions must be re-
210 jected if omissions of plant biology metadata are so se-
211 vere that the scientific conclusions cannot be reproduced
212 or understood.

213 3.1 BioSource

214 This term refers to the physical objects that were subjected
215 to metabolomic analyses, consisting of a description of the
216 species and genotype and the organ that was sampled, and
217 the bulk quantity of sample. In certain cases, more detailed
218 information may be available such as organ specifications,
219 cell types or subcellular compartments. These metadata are
220 only required and meaningful if sampling methods that
221 allowed such annotations were used. The methods used for
222 sampling should be named in such cases to enable inde-
223 pendent evaluation of data.

224 Details and explanations for required BioSource
225 metadata are given below:

Species Names of species should be described according to the NCBI
taxonomy database (Wheeler et al. 2000; Benson et al.
2000) (<http://pubmedexpress.nih.gov/Taxonomy/taxonomyhome.html/index.cgi>). Plant species need to be
named in full and not abbreviated, e.g. *Arabidopsis*
thaliana.

All necessary information on taxonomic relationships can be
derived from the correct species name and thus does not
need to be reported further.

continued

Genotype	<p>Subspecies information such as ecotype, cultivar and accession should be described according to authoritative databases such as TAIR (http://www.arabidopsis.org/). In the case of crosses or breeding results, available pedigree information must be given. In the case of transgenic or mutant organisms, name of the gene(s) that are up- or down-regulated should be reported, and the GenBank Accession number(s) for the sequence(s) of the corresponding construct(s), in addition to the parental subspecies background information, should be given.</p> <p>According to standard practice in agronomic genotype nomenclature, genotype description should comprise the author who first described or collected the cultivar, e.g. <i>Medicago truncatula</i> (Gaertn) cv. Jemalong. If available, registrations numbers for agronomic plants should be referenced, e.g. USDA GRIN. The number of backcrosses used in breeding needs to be detailed.</p> <p>In case of plant–pathogen interaction studies or other studies where information on multiple genomes is relevant, such metadata should be given.</p>
Organ	<p>Names of organs and plant structure should be described according to the authoritative database (Katica et al. 2007) maintained by the Plant Ontology Consortium to be found at http://www.plantontology.org/. All necessary information on organ relationships can be derived from the correct organ name and thus does not need to be reported further.</p>
Organ specification	<p>This should be provided <i>only if</i> such information cannot be detailed by http://www.plantontology.org/ (e.g. description of a part of an organ, the specific location of the organ or a specific tissue of an organ).</p>
Cell type	<p>This should be provided <i>only if</i> such information can be detailed in a meaningful manner, e.g. by cell sorting or dissection. Naming according to the authoritative database maintained by the Plant Ontology Consortium is to be found at http://www.plantontology.org/ under plant_structure ontology.</p> <p><i>Only if</i> such information cannot be located at this source the Cell Ontology maintained at Open Biomedical Ontologies group should be taken, which is to be found at http://lists.sourceforge.net/lists/listinfo/obo-cell-type.</p>
Subcellular location	<p>This should be described <i>only if</i> such information can be detailed in a meaningful manner, e.g. by subcellular fractionation. Naming according to the authoritative database (Gene Ontology Cellular Component) maintained by the Gene Ontology Consortium to be found at http://www.geneontology.org/</p>
BioSource amount	<p>This refers to the mass (mg fresh weight or mg dry weight), number of cells or other measurable bulk quantities (e.g. protein content).</p>

3.2 Growth environment

226

Many parameters of growth history are identical to all the plants in a given study. Researchers tend to refer to their specific growth environments as ‘standard growth conditions’ because they may not have altered these for a long time, or always use the same growth chambers and illumination conditions. However, environmental parameters are known to be very different between laboratories, and severely affect metabolite levels. Apart from obvious parameters such as (abiotic) stress conditions, even subtle alterations, such as emission spectra of light bulbs used for illumination, may cause differences in overall growth and metabolism. On the other hand, it is known that pulses of fertilizations will be reflected in changes in metabolism, and hence we suggest reporting both amount and timing of the nutritional regime. As guideline, parameters should be reported that can easily be monitored by plant researchers such as the type of growth media and light regimes; however, we intentionally do not suggest distinguishing between set points of growth conditions (such as temperature) and actually achieved parameters (which may have deviated from such set points). It is good laboratory practice to report deviations and fluctuations from controlled growth conditions; however, researchers may not be aware or may not have the instrumentation to monitor these parameters. This is an example of how ‘minimal requirements’ may be distinguished from ‘current best practice’ documents.

