Enabling systematic, harmonised and large-scale biofilms data computation: The Biofilms Experiment Workbench

Gael Pérez-Rodríguez[a], Daniel Glez-Peña[a], Nuno F. Azevedo[b], Maria Olívia Pereira[c], Florentino Fdez-Riverola[a], Anália Lourenço[a,c,]*

[a] *ESEI – Escuela Superior de Ingeniería Informática, Edificio Politécnico, Campus Universitario As Lagoas, s/n, Universidad de Vigo, 32004 Ourense, Spain[1]*

[b] *LEPABE – Department of Chemical Engineering, Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, s/n, 4200-465 Porto, Portugal*

[c] *Centre of Biological Engineering, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal*

Abstract

Background and objective: Biofilms are receiving increasing attention from the biomedical community. Biofilm-like growth within human body is considered one of the key microbial strategies to augment resistance and persistence during infectious processes. The Biofilms Experiment Workbench is a novel software workbench for the operation and analysis of biofilms experimental data. The goal is to promote the interchange and comparison of data among laboratories, providing systematic, harmonised and large-scale data computation. *Methods:* The workbench was developed with AIBench, an open-source Java desktop application framework for scientific software development in the domain of translational biomedicine. Implementation favours free and open-source third-parties, such as the R statistical package, and reaches for the Web services of the BiofOmics database to enable public experiment deposition.

Results: First, we summarise the novel, free, open, XML-based interchange format for encoding biofilms experimental data. Then, we describe the execution of common scenarios of operation with the new workbench, such as the creation of new experiments, the importation of data from Excel spreadsheets, the computation of analytical results, the on-demand and highly customised construction of Web publishable reports, and the comparison of results between laboratories.

Conclusions: A considerable and varied amount of biofilms data is being

generated, and there is a critical need to develop bioinformatics tools that expedite the interchange and comparison of microbiological and clinical results among laboratories. We propose a simple, open-source software infrastructure which is effective, extensible and easy to understand. The workbench is freely available for non-commercial use at http://sing.ei.uvigo.es/bew under LGPL license.

1.   Introduction

Infectious diseases are a major worldwide cause of morbidity and mortality, and biofilm formation stands out as a key virulence factor contributing to the emergence and dissemination of antibiotic resistance traits in chronic and nosocomial infections [1–4]. Pneumonia in cystic fibrosis patients, chronic wounds, and catheter-associated infections are meaningful examples of biofilm-based infections.

The severity and widespread dissemination of these infections, and most notably the increasing emergence of multi-drug resistant strains, have promoted dynamic research on the molecular interplay underlying these microbial agglomerates. The ultimate aim is to discover new drug targets and design drugs with alternative, and more effective modes of action [5–8]. To this end, conventional microbiological experimentation is giving place to high-throughput and multidisciplinary experimentation [9]. Cell viability, biomass formation, respiratory activity, morphological characterisation, and transcriptome and proteome profiling are among the variety of analytical methods commonly used now.

These experiments are producing a wealth of data, but experimental results could be better explored if only the interchange and comparison of data among laboratories was not so limited. Often, different laboratories report contradictory results about similar infection scenarios. The comparison of raw data and the unequivocal characterisation of experimental methods would allow the evaluation of the possible cause(s) of such nonconformity. Notably, it is of upmost importance to differentiate between procedural discrepancies and natural-occurring biological variation and thus, to be able to assure the reproducibility and ruggedness of the results. A first step in this direction is the specification of a common computer-readable and interchangeable data format, which establishes the minimum set of information necessary to guarantee the comprehensibility of the experiment (both in terms of the procedures used and the data obtained), and the representation of the data itself.

The MIABiE initiative,[2] encompassing an international body of Biofilms

experts, is working on the definition of guidelines to document biofilms experiments and the standardisation of the nomenclature in use [10]. Biofilms-centred databases such as BiofOmics[3] [11] and MorphoCol[4] [12] are already endorsing these guidelines and making experimental data publicly available. However, the community lacks computational tools to assist researchers in the preparation of experimental data files, and the individual and comparative analysis of experimental results. Typically, researchers organise and document their experiments as most convenient at the moment, and research articles show only selected results, e.g. plots and some statistical significance values, that sustain the conclusions in a more intuitive way. This *ad hoc* style of analysis can lead to human error and misinterpretation, loss of data and results, and lack of verifiability, repeatability and extensibility. Moreover, this approach prevents the experiments from being easily uploaded to public databases (namely, being available on the Web in a query-able form for the community), and being computed straightforwardly by other researchers than the authors.

To bridge this gap, this work introduces the Biofilms Experiment Workbench (BEW), the first software tool dedicated to biofilms data operation and analysis [13]. BEW aims to make analysis more systematic by consolidating data, analysis and results. Towards this end, this work also proposes the Biofilms Markup Language (BML) as a new data representation format for modelling biofilms experiments, and effectively promoting data interchange across resources and software tools. BEW endorses BML and supports main data operation and analysis functionalities, such as: (i) the customised, but standardised documentation of experiments, (ii) the statistical assessment of various analytical results, (iii) on-demand and Web-publishable experiment reporting, (iv) the deposition of experiments in public databases, and (v) the comparison of results between laboratories. Although emphasising the unique ability of BEW to work with biofilms experimental data, it is noteworthy that researchers may still put it to use for documenting more general microbiological studies. Biofilm research is built upon biofilm-specific experiments as well as experiments common to other scientific areas. So, BEW is able to manage information of microbiological studies with various purposes.

