

# ICeE an interface for *C. elegans* experiments

Frédéric Montañana<sup>1,2,3,†,‡</sup>, Renaud A Julien<sup>1,2,3,†,‡</sup>, Philippe Vaglio<sup>4</sup>, Lisa R Matthews<sup>1,2,3</sup>, Laurent Tichit<sup>5</sup>,  
and Jonathan J Ewbank<sup>1,2,3,\*</sup>

<sup>1</sup>Centre d'Immunologie de Marseille-Luminy; UM2 Aix-Marseille Université; Marseille, France; <sup>2</sup>INSERM U1104; Marseille, France; <sup>3</sup>CNRS UMR7280; Marseille, France; <sup>4</sup>Modul-Bio; Parc Scientifique Luminy Biotech II; Marseille, France; <sup>5</sup>I2M (UMR 7373); Aix-Marseille Université; Marseille, France

<sup>†</sup>Current address: Modul-Bio; Case 935; Marseille, France.

<sup>‡</sup>These authors contributed equally.

**Keywords:** Biosort, electronic notebook, database, metadata, RNAi

**Abbreviations:** ICeE, Interface for *C. elegans* experiments; qPCR, quantitative polymerase chain reaction; COPAS, Complex Object Parametric Analyzer and Sorter; RNAi, RNA interference; PHP, Hypertext Preprocessor; MySQL, My Structured Query Language; MIAME, Minimum Information About a Microarray Experiment; CGC, Caenorhabditis Genetics Center; dsRNA, double-stranded RNA; MB, megabyte; CSV, comma-separated values; LIMS, laboratory information management system; pdf, portable document format

An increasing number of laboratories are using the COPAS Biosort™ to implement high-throughput approaches to tackle diverse biological problems. While providing a powerful tool for generating quantitative data, the utility of the Biosort is currently limited by the absence of resources for data management. We describe a simple electronic database designed to allow easy storage and retrieval of Biosort data for *C. elegans*, but that has a wide potential application for organizing electronic files and data sets. ICeE is an Open Source application. The code and accompanying documentation are freely available via the web at [http://www.ciml.univ-mrs.fr/EWBANK\\_jonathan/software.html](http://www.ciml.univ-mrs.fr/EWBANK_jonathan/software.html).

## Introduction

Many different experimental approaches generate results in the form of electronic files, such as digital images, or numerical data in spreadsheets. In a typical laboratory setting, these files are generated on machines shared by multiple users and then transferred to individuals' computers for analysis and storage. To take one example, real-time qPCR machines generate spreadsheets of raw data that then must be manipulated to extract the biologically relevant information about gene expression levels. An individual researcher can quickly accumulate a substantial number of data files. In the absence of a rational and simple method to organize files, allowing the simultaneous storage of details of the biological sample and experimental protocol, data can easily be misplaced or become uninterpretable.

This situation is particularly acute for users of the COPAS™ (for Complex Object Parametric Analyzer and Sorter) platform. This specialized range of flow cytometry systems are conceived for high-throughput analysis, sorting, and dispensing of objects ranging in size from approximately 20 to 1500 microns. Five physical parameters are optically measured and recorded for each object of interest: optical density (extinction); axial length (time of flight);

and fluorescence emissions, simultaneously by 3 different detectors. They are becoming increasingly popular in laboratories engaged in functional genomic study with model organisms.

One of the COPAS™ machines, the Biosort, is adapted for use with *C. elegans*.<sup>1</sup> With its Reflex attachment, it can aspirate and analyze hundreds of individual nematodes from liquid cultures in 96-well plates. With its shortest analysis cycle, a plate can be analyzed in 36 min, so during an 8 h period more than a dozen 96-well plates can be handled.<sup>2</sup> Thus, in a day, results for more than a thousand samples, generally corresponding to worms subject to different experimental treatments, can be generated. Keeping track of this number of results, especially in the context of quantitative genome-wide RNAi screens,<sup>3</sup> is not a trivial task and incited us to design and implement a database to ensure that information was stored in a user-friendly way.

## Results

We constructed a web-based application, ICeE, using PHP and Javascript, coupled to an associated MySQL database. While ICeE is designed to be compatible with the most commonly used

© Frederic Montanana, Renaud A Julien, Philippe Vaglio, Lisa R Matthews, Laurent Tichit, and Jonathan J Ewbank

\*Correspondence to: Jonathan J Ewbank; Email: [ewbank@ciml.univ-mrs.fr](mailto:ewbank@ciml.univ-mrs.fr)

Submitted: 03/14/2014; Revised: 07/07/2014; Accepted: 07/25/2014

<http://dx.doi.org/10.4161/21624046.2014.959420>

This is an Open Access article distributed under the terms of the Creative Commons Attribution-Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The moral rights of the named author(s) have been asserted.

