The Plant Cell, Vol. 23: 2477–2482, July 2011, www.plantcell.org © 2011 American Society of Plant Biologists. All rights reserved.

LETTER TO THE EDITOR

Recommendations for Reporting Metabolite Data[W]

The last decade has witnessed dramatic advances in our ability to determine the levels of an ever broadening range of metabolites. That said, while transcriptomic approaches provide almost complete coverage (González-Ballester et al., 2010) and proteomics approaches are now capable of detecting upwards of 45% of the cellular protein complement (Baerenfaller et al., 2008), metabolomics is currently capable of determining only a small fraction of the metabolites found in any one cell (Saito and Matsuda, 2010). It has been estimated that at least 200,000 different metabolites occur across the plant kingdom, and between 7000 and 15,000 within an individual species (D'Auria and Gershenzon, 2005; Fernie, 2007). However, to date, even the most comprehensive methods detect only between 1000 and 2000 molecular features, which are putatively indicative of bona fide chemical entities synthesized in plant tissues (Aharoni et al., 2002; Giavalisco et al., 2009; Iijima et al., 2008). Metabolite measurements are further complicated by the chemical diversity of metabolites and their broad dynamic range in cellular abundance. These features currently prohibit the possibility of extracting and measuring all metabolites using single extraction and analytical procedures. Consequently, many different extraction techniques and combinations of analytical methods are employed in attempts to achieve adequate metabolite coverage (Lisec et al., 2006; De Vos et al., 2007; Kruger et al., 2008; Tohge and Fernie, 2010). This, in turn, renders the establishment of good working practices more difficult than those, for example, associated with quantitative RT-PCR (Udvardi et al., 2008). This is exacerbated by the breadth of aims of metabolite analyses that encompass targeted metabolite analysis, metabolite profiling, metabolomic-scale analyses, and metabolite fingerprinting techniques (Fiehn, 2002).

As a result of the diversity in aims and methodologies of metabolite analyses, it is particularly important to define guidelines for obtaining and reporting metabolite data, since there are so many potential sources of error or misinterpretation. Our aim is to highlight potential sources of error and provide recommendations to ensure the robustness of the metabolite data obtained and reported. Our recommendations include methods for sampling, extraction and storage, metabolite identification, processing of large sample numbers, and recommendations for reporting the methods of metabolite identification and the levels of certainty in metabolite quantification. Good suggestions for standards in reporting chemical ontology and supporting metadata have already been made by Sumner et al. (2007) and Bais et al. (2010).

SAMPLING, EXTRACTION, AND STORAGE OF METABOLITES

Compared to proteins and RNAs, many classes of metabolite, particularly intermediates in primary metabolism, have very rapid turnover times. For example, intermediates of the Calvin-Benson cycle and nucleotides turn over within fractions of a second (Stitt and Fernie, 2003; Fernie et al., 2004; Arrivault et al., 2009). For analysis of these metabolites, as well as for large-scale metabolomic analyses, it is necessary therefore to employ procedures for the immediate quenching of metabolism during extraction. For most applications, quick excision and snap-freezing in liquid nitrogen is recommended, with subsequent storage of deep-frozen tissue at constant −80°C. However, for bulky tissues, submersion in liquid nitrogen is not sufficient because the center of the tissue is cooled only slowly. Hence, for extractions from bulky tissues (i.e., those thicker than a standard leaf) and for the assay of extremely high-turnover metabolites, it is necessary to use more rapid quenching methods, such as freeze-clamping, in which the tissue is vigorously squashed flat between two prefrozen metal blocks (ap Rees et al., 1977; Badger et al., 1984). Irrespective of the method of quenching, it is also vital to avoid handling procedures that may lead to changes in the levels of the metabolites of interest in the last seconds or fractions of seconds before quenching. In addition, there are instances in which tissue handling can radically alter certain metabolites in a manner that reflects their biological characteristics. Examples of such compounds include cyanogenic glycosides and related compounds as well as certain types of volatiles. While these examples provide a strong illustration of the complexity inherent in producing comprehensive guidelines for metabolite reporting, we contest that such specialist measurements must likely be solved empirically. If material is to be freeze-dried, then this process must be performed to complete dryness and the stored material must then be sealed to prevent degradation. For example, incomplete freeze-drying will generate artifactual geometric isomers of pigments.