The next sections detail the BML, the architecture of BEW and its main functionalities. Further documentation can be found at its Web site (http://sing.ei.uvigo.es/bew), including some sample data and demos. The workbench is available free of charge for non-commercial use under LGPL license

2.   Methods

2.1.   BML data representation

The goal of the proposed BML is to serve as a software *lingua franca* supporting the encoding of biofilms experimental data such that those experiments can be exchanged and interpreted unambiguously by different software systems.

Biofilm research is built upon biofilm-specific experiments as well as experiments common to other scientific areas. The BML does not aim to overtake or compete with any existing standard or representation but rather to address biofilm-specific information requirements. Notably, the MIABiE consortium decided to delegate the documentation guidelines of non biofilm-specific data to the appropriate community initiatives [10]. For instance, data coming from transcriptome, proteome and other "omic" technologies applied to biofilm populations should be documented as suggested by MIAME [14], MIAPE [15], and similar guidelines. This would also be the case for specific techniques, such as fluorescence *in situ* hybridisation or flow cytometry, for which minimum information guidelines have also been reported [16]. While existing XML-based laboratory standards, such as the AnIML (the Analytical Instrumentation Markup Language) [17], support the documentation of laboratory workflows, high-throughput and statistical methods are not conveniently linked. As a result, BML was created with the purpose to provide a whole-experiment data representation of the biofilm study, encompassing the corresponding experimental setup, assays and methods of analysis (Fig. 1).

The Experiment structure is the highest-level construct in the BML document. It defines a grouping of components, namely the experiment name and short description, the list of methods of analysis, the list of experimental conditions, the data series (i.e. the experimental data) and the authorship information that define a given experiment. One or more component of type Experiment are allowed per instance of an BML document, depending on whether the experiment was per- formed by a single laboratory or by multiple laboratories. An Experiment must contain at least one Condition, one Method of Analysis and one Authorship data element, but BML does not impose restrictions on the total number of these.

The BML data representation format was formalised in eXtensible Markup Language (XML) [18] because of its portability and widespread acceptance as a standard data language for Bioinformatics. The language is available free of cost and restrictions to all users, developers, and other interested persons and organisations.

BEW is a desktop-based application developed with AIBench, an open-source Java desktop application framework for scientific software development in the domain of translational biomedicine [19].

AIBench provides the developer with a plugin-based and low intrusive framework which is based on sound architectural design patterns [20]. AIBench is well suited for scientific applications which follow the simple input-process-output (IPO) model. The developer has to focus only in three kinds of software objects: (i) operations, including specific data analysis algorithms and data loading/saving routines, (ii) datatypes, produced and consumed by operations, representing the application domain data and (iii) views, which are visual components rendering datatypes in an user-friendly manner. While the developer is in charge of creating these application specific objects, AIBench is in charge of providing common functionalities, such as user parameters retrieval for every operation, basic graphical user interface environment, experiment automation and reproduction thanks to an internal scripting framework, multi-thread operation execution, logging, and automatic software update installation.

BEW is an AIBench application and thus, relies on it for basic functionalities of the user interface, that is, a conventional display of objects and operations, a menu bar, a clipboard containing all generated objects during the session, and a visualisation pane for results [13]. For more specialised routines, such as data plotting and statistical analysis, BEW uses third-party libraries. The main data object or datatype in BEW is the Experiment object, which is the minimum unit of information that may be entered to or generated from BEW. An Experiment is structured into experimental conditions, methods of analysis and data series, such that the resulting data are fully contextualised and it is possible to compare results among experiments. All these objects are displayed in a customised way that enables immediate data comprehensiveness as well as customised data download or export. BML files contain a Schema used to validate the data structure used in BEW and, in a way, can be seen as a serialisation of the Experiment instances in BEW.

While the BML has been formalised to attend to biofilm- specific documentation requirements, it is important to notice that the architecture and the development strategy of BEW is quite generic. BEW supports many common analyses in micro- biology and therefore may be used in a larger set of studies, including or not biofilms data.

2.3. External software packages

The workbench integrates third-party software in support of flexible data operation and analysis. Specifically, BEW includes: a plugin to perform statistical tests in R statistical computing tool,[5] the JFreeChart library[6] to perform data plotting, the JXL library[7] to read and write Excel worksheets, and the JSoup library[8] to create customised HTML reports. It is worth of notice that BEW only incorporates free and open-source third-parties as means to not compromise future adaptation and extension of the functionalities, and allow active community participation.

All external software packages except R are included in BEW installation routine. R software has to be installed individually, but BEW provides instructions to step the user through the whole installation effortlessly.

2.4. Web services and database cross-linking

BEW works as a Web service client, i.e. BEW can directly connect to external public databases and import biofilms data. Currently, the BiofOmics database[9] is supported, as means to make experimental data readily available to the community and hence, promote data interchange between individuals and research groups as well as facilitate the collaboration of groups working on large-scale projects. A simple point-and- click interface allows users to upload experiments to their accounts in BiofOmics and issue a request for database publication (Fig. 2). Likewise, users may download for analysis any number of public experiments.