This section specifically excludes variation of growth conditions that were part of the experimental design, i.e. factors that were altered with the intention to cause metabolic differences. Such differences in study parameters should be reported as ‘treatment’. Although it cannot be made mandatory, documentation of additional metadata should be regarded as part of best plant biology practice, such as application of agrochemicals or biotic plant protection. When investigating other documents relating to the specifics of plant growth reporting requirements, a document was retrieved that was published by the International Committee for Controlled Environment Guidelines (<http://ncr101.montana.edu/>) in March 2004, detailing the ‘Minimum Guidelines for Measuring and Reporting Environmental Parameters for Experiments on Plants in Growth Rooms and Chambers’ (http://ncr101.montana.edu/min_guidelines.pdf). While we appreciated the efforts of this committee, we felt that many of the recommendations were rather demanding to be put into practice in current laboratory settings, especially in public research institutions. The intention of documents detailing minimal reporting standards, including the paper presented here, is to detail enough

275 information to re-use data and to understand the concepts and
 276 layout of experimental designs. If minimal requirements ask
 277 for a level of detail that is usually not reported by researchers,
 278 these can hardly serve as consensus, which would be
 279 endorsed and followed by the majority of active investiga-
 280 tors. Instead, we suggest such guidelines to be part of ‘best
 281 practice’ documents.

282 Details and explanations for the section ‘Plant Growth’
 283 comprise the following factors:

Growth support	Soil (type, supplier), Agar (type, supplier), Vermiculite (type, supplier), hydroponic system (type, supplier, nutrient concentrations) or other support including cell culture (media, volume, cell number per volume).
Growth location	Field trial (location), climate chamber (size m ³), greenhouse (details on accuracy of control of light, humidity and temperature conditions), other location (details on size m ³ , accuracy of control of light, humidity and temperature conditions).
Growth plot design	The way to randomize the different genotype × environment interactions. Either descriptive or using established nomenclature e.g. latin square.
Light	Light quality, light source model/type, light intensity (best reported as empirically measured at plant height), luminescence (daylight) period (h). For field trials: average light parameters in growing season. Information on time and location of the field trial enables tracking of more precise information if necessary.
Humidity	Humidity (%) at day and at night. For field trials: average humidity parameters in growing season. Information on time and location of the field trial enables tracking of more precise information if necessary.
Temperature	Temperature (°C) at day and at night. For field trials: average temperature (°C) at day and at night in growing season. Information about time and location of the field trial enables tracking of more precise information if necessary.
Watering regime	Amount and time of watering per day. For field trials: average rain fall in growing season. Information on time and location of the field trial enables tracking of more precise information if necessary. For hydroponic systems: frequency of solution change.
Nutritional regime	Amount and time of additional nutrients given to plants.
Date(s) of plant establishment	Depending on plant study, such dates could comprise: sowing, germination, transplanting, cutting, grafting or other appropriate time stamps.

continued

	Plant development stage description should accompany time stamps using established nomenclature (Boyes et al. 2001; Pujar et al. 2006; Palmquist et al. 2006).
Other specific metadata	Only if applicable. Examples comprise translocation of plants from one chamber to another, or the rotational schema of trays within a climate chamber. Examples comprise agrochemical or preventive maintenance information that is not part of ‘treatment’ factors.

