Journal of Visualized Experiments

High-throughput metabolic profiling for model refinements of microalgae

--Manuscript Draft--

Article Type:	Invited Methods Article - Author Produced Video
Manuscript Number:	JoVE61913R2
Full Title:	High-throughput metabolic profiling for model refinements of microalgae
Corresponding Author:	Kourosh Salehi-Ashtiani UNITED ARAB EMIRATES
Corresponding Author's Institution:	
Corresponding Author E-Mail:	ksa3@nyu.edu
Order of Authors:	Amnah Alzahmi
	Sarah Daakour
	Diana El Assal
	Bushra Dohai
	Amphun Chaiboonchoe
	Weiqi Fu
	David Nelson
	Alexandra Mystikou
	Kourosh Salehi-Ashtiani
Additional Information:	
Question	**Response**
Please indicate whether this article will be Standard Access or Open Access.	
Please specify the section of the submitted manuscript.	Biology
Please confirm that you have read and agree to the terms and conditions of the author license agreement that applies below:	I agree to the Author License Agreement
Please provide any comments to the journal here.	The revised video file is too large to upload, it will be emailed to Dr. Nam Nguyen, the editor handling the manuscript
Please indicate whether this article will be Standard Access or Open Access.	Open Access ($3900)
Please confirm that you have read and agree to the terms and conditions of the video release that applies below:	I agree to the Video Release

TITLE:

High-Throughput Metabolic Profiling for Model Refinements of Microalgae

AUTHORS & AFFILIATIONS:

Amnah Alzahmi[1,2], Sarah Daakour[1], Diana Charles El Assal[3], Bushra S. Dohai[3], Amphun Chaiboonchoe[3], Weiqi Fu[3,4], David R. Nelson[1], Alexandra Mystikou[1], Kourosh Salehi-Ashtiani[1,3]

[1]Center for Genomics and Systems Biology (CGSB), New York University Abu Dhabi Research Institute, Abu Dhabi, UAE
[2]Department of Biology, United Arab Emirates University, Al Ain, UAE
[3]Laboratory of Algal, Systems, and Synthetic Biology (LASSB), Division of Science and Math, New York University Abu Dhabi, Abu Dhabi, UAE
[4]Department of Marine Science, Ocean College, Zhejiang University, Zhoushan, Zhejiang, China

amnah.alzahmi@nyu.edu
sd145@nyu.edu
dce1@nyu.edu
bsd4@nyu.edu
amphun@yahoo.com
wf21@nyu.edu
drn2@nyu.edu
am8406@nyu.edu

Corresponding Author; Kourosh Salehi-Ashtiani
Email Address: ksa3@nyu.edu

KEYWORDS:

Microalgae, *Chlamydomonas reinhardtii*, phenotype microarray, flux balance analysis, metabolic network reconstruction, metabolic model refinement

SHORT ABSTRACT:

This protocol demonstrates the use of a phenotype microarray (PM) technology platform to define metabolic requirements of *Chlamydomonas reinhardtii,* a green microalga, and refine an existing metabolic network model.

LONG ABSTRACT:

Metabolic models are reconstructed based on an organism's available genome annotation and provide predictive tools to study metabolic processes at a systems-level. Genome-scale metabolic models may include gaps as well as reactions that are unverified experimentally. Reconstructed models of newly isolated microalgal species will result in weaknesses due to these gaps, as there is usually sparse biochemical evidence available for the metabolism of such isolates. The phenotype microarray (PM) technology is an effective, high-throughput method that functionally determines cellular metabolic activities, in response to a wide array of entry metabolites. Combining the high throughput phenotypic assays with metabolic modeling can

allow existing metabolic network models to be rapidly reconstructed or optimized by providing biochemical evidence to support and expand genomic evidence. This work will show the use of PM assays for the study of microalgae by using the green microalgal model species *Chlamydomonas reinhardtii* as an example. Experimental evidence for over 254 reactions obtained by PM was used in this study to expand and refine a genome-scale *C. reinhardtii* metabolic network model, *i*RC1080, by approximately 25 percent. The protocol created here can be used as a basis for functionally profiling the metabolism of other microalgae, including known microalgae mutants and new isolates.

INTRODUCTION:

Optimizing algal metabolism for enhanced and stable production of targeted metabolites requires the development of complex metabolic engineering strategies through systems-level analyses of metabolic networks. Metabolic network models can guide the rational designs for the rapid development of optimization strategies[1-4]. Although approximately 160 microalgal species have been sequenced[5], there are, to our knowledge, only 44 algal metabolic models available[4,6,7]. Due to the difficulty in obtaining high-throughput metabolic phenotypic data for experimental validation of genomic information, the reconstruction of high-quality network models lags behind the rapid development of algal genome sequencing.

C. reinhardtii is an attractive model system for algal-based studies. This species can grow photoautotrophically or heterotrophically and has been widely used as a model organism in basic and applied research. Its genome sequence was published in 2007[8] with genome-scale metabolic models subsequently reconstructed for the species[9-11]. The genome-scale model for *C. reinhardtii* (*i*RC1080) was reconstructed by Chang *et al.*[10] based on genomic and literature evidence (entailing~250 sources). It has 1,706 metabolites with 2,190 reactions[10]; however, the completeness of the model could not be verified beyond the available published experimental evidence at the time.

The phenotype microarrays (PMs) technology is a high-throughput platform that can provide metabolic profiling information for heterotrophic microorganisms as well as tissue-culture cells. In particular, it can be used to address the phenotype-to-genotype knowledge gap in microalgae, as first reported for *Chlamydomonas reinhardtii*[12] and subsequently for a species of Chloroidium[13] and Chlorella[14]. By studying cell responses to thousands of metabolites, signaling molecules, osmolytes, and effector molecules, the PM assays can provide functional metabolic profiling and offer insights into the function, metabolism, and environmental sensitivity[15-17]. Specifically, PM assays detect cells metabolite utilization in 96-well microplates with different nutrients, metabolites, or osmolytes contained in each well. Moreover, it is also possible to assay bioactive molecules, such as antibiotics and hormones. As determined by the intensity of color production by the NADH reduction of a tetrazolium-based redox dye, the metabolic utilization of substrates is evaluated in terms of cell respiration[15-17]. The experiments in 96-well microplates can be monitored and determined automatically over time with the phenotype microarray instrument (PMI) platform. Twenty 96-well microplates are designed to represent the common set metabolites to study cellular phenotypes to utilize carbon, nitrogen, sulfur, and phosphorus sources, along with different osmotic/ion and pH effects. The PM technology

has been successfully used for updating and upgrading a number of existing genome-scale metabolic models for microorganisms[15-18].

The protocol and data shown here are based on previously published work by Chaiboonchoe *et al.*[12] .The presented work details the use of the PM assay method to characterize the metabolic phenotypes of microalgae and to expand an existing algal metabolic model of *C. reinhardtii* as well as to guide the reconstruction of new metabolic models.

PROTOCOL:

1. Phenotype Microarray Experiments

1.1 Obtain *C. reinhardtii* strain CC-503 from the Chlamydomonas Resource Center at the University of Minnesota, USA (https://www.chlamycollection.org).

1.2 Grow the cells in fresh Tris-Acetate-Phosphate (TAP) media[19] with final concentrations of 400 µg/mL timentin, 50 µg/mL ampicillin, and 100 µg/mL kanamycin (to inhibit bacterial growth) under 400 micromol photons/m^2s, at 25 °C, for two days to mid-log phase.

1.3 Spin down the culture at 2,000 x g for 10 min at 22 °C and discard the supernatant without disturbing the pellet.

1.4 Prepare fresh TAP media containing 0.1% tetrazolium violet dye “D”.

NOTE: Modify TAP media in this step to exclude some nutrients depending on the metabolite category tested in each plate (e.g., exclude ammonium chloride for nitrogen source plates).

1.5 Resuspend the pellet in fresh TAP media prepared (from step 1.2) to a final concentration of 1 x 10^6 cells/mL.

1.6 Use chemical compound array assay plates (carbon sources, nitrogen sources, phosphorus and sulfur sources plates, and the peptide nitrogen sources).

1.7 Inoculate a 100 µL aliquot of cell-containing media into each well of the assay plates.

NOTE: Make sure to duplicate the assays.

1.8 Streak cells on yeast extract/peptone plates and perform gram staining, as in Smith *et al.*[20] before and after the assay to monitor bacterial contamination.

1.9 Insert the chemical compound array assay plates into the microplate reader system.

1.10 Incubate all the plates at 30 °C for up to 7 days and program the microplate reader system to read the dye color change every 15 min.

NOTE: As most microplate readers do not provide a source of continuous light during incubation, the algae should be able to carry out heterotrophic respiration.

2. Data Analysis

2.1 Export the raw kinetic data from the microplate reader as CSV files, which will subsequently be used as input to the Omnilog Phenotype Microarray (OPM) package in R. Add the biological information as metadata (e.g., strain designation, growth media, temperature, etc.).

2.1.1 Using the PM Kinetic data converter software; load the D5E data files, and convert them to OKA files using the following command lines in the PM kinetic analysis software:
Load | Import (locate the folder of the OKA files) | Populate Filters | Import | Add All Plates | Close.
Export | choose read data **(Kinetic),** choose format **(CSV) (Tabulate Header)**, and choose plates **(every plate (individual Files)) | Export data | Save.**

2.1.2 To carry out the Phenotype Microarray (PM) data analysis, use the OPM software package[21,22] that runs within the R software environment. The package, tutorial, and reference documentations are available at: http://www.goeker.org/opm/. In RStudio, a graphical user interface for R, install the opm package and its dependencies using the following commands:
source (http://www.goeker.org/opm/install_opm.R)
library (opm)

2.1.3 Navigate to the directory that contains the CSV files of the kinetic data and import the data using the read_*opm* function:
x <- read_opm(".", convert="grp", include=list ("csv:"))

2.1.4 Aggregate and discretize the kinetic data using curve-parameter estimation.

```
For (i in 1:length(x)) {
 x[[i]] <I do _aggre(x[[i]], boot = 0L, cores = 1L, method = "splines", options = set_spline_options("smooth.spline"))
 x[[i]] <- do_disc(x[[i]], cutoff = FALSE)
}

#Collection of the metadata
metadata <- collect_template(".")
metadata$Strain <- c("BLANK", "CC-503")
for (I in 1:length(x)) {x[[i]] <- include_metadata(x[[i]], md = metadata, replace = TRUE)}
```

2.1.5 Use the function *xy_plot* to map the respiration (or growth) measurements (y-axis) as a function of time (x-axis) for the assayed 96-well plates.

```
print (xy_plot(x[[1]], include = "Strain", theor.max = FALSE))
```

2.1.6 Visualize the data as a heat map using the function *level_plot* to allow for a quick comparative overview of the kinetic data.

level_plot(x, main = list(), colors = opm_opt("color.borders"), panel.headers = metadata$Strain, cex = NULL, strip.fmt = list(), striptext.fmt = list(), legend.sep = " ", space = "Lab", bias = 0.7, num.colors = 200L)

2.1.7 Extract important biological information, the *curve parameters,* from the raw kinetic curves and include the lag phase (λ), the growth rate (μ), the maximum cell respiration (A), and the area under the curve (AUC)[21] . To identify positive metabolites, use the A values of the negative control, which represents the abiotic reactivity of the dye with the medium, in addition to the blank of each microwell plate as background subtraction values. The *extract* function is used to obtain the A parameter.

opm_opt("curve.param")
param <- extract (x, as.labels = list("Strain")))

3. Identification of Reactions and Genes Associated with New Metabolites

3.1 Search KEGG (Kyoto Encyclopedia of Genes and Genomes) (http://www.genome.jp/kegg/) and MetaCyc (http://metacyc.org/) to identify Enzyme Commission numbers (ECs) for reactions using metabolites found from chemical compound arrays[23,24].