BiofOmics Web services are also used to keep an updated and harmonised version of the biofilms vocabulary to be used in the description of experiments (Fig. 3). Specifically, the user may update the methods of analysis and common test conditions, such as organisms, growth media, antimicrobial agents and adhesion materials (http://biofomics.org/ pages/apiExplanation). Although BEW always includes a release of BiofOmics harmonised vocabulary, this functionality enables continuous update and active vocabulary enrichment, i.e. users may introduce new methods of analysis or test conditions in the local vocabulary and then, propose them to BiofOmics curators.

Notably, part of the vocabulary is cross-linked to domain databases, e.g. organisms are linked to NCBI Taxonomy [21] and other collections, drugs are linked to DrugBank [22] and natural peptides to CAMP [23]. Cross-linking is important to ensure the use of harmonised vocabulary coming from other scientific domains.

3. Results

3.1. User directives and case studies

Given the variety of biological scenarios emulated in biofilms experiments and the wide variety of methods of analysis used, it was important to have a number of different experiments as cases studies. It was crucial to ensure the ability to represent different test conditions (nominal such as organism species or adhesion materials, or numeric such as pH and temperatures values), and design data tables contemplating different ways of combining conditions, and unequivocally expressing data replicates and experimental reproductions of those tests (Table 1). We used a total of 22 experiments, all publicly available at BiofOmics.

Given that this is the very first software to attend to Biofilms bioinformatics needs, the evaluation of user-system interaction is considered crucial to promote and disseminate this tool among researchers. Therefore, BEW was submitted to an interactive evaluation involving 8 testers, with different levels of expertise and biofilms data operation routines. Moreover, BEW was tested in multiple versions of Windows, Linux and Mac operating systems as means to guarantee a satisfactory installation and user experience.

The next sections detail the creation, population and analysis of experimental data in BEW step-by-step using one of the tested case studies – the study of the adhesion of water stressed *Helicobacter pylori* to abiotic surfaces (http://biofomics. org/BOID 1) [23].

3.2. Experiment management

BEW manages projects, which encompass one or multiple experiments and all associated analyses. Users specify projects according to their individual or collaborative goals such that the project encapsulates all data and metadata required to interpret the data and the analyses unambiguously.

Data interoperation is supported by two experiment file for- mats: MS Excel^{TM} worksheets (.xls) and the Biofilms Markup Language (.xml). The BML data format is the default data format of BEW, but spreadsheets are the most common file format among researchers. So, it was important to support this format as well, and enable both data importation and exportation from/to spreadsheets [12]. As such, users may introduce old, manually created data files in BEW, and generate data in BML format for database submission and general interchange.

Fig. 4 illustrates the BML data file for the present case study, which is about a

single experiment that evaluates two conditions – the adhesion material (6 possible values) and the time of growth (7 possible values) – by counting the number of colony forming units.

Regarding user-system interaction, experiment importation and new experiment wizards are similar, the main difference being the automatic introduction of data in importation (Fig. 5). For each method of analysis, the user has to identify or verify the conditions that are tested and indicate the units of measure of the introduced data. Also, the user is urged to employ harmonised vocabulary or clearly state the use of other vocabulary.

After creating the experiment, the user can save the work session at any point. Experiment settings and data are saved as a whole in markup language format, while plots and statistical test results should be downloaded individually.

3.3. Descriptive and statistical analysis

BEW is equipped with a powerful analytical component that supports on-demand and customised construction of data plots and statistical data testing. Both data plotting and testing were made as flexible as possible in order to accommodate the analysis of the results produced by virtually any combination of test conditions (Fig. 6). This ability is crucial given the variety of goals of analysis and the methods of analysis employed to meet them.

Currently, BEW supports the creation of 2D scatter and bar plots, performs the detection of outliers, examines data normality and homoscedasticity, and analyses data variance. For example, using the data of the case study presented above, Fig. 6 shows a bar plot where the impact of the adhesion material on biofilm growth can be assessed, while Fig. 7 illustrates statistical significance of such data.

These were the analytical abilities most requested by the contacted users, but BEW can easily incorporate any other 2D plot or statistical test upon request. Moreover, individual plots may be exported as publishable-quality images in PNG file for- mat, and may be included in experiment reports on demand.

3.4. Experiment report and database submission

The documentation of Biofilms experiments has two practical purposes: the deposition of the experiments in public databases and the description of the results reported in scientific manuscripts. BEW enables the construction of customised Web publishable reports, which may include details on experimental setup, data summary and statistical results. These reports may be

used as experiment profile records that complement the actual data files. In Fig. 8, examples of these records for our case study are presented. These can highlight, for instance, the main points of the experimental setup, by describing the authors, the publication details and providing a summary of the work; or can provide a more detailed look on a subset of data that might be of a particular interest to the user (in this case the number of CFUs for all conditions is depicted).

Experiment submission to and download from BiofOmics database is a key functionality regarding the reinforcement of data interchange across the community (Fig. 2). After users register in BiofOmics, they are entitled to a private area where experiment access is limited to the account owner and BiofOmics curators. Data becomes public only after complying with the minimum information requirements. Meanwhile, BEW enables account owners to download, upload or update data files, and to download any public data files.