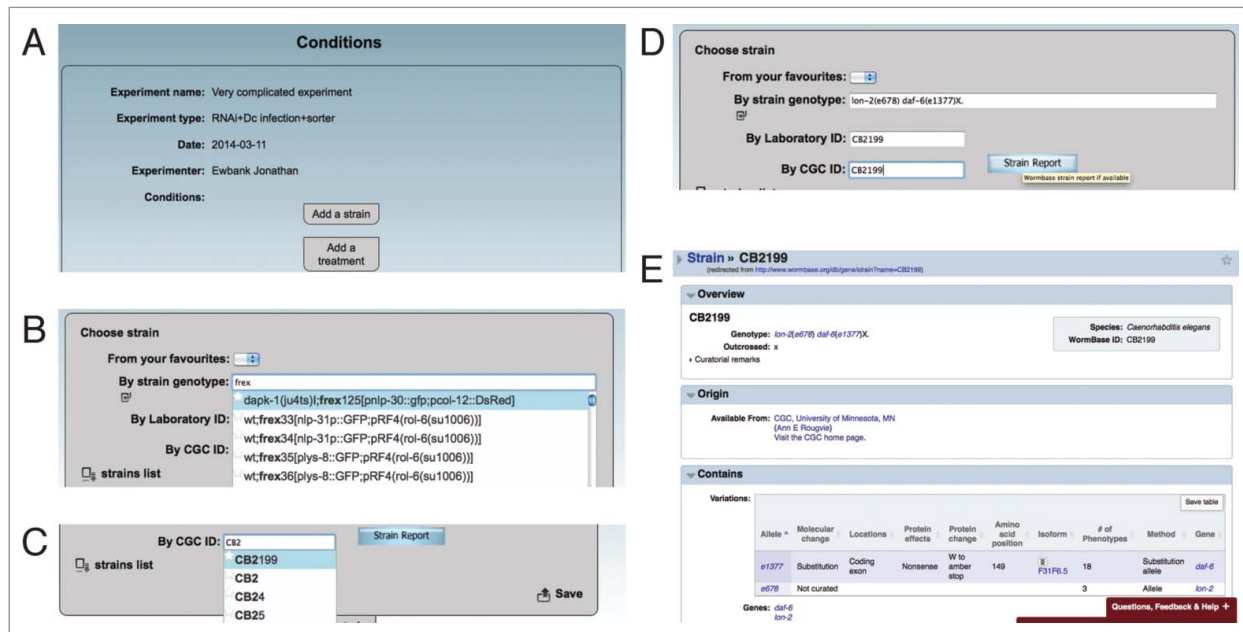
web browsers, the use of Firefox™ is recommended. ICeE permits users to specify a series of pre-defined parameters to provide a complete description of an experiment, inspired by the MIAME guidelines ([www.mged.org](http://www.mged.org)), and to associate these experimental criteria with a set of data. First, there are general identifiers, such as the date and the experimenter's name. Then, there is the rubric "protocol," which can contain free-text, or be linked to one or more separate files. Under the section, "Conditions," the experimenter adds precise details of the experimental conditions. Since it is designed for *C. elegans*, the application comes pre-loaded with a list of the nematode strains available from the *Caenorhabditis* Genetics Center (<http://www.cbs.umn.edu/CGC>). ICeE offers the possibility of updating this list and of importing a custom list of strains from a tab-delimited text file. Many laboratories use proprietary software, such as Filemaker™ to record their strain collections. Most allow data export in a tab-delimited format and so are easily compatible with ICeE. Strains can be picked from a pull-down menu; the first few characters of a strain's name or genotype will be auto-completed to give the user a more and more refined list of the corresponding strains. As individual users may frequently use the same nematode strains, these can be added simply to an easily accessible list of "favorites." The database entries for strains from the CGC are directly linked to Wormbase; users can access the corresponding Wormbase strain report page with a click (Fig. 1).