3.2 Use the identified EC numbers as a search basis in multiple available algal annotation resources such as Joint Genome Institute (JGI), Phytozome (http://www.phytozome.net), and peer-reviewed publications[23,25-27].

3.3 If a query returns no genetic evidence for a given EC number, identify the relevant associated proteins in other organisms, starting with species closest to the *C. reinhardtii*, then carry out a profile-based search using the NCBI PSI-BLAST server with default settings and use non-redundant proteins (nr) in *C. reinhardtii* (taxid:3055) to identify candidate genes associated to the reaction[12].

3.4 Manually curate PSI-BLAST hits with E values of $\leq$ 0.05 for relevance to the searched EC number through querying those BLAST hits through EMBL-EBI Pfam (http://pfam.xfam.org/search), or InterPro (http://www.ebi.ac.uk/interpro/) protein domain prediction servers. Note that the latter two scans are critical steps to ensure the identification of the correct enzymatic activity for the protein.

4. Model Refinement and Evaluation

4.1 Use the latest COBRA Toolbox v.3.0[28] in MATLAB[29,30] platform to carry out the following steps for model refinement. The COBRA Toolbox can be installed by following the steps in: https://opencobra.github.io/cobratoolbox/stable/installation.html . Alternatively, note that the

COBRA Toolbox is also implemented across other open-source programming languages, such as Python (COBRApy[31]) and is available at: https://opencobra.github.io/cobrapy/ .

4.1.1 After installing the COBRA Toolbox v.3.0, open MATLAB and execute the following command to initialize the toolbox:

```
initCobraToolbox;
```

4.1.2 Add the identified reactions with their associated genes to the metabolic model, such as *i*RC1080, using the COBRA Toolbox functions *addReaction* and *changeGeneAssociation*. Navigate to the directory that contains the *i*RC1080 model, downloaded from http://bigg.ucsd.edu/models/iRC1080 and execute the following commands to load the model, rename it, and add a new reaction and its associated gene.

```
Load('iRC1080.mat')
modelNew = iRC1080;
modelNew = addReaction(modelNew, 'D-ALA2', …
 {'d-ala[c]','atp[c]', …
 'D-aladata[c]', 'adp[c]','pi[c]', …
 'h[c]'},[-2 -1 1 1 1 1],false);
modelNEW = changeGeneAssociation(modelNew, …
 'D-ALA2','au.g14655_t1');
```

4.1.3 In some cases when the metabolite is not produced intracellularly but is taken up from the medium, add transport reactions for the new metabolites to the model. These transport reactions represent passive diffusion of a metabolite from the extracellular medium to the cytosol. In addition, add a corresponding artificial exchange reaction using the *addExchangeRxn* function to input or output the metabolite into the extracellular medium.

```
modelNew = addReaction(modelNew, 'CYCPt',…
 {'cycp[e]','cycp[c]'},[-1 1],true)'
modelNew = addExchangeRxn(modelNew, 'cycp[e]', -1000, 1000);
```

4.1.4 Test the behavior of the new resultant model, e.g., *i*BD1106, by carrying out flux balance analysis (FBA) using the function *optimizeCbModel* under light and dark conditions for the maximization of biomass as the objective function. For light growth, set the lower and upper bounds of the PRISM solar litho' light reactions to 646.07 (maximum rate). For dark growth, set the bounds of all PRISM light reactions to zero. Use the Biomass function defined previously[10] for growth under dark and light conditions. The FBA solution will output two vectors corresponding to reaction fluxes (solution.v) and reduced cost (solution.w), as well as one vector corresponding to metabolites' shadow prices (solution.y).

```
%Simulate growth under light condition:
modelNew = changeRxnBounds(modelNew,{…
 %'PRISM_solar_litho',…
 'PRISM_solar_exo',…
 'PRISM_incandescent_60W',…
 'PRISM_fluorescent_cool_215W',…
```

'PRISM_metal_halide',...
'PRISM_high_pressure_sodium',...
'PRISM_growth_room',...
'PRISM_white_LED',...
'PRISM_red_LED_array_653nm',...
'PRISM_red_LED_674nm',...
'PRISM_fluorescent_warm_18W',...
'PRISM_design_growth',...
},0,'b');
modelNew = changeObjective(modelNew, 'BIOMASS_Chlamy_mixo');
FBAsolutionNew = optimizeCbModel(modelNew, 'max');

4.1.5 Repeat step 4.1.4 for *i*RC1080 to compare FBA solutions obtained for *i*BD1106 with those obtained for *i*RC1080.

4.1.6 There is a range of COBRA methods available that can be used to compare models (e.g., flux variability analysis, gene deletion studies, robustness analyses, flux split predictions, FBA, sampling, etc.). Detailed tutorials can be found at https://opencobra.github.io/cobratoolbox/stable/tutorials/index.html. Here, an example is provided where the *i*RC1080 model is compared with its refined version, *i*BD1106, by obtaining the shadow prices (sensitivity of the biomass objective function to changes in system variable) of the metabolites accounted for in each model.

Obtain the shadow prices for the metabolites:
shadowPrices = table(modelNew.mets, ...
modelNew.metNames, FBAsolutionNew.y);

REPRESENTATIVE RESULTS:

Phenotype Microarray screening of model alga *Chlamydomonas reinhardtii*

The PM assays test the ability of the alga to utilize various sources of carbon, nitrogen, sulfur, and phosphorus in a minimal medium. In this methods description, we demonstrated how PM assays were used to identify carbon and nitrogen metabolism. Carbon and nitrogen utilization kinetics were measured with a microplate reader. Data were analyzed using PMI software. The summary kinetics of selected PM assay plates (PM01 and PM03) are shown in **Figure 1**. The "xy plots" display the respiration measurements over time plotted for the 96-well plates' assays, where the y-axis and x-axis represent the values of raw measurements and time, respectively. The data was converted to a heat-map pattern to comparatively analyze the assembly of the kinetic data.

The pipeline of refining genome-scale metabolic network using PM data (**Figure 2**) illustrates the integration of the high-throughput PM assays with experimental evidence provided by genomic searches can expand a metabolic network model.

To determine the reproducibility of the PM data obtained from PM01 - 04 and PM10 plates, a linear regression was analyzed to plot the data from two independent replicate experiments

against each another (**Figure 3**). **Figure 3** shows that the majority of the data were almost similar as they fall on the 45° line, with only a few outliers being present, and their coefficient of determination R^2 was 0.9. The consistency and reproducibility of the experiments for the alga are verified by this plot.

Identification of new metabolites

The PM assay identified 662 metabolites in seven plates; PM01-PM04 and PM06-PM08, while Gas Chromatography Time-Of-Flight (GC-TOF) had identified 77 metabolites[32] (**Figure 4**). When comparing these two sets with the 1068 metabolites accounted in the *i*RC1080, only six metabolites overlapped between the three sets, and 149 overlapped between the *i*RC1080 and the PM. This result demonstrates that the metabolic profiling platform can be a significant source of new metabolic information.

Acetic acid was the only carbon source detected in plate PM01 as a supporting carbon after subtracting the background signal. This finding is consistent with the literature[33] and shows the specificity of the PM assays. The PM assays revealed new sulfur, phosphorus, and nitrogen sources that *C. reinhardtii* can utilize for growth. The sulfur metabolites were sulfate, thiosulfate, tetrathionate, and DL-Lipoamide. The phosphorus sources were thiophosphate, dithiophosphate, D-3-phospho-glyceric acid, and cysteamine-S-phosphate. The nitrogen source metabolites were L-amino and D-amino acids, including less common amino acids; L-homoserine, L-pyroglutamic, methylamine, ethylamine, ethanolamine, and D,L-α-amino-butyric, and 108 Di-peptides and five Tri-peptides (**Table 1**). All the 128 newly identified metabolites were searched in KEGG and MetaCyc for their associated reactions, EC numbers, and pathways.

The new 128 metabolites were associated with 49 unique EC numbers. Of these, 15 ECs were linked to their genomic evidence using five sources including; Phytozome Version 10.0.2[34] JGI Version 4[35], AUGUSTUS 5.0 and 5.2[10] annotations from Manichaikul *et al.*[36] and KEGG[13]. Metabolites without genomic evidence were entered into the Universal Protein Resource website (UniProt, http://www.uniprot.org/)[37,38] where their related sequences were found in other organisms. Homologous sequences in *C. reinhardtii* were identified by running Position-Specific Iterated BLAST (PSI-BLAST, https://blast.ncbi.nlm.nih.gov/Blast.cgi) from the NCBI website considering only sequences that produced significant alignments (E-value <0.005).

Model refinement

Reactions associated with the new 128 metabolites, along with their encoded genes, were added to the *i*RC1080 model, expanding the network. The resulting model *i*BD1106, accounts for 2,444 reactions, 1,959 metabolites, and 1,106 genes (**Table 2**). The new 254 added reactions were 20 amino acid oxidation reactions, 108 di-peptide hydrolysis reactions, five tripeptide hydrolysis reactions, and 120 transport reactions, encoded by four genes (Cre02.g096350.t1.3, au.g14655_t1, e_gwW.1.243.1, Cre12.g486350.t1.3).

A total of 113 added new reactions account for the hydrolysis of di-peptides and tri-peptides. The hydrolysis of di-peptides and tri-peptides are associated with two genes, one for di-peptides (Cre02.g078650.t1.3), and one for tri-peptides (Cre16.g675350.t1.3).

Concerning sources of phosphorus, a reaction for hydrolysis of cysteamine-S-phosphate into cysteamine and phosphate was added associated with the gene JLM_162926.

The WoLF PSORT tool[39] (http://www.genscript.com/psort/wolf_psort.html) and results reported by Ghamsari *et al.*[35] were applied to obtain the specification of the cellular compartments where the new reactions take place. By analyzing protein sequences associated with the new reactions, WoLF PSORT predicted cytosol as the cellular compartment for the reactions.