Samples should be stored in an appropriate manner both before and after extraction (Bais et al., 2010). Storage at temperatures between 0 and 40°C is especially problematic because substances can be concentrated in a residual aqueous phase. Short-term storage of liquid aqueous or organic solvent extracts, even at low temperatures (−20°C), is not recommended. The best approach to storage for many metabolic analyses is the removal of aqueous or organic solvent to create a dry residue. Deep-frozen samples should be processed as quickly as experimentally feasible; storage for weeks or months should be avoided or performed in liquid nitrogen. However, the appropriate means of storage is strictly dependent on the stability of the class of targeted metabolites or of the profiled metabolite fraction under study. Notably, the strategies mentioned above are not appropriate for volatile or even semivolatile metabolites—the analysis of which should preferably be performed on fresh material.

In cases where the standardized material is required throughout the entire study period (some studies can extend into

[W] Online version contains Web-only data.
www.plantcell.org/cgi/doi/10.1105/tpc.111.086272

Downloaded from https://academic.oup.com/plcell/article/23/7/2477/6097180 by U.S. Department of Justice user on 16 August 2022

months and even years), aliquots of a chemically defined repeatable standard mixture or of a standardized biological reference sample should be stored alongside samples.

REPLICATION

Another important issue is the nature and appropriate number of biological, technical, and analytical replicates. Biological replication can be misconstrued; for example, aliquots from a bulk preparation are not biological replicates. Biological replicates are normally from independent sources of the same genotype, grown under identical conditions; however, this definition is not directly applicable for ecological or evolutionary biology studies in which growth circumstance is often taken as a variable. In the case of transgenic primary transformants, this may be problematic as typically there are no replicates and often independent sampling of the plant, or vegetative clone of it, must be used. This problem can be circumvented by the analysis of multiple independent primary transformants and the appropriate statistical analysis; however, increased sampling is preferable. Technical replication involves independent performance of the complete analytical process rather than repeat injections of the same sample, the latter being an analytical replicate. While analytical replicates are useful in assessing machine performance, technical replicates encompassing the entire experimental procedure allow a far more comprehensive assessment of experimental variance in data generation. That said, biological replication is significantly more important than technological replication and should involve at least three and preferably more replicates. Care should be taken that these are harvested from as similar part of the plant as possible and at the same time of day. Also, a full and independent repeat of a biological experiment may be required to assess the robustness of results from metabolic studies (Sanchez et al., 2010). Whether technical replication is required is very much dependent on the precision of the analytical methods employed. In instances where technical variation is considerably lower than biological variation, it is sensible to sacrifice the former to increase the latter.

An essential yet less commonly adopted practice is the careful spatiotemporal randomization of biological replicates (with respect to experimental class) throughout biological experiments, sample preparation workflows, and instrumental analyses to minimize the influence of uncontrolled variables. For example, if a set of samples is analyzed in a nonrandom order, treatment and control samples can end up being analyzed at very different times and the resulting statistics can be greatly influenced by sample age or shifting machine performance, occluding the true biological interpretation of the data (Scholz et al., 2004). A simple and effective sample randomization approach is randomized-block design, which is equally applicable to field trials, sample processing, and instrumental analysis. The adoption of this approach in large-scale experiments is therefore strongly recommended.

INSTRUMENT PERFORMANCE AND DATA QUALITY

Another major, but easily preventable, problem in metabolomics is the publication of data sets without any information with which to assess instrument performance or data quality. Failure of many laboratories to adopt instrument performance tests, and report them, has likely resulted in the publication of a significant number of low-quality data sets that needlessly suffer from symptoms of poor instrument performance (e.g., missing or low signal-to-noise peaks). These issues could be ameliorated through the routine analysis of global-standard positive-control samples to verify the satisfactorily sensitive detection of expected, relevant metabolites. These samples, which could be incorporated into block-randomized analytical sequences as an additional experimental class, could be mixtures of authentic metabolite standards at defined concentrations. However, dry-stored aliquots of a broadly shared appropriate well-characterized global-standard biological extract (e.g., for example in experiments with *Arabidopsis thaliana*, samples from a multiple kilogram extraction of Columbia-0 plants grown both in optimal and stressed conditions to allow detection of the highest possible number or metabolites; Hannah et al., 2010; Vogel et al., 2010) are likely to be the most useful reference because they will provide a quantitative global reference for many metabolites that are of unknown structure or for which standards are not readily available. Such practice would enable enormous improvements in methods of data assessment, allow the establishment of objective minimum quality metrics, and would allow more meaningful quantitative comparisons to be made between different data sets, greatly enhancing the utility of emerging databases of plant metabolic phenotypes (Kopka et al., 2005; Scholz and Fiehn, 2007; Bais et al., 2010; Carroll et al., 2010). The use of and reporting on instrument performance tests is also strongly recommended, together with the incorporation of global standard reference extracts.