3.3 Treatment

284

Plant biology study designs can be broadly classified according to *Genotype* × *Environment* interactions, or, for the sake of clarity, alterations of parameters that are denoted here as *BioSource* × *Treatment*. Publicly available and authoritative ‘treatment’ databases that label and detail the variety of treatment factors and their relative hierarchy and dependencies are not yet available in repositories like TAIR or PlantOntology. Hence, without further work on ontologies for such terms, it can only be recommended that terminology is used that is frequently found in plant research journals. In addition, we recommend a broad classification of treatment types (biotic, abiotic and intervention), which need to be complemented by information as to the dose or intensity levels, and time intervals or durations in which treatments were given. However, specific treatments (such as use of elicitors like methyljasmonate, abscisic acid or salicylic acid) is often termed as biotic stress treatment, and hence, there is yet some degree of ambiguity in nomenclature.

285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303

Treatment factors	Biotic treatment	E.g. infection (species), herbivore attack (species), competition with other plants (species) or other factors
	Abiotic treatment	E.g. light intensity variations, cold acclimation (temperature), heat stress, drought (description of residual growth support moisture, or quantitative description of reduction in watering regime), water stress, saline stress or other factors
	Intervention treatment	E.g. application of agrochemicals, enzyme inhibitors, hormones, elicitors or other factors
Treatment dose or intensity levels		Depending on treatment factors
Treatment time, time intervals and duration before harvest		Depending on treatment factors and treatment time

304 3.4 Harvest

305 The harvest determines the set point for stopping metabo-
 306 lism, analogous to sampling time points in related docu-
 307 ments on biology context metadata. However, apart from a
 308 simple time stamp, further metadata are required. For
 309 example, if harvests are reported from different research
 310 groups that were using identical plants and growth condi-
 311 tions, and sampling at the same hour of the day, results can
 312 still be different: one laboratory may have used a 16-h light
 313 periods beginning the illumination period at 06:00 a.m.,
 314 whereas the other laboratory may have begun illumination
 315 at 08:00 a.m. Hence, plants in the first study would be 2 h
 316 ahead in their daily period of photosynthesis and starch
 317 accumulation, which is known to cause metabolic altera-
 318 tions. Therefore, the time point and duration of harvest
 319 should be given relative to the photoperiod.

320 Another parameter is the age of the plants under study,
 321 e.g. the time between seed germination and date of harvest,
 322 may not necessarily convey similarity or comparability of
 323 growth and thus metabolic status. Even in controlled envi-
 324 ronments, some plants may flower earlier or grow faster
 325 than others, which refer to important turning points in the
 326 life cycle of a plant. Therefore, plant growth stages need to
 327 be defined relative to standardized growth stages. For some
 328 model species, like *Arabidopsis*, such growth stages have
 329 been defined in the literature; for other species, nomencla-
 330 ture should be used according to established terminology in
 331 plant journals. Most recently, an ontology for standard
 332 growth stages has been developed for angiosperms (Pujar
 333 et al. 2006), and it is recommended to exploit this resource
 334 for detailing plant metabolomic experiments.

335 Analogous to other 'biology context' documents, the
 336 method and time at which metabolism was stopped are also
 337 important to denote. Metabolites differ vastly in their turn-
 338 over rates, and some (such as glutathione or NADH) are
 339 extremely sensitive to oxidation. Therefore, details on the
 340 harvest methods need to be provided to enable the assess-
 341 ment and validation of metabolomic data acquisitions.

Harvest date, time	Harvest time relative to the luminescence cycle. Duration of harvest if relevant to the plant study (e.g. for volatile analysis).
Plant growth stage	It is advised to refer to established literature, e.g. for <i>Arabidopsis</i> (Boyes et al. 2001) and <i>Medicago truncatula</i> (Palmquist et al. 2006); and for general growth stage ontology (Pujar et al. 2006).
Metabolism quenching method	Time after harvest before stopping cellular metabolism. (may be greater than weeks for certain post-harvest physiology experiments, may be less than seconds for assessing high turnover metabolites). Method to stop cellular metabolism

continued

Harvest method	Details of operation to gather the plant organ (sample) Details of pooling of plant tissues for analysis
Sample storage	Operations to store sample (e.g. freeze-drying, grinding) prior to preparation for metabolomic analysis. Duration and temperature of storage before extraction for analysis.