Having chosen the strain, the particular experimental conditions are defined. These include different types of "treatment," such as "chemical exposure." The details of the chemical treatment, the compound(s) used and their concentration, the time of exposure and the developmental stage can be defined from

specific menus. Users have the possibility of adding compounds to the menu. Since researchers frequently apply RNAi to inactivate the expression of specific genes in *C. elegans*, "RNAi" is included as a possible treatment. RNAi is often performed by feeding nematodes with individual bacterial clones expressing a defined dsRNA. The clones generally come from large collections (e.g. reference 4). The main ones come pre-loaded in the database, and can be updated if required. Otherwise, users also have the possibility of adding their own clones, singly or by importing descriptions of multi-well plates (Fig. 2).

To provide flexibility, multiple treatments can be associated with a single experimental condition (e.g., following an RNAi treatment, worms are cultured in the presence of a high concentration of salt). Also, experiments may use a single or multiple nematode strains. For the latter, lists of strains can be constituted and associated with the same experimental treatment(s), facilitating data entry (Fig. 3).

Once the experimental conditions have been defined, the user can upload the results file(s). There are no constraints for the format of the result files, which can be up to 10 MB in the released version. Multiple result files of the same or different types can be uploaded for a single experiment (Fig. 4). At this stage, an additional rubric allows any further comments to be associated with the experiment. The user is then prompted to add the entire experiment (protocol, conditions, results and comments) to the database. In line with common practices, once an experiment has been entered, it can only be deleted by the database administrator. Users on the other hand can modify entries. In this case, both the original and edited versions are recorded in the database.



**Figure 1.** Initial data entry in ICeE. Screen-grabs of data entry for an imaginary experiment. Having specified a name for the experiment and defining its nature (A), users are prompted to list the strains used, by entering either the strain genotype (B) or strain name (C). In both cases, pull-down lists of strains containing the typed text appear. Once a strain is chosen, its name and details of its genotype appear. (D) For strains that are at the CGC, clicking on the "Strain Report" button links to the relevant strain report page at Wormbase (E).

In the above, we have described the typical sequence for entering an experiment, but ICeE is designed to allow the different elements to be entered in the most appropriate order. Indeed, to upload result files, it is sufficient to have defined the experiment's name. All other details can be added subsequently, and there is the possibility to edit any and all fields up until the point when the experiment is added to the database.

Since researchers may perform the same type of experiment repeatedly, an existing experimental protocol can be duplicated as a template for a new experiment, simply by clicking "Insert similar" (Fig. 5), alleviating the need to re-enter all the experimental details. In this way, data entry becomes increasingly simple and rapid as more types of experimental protocols are included. Further, once an experiment has been entered into ICeE, users can return to the corresponding entry in the database and add additional files at any time. Thus, for example, having initially entered the raw data, a user can subsequently upload any number of files of data analysis; an initial raw text or CSV data file from the Biosort can be complemented with graphical representations or statistical analyses generated in software such as Open Office™, or using dedicated tools.<sup>5</sup> Comments can be added to the stored experiment to explain the contents of the different files.

In addition to serving as a data repository, ICeE gives users the possibility of performing searches to retrieve experiments and the associated data sets. An intuitive interface provides the possibility of searching globally or in defined fields, using exact or inexact queries. Having selected the experiment of interest, all the constituent files (protocol and results) can be opened or downloaded.

A complete description of how to use ICeE is provided in the instruction manual, accessible from all pages within the application (under the rubric "Help").

**A**

Search RNAi : sup-9 Show 10 clones.

sjj_F34D6.3	sup-9	Julie Ahringer	<input checked="" type="checkbox"/>
-------------	-------	----------------	-------------------------------------

Input a plate list (specific for screen):

Exposure Time : 12 Hour(s) 30 Minute(s)

Age exposed : L1 L2 L3 L4 Young Adult

Synchronized

**B**

Experiment name: Very complicated experiment

Experiment type: RNAi+Dc infection+sorter

Date: 2014-03-11

Experimenter: Ewbank Jonathan

Conditions:

Genotype : lon-2(e678) daf-6(e1377)X. Lab ID : CB2199 CGC ID : CB2199

RNAi using the clone : sjj\_F34D6.3 L4

Add a strain

Add a treatment

**C**

Genotype : lon-2(e678) daf-6(e1377)X. Lab ID : CB2199 CGC ID : CB2199













RNAi using the clone : sjj\_F34D6.3 L4

Chemical exposure of Tunicamycin on Plates L4

**Figure 2.** Defining experimental conditions. Screen-grabs of data entry for an imaginary experiment. **(A)** Having chosen "RNAi" by clicking on the left-hand icon, the user can search for a specific clone within the pre-loaded libraries. Here, the query "sup-9" returns one clone. The details of the experiment are then added (time of exposure and starting stage). The other icons correspond to (from left to right), exposure to a chemical, infection, wounding, heat-shock and control. For each, dedicated sub-menus appear. **(B)** Once a treatment has been defined, a summary is shown. **(C)** Any number of treatments can be added. In the example here, RNAi treatment is associated with exposure to tunicamycin.