3.5. *Demo and help*

Due to the unfamiliarity of the potential users to the software and their relatively small bioinformatics experience, we are committed to provide as much user support as possible. To begin with, the installation of BEW has been made as simple as possible so that the users may download and try out the software as effortlessly as possible.

In the workbench, most interfaces have a "help" button to provide information about the interface. To further expedite software exploration, BEW includes a directory of sample data, with examples of single and multi-lab experiments. Also, at the Web site (http://sing.ei.uvigo.es/bew), users may find step-by-step tutorials and demos on both installation and operation.

4. Discussion

Biofilms is a key domain in Clinical Microbiology that is now engaged in high-throughput and systems-level research, being quite prolific in the generation of experimental data. Such activity has raised the need to find an adequate and scientifically sound way to control the quality of the data and the documentation accompanying the public deposition of biofilm-related experiments.

The international consortium MIABiE has initiated the identification and organisation of a set of modules containing the minimum information that needs to be reported to guarantee the interpretability and independent verification of experimental results, and their integration with knowledge coming from other

fields [10]. To prevent *ad hoc* analysis and to equip researchers with suitable computational means is an immediate necessity of this community effort. Experiments need to be documented consistently and systematically, following common guidelines and using harmonised vocabulary. A common language and data structure will potentiate the use of public databases and expedite laboratory data interchange. Currently, most laboratories do not follow a standardised documentation approach. So, an intrinsic aspect of BEW and BML development is the adoption of a participative, community-oriented, and informal approach that follows MIABiE guidelines. Typically, we take advantage of expert meetings, such as the Eurobiofilms 2013 meeting in Ghent, as well as our long-term collaboration with other laboratories, to raise discussion and acquire user feedback. Indeed, much of the BEW software has grown out of years of collaborating with researchers on biofilms analysis, aiming to promote data exploration and lead to a greater understanding of data relationships. Notably, BEW replicates basic routines of data preparation and analysis, encourages data submission to public databases, and supports collaborative data analysis among laboratories. In the future, the analytical abilities of BEW will be extended to the construction of application-specific decision making models and biological interactions networks.

By using BEW, the daily routine of researchers will be functionally upgraded, paving the way to experiment reproduction across laboratories. Likewise, this will also have an important effect on scientific peer reviewing, raising the field to a new level of confidence and consideration by other scientists in Life Sciences.

5.   Conclusions

Biofilms are a prominent subject within Biomedical research. In this paper we present BEW, the very first bioinformatics tool developed to meet the data processing and analysis requirements of the Biofilms research domain. In complement, we introduce the BML for the formal computerised representation of biofilm experiments, in the broader benefit of the development and use biofilms resources and software packages. With greater interaction between tools, and a common data for- mat for publications and databases, researchers will be able to perform systematic experiment comparison and data inter-change.

Ongoing work is mostly centred in refining the analytical capacities of the workbench, providing new advanced and specialised functionalities. For example, we are working on multidimensional data analysis and the processing and analysis of image data. Based on users' feedback, we devote additional

efforts to satisfy domain-specific requests (both in terms of data management and analysis) and to improve the interface (easy-to-use and the intuitiveness) of the work- bench. In close cooperation with the MIABiE consortium, we are also planning some hands-on courses to introduce the software to biofilms practitioners.

The software, the documentation, and example datasets are publicly available at http://sing.ei.uvigo.es/bew under the LGPL license.

Conflict of interests

The authors do not have any competing interests.

Acknowledgements

References

[1] C.J. Sanchez, K. Mende, M.L. Beckius, K.S. Akers, D.R. Romano, J.C. Wenke, et al., Biofilm formation by clinical isolates and the implications in chronic infections, BMC Infect. Dis. 13 (2013) 47, http://dx.doi.org/10.1186/ 1471-2334-13-47.

[2] T. Bjarnsholt, The role of bacterial biofilms in chronic infections, APMIS (Suppl.) (2013) 1–51, http://dx.doi.org/ 10.1111/apm.12099.

[3] U. Römling, C. Balsalobre, Biofilm infections, their resilience to therapy and innovative treatment strategies, J. Intern. Med. 272 (2012) 541–561,

http://dx.doi.org/10.1111/ joim.12004.

[4] B.M. Peters, M.A. Jabra-Rizk, G.A. O'May, J.W. Costerton, M.E. Shirtliff, Polymicrobial interactions: impact on pathogenesis and human disease, Clin. Microbiol. Rev. 25 (2012) 193–213, http://dx.doi.org/10.1128/CMR.00013-11.

[5] S. de Bentzmann, C. Giraud, C.S. Bernard, V. Calderon, F. Ewald, P. Plésiat, et al., Unique biofilm signature, drug susceptibility and decreased virulence in Drosophila through the *Pseudomonas aeruginosa* two-component system PprAB, PLoS Pathog. 8 (2012) e1003052, http://dx.doi.org/ 10.1371/journal.ppat.1003052.