A generated metabolic model may contain gaps when the biochemical information is incomplete. In such cases, *gapFind*, A COBRA command, is used. It lists root gaps and allows the identification of new gaps introduction in the new model. The metabolites that cannot be produced in a metabolic model are referred to as root gaps[40,41]. Analyzing the root gap indicated that both *i*RC1080 and *i*BD1106 models contain the same 91 gaps. This shows that adding the new metabolites and their associated reactions did not introduce any additional root gaps. It should be noted that the phenotyping method used in this protocol does not close root gaps, because the original root gap metabolites lack transport or production mechanisms, which were not addressed in the phenotyping assays. Flux balance analysis was carried out to test the metabolic behavior of *i*BD1106 under light and dark conditions; (no acetate) and (with acetate), respectively. The algorithm maximizes the biomass precursor reactions for an objective function (biomass growth). To evaluate the involvement of each metabolite to the set objective function, “shadow prices” for all metabolites were calculated. The change in the objective function concerning flux changes of the metabolite defines the shadow price of a metabolite[30,42]. The indication of whether a metabolite is in “excess” or is “limiting” the objective function can be determined by shadow price analysis, e.g., biomass production. Negative or positive shadow price values reveal metabolites that, upon addition, will decrease or increase the objective function. Zero values of shadow prices reveal metabolites that will not affect the objective function. The comparison of shadow prices between *i*BD1106 and *i*RC1080 in **Figure 5** shows that, for most metabolites, a significant change is not observed; though, differences are found in 105 and 70 cases under light and dark growth conditions, respectively. **Table 4** includes examples of such metabolites.

FIGURE AND TABLE LEGENDS:

Figure 1: Phenotypic microarray profiling of *C. reinhardtii*. Respiration XY-plots and level plots of the PM01 (Carbon sources; A, C) and PM03 (Nitrogen sources; B, D) assay plates are shown. The figure is an 8x12 array where each cell represents a well-plate and, thus, a given metabolite or growth environment. Within each cell or well representation, curves represent dye conversion by reduction (y-axis) as a function of time (x-axis). PM respiration curves from the CC-503 and blank wells are shown in each cell and are indicated by color (teal color represents blank wells and purple color represents CC-503). The level-plot represents each respiration

curve as a thin horizontal line changing color (or remaining unchanged) over time. Heatmap color changes are from light yellow (little to no respiration has taken place) to dark orange or brown (significant respiration has taken place). Metabolites utilized by *C. reinhardtii* (CC-503) and the blank plates are shown. This figure is from a previously published work by Chaiboonchoe *et al.* [12].

Figure 2: Genome-scale metabolic network refinement pipeline using PM data. After a new compound tests positive in a PM assay, its Enzyme Commission number (EC), reaction, and pathway are identified from available databases, e.g., KEGG and MetaCyc. Genomic evidence is then extracted from genomic and annotation resources when available and constitutes a link between genotype and phenotype. When direct genomic evidence is unavailable, the protein sequence is identified from the EC numbers, and genetic evidence is identified via PSI-Blast. The reconstructed metabolic network is then refined based on newly identified compounds, but only after a quality control step that entails querying the protein domains using relevant databases. This figure has been modified from previously published work by Chaiboonchoe *et al.* [12].

Figure 3: Reproducibility of PM tests. The PMI values were collected over a 168 hours period, and the maximum PMI values were plotted for two replicate studies. Each axis represents the maximum PMI values for each study (the x-axis being one replicate study and the y-axis another). Reproduced values are equidistant from each axis. Each point represents a single maximum value. The linear regression was performed by excel and the coefficient of determinant (R^2) is 0.9. This figure has been modified from previously published work by Chaiboonchoe *et al.* [12].

Figure 4: Venn diagram of metabolites. The Venn diagram enumerates metabolites identified by PM plates, the *i*RC1080 metabolic model, and Gas Chromatography Time of Flight (GC-TOF) experiments. Each circle indicates the total number of metabolites that exist in each respective method of study. At the same time, the overlapping regions represent the number of metabolites shared between those methods. The *i*RC1080 metabolic model contains a total of 1,068 unique metabolites, the GC-TOF identified a total of 77 metabolites[32], while there are a total of 662 metabolites identified using the PM plates. This figure is from previously published work by Chaiboonchoe *et al.* [12]

Figure 5: Shadow prices of metabolites in *i*RC1080 and *i*BD1106 under different conditions for biomass maximization. Each circle on the "radar plots" corresponds to a shadow price value, while each line extending from the center of a plot indicates a metabolite. **(A)** Shadow prices and metabolic behaviors of *i*RC1080 and *i*BD1106 under light growth condition; **(B),** different metabolic behaviors of *i*RC1080 and *i*BD1106 under dark growth condition. This figure is from previously published work by Chaiboonchoe *et al.* [12]

Table 1: List of identified positive substrate utilization metabolites (C, P, S, N) not present in the *i*RC1080 metabolic model.

*Reaction was not included if no gene was identified.

[1]Phytozome version 10.0.2 (http://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Creinhardtii).
[2]JGI version 4 [35].
[3]Augustus version 5[10].
[4]KEGG (http://www.genome.jp/kegg/kegg1.html).
[5]JGI version 3.1[36].
This table is from previously published work by Chaiboonchoe *et al.* [12]

Table 2: Contents of *i*RC1080 and *i*BD1106.
This table is from previously published work by Chaiboonchoe *et al.* [12]

Table 3: Summary of new reactions in *i*BD1106
This table is from previously published work by Chaiboonchoe *et al.* [12]

Table 4: Example of significant shadow prices for *i*RC1080 and *i*BD1106
This table is from previously published work by Chaiboonchoe *et al.* [12]

DISCUSSION:

Metabolic phenotyping of the green microalga, *C. reinhardtii*, was described here using high throughput PM assay plates and an unmodified PMI. The assays were utilized for a total of 190 carbon sources (PM01 and PM02), 95 nitrogen sources (PM03), 59 phosphorus sources, and 35 sulfur sources (PM04), along with peptide nitrogen sources (PM06-08). Positive respiration was observed for 148 nutrients (one positive assay for C-source utilization, four positive assays for each the S-source and P-source utilizations, and 139 positive assays for N-source utilization). The substrates or nutrients (carbon, nitrogen, phosphorus, or sulfur) component of the media should not be added to the defined medium when applied to the relevant PM microplates that test for each of those sources.

The method shown here is effective for characterizing metabolic microalgae phenotypes that can be used to extend existing metabolic network models or direct the reconstruction of new models. Further, as nutritional requirements of most microalgae are not known, this platform can be used to define these rapidly. Nelson *et al.*[43] had successfully applied these methods to identify new compounds that support the growth of the microalgae *Chloroidium* sp. UTEX 3007 and used the obtained information to define the species entry metabolites, which, unlike Chlamydomonas, includes 40 different carbon sources.

One major limitation of the PM for profiling microalgae is that the PMI has no illumination in the incubation chamber, and the microalgae need to be able to carry out heterotrophic metabolism. The absence of light could affect the interpretation of models that incorporate light to calculate metabolic fluxes. Gene pairs with coordinating functions have co-evolved to constitute metabolic network hubs, and distinction between photosynthetic and non-photosynthetic network hubs can be made[44]. In general, photosynthetic network hubs (i.e., highly-connected nodes in the model) would be left out of heterotrophic models. For practical purposes, modeling heterotrophism in mixotrophic species should omit reactions known to be

driven by light and account for the energy balance differences between conditions. Thus, modeling light-dependent and light-independent metabolism is standard practice in Chlamydomonas metabolic modeling[6,45].

Some green microalgae, like Trebouxiophytes, are known to assimilate a variety of carbon molecules for growth, and this is thought to have arisen from their long evolutionary history as members of lichens[46]. While chlorophytes like *Chlamydomonas* can use acetate for growth, the brown marine microalga *Tisochrysis lutea,* known for its potential to commercially produce very-long-chain polyunsaturated fatty acids (VLC-PUFAs), cannot use acetate but can use glycerol for growth[47]. Biomass concentration of more than 100 g l^{-1} dry cell weight has been achieved with *Chlorella* with optimized addition of organic carbon sources in a fed-batch mode[48]. Further, the addition of sugar to *Chlorella vulgaris* can elevate its sequestration of CO_2, thus providing an additive benefit during photosynthetic growth[49]. Most heterotrophic microalgae can also grow mixotrophically, but the Chlorophyte *Chromochloris zofingiensis* has been shown to shut off photosynthesis upon the addition of sugar[50].

Diatoms, belonging to the division Bacillariophyta, are a major group of phytoplankton. Although most of the diatoms can only grow photoautotrophically, some of them can be cultivated mixotrophically or heterotrophically[51]. For example, glycerol was found to support growth in the light in the absence of CO_2 in some diatoms, including the model species *Phaeodactylum tricornutum*[52]. Also, some benthic diatoms like *Nitzschia linearis* can grow on carbohydrates in the dark[53]. It is likely to extend the PM assays to diatoms and other algal groups by supplementing suitable organic carbon sources to enable the cells to grow heterotrophically, and a mixotrophy strategy can also be potentially used for the obligate autotrophic microalgae providing a minimally required light supply.

To assess the reproducibility of the data, it is highly recommended to carry out duplicate assays for all plates. An assay may be considered positive only if, after subtraction from the negative control and the respective blank wells, the absorbance (PMI value) is positive. This description, in the presence of the tested compound, is a reflection of the dye's abiotic reaction with the medium.

ACKNOWLEDGMENTS:

Major support for this work was provided by New York University Abu Dhabi Institute grant (73 71210 CGSB9) and NYU Abu Dhabi Faculty Research Funds (AD060). W.F. was additionally supported by the Hundred Talents Program of Zhejiang University. We thank Ashish Jaiswal for help in recording the video. We thank Hong Cai for generating the metabolic phenotype data.

DISCLOSURES:

The authors have nothing to disclose.

REFERENCES:

1 Oberhardt, M. A., Palsson, B. Ø., Papin, J. A. J. M. s. b. Applications of genome-scale metabolic reconstructions. *Molecular Systems Biology.* **5** (1), 320 (2009).