METABOLITE IDENTIFICATION

For metabolite identification, NMR represents the gold standard in structural identification (for example, certain chemical phytochemical journals demand it before accepting the description of a novel plant metabolite). NMR, largely due to its reliance on purely physical criteria, is highly reproducible such that instrument description, configuration, extract type, and data acquisition parameters alongside spectroscopic data outputs should suffice to allow unambiguous definition of a metabolite. That said, even NMR can produce ambiguous identifications, and total synthesis, x-ray chromatography, and methods capable of assigning absolute stereochemistry may ultimately be required. Identification for hyphenated mass spectrometry (MS) protocols (e.g., gas chromatography [GC]-MS and liquid chromatography [LC]-MS) can be more ambiguous, given that the chromatographic separation and subsequent mass determination often fail to distinguish between highly related metabolites or exact chemical isomers, which frequently exist in biological material. Sufficient information should be provided on the means of separation and the retention times (or better, retention time indices), as well as detailed mass data, when such methods are used for metabolite identification. The information needed will depend on the analytical equipment used (see Supplemental Tables 1 and 2 online). In the case of GC-MS, full information is needed about

Downloaded from https://academic.oup.com/plcell/article/23/7/2477/6097180 by U.S. Department of Justice user on 16 August 2022

the derivatization chemistry that was employed. In the case of LC-MS and capillary electrophoresis (CE)-MS, any measures taken to ameliorate problems of ion suppression should be noted. It is essential that the basis of identification is documented for GC-MS, LC-MS, CE-MS, or Fourier transform–ion cyclotron reactor-MS.

Specific criteria that should be considered are (1) annotation based on comparison to an authentic standard. The gold standard here would be an internal isotopically labeled standard (Fiehn et al., 2000; Birkemeyer et al., 2005; Lunn et al., 2006; Feldberg et al., 2009; Giavalisco et al., 2009), although control samples with spiked-in nonlabeled standards also represent a valid alternative if stable isotope–labeled standards are not available (Kopka et al., 2004). (2) If authentic standards are not available (which is often the case for secondary metabolites that are hard to purify or expensive to synthesize), the most probable metabolite identity or metabolite class can be matched to publically available databases or libraries (for example, see Carroll et al., 2010; also, the Golm Metabolome Database for GC-MS reference data; Kopka et al., 2005; Hummel et al., 2010). (3) In the case of accurate MS, a best hit to a sum formula can be derived. Multidimensional MS (MSn) data might be acquired and used for identification or support of the proposed structure hypothesis (Matsuda et al., 2009). If MSn is employed, fragment ion patterns of the multidimensional MS that forms the basis of the identification should be provided as supplemental data for publication together with the exact ionization and fragmentation settings of the instrumentation (which can vary dramatically within a single analytical system or between the different manufacturers' MS systems. (4) HPLC UV/Vis systems employed for metabolite determinations should demonstrate that extracts are of sufficient purity to enable the spectral quality and chromatographic resolution necessary for accurate quantification at their wavelength maxima. If necessary, multiple chromatographic systems should be used for individual components within a biosynthetic class of compounds. Analysts should demonstrate that standards and samples alike are free of interfering and/or coeluting compounds that may corrupt metabolite-specific and selective quantification by means of UV/Vis or fluorescence spectroscopy. (5) Enzyme-based determinations of metabolites should, on their first application, be supported by ample evidence that they are specific to the metabolite in question.

Application: Metabolomics of a Large Collection of Genotypes

While quality control of small-scale experiments is relatively facile regarding experiments involving mass genetic variance, such as profiling of recombinant inbred lines, introgression lines or libraries of ecotypes or cultivars represent a considerably different prospect. Work of this scale has recently been conducted in several plant species (Keurentjes et al., 2006; Schauer et al., 2006; Lisec et al., 2008; Rowe et al., 2008). Studies in tomato benefited from thorough optimization studies, which included metabolic recombination experiments (Roessner-Tunali et al., 2003; see Recovery Experiments below). In more recent studies (Sulpice et al., 2009, 2010; Do et al., 2010), aliquots from a master mix prepared from control genotypes were injected with each machine run resulting in a far better data quality. Given that the vast majority of metabolomic approaches provide relative rather than absolute levels, the adoption of similar approaches for large-scale biological experiments is highly recommended. For translational biology, it is important to note that empirical tests are required to optimize methods for the tissue in question (see Recovery Experiments below). In a recent survey, we demonstrated the potential pitfalls in measuring Calvin-Benson cycle intermediates in *Chlamydomonas* and secondary metabolites in the crop species tomato and rice (Tohge et al., 2011); such preliminary experiments are essential before embarking on large-scale screenings.