## Discussion

The database described here simplifies and rationalizes the storage of disparate electronic data files. While its conception was largely influenced by the needs of a laboratory using *C. elegans* and a COPAS™ Biosort (see for example, refs.

<input checked="" type="checkbox"/> Genotype : acbp-3(sv73)IV Lab ID : FE0002 CGC ID : FE0002	  
<input checked="" type="checkbox"/> Genotype : pmk-1(km25)IV;frEx506[dcar-1p::GFP;col-1 Lab ID : IG1428 CGC ID : not specified	  
<input checked="" type="checkbox"/> Genotype : hmt-1(gk161) III. Lab ID : VC287 CGC ID : VC287	  
<input checked="" type="checkbox"/> Genotype : mes-2(ok2480)/mT1 II; +/mT1[dpy-10(e128) Lab ID : VC2409 CGC ID : VC2409	  

**Figure 3.** Defining sets of strains. Screen-grab of data entry for an imaginary experiment. To facilitate data entry, multiple strains can be chosen and each associated with a single set of experimental conditions.

6–10), it is designed to be readily adaptable; the code is Open Source. Currently, ICeE is a stand-alone tool, but it can be integrated into a broader system, for example, with a laboratory information management system (LIMS) to ensure sample traceability (unpublished results). It allows a summary document to be produced automatically and could clearly be modified if minimal reporting standards, such as those used for microarrays and proteomics, become more generally

applied, to facilitate submission to the appropriate public data repositories.

By default, the metadata and data entered into ICeE is accessible to all users. With the system's search facilities, this means that all members of a laboratory can recall the details and results of any experiment (Fig. 5). This allows one to determine rapidly whether a given type of experiment has already been conducted, and to know the results. This should reduce unnecessary duplication of experiments and allow cross-investigator comparisons.


One challenge faced by any laboratory that aims to maintain long-term accessibility to electronic data is that of user compliance. Any formalized data storage procedure can be perceived as an unnecessary drain on time. ICeE was designed to require a minimal effort to record the essential metadata associated with a set of results. Further, the more that ICeE is used, the more likely it is that a user can simply duplicate a pre-existing set of experimental conditions and apply limited modifications to suit a particular experiment. In practice, in order to guarantee compliance, it may be necessary to implement specific procedures for the use of ICeE. This might

6–10), it is designed to be readily adaptable; the code is Open Source. Currently, ICeE is a stand-alone tool, but it can be integrated into a broader system, for example, with a laboratory information management system (LIMS) to ensure sample traceability (unpublished results). It allows a summary document to be produced automatically and could clearly be modified if minimal reporting standards, such as those used for microarrays and proteomics, become more generally

## Results

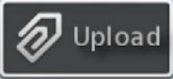
**Experiment name:** Very complicated experiment

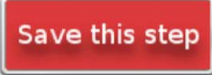
**Results files:**

File(s) already uploaded	Del
ChV-6-2-1-04-03-2010.xls	

Individual results files must not exceed 10MB in size. Files with the extensions xlsx, pptx and docx should be avoided as they may not be fully compatible with ICeE. Some characters may be replaced automatically (accents, spaces, etc)."

Browse... No file selected.





**Figure 4.** Uploading result files. Screen-grab of data entry for an imaginary experiment. Once a minimum amount of information has been stored (experiment name and date), users can upload one or more result files, in this example, a single Microsoft Excel™ (xls) file.



**Figure 5.** Retrieving information. Screen-grab of the data associated with an experiment. The collection of experiments can be queried and the data and metadata retrieved. In this example, the entry for “time course D. C. infection for RNA seq” has been edited since its original creation in ICeE. Clicking on the pull-down menu at the top right gives access to previous versions. In this specific example, RNA samples were prepared from worms infected with the fungus *Drechmeria coniospora* for subsequent RNaseq analysis (S. Omi, unpublished results). In addition to the report on the RNA quality (the pdf file at the bottom of the list), the raw data for qRT-PCR tests on selected genes (the xls file at the top) and a graphical summary (the xlsx file) conducted to characterize the samples, are included.

involve an administrator regularly verifying correct use of the database, but could also rely on a more formalized system. As an example, electronic files of Biosort results could be provided to users only when the appropriate metadata has been entered into ICeE. But it is to be hoped that all users would be able to see the advantage of having a common searchable metadata and data repository and embrace its use without the need for coercion.