2 Schmidt, B. J., Lin-Schmidt, X., Chamberlin, A., Salehi-Ashtiani, K., Papin, J. A. Metabolic systems analysis to advance algal biotechnology. *Biotechnology Journal.* **5** (7), 660-670 (2010).
3 Koskimaki, J. E., Blazier, A. S., Clarens, A. F., Papin, J. A. J. I. B. Computational models of algae metabolism for industrial applications. *Industrial Biotechnology.* **9** (4), 185-195 (2013).
4 Koussa, J., Chaiboonchoe, A., Salehi-Ashtiani, K. J. B. r. i. Computational approaches for microalgal biofuel optimization: a review. *BioMed Research.* **2014**, 649453 (2014).
5 Nelson, D. R. *et al.* Large-scale genome sequencing reveals the driving forces of viruses in microalgal evolution. *Cell Host & Microbe.* **29** (2), 250-266. e258 (2021).
6 Shene, C., Asenjo, J. A., Chisti, Y. Metabolic modelling and simulation of the light and dark metabolism of Chlamydomonas reinhardtii. *The Plant Journal.* **96** (5), 1076-1088 (2018).
7 Tibocha-Bonilla, J. D., Zuñiga, C., Godoy-Silva, R. D., Zengler, K. Advances in metabolic modeling of oleaginous microalgae. *Biotechnology for Biofuels.* **11** (1), 241 (2018).
8 Merchant, S. S. *et al.* The Chlamydomonas genome reveals the evolution of key animal and plant functions. *Science.* **318** (5848), 245-250 (2007).
9 May, P., Christian, J.-O., Kempa, S., Walther, D. J. B. G. ChlamyCyc: an integrative systems biology database and web-portal for Chlamydomonas reinhardtii. *BMC Genomics.* **10** (1), 209 (2009).
10 Chang, R. L. *et al.* Metabolic network reconstruction of Chlamydomonas offers insight into light-driven algal metabolism. *Molecular Systems Biology.* **7** (1) (2011).
11 de Oliveira Dal'Molin, C. G., Quek, L.-E., Palfreyman, R. W., Nielsen, L. K. AlgaGEM – a genome-scale metabolic reconstruction of algae based on the Chlamydomonas reinhardtii genome. *BMC Genomics.* **12** (S5) (2011).
12 Chaiboonchoe, A. *et al.* Microalgal metabolic network model refinement through high-throughput functional metabolic profiling. *Frontiers in Bioengineering and Biotechnology.* **2**, 68 (2014).
13 Kanehisa, M. *et al.* Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Research.* **42** (D1), D199-D205 (2014).
14 Zuñiga, C. *et al.* Genome-scale metabolic model for the green alga Chlorella vulgaris UTEX 395 accurately predicts phenotypes under autotrophic, heterotrophic, and mixotrophic growth conditions. *Plant Physiology.* **172** (1), 589-602 (2016).
15 Bochner, B. R. New technologies to assess genotype-phenotype relationships. *Nature Reviews Genetics.* **4** (4), 309-314 (2003).
16 Bochner, B. R. Global phenotypic characterization of bacteria. *FEMS Microbiology Reviews.* **33** (1), 191-205 (2009).
17 Bochner, B. R., Gadzinski, P., Panomitros, E. Phenotype microarrays for high-throughput phenotypic testing and assay of gene function. *Genome Research.* **11** (7), 1246-1255 (2001).
18 Bartell, J. A., Yen, P., Varga, J. J., Goldberg, J. B., Papin, J. A. Comparative metabolic systems analysis of pathogenic Burkholderia. *Journal of Bacteriology.* **196** (2), 210-226 (2014).
19 Gorman, D. S., Levine, R. J. P. o. t. N. A. o. S. Cytochrome f and plastocyanin: their

sequence in the photosynthetic electron transport chain of Chlamydomonas reinhardi. *PNAS.* **54** (6), 1665-1669 (1965).

20 Smith, A. C., Hussey, M. A. Gram stain protocols. *American Society for Microbiology.* 1-9 (2005).

21 Vaas, L. A. I. *et al.* opm: an R package for analysing OmniLog phenotype microarray data. *Bioinformatics.* **29** (14), 1823-1824 (2013).

22 Vaas, L. A. I., Sikorski, J., Michael, V., Göker, M., Klenk, H.-P. Visualization and Curve-Parameter Estimation Strategies for Efficient Exploration of Phenotype Microarray Kinetics. *PLoS ONE.* **7** (4), e34846 (2012).

23 Caspi, R. *et al.* The MetaCyc database of metabolic pathways and enzymes-a 2019 update. *Nucleic Acids Research.* **48** (D1), D445-D453 (2020).

24 Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M., Tanabe, M. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Research* (2020).

25 Lopez, D., Casero, D., Cokus, S. J., Merchant, S. S., Pellegrini, M. Algal Functional Annotation Tool: a web-based analysis suite to functionally interpret large gene lists using integrated annotation and expression data. *BMC Bioinformatics.* **12** (1), 282 (2011).

26 Caspi, R. et al. The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Research*. **46** (D1), 633-639 (2018).

27 Sahoo, S. *et al.* dEMBF v2. 0: An Updated Database of Enzymes for Microalgal Biofuel Feedstock. *Plant and Cell Physiology.* **61** (5), 1019-1024 (2020).

28 Heirendt, L. *et al.* Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v. 3.0. *Nature Protocols.* **14** (3), 639-702 (2019).

29 Heirendt, L. *et al.* Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v. 3.0. *Nature Protocols.* **1** (2019).

30 Orth, J. D., Thiele, I., Palsson, B. Ø. What is flux balance analysis? *Nature Biotechnology.* **28** (3), 245 (2010).

31 Ebrahim, A., Lerman, J. A., Palsson, B. O., Hyduke, D. R. COBRApy: constraints-based reconstruction and analysis for python. *BMC Systems Biology.* **7** (1), 74 (2013).

32 Bölling, C., Fiehn, O. Metabolite profiling of Chlamydomonas reinhardtii under nutrient deprivation. *Plant Physiology.* **139** (4), 1995-2005 (2005).

33 Harris, E. H. *The Chlamydomonas sourcebook: introduction to Chlamydomonas and its laboratory use*. Vol. 1 (Academic Press, 2009).

34 Goodstein, D. M. *et al.* Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research.* **40** (D1), D1178-D1186 (2012).

35 Ghamsari, L. *et al.* Genome-wide functional annotation and structural verification of metabolic ORFeome of Chlamydomonas reinhardtii. *BMC Genomics.* **12** (1), S4 (2011).

36 Manichaikul, A. *et al.* Metabolic network analysis integrated with transcript verification for sequenced genomes. *Nature Methods.* **6** (8), 589-592 (2009).

37 Apweiler, R. *et al.* UniProt: the Universal Protein knowledgebase. *Nucleic Acids Research.* **32** (Database issue), D115-119 (2004).

38 Consortium, T. U. Activities at the universal protein resource (UniProt). *Nucleic Acids Research.* **42** (11), 7486-7486 (2014).

39 Horton, P. *et al.* WoLF PSORT: protein localization predictor. *Nucleic Acids Research.* **35**

(suppl_2), W585-W587 (2007).
40 Becker, S. A. *et al.* Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nature Protocols.* **2** (3), 727-738 (2007).
41 Schellenberger, J. *et al.* Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2. 0. *Nature Protocols.* **6** (9), 1290 (2011).
42 Varma, A., Boesch, B. W., Palsson, B. O. Stoichiometric interpretation of Escherichia coli glucose catabolism under various oxygenation rates. *Applied and Environmental Microbiology.* **59** (8), 2465-2473 (1993).
43 Nelson, D. R. *et al.* The genome and phenome of the green alga Chloroidium sp. UTEX 3007 reveal adaptive traits for desert acclimatization. *eLife.* 10.7554/eLife.25783 (2017).
44 Chaiboonchoe, A. *et al.* Systems level analysis of the Chlamydomonas reinhardtii metabolic network reveals variability in evolutionary co-conservation. *Molecular BioSystems.* **12** (8), 2394-2407 (2016).
45 Chang, R. L. *et al.* Metabolic network reconstruction of Chlamydomonas offers insight into light-driven algal metabolism. *Molecular Systems Biology.* **7** (1), 518 (2011).
46 Rajendran, A., Hu, B. Mycoalgae biofilm: development of a novel platform technology using algae and fungal cultures. *Biotechnology for Biofuels.* **9** (1), 112 (2016).
47 Hu, H. *et al.* Effect of cultivation mode on the production of docosahexaenoic acid by Tisochrysis lutea. *AMB Express.* **8** (1), 50 (2018).
48 Bumbak, F., Cook, S., Zachleder, V., Hauser, S., Kovar, K. Best practices in heterotrophic high-cell-density microalgal processes: achievements, potential and possible limitations. *Applied Microbiology and Biotechnology.* **91** (1), 31 (2011).
49 Fu, W. *et al.* Sugar-stimulated CO2 sequestration by the green microalga Chlorella vulgaris. *Science of the Total Environment.* **654**, 275-283 (2019).
50 Roth, M. S. *et al.* Regulation of oxygenic photosynthesis during trophic transitions in the green alga Chromochloris zofingiensis. *The Plant Cell* (2019).
51 Villanova, V. *et al.* Investigating mixotrophic metabolism in the model diatom Phaeodactylum tricornutum. *Philosophical Transactions of the Royal Society B: Biological Sciences.* **372** (1728), 20160404 (2017).
52 Cerón-García, M. *et al.* Mixotrophic growth of Phaeodactylum tricornutum on fructose and glycerol in fed-batch and semi-continuous modes. *Bioresource Technology.* **147**, 569-576 (2013).
53 Tuchman, N. C., Schollett, M. A., Rier, S. T., Geddes, P. in *Advances in Algal Biology: A Commemoration of the Work of Rex Lowe* 167-177 (Springer, 2006).