QUANTIFICATION

With quantification of metabolites, a further set of issues arises. Although many metabolite data are currently presented as relative values, the following comments are equally applicable to such data as to those resulting from absolute quantifications. (1) One important but frequently overlooked recommendation is to ensure that the levels of metabolites measured all lie within their linear ranges of detection. This is often impossible to achieve for all metabolites in a complex sample in a single analytical run due to the large dynamic range of metabolite levels in any biological sample (Sumner et al., 2003). This problem can be solved by running several independent dilutions of each extract, as has been performed in the analysis of *Arabidopsis* grown under different environmental conditions, transgenic tomato, and an introgression line population of tomato (Roessner-Tunali et al., 2003; Schauer et al., 2006; Arrivault et al., 2009). While this experimental solution should not be a prerequisite for publication, reporting that the measurements lie within the linear range of the analytical methods in the supplemental data is strongly recommended.

(2) Incomplete tissue disruption is one of the major sources of variation in a metabolite profiling workflow. For example, tomato skin is notoriously difficult to completely homogenize compared to fruit pericarp tissue but contains many important health-associated flavonoid compounds. Therefore, it is important to ensure that complete disruption of tissue has been achieved during extraction.

(3) Another problem is the evaluation of the stability of metabolites through the extraction, storage, and measurement processes. This can vary greatly from metabolite to metabolite, from extraction process to extraction process, and from tissue to tissue. This can be checked by recovery experiments, wherein known quantities of authentic standards can be added to an aliquot of frozen sample prior to extraction (typically at amounts equivalent to those found in the tissue), and their quantitative recovery can be assessed in comparison to an equivalent aliquot to which the standard was not added (see Application section below). In cases where recoveries are poor, it is often possible to define the step(s) at which the problems arise(s) by adding the standards at various stages in the pipeline. Following identification, problem stages in an extraction protocol can be optimized to alleviate or at least minimize the problem. Recovery experiments also provide an excellent cross-check of whether there is enough biological and technical replication; if this is inadequate, then the values will show a large spread both above and below 100%.

Downloaded from https://academic.oup.com/plcell/article/23/7/2477/6097180 by U.S. Department of Justice user on 16 August 2022

(4) In the case of poorly characterized tissues, estimations of the limits of detection, limits of identification, and limits of quantification are also useful, especially of the key classes of metabolites in the extracts.

Application: Recovery Experiments

Recovery experiments were previously vigorously championed by ap Rees and Hill (1994) and Dancer and ap Rees (1989) and can provide persuasive evidence that the data reported represent a valid reflection of cellular metabolite compositions. Recent examples of their application can be found in Roessner et al. (2000), Lunn et al. (2006), and Arrivault et al. (2009). However, the metabolomics community has been relatively slow in adopting these control procedures. One reason is that such experiments are possible only for compounds that are commercially available and/or easy to synthesize chemically. Another is that metabolomics, by definition, looks at a very wide range of metabolites for most of which there are no standards available. This is obviously always the case for unknown compounds, for which this approach is impossible. However, there is an alternative approach that does not suffer from this practical limitation. This is to combine a novel plant tissue with one that has been previously very well characterized, such as *Arabidopsis* Columbia-0 leaves. Such experiments also allow a quantitative assessment of the reliability of known peaks (Roessner-Tunali et al., 2003). A schematic representation of recovery and metabolic recombination experiments is presented in Supplemental Table 1 online. The approach of combining a novel tissue with a well-characterized standard tissue could in some cases run into practical problems in that it may result in such a complex chemical matrix that analysis becomes difficult or impossible, and/or many of the studied metabolites may be absent in the standard tissue. Nonetheless, this may be partially circumvented by choice of the appropriate reference tissue. In most cases, it should at least be possible to use metabolic recombination within a given experiment. For example, to support a claim that metabolites change as the result of a treatment by a control experiment, samples from most extreme treatments, be they environmentally or genetically determined, could be mixed and then extracted and analyzed in parallel with the unmixed samples. We suggest that for known metabolites, recovery or metabolic recombination experiments are performed for each new tissue/species type under study. While it is clear that for any metabolomics-scale study, certain metabolites will have poor recoveries, while this does not preclude the reporting of their values it is important that this is documented to allow the reader's discretion in interpretation of such data. For unknown metabolites, precise documentation of chromatographic and spectral properties should suffice. For both recovery and metabolic recombination experiments, technical repetition only is sufficient and recoveries of between 80 and 120% are acceptable (values above 100% will be achieved as a simple result of variance associated with biological material and the analytic procedures). Anything deviating beyond this range represents a metabolite whose quantification should be deemed questionable or unreliable.