[6] C.G. Pierce, J.L. Lopez-Ribot, Candidiasis drug discovery and development: new approaches targeting virulence for discovering and identifying new drugs, Expert Opin. Drug Discov. (2013) 1–10, http://dx.doi.org/10.1517/ 17460441.2013.807245.

[7] L.S. Tavares, C.S.F. Silva, V.C. de Souza, V.L. da Silva, C.G. Diniz, M.O. Santos, Strategies and molecular tools to fight antimicrobial resistance: resistome, transcriptome, and antimicrobial peptides, Front. Microbiol. 4 (2013) 412, http://dx.doi.org/10.3389/fmicb.2013.00412.

[8] A. Gautam, K. Chaudhary, S. Singh, A. Joshi, P. Anand, A. Tuknait, et al., Hemolytik: a database of experimentally determined hemolytic and non-hemolytic peptides, Nucleic Acids Res. (2013) 1–6, http://dx.doi.org/10.1093/nar/gkt1008.

[9] N.F. Azevedo, S.P. Lopes, C.W. Keevil, M.O. Pereira, M.J. Vieira, Time to "go large" on biofilm research: advantages of an omics approach, Biotechnol. Lett. 31 (2009) 477–485, http://dx.doi.org/10.1007/s10529-008-9901-4.

[10] A. Lourenço, T. Coenye, D.M. Goeres, G. Donelli, A.S. Azevedo, H. Ceri, et al., Minimum information about a biofilm experiment (MIABiE): standards for reporting experiments and data on sessile microbial communities living at interfaces, Pathog. Dis. 70 (2014) 250–256, http://dx.doi.org/10.1111/2049-632X.12146.

[11] A. Lourenço, A. Ferreira, N. Veiga, I. Machado, M.O. Pereira, N.F. Azevedo, BiofOmics: a Web platform for the systematic and standardized collection of high-throughput biofilm data, PLOS ONE 7 (2012) e39960, http://dx.doi.org/ 10.1371/journal.pone.0039960.

[12] A.M. Sousa, A. Ferreira, N.F. Azevedo, M.O. Pereira, A. Lourenço, Computational approaches to standard-compliant biofilm data for reliable analysis and integration, J. Integr. Bioinform. 9 (2012) 203, http://dx.doi.org/10.2390/biecoll-jib-2012-203.

[13] G.P. Rodríguez, D. Glez-Peña, N.F. Azevedo, M.O. Pereira, F. Fdez-Riverola, A. Lourenço, BEW: Bioinformatics Workbench for Analysis of Biofilms

Experimental Data, in: 8th Int. Conf. Pract. Appl. Comput. Biol. Bioinforma (PACBB 2014), Adv. Intell. Syst. Comput. 294 (2014) 49–56, http://dx.doi.org/ 10.1007/978-3-319-07581-5 6.

[14] A. Brazma, P. Hingamp, J. Quackenbush, G. Sherlock, P. Spellman, C. Stoeckert, et al., Minimum information about a microarray experiment (MIAME)-toward standards for microarray data, Nat. Genet. 29 (2001) 365–371, http://dx.doi.org/10.1038/ng1201-365.

[15] C.F. Taylor, N.W. Paton, K.S. Lilley, P.-A. Binz, R.K. Julian, A.R. Jones, et al., The minimum information about a proteomics experiment (MIAPE), Nat. Biotechnol. 25 (2007) 887–893, http://dx.doi.org/10.1038/nbt1329.

[16] E.W. Deutsch, C.A. Ball, J.J. Berman, G.S. Bova, A. Brazma, R.E. Bumgarner, et al., Minimum information specification for in situ hybridization and immunohistochemistry experiments (MISFISHIE), Nat. Biotechnol. 26 (2008) 305–312, http://dx.doi.org/10.1038/nbt1391.

[17] B.A. Schaefer, D. Poetz, G.W. Kramer, Documenting laboratory workflows using the Analytical Information Markup Language, JALA 9 (2004) 375–381.

[18] T. Bray, J. Paoli, C.M. Sperberg-McQueen, Extensible markup language (XML), 1998.

[19] D. Glez-Peñ a, M. Reboiro-Jato, P. Maia, M. Rocha, F. Díaz, F. Fdez-Riverola, AIBench: a rapid application development framework for translational research in biomedicine, Comput. Methods Programs Biomed. 98 (2010) 191–203, http://dx.doi.org/10.1016/j.cmpb.2009.12.003.

[20] F. Fdez-Riverola, D. Glez-Peñ a, H. López-Fernández, M. Reboiro-Jato, J.R. Méndez, Java application framework for scientific software development, Softw. Pract. Exp. 42 (2012) 1015–1036, http://dx.doi.org/10.1002/spe.1108.

[21] NCBI Resource Coordinators, Database resources of the National Center for Biotechnology Information, Nucleic Acids Res. 41 (2013) D8–D20, http://dx.doi.org/10.1093/ nar/gks1189.

[22] C. Knox, V. Law, T. Jewison, P. Liu, S. Ly, A. Frolkis, et al., DrugBank 3.0: a comprehensive resource for "omics" research on drugs, Nucleic Acids Res. 39 (2011) D1035–D1041, http://dx.doi.org/10.1093/nar/gkq1126.