Figure 1

Click here to access/download;Figure;Figure1.jpg

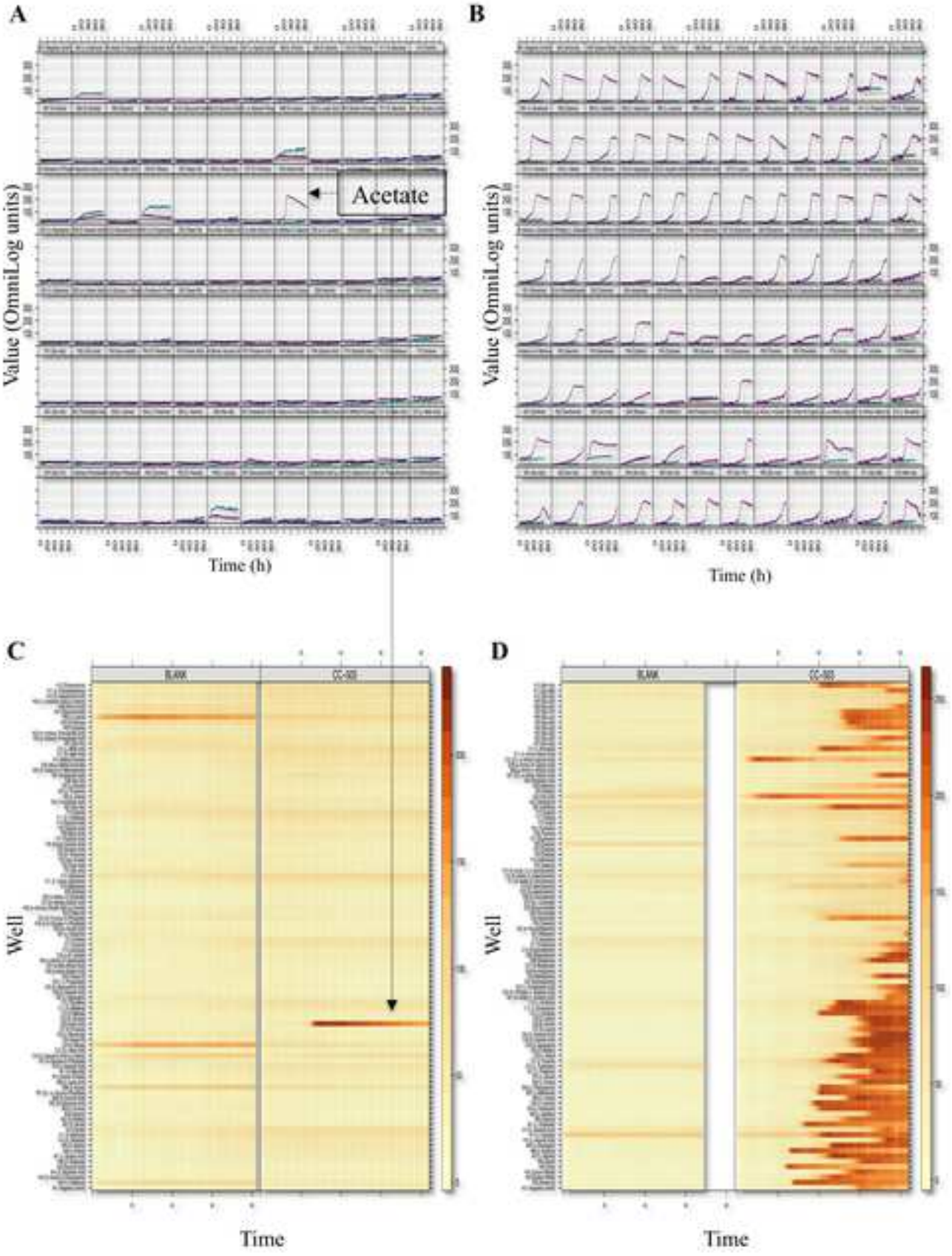


Figure 2

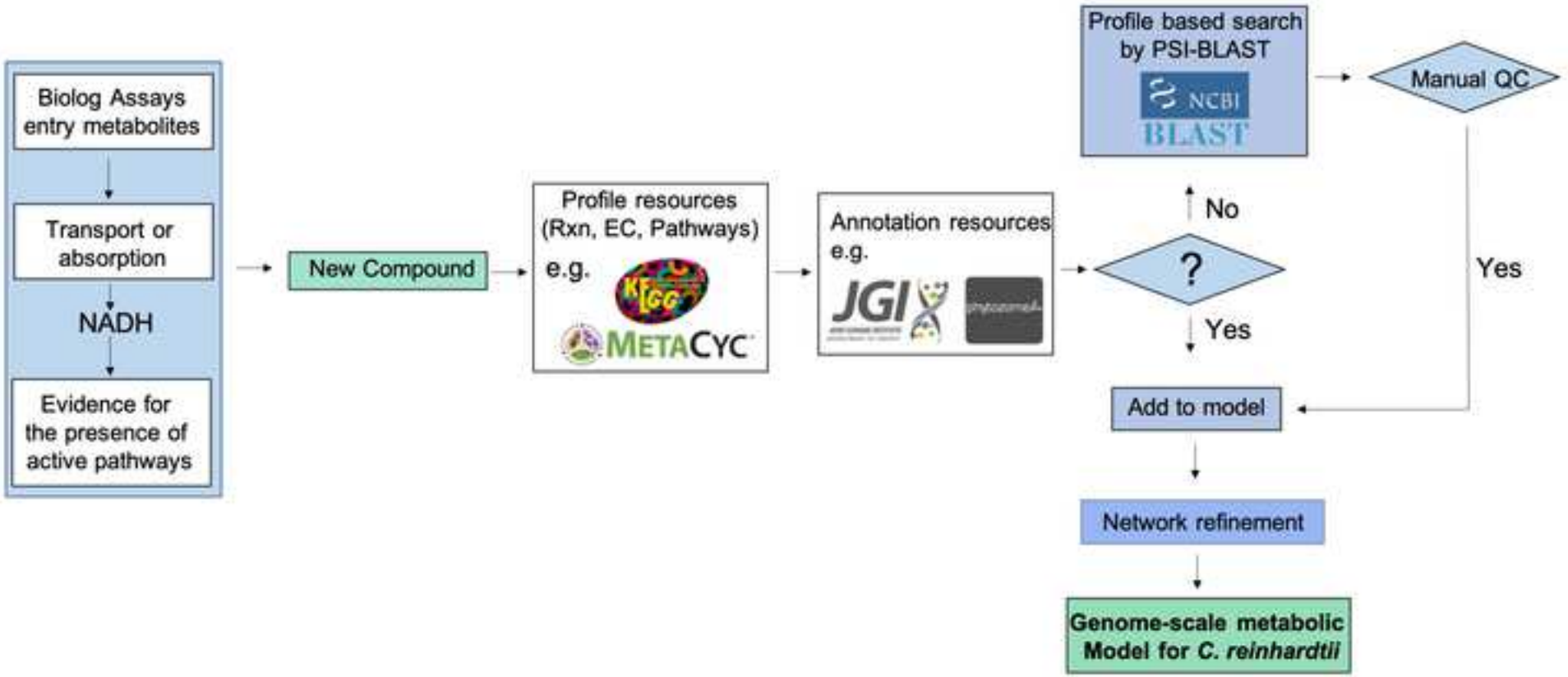


Figure 3

Click here to access/download;Figure;Figure3.jpg

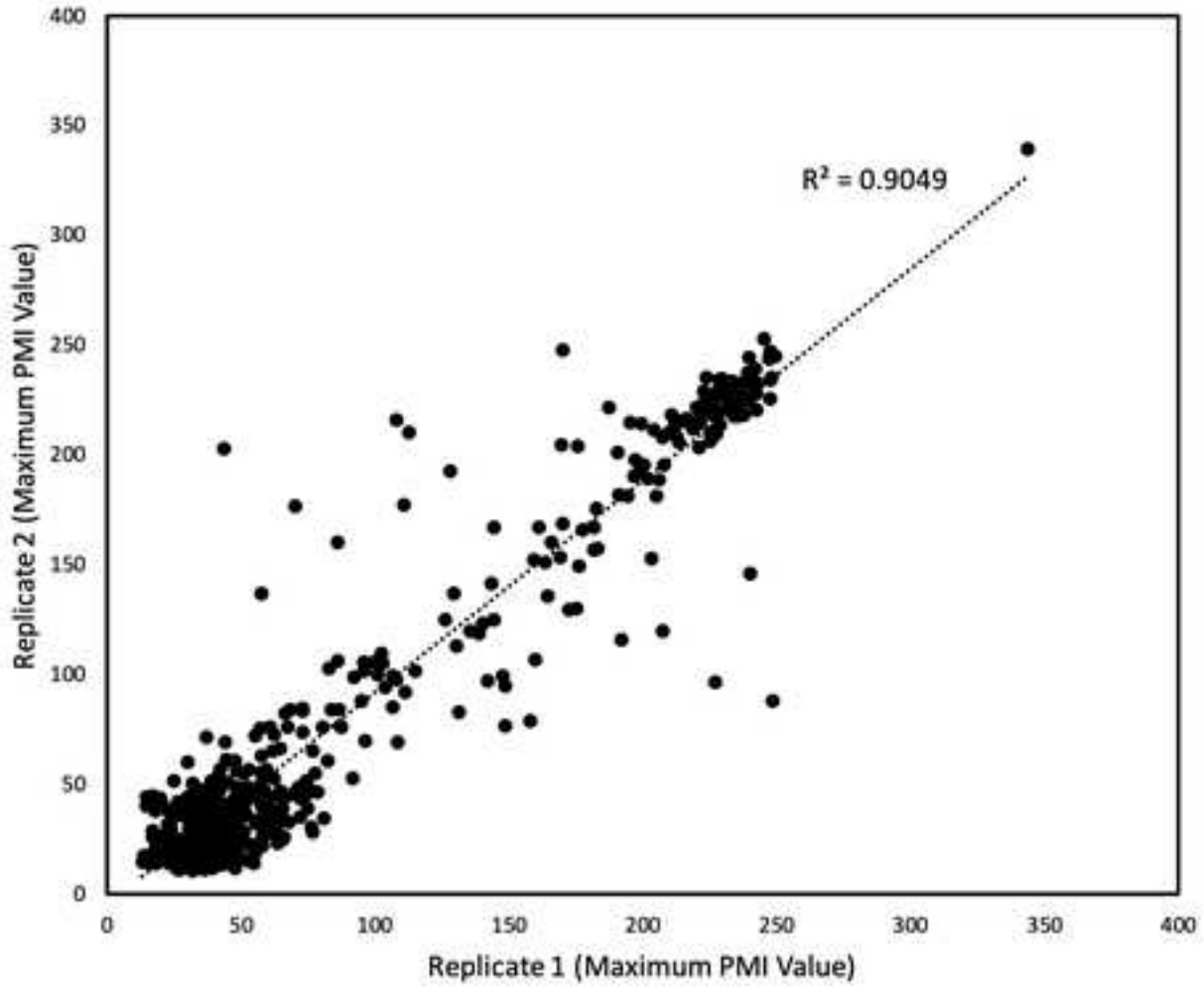


Figure 4

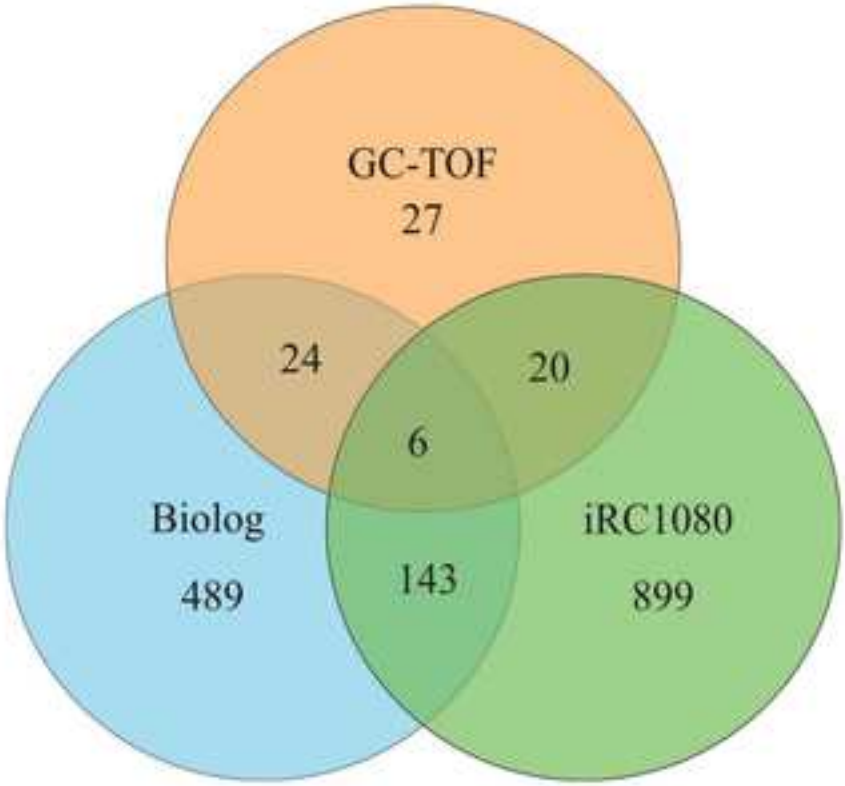
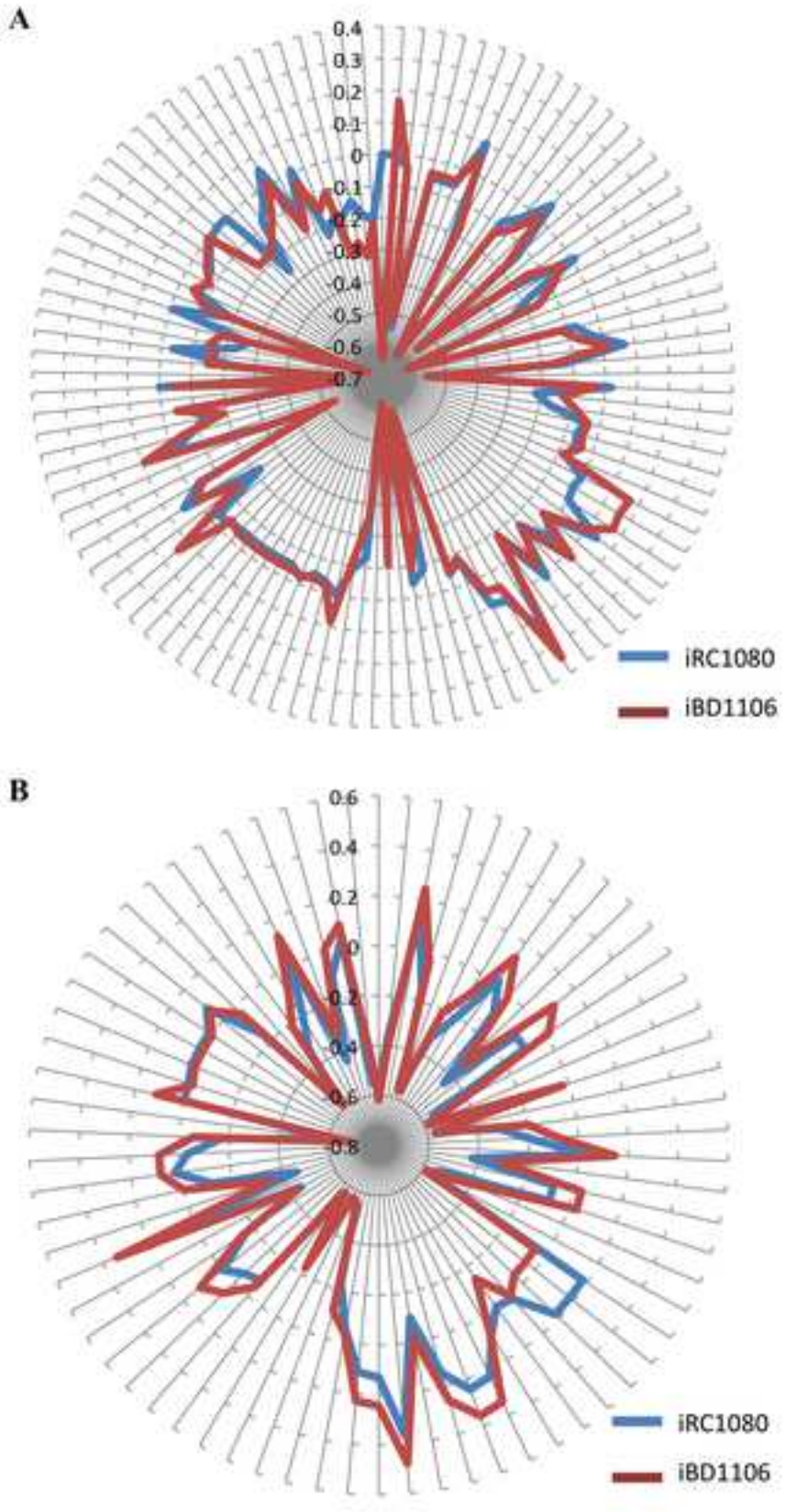


Figure 5



Biolog Chemical	EC*	Gene Annotation
Cysteamine-S-Phosphate	3.1.3.1	JLM_162926 [1,2,3,4]
Tetrathionate	1.8.2.2	
	1.8.5.2	
D-Alanine	1.4.1.1	
	1.5.1.22	
	2.1.2.7	
	1.4.3.3	Cre02.g096350.t1.3[5]
	2.3.2.10	
	2.3.2.14	
	2.3.2.16	
	2.3.2.17	
	2.3.2.18	
	2.6.1.21	
	3.4.13.22	
	3.4.16.4	Chlre2_kg.scaffold_14000039[1,2,3]
	3.4.17.8	
	3.4.17.13	
	3.4.17.14	
	4.5.1.2	
	6.1.1.13	
	6.1.2.1	
	6.3.2.4	au.g14655_t1[1,2,3]
	6.3.2.10	
	6.3.2.16	
	6.3.2.35	
D-Asparagine	1.4.5.1	
	1.4.3.3	Cre02.g096350.t1.3[5]
	3.1.1.96	
	2.3.1.36	
	1.4.99.1	
	3.5.1.77	e_gwW.1.243.1 [1,2]
	3.5.1.81	
	5.1.1.10	
D-Aspartic Acid	6.3.1.12	
	1.4.3.3	Cre02.g096350.t1.3[5]
D-Glutamic Acid	1.4.3.7	
	1.4.3.3	

D-Lysine	5.4.3.4	
	1.4.3.3	Cre02.g096350.t1.3[5]
	6.3.2.37	
D-Serine	2.7.11.8	
	2.7.11.17	Cre12.g486350.t1.3[1,2,3,4]
	3.4.21.78	
	3.4.21.104	
	4.3.1.18	g6244.t1[4]
	6.3.2.35	
	6.3.3.5	
	1.4.3.3	Cre02.g096350.t1.3[5]
D-Valine	1.21.3.1	
	6.3.2.26	
	1.4.3.3	Cre02.g096350.t1.3[5]
L-Pyroglutamic Acid		
Thiophosphate		
Dithiophosphate		
Ethylamine	6.3.1.6	
D,L-a-Amino-Butyric Acid	2.1.1.49	
	1.4.3.3	Cre02.g096350.t1.3[5]
Di-peptide	3.4.13.18	Cre02.g078650.t1.3[1]
Tri-peptide	3.4.11.4	Cre16.g675350.t1.3[1]

PSI-BLAST
insignificant E-value
insignificant E-value
XP_001700222.1
failed manual QC
insignificant E-value
insignificant E-value
insignificant E-value
insignificant E-value
insignificant E-value
insignificant E-value
failed manual QC
XP_001698572.1, XP_001693532.1, XP_001701890.1, XP_001700930.1
failed manual QC
insignificant E-value
insignificant E-value
insignificant E-value
failed manual QC
failed manual QC
failed manual QC
insignificant E-value
insignificant E-value
insignificant E-value
insignificant E-value
insignificant E-value
XP_001692123.1
insignificant E-value
failed manual QC
insignificant E-value
insignificant E-value
insignificant E-value

insignificant E-value
failed manual QC
insignificant E-value
failed manual QC
failed manual QC
failed manual QC
insignificant E-value
insignificant E-value
failed manual QC
failed manual QC
insignificant E-value

Table 2

Model	Reactions	Metabolites	Genes
iRC1080	2,191	1,706	1,086
iBD1106	2,445	1,959	1,106

Category or Class of reactions	Number of reactions
Amino Acids	20
Dipeptides	108
Tripeptides	5
Transport reaction	120

Table 4 Click here to access/download;Table;Table 4.xlsx

Growth condition	Metabolite	Name	iRC1080	iBD1106
Light	4r5au	4-(1-D-Ribitylamino)-5-aminouracil	0	0.168
	5aprbu	5-Amino-6-(5'-phosphoribitylamino)uracil	-0.009	0.158
	pa1819Z18111Z	1-(9Z)-octadecenoyl,2-(11Z)-octadecenoyl-sn-glycerol3-phosphate	-0.009	-0.65
Dark	4abut	4-aminobutanoate	0.18	-0.05