DOCUMENTING NOVELTIES

It is our contention that any study reporting the application of a given protocol for the first time on a novel species or tissue type, any first study of a genotype exhibiting a dramatically altered chemotype, or an environmental or physiological treatment genotype exhibiting a dramatically altered chemotype should, if using a profiling approach, carry out such experiments and document their results. An obvious alternative route would be the development of a targeted protocol for certain metabolites; however, equally vigorous controls for such a protocol should also be observed. We realize that the best validation strategy will depend on the experiment, the biological material, the analytical platform, and the kinds of metabolites that are being studied and that there is always a balance to be struck between perfection and practicality. Nevertheless, attempting to apply these practices would greatly enhance the reliability of quantitative aspects of metabolite data.

Application: Tissue Distribution of Previously Uncharacterized Metabolites

Another interesting aspect derived from metabolomics is the discovery of novel compounds and elucidation of their biosynthesis. Such novel discoveries can be divided into two cases: (1) discovery of a novel compound and (2) finding a new metabolite in analyzed plant species. The first case refers to the discovery of novel metabolite and indicates a discovery of an entirely new metabolic pathway that includes finding of novel regulators and enzymatic genes. The second case represents the novel observation of a compound in a particular tissue and/or plant species. In some cases, the metabolite visualization database based on the eFP Browser (AtMetExpress development; Matsuda et al., 2010) is quite helpful for comparing metabolite abundance with gene expression data. Several examples of the discovery of uncharacterized metabolites function have been shown in comparison of the difference between tissue types and wild accessions. The flavonol-3-*O*-arabinosyltransferase (At5g17030) was characterized according to a discovery of flavonol-3-*O*-arabinoside-7-*O*-rhamnosides in *Arabidopsis* flower by comprehensive flavonol profiling (Yonekura-Sakakibara et al., 2008). Moreover, the detection of *Arabidopsis* accession-specific flavonol glycoside revealed several candidate genes following integration with microarray data (Tohge and Fernie, 2010). The 2-oxoglutarate dependent dioxygenases (AOP2 and AOP3) of glucosinolate production were identified on the basis of metabolite profiling comparisons between various *Arabidopsis* accession and tissues (Kliebenstein et al., 2003), while loganic acid methyltransferase was characterized using the metabolomic differences of tissues by the leaf epidermome of *Catharanthus* (Murata et al., 2008). However, there are some important considerations for accomplishing this approach. First, unstable metabolites and their associated breakdown compounds, for example, compound derived from enzymatic breakdown, and pigments that are stereo-isomerized by light irradiation, should be taken into account. Second, the accuracy of peak prediction needs to be carefully considered. As we discussed above, metabolic recombination experiments

Downloaded from https://academic.oup.com/plcell/article/23/7/2477/6097180 by U.S. Department of Justice user on 16 August 2022

should ideally be performed. Additional targeted experiments, such as fractionation, enzymatic assay, hydrolysis, and a test of derivatization of certain moieties, will additionally aid in elucidation of the accurate chemical structure of novel metabolites.

CONCLUSION

To simplify the adoption of these recommendations, we supply Supplemental Tables 1 and 2 online as an Excel spreadsheet. Supplemental Table 1 contains a list of simple questions regarding the reporting of metabolites data, which we recommend be completed on the submission of future manuscripts. Supplemental Table 2 provides recommendations for supplemental data to be included with the presentation of a typical LC-MS experiment. We suggest that the use of these tables will improve the reporting of metabolite data and will enhance community efforts to improve the annotation of plant metabolomes.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Table 1. Metabolite Reporting Checklist.

Supplemental Table 2. Recommendations for GC- and LC-MS.

Received April 10, 2011; revised June 14, 2011; accepted June 30, 2011; published July 19, 2011.

Alisdair R. Fernie
Max-Planck-Institut für Molekulare Pflanzenphysiologie
14476 Golm, Germany

Asaph Aharoni
Department of Plant Sciences
Weizmann Institute of Sciences
Rehovot 76100, Israel

Lothar Willmitzer
Max-Planck-Institut für Molekulare Pflanzenphysiologie
14476 Golm, Germany

Mark Stitt
Max-Planck-Institut für Molekulare Pflanzenphysiologie
14476 Golm, Germany

Takayuki Tohge
Max-Planck-Institut für Molekulare Pflanzenphysiologie
14476 Golm, Germany

Joachim Kopka
Max-Planck-Institut für Molekulare Pflanzenphysiologie
14476 Golm, Germany