4 Discussion

342 This document presents a first attempt to collate the min- 343
 344 imum required metadata for reporting plant metabolomic 345
 346 data. Many of the factors are regarded as classic descriptors 347
 348 for any study in plant biology. However, such metadata are 349
 350 often not explicitly detailed but rather used in reference to 351
 352 previous studies, or as in typical plant research reports, the 353
 354 information may be given scattered in text strings 355
 356 throughout the document. The aim here is to provide a 357
 358 guide for gross descriptors of plant studies, in order to 359
 360 allow comparisons of study designs, and to categorize 361
 362 experiments that will be reported in the literature or in 363
 364 databases. Parameters given here are still rather vague and 365
 366 allow for a number of exceptions and deviations. Specifi- 367
 368 cally, controlled vocabularies, ontologies, and exact defi- 369
 370 nitions for value units and string text restrictions need to be 371
 372 developed in order to implement such minimum require- 373
 374 ments into executable, queryable and MSI-compliant 375
 376 databases. For this reason, no UML schema can be given at 377
 378 this point in time. To this end, only for some factors such as 379
 380 'species', 'plant organs' 'growth stages' and 'cellular 381
 382 compartments', clearly defined hierarchies and repositories 383
 384 can be used. It will be important to continue and extend 385
 386 efforts on terminology requirements that are compiled by 387
 388 the MSI Ontology working group (Sansone et al. this 389
 390 issue), and to collaborate further with groups within the 391
 392 Open Biomedical Ontology consortium (OBO, <http://obo.sourceforge.net>). As one of the short-term goals, the list presented here will be incorporated into a 'minimum information' checklist (MIcheck, <http://micheck.sourceforge.net/>), which is envisioned to become a general repository for reporting standards in functional genomics and could be a resource for journal editorial guidelines.

393 It needs to be emphasized that the list presented here does 394
 395 not imply a sufficient or exhaustive description of plant studies 396
 397 in general, or even for plant metabolomics. Many parameters 398
 399 may require much greater detail if a full reproduction of an 399
 400 experiment is desired. For example, air circulation has not 401
 402 been listed as a parameter, despite its great role in evaporation 403
 404 and transpiration, and hence water availability to plants. In 405
 406 addition, air flow can also be a mechanical stress factor. In this 407
 408 discussion, we have opted against exhaustive lists of param- 409

384 eters because journal reports are also accepted without such
385 level of detail. Nevertheless, plant metabolomics researchers
386 are highly encouraged to collect additional details about their
387 specific studies, which will eventually translate into a better
388 understanding of plant metabolism by exchange and re-use
389 across laboratories and studies.

390 **Acknowledgments** Work on this document was partially supported
391 by a research grant MCB-0520140 from the U.S. National Science
392 Foundation.

393 References

394 Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., Rapp, B.
395 A., & Wheeler, D. L. (2000). GenBank. *Nucleic Acids Research*,
396 28, 15–18.
397 Bino, R. J., Hall, R. D., Fiehn, O., Kopka, J., Saito, K., Draper, J.,
398 Nikolau, B. J., Mendes, P., Roessner-Tunali, U., Beale, M. H.,
399 Trethewey, R. N., Lange, B. M., Wurtele, E. S., & Sumner, L.
400 W. (2004). Potential of metabolomics as a functional genomics
401 tool. *Trends in Plant Science*, 9, 418–425.
402 Boyes, D. C., Zayed, A. M., Ascenzi, R., McCaskill, A. J., Hoffman,
403 N. E., Davis, K. R., & Grolach, J. (2001). Growth stage-based
404 phenotypic analysis of Arabidopsis: A model for high throughput
405 functional genomics in plants. *Plant Cell*, 13, 1499–1510.
406 Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman,
407 P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C. A., Causton, H.
408 C., Gaasterland, T., Glenisson, P., Holstege, F. C. P., Kim, I. F.,
409 Markowitz, V., Matese, J. C., Parkinson, H., Robinson, A.,
410 Sarans, U., Schylze-Kremer, S., Stewart, J., Taylor, R., Violo, J.,
411 & Vingron, M. (2001). Minimum information about a micro-
412 array experiment (MIAME) – towards standards for microarray
413 data. *Nature Genetics*, 29, 365–371.
414 Fostel, J., Choi, D., Zwickl, C., Morrison, N., Rashid, A., Hasan, A.,
415 Bao, W., Richard, A., Tong, W., Bushel, P. R., Brown, R.,
416 Bruno, M., Cunningham, M. L., Dix, D., Eastin, W., Frade, C.,
417 Garcia, A., Heinloth, A., Irwin, R., Madenspacher, J., Merrick,
418 B. A., Papoian, T., Paules, R., Rocca-Serra, P., Sansone, A. S.,
419 Stevens, J., Tomer, K., Yang, C., & Waters, M. (2005). Chemical
420 effects in biological systems—data dictionary (CEBS-DD): a