Click here to access/download

Table of Materials

Materials.xlsx

Dear Dr. Nguyen,

We thank the editor(s) and reviewers for their thoughtful and constructive comments and recommendations; we are pleased that our manuscript was found to be of interest. We have addressed all the reviewers' concerns by modifying the text and the video, and providing clarifications. We believe our manuscript has been improved greatly by making these changes based on the reviewers' comments. All the changes in the manuscript have been tracked. Below we provide a detailed point-by-point response to address the reviewers' comments.

Editorial and production comments:
Changes to be made by the Author(s) regarding the written manuscript:
1. Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammar issues.

Response: *We have proofread the manuscript and corrected any spelling or grammar issues.*

2. Please revise the following lines to avoid previously published work: 32-38, 74-78, 138-142, 154-157, 194-197, 198-205, 252-271, 335-348, 380-384

Response: *We have revised the text in the specified lines.*

3. Please obtain explicit copyright permission to reuse any figures from a previous publication. Explicit permission can be expressed in the form of a letter from the editor or a link to the editorial policy that allows re-prints. Please upload this information as a .doc or .docx file to your Editorial Manager account. The Figure must be cited appropriately in the Figure Legend, i.e. "This figure has been modified from [citation]."

Response: *Frontiers articles are published under the Creative Commons Attribution License and can be reproduced contingent on crediting the source. The link to the Frontier Copyright Policy is: https://zendesk.frontiersin.org/hc/en-us/articles/201904552 The figure legends have been revised to provide the citations of the sources accordingly.*

4. If the figures are not reused from a previous publication, please provide new figure legends as they are from previously published work: https://www.frontiersin.org/articles/10.3389/fbioe.2014.00068/full

Response: *The figures are reused from a previous publication and have been cited in the figures legends in the revised version.*

5. JoVE cannot publish manuscripts containing commercial language. This includes trademark symbols (™), registered symbols (®), and company names

before an instrument or reagent. Please remove all commercial language from your manuscript and use generic terms instead. All commercial products should be sufficiently referenced in the Table of Materials.

Response: *We have removed all commercial language from the manuscript and used generic terms in the revised version. All commercial products are referenced in the Table of Materials.*

6. Please add more details to your protocol steps. Please ensure you answer the "how" question, i.e., how is the step performed? Alternatively, add references to published material specifying how to perform the protocol action.