Adam J. Carroll
Plant Energy Biology–Australian Research Council Centre of Excellence
Research School of Biology
Australian National University
Canberra, Australian Capital Territory 0200, Australia

Kazuki Saito
RIKEN Plant Science Center
Yokohama, Kanagawa 230-0045
Japan Graduate School of Pharmaceutical Sciences
Chiba University
Chiba 263-8522, Japan

Paul D. Fraser
School of Biological Sciences
Royal Holloway, University of London
Egham TW20 0EX
United Kingdom

Vincenzo DeLuca
Biology Department, Brock University
St. Catharines
Ontario L2S3A1, Canada

REFERENCES

Aharoni, A., Ric de Vos, C.H., Verhoeven, H.A., Maliepaard, C.A., Kruppa, G., Bino, R., and Goodenowe, D.B. (2002). Nontargeted metabolome analysis by use of Fourier Transform Ion Cyclotron Mass Spectrometry. OMICS **6:** 217–234.

ap Rees, T., Fuller, W.A., and Wright, B.W. (1977). Measurements of glycolytic intermediates during the onset of thermogenesis in the spadix of *Arum maculatum.* Biochim. Biophys. Acta **461:** 274–282.

ap Rees, T., and Hill, S.A. (1994). Metabolic control analysis of plant metabolism. Plant Cell Environ. **17:** 587–599.

Arrivault, S., Guenther, M., Ivakov, A., Feil, R., Vosloh, D., van Dongen, J.T., Sulpice, R., and Stitt, M. (2009). Use of reverse-phase liquid chromatography, linked to tandem mass spectrometry, to profile the Calvin cycle and other metabolic intermediates in Arabidopsis rosettes at different carbon dioxide concentrations. Plant J. **59:** 826–839.

Badger, M.R., Sharkey, T.D., and von Caemmerer, S. (1984). The relationship between steady–state gas exchange of bean leaves and the levels of carbon–reduction-cycle intermediates. Planta **160:** 305–313.

Baerenfaller, K., Grossmann, J., Grobei, M.A., Hull, R., Hirsch-Hoffmann, M., Yalovsky, S., Zimmermann, P., Grossniklaus, U., Gruissem, W., and Baginsky, S. (2008). Genome-scale proteomics reveals *Arabidopsis thaliana* gene models and proteome dynamics. Science **320:** 938–941.

Bais, P., et al. (2010). PlantMetabolomics.org: A web portal for plant metabolomics experiments. Plant Physiol. **152:** 1807–1816.

Birkemeyer, C., Luedemann, A., Wagner, C., Erban, A., and Kopka, J. (2005). Metabolome analysis: The potential of in vivo labeling with stable isotopes for metabolite profiling. Trends Biotechnol. **23:** 28–33.

Carroll, A.J., Badger, M.R., and Harvey Millar, A. (2010). The MetabolomeExpress Project: Enabling web-based processing, analysis and transparent dissemination of GC/MS metabolomics datasets. BMC Bioinformatics **11:** 376.

Dancer, J.E., and ap Rees, T. (1989). Phosphoribosyl pyrophosphate and the measurement of inorganic pyrophosphate in plant tissues. Planta **95:** 261–264.

D'Auria, J.C., and Gershenzon, J. (2005). The secondary metabolism of *Arabidopsis thaliana*: Growing like a weed. Curr. Opin. Plant Biol. **8:** 308–316.

De Vos, R.C., Moco, S., Lommen, A., Keurentjes, J.J., Bino, R.J., and Hall, R. D. (2007). Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. Nat. Protoc. **2:** 778–791.

Do, P.T., Prudent, M., Sulpice, R., Causse, M., and Fernie, A.R. (2010). The influence of fruit load on the tomato pericarp metabolome in a *Solanum chmielewskii* introgression line population. Plant Physiol. **154:** 1128–1142.

Fernie, A.R. (2007). The future of metabolic phytochemistry: Large numbers or metabolites, higher resolution, greater understanding. Phytochemistry **68:** 2861–2880.

Fiehn, O. (2002). Metabolomics—The link between genotypes and phenotypes. Plant Mol. Biol. **48:** 155–171.

Fiehn, O., Kopka, J., Dörmann, P., Altmann, T., Trethewey, R.N., and Willmitzer, L. (2000). Metabolite profiling for plant functional genomics. Nat. Biotechnol. **18:** 1157–1161.

Feldberg, L., Venger, I., Malitsky, S., Rogachev, I., and Aharoni, A. (2009). Dual labeling of

Downloaded from https://academic.oup.com/plcell/article/23/7/2477/6097180 by U.S. Department of Justice user on 16 August 2022

metabolites for metabolome analysis (DLEM-MA): A new approach for the identification and relative quantification of metabolites by means of dual isotope labeling and liquid chromatography-mass spectrometry. Anal. Chem. **81:** 9257–9266.