#### Disclosure of Potential Conflicts of Interest

I declare that part of the work for this project was undertaken in the context of a consortium, funded by the Region Provence-Alpes-Côte d’Azur, involving my laboratory, and the private companies Modul-Bio and Union Biometrica.

#### References

1. Pulak R. Techniques for analysis, sorting, and dispensing of *C. elegans* on the COPAS flow-sorting system. *Methods Mol Biol* 2006; 351:275-86; PMID:16988441
2. Duverger Y, Belougne J, Scaglione S, Brandli D, Beclin C, Ewbank JJ. A semi-automated high-throughput approach to the generation of transposon insertion mutants in the nematode *Caenorhabditis elegans*. *Nucleic Acids Res* 2007; 35:e11; PMID:17164286; <http://dx.doi.org/10.1093/nar/gkl1046>
3. Squiban B, Belougne J, Ewbank J, Zugasti O. Quantitative and automated high-throughput genome-wide RNAi screens in *C. elegans*. *J Vis Exp* 2012; 60:e3448; PMID:22395785
4. Kamath RS, Ahringer J. Genome-wide RNAi screening in *Caenorhabditis elegans*. *Methods* 2003; 30:313-21; PMID:12828945; [http://dx.doi.org/10.1016/S1046-2023\(03\)00050-1](http://dx.doi.org/10.1016/S1046-2023(03)00050-1)
5. Morton E, Lamitina T. A suite of MATLAB-based computational tools for automated analysis of COPAS Biosort data. *Biotechniques* 2010; 48:xxv-xxx; PMID:20569218; <http://dx.doi.org/10.2144/000113427>
6. Cardoso C, Couillaud C, Mignon-Ravix C, Miller A, Ewbank JJ, Fontés M, Pujol N. XNP-1/ATR-X acts with RB, HP1 and the NuRD complex during larval development in *C. elegans*. *Dev Biol* 2005; 278:49-59; PMID:15649460; <http://dx.doi.org/10.1016/j.ydbio.2004.10.014>
7. Ziegler K, Kurz CL, Cypowyj S, Couillaud C, Pophillat M, Pujol N, Ewbank JJ. Antifungal innate immunity in *C. elegans*: PKCdelta links G protein signaling and a conserved p38 MAPK cascade. *Cell Host Microbe* 2009; 5:341-52; PMID:19380113; <http://dx.doi.org/10.1016/j.chom.2009.03.006>
8. Dierking K, Polanowska J, Omi S, Engelmann I, Gut M, Lembo F, Ewbank JJ, Pujol N. Unusual regulation of a STAT protein by an SLC6 family transporter in *C. elegans* epidermal innate immunity. *Cell Host Microbe* 2011; 9:425-35; PMID:21575913; <http://dx.doi.org/10.1016/j.chom.2011.04.011>
9. Labeid SA, Omi S, Gut M, Ewbank JJ, Pujol N. The pseudokinase NIP1-4 is a novel regulator of antimicrobial peptide gene expression. *PLoS One* 2012; 7:e33887; PMID:22470487; <http://dx.doi.org/10.1371/journal.pone.0033887>
10. Zugasti O, Bose N, Squiban B, Belougne J, Kurz CL, Schroeder FC, Pujol N, Ewbank JJ. Activation of a G protein-coupled receptor by its endogenous ligand triggers the innate immune response of *Caenorhabditis elegans*. *Nat Immunol* 2014; 15:833-8.

#### Acknowledgments

We thank Jérôme Belougne, Frederic Garcia, Juliette Hayer, Mathilde Pellerin, Thierry Molina and Eric Vercherat for their active participation in the elaboration and implementation of ICeE, Ian Hope for constructive criticisms, and Don Moerman, Mike Simpson, Mark Edgley, Angela Miller as well as members of the Ewbank lab for beta-testing.

#### Funding

This work was supported by institutional funding from INSERM and CNRS, and project grants from the ANR (MIME-2007 and ANR-12-BSV3-0001-01), RNG and the PACA Region.