Response: *We have added more steps to our protocol in the revised version. We have added a reference to gram staining protocol step 1.8 in the revised manuscript.*

7. 1.2: What are the growth conditions?

Response: *We have added the growth conditions to step 1.2. The growth conditions are to grow the cells in fresh Tris-Acitate-Phosphate (TAP) media with final concentrations of 400μg/mL timentin, 50μg/mL ampicillin, and 100μg/mL kanamycin to mid-log phase.*

8. 1.7: Please expand on the standard assay plates. What are they?

Response: *We have explained what are the standard assay plates in the revised manuscript. The standard assay plates are carbon sources, nitrogen sources, phosphorus and sulfur sources plates, and peptide nitrogen sources.*

9. 1.9: How is the gram staining done?

Response: *We have added a reference to the gram staining protocol step 1.8 in the revised manuscript.*

10. Please include all user input commands (File | Save | etc. or run xxx -t etc) so that the steps 2-4 are explicitly detailed. We need these details so that others can faithfully replicate the protocol.

Response: *We have added all the user input commands to steps 2-4 in the revised manuscript..*

Changes to be made by the Author(s) regarding the video:
1. Please revise the narration to be more homogenous with the written manuscript. Ideally, the narration is a word for word reading of the written protocol.

Response: *The narration was revised as requested, and we added more detailed descriptions for the protocol steps 2 and 4 as described in the manuscript, with screen recording representing the different steps of the computational data analysis.*

2. Please reduce the frame rate to 30 frames/second.

Response: *All video parts were retaken at 30 frames/second.*

3. Please stabilize the video and avoid the fast zooms.

Response: *The videos were taken using a fixed tripod, and zoomed shots were taken separately to avoid zooming while filming.*

4. Add some transitions to the footage, it will smooth out the video and make it easier to follow

Response: *We added the titles for each part of the protocol as described in the manuscript, in addition to the previously present parts (results and conclusion).*

5. 2:17: Please use the Greek symbol mu for the microgram abbreviation. Please capitalize the L in the mL abbreviation.

Response: *The Greek symbol mu was used as requested for the antibiotics' concentrations.*

6. 2:39: Please avoid commercialism in the video. Remove the Biolog feature. Use generic terms whenever possible.

Response: *We removed the Biolog feature in the video, and made sure the logo doesn't show on the instrument.*

7. 3:03: Please do not feature the Omnilog system. Use generic terms whenever possible.

Response: *We removed the Omnilog system and replaced it with Phenotype Microarray assay system (when possible), except for one section in the data analysis part where the software used includes the Omnilog mention in the R-package software.*

8. Add a white background behind the graphics and charts

Response: *White background was added behind graphics and charts as requested.*

Please upload a revised high-resolution video here:
https://www.dropbox.com/request/HJGfhygJyxuTIMInMc9T?oref=e

Response: *The revised video is uploaded.*

Reviewers' comments:
Reviewer #1:
Manuscript Summary:
The authors use commercially available Biolog plates to metabolically phenotype a model microalgae (C. reinhardtii). They use the resulting substrate utilization patterns as evidence for the presence/absence of metabolic pathways via a computational pipeline. After their verification, metabolic pathway information was used to update an existing metabolic model. This new model was then benchmarked against the previous model by addition of data from another method (GC-TOF) and via a final model test by running shadow cost analysis for light and dark reactions for both models.

Major Concerns:
The authors developed a satisfactory pipeline for addressing gaps within metabolic models and I have few minor comments that need to be addressed before acceptance of the manuscript.
Bacterial growth. The authors do not control for bacterial growth in cultures of C. reinhardtii during the primary cultivation step. Three types of antibiotics are added after microalgae have been grown to a high enough density for subsequent Biolog incubation. How have the authors ensured that this antibiotic treatment truly inhibited all bacterial growth (as claimed in line 104-105) and that their signal is not an artefact from bacterial metabolism? I am not fundamentally questioning the validity of their data but would like clarification on this and also like the authors to suggest the addition of antibiotics at the initial incubation step (line 96).

Response: *We have added descriptions of antibiotic addition at the incubation step as recommended by the reviewer. We now also state in the protocol, step 1.8, that gram staining was performed before and after the assays to monitor bacterial contamination.*

Light conditions. The authors acknowledge (and state repeatedly) that the OmniLog system does not allow for the illumination of microalgae samples and therefore all metabolism occurs under prolonged darkness (up to 7 days). This shortcoming is discussed by the authors but only by stating what other strains also show heterotrophic growth. What I a missing in this discussion is a section on how the absence could affect their model interpretation.

Response: *We have expanded the discussion on how the absence of light could affect the model interpretation in the revised manuscript (lines 422-432).*

Minor Concerns:
Line 96: Please provide additional details on how C. reinhardtii was grown (i.e. light conditions, temperature, and duration to 'mid-log' phase).

Response: *We have added the growth conditions in the protocol step 1.2 in the revised manuscript.*

Line 111: Please provide a volume for these cell numbers

Response: *We have provided the volume for the cell number in the revised manuscript; the volume is in mL.*

Line 120: Why was bacterial contamination checked here and not before growth? Further, why was C. reinhardtii not grown with antibiotics added (step 1.5 in the protocol describes the addition of antibiotics after the cells were already grown)

Response: *Cultures were determined to be axenic before inoculation and grown with antibiotics. We have added descriptions of antibiotic addition at the incubation step. We now also state in the protocol at step 1.8 to perform gram staining before and after the assays to monitor bacterial contamination.*

Line 203: 'Figure 3 shows that the majority of the data were nearly identical as they fall on the 45-degree line'. Please perform a linear regression and display the resulting R2 value onto the figure and state it in the text.

Response: *We have performed a linear regression and displayed the resulting R^2 value of 0.9 onto the figure and stated it in the text in the revised manuscript (lines 691-692).*

Line 256-257: The author state that the addition of new metabolites did not introduce additional root gaps in previously published (and now augmented) metabolic models. Why did your phenotyping method not close some of these root gaps?

Response: *In this revised version, we have stated that it should be noted that the phenotyping method used in this protocol does not close root gaps, because the original root gap metabolites lack transport or production mechanisms, which were not addressed in the phenotyping assays.*

Line 276: Respiration is not growth. Please delete growth

Response: *We thank the reviewer for pointing this out; we have deleted "**growth"** in the revised manuscript (line 663).*

Line 301: Please indicate which test values (i.e. pathways/compounds) deviate away from the supposed linear relationship. Please indicate how this deviation was measured (standard deviation/error?).

Response: *Values of compounds, a linear regression was performed and the obtaoned coefficient of determination (R^2) of 0.9 is stated in the text of the revised manuscript (lines 691-692) and indicated in the revised figure 3.*

Table 1: Please change the layout/size of this table so the PSI-blast values/ annotations are visible together with the EC numbers on the same page.
Table of materials: Same comment as above, ensure readability by placing table contents onto same page

Response: *We have changed the layout of the tables in the revised version.*

Reviewer #2:
Manuscript Summary:
This manuscript aims to detail a protocol for using Biolog data for improving genome-scale models of metabolism, particularly for the organism Chlamydamonas reinhardtii. The paper is generally fine until section 3 dealing with identification and genes associated with new metabolites. From that point on, it is confusing and not detailed enough.
Response:

Major Concerns:
Statement 3.1 doesn't make sense. Compounds don't have EC numbers, enzymes do. The biolog data does not identify enzymes.

Response: *We have corrected this error and have edited this sentence in the revised manuscript. We now state that we search KEGG (http://www.genome.jp/kegg/) and MetaCyc (http://metacyc.org/) to identify the Enzyme Commission.*

In section 3.3, the authors need to detail what should happen if no candidate genes are found for the EC. Also, they need to expand on the term "other organisms" does it refer to other algae, specific family of bacteria, or all sequenced genomes. The statement is vague.

Response: *The reviewer is referring to protocol step 3.3. We have clarified this to describe a search for homologs in other organisms, starting with species closest to C. reinhardtii, to identify the relevant protein for the query reaction.*

If no corresponding protein can be identified in C. reinhardtii, the reaction can be added without an associated enzyme.

For section 3.4, E value of 0.05 is large. The number of possible genes that could be identified could be large. How should they be further pared?

Use of Pfam and Interpro should be detailed.

Response*: We agree with the reviewer that, in general, an E value of 0.05 is too relax; however, in our experience, we did not find the problem of having too many hits. Further, we like to draw attention that this step is followed by additional QC steps to evaluate the obtained candidate proteins for having the correct enzymatic activity associated with them. Namely, the obtained hit will be evaluated by Interpro and PFam scans. Last, the Pfam and Interpro tools described here are web-based tools, and their use does not require any informatics skills or detailed explanation.*

In section 4, there are multiple versions of COBRA toolbox available. It should be noted; particularly since due to the pay nature of matlab, many researchers are switching to cobrapy.

Response: *We have mentioned the availability of a version of COBRA toolbox that runs in python (COBRApy) in the revised version and does not require Matlab.*

Section 4.1.2, there is a hierarchy to transport of metabolites. The way the sentence is written it seems as though one can directly transport metabolites from external environment to the mitochondria and other compartments surrounded by cytosol. This is not proper modeling.

Response: *We have added more details and command line in step 4.1.3 in the revised version.*

Line 178 and 179, states "two columns correspond to reactions; flux and reduced cost, and one column corresponds to metabolites' shadow prices." It would be lot less confusing to say 3 columns. This only serves to confuse a novice user.

Response: *We have added more details and command line in step 4.1.4 in the revised manuscript.*

4.1.5 is too simplified. A much more detailed comparison of the models can be conducted beyond just comparing shadow prices.

Response: *We have added more details and command line to step 4.1.6 in the revised manuscript.*

Why wait until representative results to refer to figures? The figures would drastically help explain concepts earlier in the manuscript.

Response: *We thank the reviewer for his/her suggestion. We have arranged the manuscript's material based on the journal's template; the revised accompanying video does integrate results and analyses and should be helpful in this regard.*

Also it is very confusing to say in this study we ...., when what has been described is a protocol.

Response: *We have modified it to "in this methods description".*

The authors need to explain why plates PM01 and PM03 where chosen.

Response: *We used PM01 and PM03 plates only as examples of our results. Any phenotype microarray plates could be used for the analysis.*

On line 198, what pipeline are the authors referring to?

Response: *The pipeline of refining genome-scale metabolic network using PM data. Line 465 in the revised manuscript.*

They found 149 metabolites that overlap between biology and old model. However, later they mention 128, where did this later number come from?

Response: *149 metabolites were overlapped between the 662 PM metabolites and the 1068 iRC1080 metabolites, while only 128 metabolites were identified from literature.*

In conclusion, the paper as written is confusing for both experimentalist and theoretical scientists. There is not enough detail on protocols after section 3.

Response: *We have added more details, steps, and command lines in steps 3 and 4 in the revised manuscript.*

Minor Concerns:
1. In the abstract: systems-level is a complex word and needs a hyphen.