Fernie, A.R., Trethewey, R.N., Krotzky, A.J., and Willmitzer, L. (2004). Metabolite profiling: From diagnostics to systems biology. Nat. Rev. Mol. Cell Biol. **5:** 763–769.

Giavalisco, P., Köhl, K., Hummel, J., Seiwert, B., and Willmitzer, L. (2009). ^{13}C isotope-labeled metabolomes allowing for improved compound annotation and relative quantification in liquid chromatography-mass spectrometry-based metabolomic research. Anal. Chem. **81:** 6546–6551.

González-Ballester, D., Casero, D., Cokus, S., Pellegrini, M., Merchant, S.S., and Grossman, A.R. (2010). RNA-seq analysis of sulfur-deprived *Chlamydomonas* cells reveals aspects of acclimation critical for cell survival. Plant Cell **22:** 2058–2084.

Hannah, M.A., Caldana, C., Steinhauser, D., Balbo, I., Fernie, A.R., and Willmitzer, L. (2010). Combined transcript and metabolite profiling of Arabidopsis grown under widely variant growth conditions facilitates the identification of novel metabolite-mediated regulation of gene expression. Plant Physiol. **152:** 2120–2129.

Hummel, J., Strehmel, N., Selbig, J., Walther, D., and Kopka, J. (2010). Decision tree supported substructure prediction of metabolites from GC-MS profiles. Metabolomics **6:** 322–333.

Iijima, Y., et al. (2008). Metabolite annotations based on the integration of mass spectral information. Plant J. **54:** 949–962.

Keurentjes, J.J.B., Fu, J.Y., de Vos, C.H.R., Lommen, A., Hall, R.D., Bino, R.J., van der Plas, L.H.W., Jansen, R.C., Vreugdenhil, D., and Koornneef, M. (2006). The genetics of plant metabolism. Nat. Genet. **38:** 842–849.

Kliebenstein, D.J., Lim, J.E., Landry, L.G., and Last, R.L. (2003). Arabidopsis UVR8 regulates ultraviolet-B signal transduction and tolerance and contains sequence similarity to human regulator of chromatin condensation 1. Plant Physiol. **130:** 234–243.

Kopka, J., Fernie, A., Weckwerth, W., Gibon, Y., and Stitt, M. (2004). Metabolite profiling in plant biology: Platforms and destinations. Genome Biol. **5:** 109.

Kopka, J., et al. (2005). GMD@CSB.DB: The Golm Metabolome Database. Bioinformatics **21:** 1635–1638.

Kruger, N.J., Troncoso-Ponce, M.A., and Ratcliffe, R.G. (2008). 1H NMR metabolite fingerprinting and metabolomic analysis of perchloric acid extracts from plant tissues. Nat. Protoc. **3:** 1001–1012.

Lisec, J., Meyer, R.C., Steinfath, M., Redestig, H., Becher, M., Witucka-Wall, H., Fiehn, O., Törjék, O., Selbig, J., Altmann, T., and Willmitzer, L. (2008). Identification of metabolic and biomass QTL in Arabidopsis thaliana in a parallel analysis of RIL and IL populations. Plant J. **53:** 960–972.

Lisec, J., Schauer, N., Kopka, J., Willmitzer, L., and Fernie, A.R. (2006). Gas chromatography mass spectrometry-based metabolite profiling in plants. Nat. Protoc. **1:** 387–396.

Lunn, J., Feil, R., Hendriks, J.H., Gibon, Y., Morcuende, R., Osuna, D., Scheible, W.R., Carillo, P., Hajirezaei, M.R., and Stitt, M. (2006). Sugar-induced increases in trehalose-6-phosphate are correlated with redox activation of ADPglucose pyrophosphorylase and higher rates of starch synthesis in *Arabidopsis thaliana*. Biochem. J. **397:** 139–148.

Matsuda, F., Hirai, M.Y., Sasaki, E., Akiyama, K., Yonekura-Sakakibara, K., Provart, N.J., Sakurai, T., Shimada, Y., and Saito, K. (2010). AtMetExpress development: A phytochemical atlas of Arabidopsis development. Plant Physiol. **152:** 566–578.

Matsuda, F., Yonekura-Sakakibara, K., Niida, R., Kuromori, T., Shinozaki, K., and Saito, K. (2009). MS/MS spectral tag-based annotation of non-targeted profile of plant secondary metabolites. Plant J. **57:** 555–577.