[23] P. Taylor, P. Jorge, A. Lourenço, M.O. Pereira, New trends in peptide-based anti-biofilm strategies: a review of recent achievements and bioinformatic approaches, Biofouling 28 (2012) 1033–1061, http://dx.doi.org/10.1080/08927014. 2012.728210.

(a)
```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<ns0:bml xmlns:ns0="http://www.example.org/bml" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.example.org/bml http://miabie.org/xml/bml.xsd">
  <experiment bioID="bio_1" contact="nazevedo@fe.up.pt" date="6/5/11"
   experimentName="Adhesion of water stressed Helicobacter pylori to"
   organization="Laboratory for Process, Environmental and Energy Engineering, University of Porto"
   publication="16907822">
    <authors>
        <![CDATA[Nuno  Azevedo]]>
    </authors>
    <notes>
            <![CDATA[The main aim of this work was to study...]]>
    </notes>
    <methods>
        # Method list for this experiment
    </methods>
    <constantConditions>
      <constantCondition condition="organism" conditionUnits="-" conditionValue="*helicobacter pylori*"/>
      # More constant conditions...
    </constantConditions>
  </experiment>
</ns0:bml>
```

(b)
```
<methods>
    <method methodName="CFU" methodUnits="cells cm-2">
      <conditions>
        <condition conditionName="time" conditionUnits="NaN">2#6#13#24#...</condition>
        <condition conditionName="material" conditionUnits="NaN">*stainless steel*#...</condition>
        # More conditions
      </conditions>
      <dataSeriesSet>
        <dataSerie>
          <conditionValues>2#*stainless steel*</conditionValues>
          <measurements>120000#40000#0#280000#...</measurements>
        </dataSerie>
        # More dataSeries
      </dataSeriesSet>
    </method>
</methods>
```

Fig. 1 – Structure of a biofilm experiment represented in BML format, namely the description of (a) all possible top-level elements and (b) the methods data.