Response: *We have added a hyphen to* ***"systems-level"*** *word in the revised version.*

2. line 50, don't need to add "or pathways".

Response: *We have deleted* ***"or pathway"*** *in the revised manuscript.*

3. line 53, This sentence seems dated. It references a paper from 6 years ago. There have been more algal models developed since then. Perhaps the authors should enumerate the number of current models available and contrast it to the

number of algal genomes that have been sequences. This vague sentence does a disservice to the paper.

Response: *We have added a reference to the number of sequenced algal species and have stated in the revised version that "Although approximately 160 microalgal species have been sequenced5, there are, to our knowledge, only 44 algal metabolic models available."*

4. line 62, run on sentence. Change to (entailing ~~-~~250 sources). It has 1,706

Response: *We have edited the sentence in the revised manuscript.*

5. The entire section 3 is missing relevant references.

Response: *We have added the references to section 3 in the revised manuscript.*

6. line 154, genetic instead of genic

Response: *We have corrected the word to genetic in the revised manuscript.*

jove JOURNAL OF VISUALIZED EXPERIMENTS
1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT

Title of Article:	High-throughput metabolic profiling for model refinements of microalgae
Author(s):	Amnah Alzahmi, Sarah Daakour, Amphun Chaiboonchoe, Bushra S. Dohai, Weiqi Fu, David R. Nelson, Alexandra Mystikou, Kourosh Salehi-Ashtiani

Item 1: The Author elects to have the Materials be made available (as described at http://www.jove.com/publish) via:

☐ Standard Access ☒ Open Access

Item 2: Please select one of the following items:

☒ The Author is **NOT** a United States government employee.

☐ The Author is a United States government employee and the Materials were prepared in the course of his or her duties as a United States government employee.

ARTICLE AND VIDEO LICENSE AGREEMENT

1. **Defined Terms.** As used in this Article and Video License Agreement, the following terms shall have the following meanings: "**Agreement**" means this Article and Video License Agreement; "**Article**" means the article specified on the last page of this Agreement, including any associated materials such as texts, figures, tables, artwork, abstracts, or summaries contained therein; "**Author**" means the author who is a signatory to this Agreement; "**Collective Work**" means a work, such as a periodical issue, anthology or encyclopedia, in which the Materials in their entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole; "**CRC License**" means the Creative Commons Attribution-Non Commercial-No Derivs 3.0 Unported Agreement, the terms and conditions of which can be found at: http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode; "**Derivative Work**" means a work based upon the Materials or upon the Materials and other pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Materials may be recast, transformed, or adapted; "**Institution**" means the institution, listed on the last page of this Agreement, by which the Author was employed at the time of the creation of the Materials; "**JoVE**" means MyJove Corporation, a Massachusetts corporation and the publisher of The Journal of Visualized Experiments; "**Materials**" means the Article and / or the Video; "**Parties**" means the Author and JoVE; "**Video**" means any video(s) made by the Author, alone or in conjunction with any other parties, or by JoVE or its affiliates or agents, individually or in collaboration with the Author or any other parties, incorporating all or any portion of the Article, and in which the Author may or may not appear.

2. **Background.** The Author, who is the author of the Article, in order to ensure the dissemination and protection of the Article, desires to have the JoVE publish the Article and create and transmit videos based on the Article. In furtherance of such goals, the Parties desire to memorialize in this Agreement the respective rights of each Party in and to the Article and the Video.

3. **Grant of Rights in Article.** In consideration of JoVE agreeing to publish the Article, the Author hereby grants to JoVE, subject to **Sections 4** and **7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Article in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Article into other languages, create adaptations, summaries or extracts of the Article or other Derivative Works (including, without limitation, the Video) or Collective Works based on all or any portion of the Article and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and(c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. If the "Open Access" box has been checked in **Item 1** above, JoVE and the Author hereby grant to the public all such rights in the Article as provided in, but subject to all limitations and requirements set forth in, the CRC License.

4. **Retention of Rights in Article.** Notwithstanding the exclusive license granted to JoVE in **Section 3** above, the Author shall, with respect to the Article, retain the non-exclusive right to use all or part of the Article for the non-commercial purpose of giving lectures, presentations or teaching classes, and to post a copy of the Article on the Institution's website or the Author's personal website, in each case provided that a link to the Article on the JoVE website is provided and notice of JoVE's copyright in the Article is included. All non-copyright intellectual property rights in and to the Article, such as patent rights, shall remain with the Author.

5. **Grant of Rights in Video – Standard Access.** This **Section 5** applies if the "Standard Access" box has been checked in **Item 1** above or if no box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby acknowledges and agrees that, Subject to **Section 7** below, JoVE is and shall be the sole and exclusive owner of all rights of any nature, including, without limitation, all copyrights, in and to the Video. To the extent that, by law, the Author is deemed, now or at any time in the future, to have any rights of any nature in or to the Video, the Author hereby disclaims all such rights and transfers all such rights to JoVE.

6. **Grant of Rights in Video – Open Access.** This **Section 6** applies only if the "Open Access" box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby grants to JoVE, subject to **Section 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Video in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Video into other languages, create adaptations, summaries or extracts of the Video or other Derivative Works or Collective Works based on all or any portion of the Video and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. For any Video to which this **Section 6** is applicable, JoVE and the Author hereby grant to the public all such rights in the Video as provided in, but subject to all limitations and requirements set forth in, the CRC License.

7. **Government Employees.** If the Author is a United States government employee and the Article was prepared in the course of his or her duties as a United States government employee, as indicated in **Item 2** above, and any of the licenses or grants granted by the Author hereunder exceed the scope of the 17 U.S.C. 403, then the rights granted hereunder shall be limited to the maximum rights permitted under such statute. In such case, all provisions contained herein that are not in conflict with such statute shall remain in full force and effect, and all provisions contained herein that do so conflict shall be deemed to be amended so as to provide to JoVE the maximum rights permissible within such statute.

8. **Protection of the Work.** The Author(s) authorize JoVE to take steps in the Author(s) name and on their behalf if JoVE believes some third party could be infringing or might infringe the copyright of either the Author's Article and/or Video.

9. **Likeness, Privacy, Personality.** The Author hereby grants JoVE the right to use the Author's name, voice, likeness, picture, photograph, image, biography and performance in any way, commercial or otherwise, in connection with the Materials and the sale, promotion and distribution thereof. The Author hereby waives any and all rights he or she may have, relating to his or her appearance in the Video or otherwise relating to the Materials, under all applicable privacy, likeness, personality or similar laws.

10. **Author Warranties.** The Author represents and warrants that the Article is original, that it has not been published, that the copyright interest is owned by the Author (or, if more than one author is listed at the beginning of this Agreement, by such authors collectively) and has not been assigned, licensed, or otherwise transferred to any other party. The Author represents and warrants that the author(s) listed at the top of this Agreement are the only authors of the Materials. If more than one author is listed at the top of this Agreement and if any such author has not entered into a separate Article and Video License Agreement with JoVE relating to the Materials, the Author represents and warrants that the Author has been authorized by each of the other such authors to execute this Agreement on his or her behalf and to bind him or her with respect to the terms of this Agreement as if each of them had been a party hereto as an Author. The Author warrants that the use, reproduction, distribution, public or private performance or display, and/or modification of all or any portion of the Materials does not and will not violate, infringe and/or misappropriate the patent, trademark, intellectual property or other rights of any third party. The Author represents and warrants that it has and will continue to comply with all government, institutional and other regulations, including, without limitation all institutional, laboratory, hospital, ethical, human and animal treatment, privacy, and all other rules, regulations, laws, procedures or guidelines, applicable to the Materials, and that all research involving human and animal subjects has been approved by the Author's relevant institutional review board.

11. **JoVE Discretion.** If the Author requests the assistance of JoVE in producing the Video in the Author's facility, the Author shall ensure that the presence of JoVE employees, agents or independent contractors is in accordance with the relevant regulations of the Author's institution. If more than one author is listed at the beginning of this Agreement, JoVE may, in its sole

discretion, elect not take any action with respect to the Article until such time as it has received complete, executed Article and Video License Agreements from each such author. JoVE reserves the right, in its absolute and sole discretion and without giving any reason therefore, to accept or decline any work submitted to JoVE. JoVE and its employees, agents and independent contractors shall have full, unfettered access to the facilities of the Author or of the Author's institution as necessary to make the Video, whether actually published or not. JoVE has sole discretion as to the method of making and publishing the Materials, including, without limitation, to all decisions regarding editing, lighting, filming, timing of publication, if any, length, quality, content and the like.

12. **Indemnification.** The Author agrees to indemnify JoVE and/or its successors and assigns from and against any and all claims, costs, and expenses, including attorney's fees, arising out of any breach of any warranty or other representations contained herein. The Author further agrees to indemnify and hold harmless JoVE from and against any and all claims, costs, and expenses, including attorney's fees, resulting from the breach by the Author of any representation or warranty contained herein or from allegations or instances of violation of intellectual property rights, damage to the Author's or the Author's institution's facilities, fraud, libel, defamation, research, equipment, experiments, property damage, personal injury, violations of institutional, laboratory, hospital, ethical, human and animal treatment, privacy or other rules, regulations, laws, procedures or guidelines, liabilities and other losses or damages related in any way to the submission of work to JoVE, making of videos by JoVE, or publication in JoVE or elsewhere by JoVE. The Author shall be responsible for, and shall hold JoVE harmless from, damages caused by lack of sterilization, lack of cleanliness or by contamination due to the making of a video by JoVE its employees, agents or independent contractors. All sterilization, cleanliness or decontamination procedures shall be solely the responsibility of the Author and shall be undertaken at the Author's expense. All indemnifications provided herein shall include JoVE's attorney's fees and costs related to said losses or damages. Such indemnification and holding harmless shall include such losses or damages incurred by, or in connection with, acts or omissions of JoVE, its employees, agents or independent contractors.

13. **Fees.** To cover the cost incurred for publication, JoVE must receive payment before production and publication of the Materials. Payment is due in 21 days of invoice. Should the Materials not be published due to an editorial or production decision, these funds will be returned to the Author. Withdrawal by the Author of any submitted Materials after final peer review approval will result in a US$1,200 fee to cover pre-production expenses incurred by JoVE. If payment is not received by the completion of filming, production and publication of the Materials will be suspended until payment is received.

14. **Transfer, Governing Law.** This Agreement may be assigned by JoVE and shall inure to the benefits of any of JoVE's successors and assignees. This Agreement shall be governed and construed by the internal laws of the Commonwealth of Massachusetts without giving effect to any conflict of law provision thereunder. This Agreement may be executed in counterparts, each of which shall be deemed an original, but all of which together shall be deemed to me one and the same agreement. A signed copy of this Agreement delivered by facsimile, e-mail or other means of electronic transmission shall be deemed to have the same legal effect as delivery of an original signed copy of this Agreement.