compendium of terms for the capture and integration of 421
biological study design description, conventional phenotypes, 422
and 'omics data. *Toxicological Sciences*, 88, 585–601. 423
Jenkins, H., Hardy, N., Beckmann, M., Draper, J., Smith, A. R., 424
Taylor, J., Fiehn, O., Goodacre, R., Bino, R. J., Hall, R., Kopka, 425
J., Lane, G. A., Lange, B. M., Liu, J. R., Mendes, P., Nikolau, B. 426
J., Oliver, S. G., Paton, N. W., Rhee, S., Roessner-Tunali, U., 427
Saito, K., Smedsgaard, J., Sumner, L. W., Wang, T., Walsh, S., 428
Wurtele, E. S., & Kell, D. B. (2004). A proposed framework for 429
the description of plant metabolomics experiments and their 430
results. *Nature Biotechnology*, 22, 1601–1606. 431
Jenkins, H., Johnson, H., Kular, B., Wang, T., & Hardy, N. (2005). 432
Toward supportive data collection tools for plant metabolomics. 433
Plant Physiology, 138, 67–77. 434
Ilic, K., Kelloogg, E. A., Jaiswal, P., Zapata, F., Stevens, P. F., 435
Vincent, L. P., Avraham, S., Reiser, L., Pujar, A., Sachs, M. M., 436
Whitman, N. T., McCouch, S. R., Schaeffer, M. L., Ware, D. H., 437
Stein, L. D., & Rhee, S. Y. (2007) The plant structure ontology. a 438
unified vocabulary of anatomy and morphology of a flowering 439
plant. *Plant Physiology*, 143, 1–13. 440
Lindon, J., Nicholson, J., Holmes, E., Keun, H., Craig, A., Pearce, J., 441
Bruce, S., Hardy, N., Sansone, S., Antti, H., Jonsson, P., Daykin, 442
C., Navarange, M., Beger, R., Verheij, E., Amberg, A., 443
Baunsgaard, D., Cantor, G., Lehman-McKeeman, L., Earll, M., 444
Wold, S., Johansson, E., Haselden, J., Kramer, K., Thomas, C., 445
Lindberg, J., Schuppe-Koistinen, I., Wilson, I., Reily, M., 446
Robertson, D., Senn, H., Krotzky, A., Kochhar, S., Powell, J., 447
van der Ouderaa, F., Plumb, R., Schaefer, H., & Spraul, M. 448
(2005). Summary recommendations for standardization and 449
reporting of metabolic analyses. *Nature Biotechnology*, 23, 450
833–838. 451
Palmquist, D., & Vance, C. P. (2006). A standardized method for 452
analysis of *Medicago truncatula* phenotypic development. *Plant* 453
Physiology, 142, 207–219. 454
Pujar, A., Jaiswal, P., Kellogg, E. A., Ilic, K., Vincent, L., Avraham, 455
S., Stevens, P., Zapata, Z., Reiser, L., Rhee, S. Y., et al. (2006). 456
Whole plant growth stage ontology for angiosperms and its 457
application in plant biology. *Plant Physiology*, 142, 414–428. 458
Wheeler, D. L., Chappey, C., Lash, A. E., Leipe, D. D., Madden, T. 459
L., Schuler, G. D., Tatusova, T. A., & Rapp, B. A. (2000). 460
Database resources of the National Center for Biotechnology 461
Information. *Nucleic Acids Research*, 28, 10–14. 462
463

UNCORRECTED