Murata, J., Roepke, J., Gordon, H., and De Luca, V. (2008). The leaf epidermome of *Catharanthus roseus* reveals its biochemical specialization. Plant Cell **20:** 524–542.

Roessner, U., Wagner, C., Kopka, J., Trethewey, R.N., and Willmitzer, L. (2000). Technical advance: Simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry. Plant J. **23:** 131–142.

Roessner-Tunali, U., Hegemann, B., Lytovchenko, A., Carrari, F., Bruedigam, C., Granot, D., and Fernie, A.R. (2003). Metabolic profiling of transgenic tomato plants overexpressing hexokinase reveals that the influence of hexose phosphorylation diminishes during fruit development. Plant Physiol. **133:** 84–99.

Rowe, H.C., Hansen, B.G., Halkier, B.A., and Kliebenstein, D.J. (2008). Biochemical networks and epistasis shape the *Arabidopsis thaliana* metabolome. Plant Cell **20:** 1199–1216.

Saito, K., and Matsuda, F. (2010). Metabolomics for functional genomics, systems biology, and biotechnology. Annu. Rev. Plant Biol. **61:** 463–489.

Sanchez, D.H., Szymanski, J., Erban, A., Udvardi, M.K., and Kopka, J. (2010). Mining for robust transcriptional and metabolic responses to long-term salt stress: A case study on the model legume *Lotus japonicus.* Plant Cell Environ. **33:** 468–480.

Schauer, N., et al. (2006). Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. Nat. Biotechnol. **24:** 447–454.

Scholz, M., and Fiehn, O. (2007). Setup X–A public study design database for metabolomic projects. Pac. Symp. Biocomput.**12:** 169–180.

Scholz, M., Gatzek, S., Sterling, A., Fiehn, O., and Selbig, J. (2004). Metabolite fingerprinting: Detecting biological features by independent component analysis. Bioinformatics **20:** 2447–2454.

Stitt, M., and Fernie, A.R. (2003). From measurements of metabolites to metabolomics: An 'on the fly' perspective illustrated by recent studies of carbon-nitrogen interactions. Curr. Opin. Biotechnol. **14:** 136–144.

Sulpice, R., et al. (2009). Starch as a major integrator in the regulation of plant growth. Proc. Natl. Acad. Sci. USA **106:** 10348–10353.

Sulpice, R., Sienkiewicz-Porzucek, A., Osorio, S., Krahnert, I., Stitt, M., Fernie, A.R., and Nunes-Nesi, A. (2010). Mild reductions in cytosolic NADP-dependent isocitrate dehydrogenase activity result in lower amino acid contents and pigmentation without impacting growth. Amino Acids **39:** 1055–1066.

Sumner, L.W., et al. (2007). Proposed minimum reporting standards for chemical analysis. Metabolomics **3:** 211–221.

Sumner, L.W., Mendes, P., and Dixon, R.A. (2003). Plant metabolomics: Large-scale phytochemistry in the functional genomics era. Phytochemistry **62:** 817–836.

Tohge, T., and Fernie, A.R. (2010). Combining genetic diversity, informatics and metabolomics to facilitate annotation of plant gene function. Nat. Protoc. **5:** 1210–1227.

Tohge, T., Mettler, T., Arrivault, S., Carroll, A.J., Stitt, M., and Fernie, A.R. (2011). From models to crop species: Caveats and solutions for translational metabolomics. Front. Plant Phys., in press.

Udvardi, M.K., Czechowski, T., and Scheible, W.R. (2008). Eleven golden rules of quantitative RT-PCR. Plant Cell **20:** 1736–1737.

Vogel, J.T., Walter, M.H., Giavalisco, P., Lytovchenko, A., Kohlen, W., Charnikhova, T., Simkin, A.J., Goulet, C., Strack, D., Bouwmeester, H.J., Fernie, A.R., and Klee, H.J. (2010). SlCCD7 controls strigolactone biosynthesis, shoot branching and mycorrhiza-induced apocarotenoid formation in tomato. Plant J. **61:** 300–311.

Yonekura-Sakakibara, K., Tohge, T., Matsuda, F., Nakabayashi, R., Takayama, H., Niida, R., Watanabe-Takahashi, A., Inoue, E., and Saito, K. (2008). Comprehensive flavonol profiling and transcriptome coexpression analysis leading to decoding gene-metabolite correlations in *Arabidopsis*. Plant Cell **20:** 2160–2176.

Downloaded from https://academic.oup.com/plcell/article/23/7/2477/6097180 by U.S. Department of Justice user on 16 August 2022