Fig. 2 – The use of BiofOmics Web services for data submission and download. Examples of data submission and http petition/response, where (a) the user uploads the experiment named "Adhesion of water stressed Helicobacter pylori to abiotic surfaces" to BiofOmics, and (b) the user searches BiofOmics and downloads the experiment named "Adhesion influence of water-stressed Helicobacter pylori to SS304 and polypropylene".
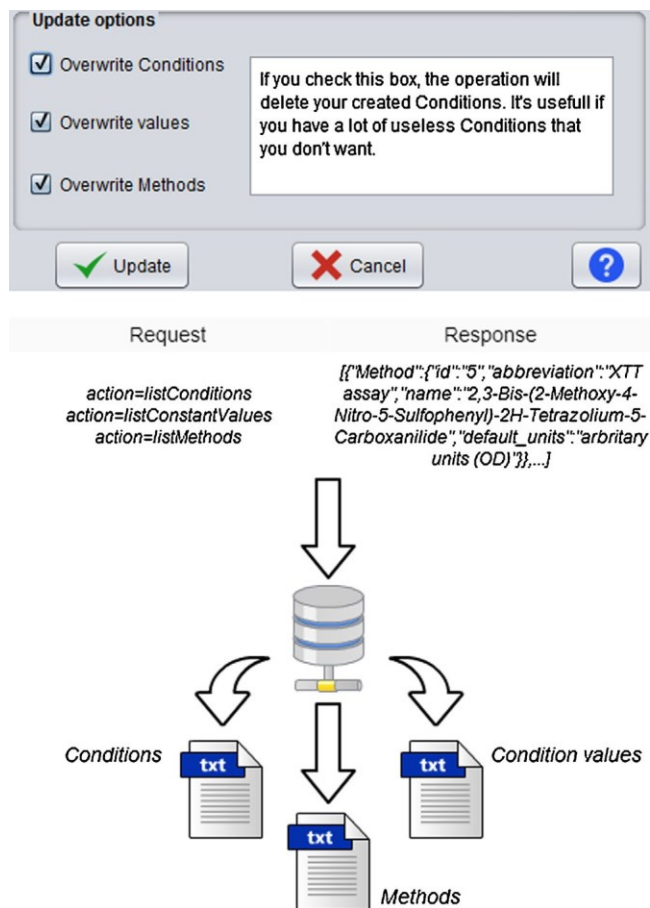
Fig. 3 − The use of BiofOmics Web services to refresh metadata. User may download vocabulary on conditions, condition values and methods of analysis.

```xml
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<ns0:bml xmlns:ns0="http://www.example.org/bml" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.example.org/bml http://miabie.org/xml/bml.xsd">
  <experiment bioID="bio_1" contact="nazevedo@fe.up.pt" date="6/5/11"
  experimentName="Adhesion of water stressed Helicobacter pylori to"
  organization="Laboratory for Process, Environmental and Energy Engineering, University of Porto"
  publication="16907822">
   <authors><![CDATA[Nuno  Azevedo]]></authors>
   <notes><![CDATA[The main aim of this work was to study...]]></notes>
   <methods>
     <method methodName="CFU" methodUnits="cellsÂ cmâ'2">
       <conditions>
         <condition conditionName="time" conditionUnits="NaN">2#6#13#24#48#96#192</condition>
         <condition conditionName="material" conditionUnits="NaN">
             *stainless steel grade ss304*
             #*stainless steel grade ss316*#copper#polyvinyl chloride#*polipropilene*#glass
             #*stainless steel grade ss304*#*stainless steel grade ss316*#copper#polyvinyl chloride
             #*polipropilene*#glass#*stainless steel grade ss304*#*stainless steel grade ss316*#copper
             #polyvinyl chloride#*polipropilene*#glass#*stainless steel grade ss304*
             #*stainless steel grade ss316*#copper#polyvinyl chloride#*polipropilene*#glass
             #*stainless steel grade ss304*#...
         </condition>
       </conditions>
       <dataSeriesSet>
         <dataSerie>
           <conditionValues>2#*stainless steel grade ss304*</conditionValues>
           <measurements>
               120000#40000#0#280000#320000#0#240000
               #200000#360000#0#120000#160000#200000#240000#40000
               #80000#200000#320000#160000#0#0#240000#400000#120000
               #120000#280000#40000#320000#120000#160000#320000#80000
               #0#120000#200000#240000#0#0#200000#200000#120000#80000
               #200000#160000#200000#200000#120000#80000#240000#320000
               #0#NaN#NaN#NaN#NaN#NaN#NaN#NaN#NaN#NaN
           </measurements>
         </dataSerie>
         # ...
       </dataSeriesSet>
     </method>
   </methods>
   <constantConditions>
     <constantCondition condition="organism" conditionUnits="-" conditionValue="*helicobacter pylori*"/>
     <constantCondition condition="device" conditionUnits="-"
     conditionValue="6-well microtiter plate with coupons"/>
     # ...
   </constantConditions>
  </experiment>
</ns0:bml>
```
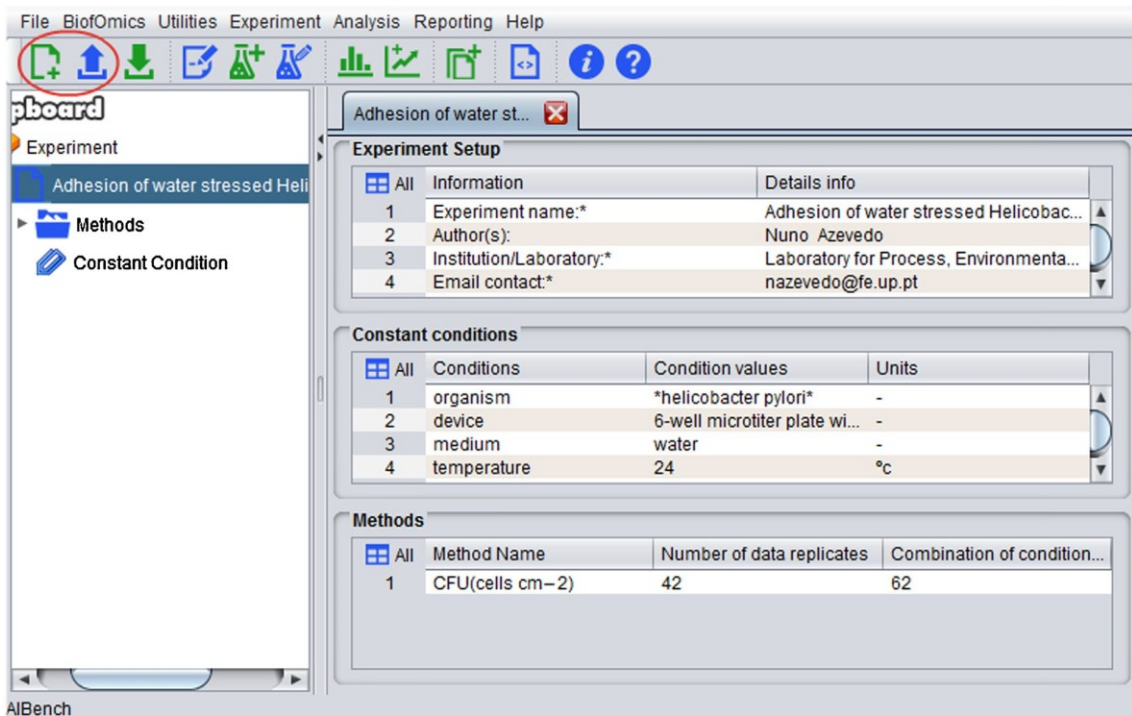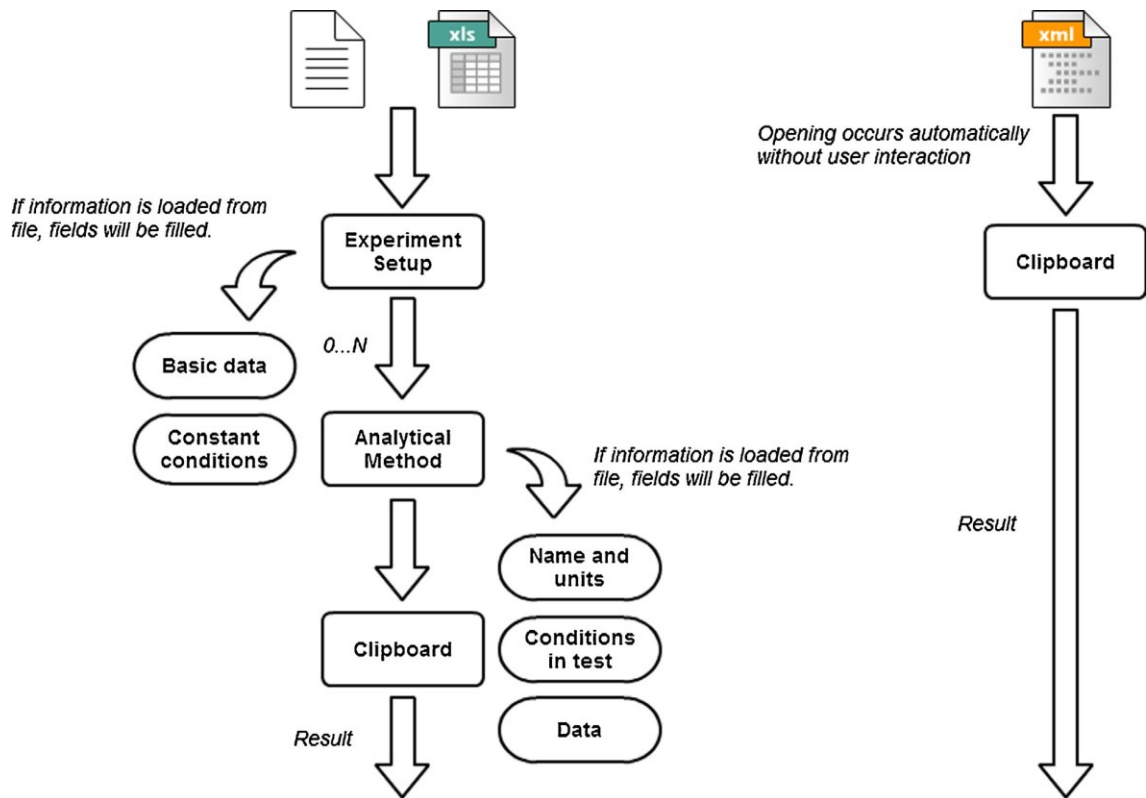
Fig. 4 − BML structure for the case study.

Fig. 5 – Experiment management in BEW: (left) steps to create or import an Experiment from XLS; (right) steps to load an Experiment from XML.
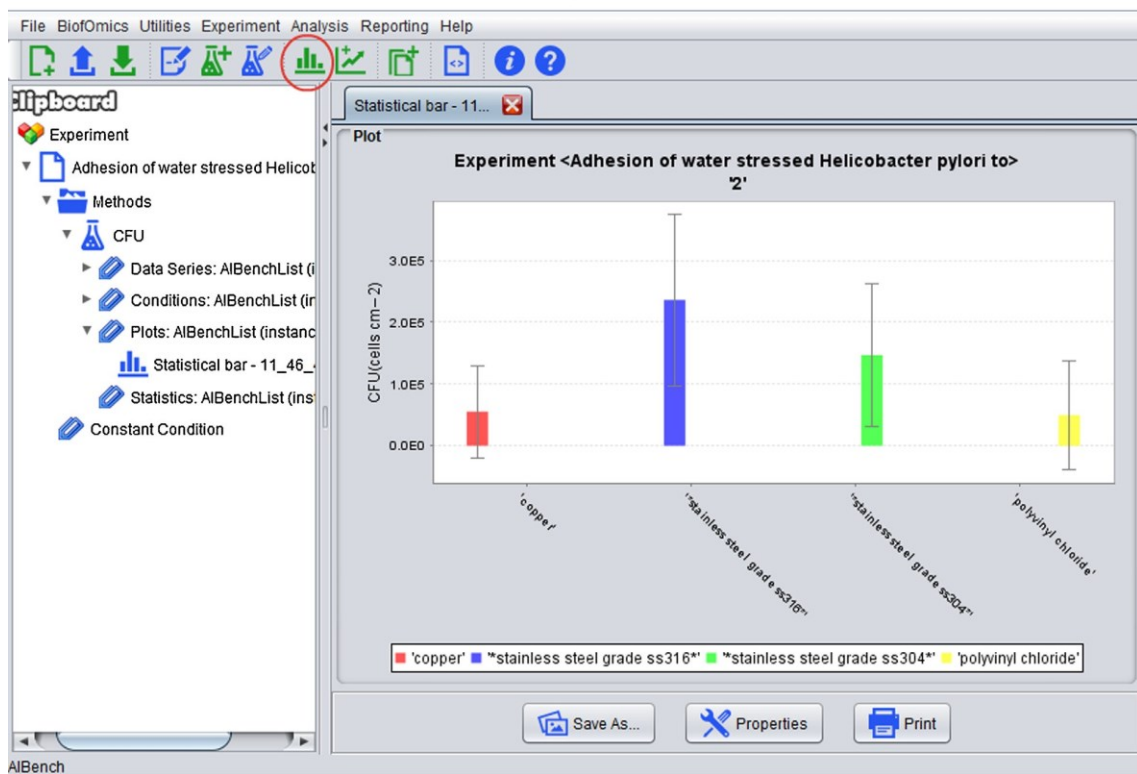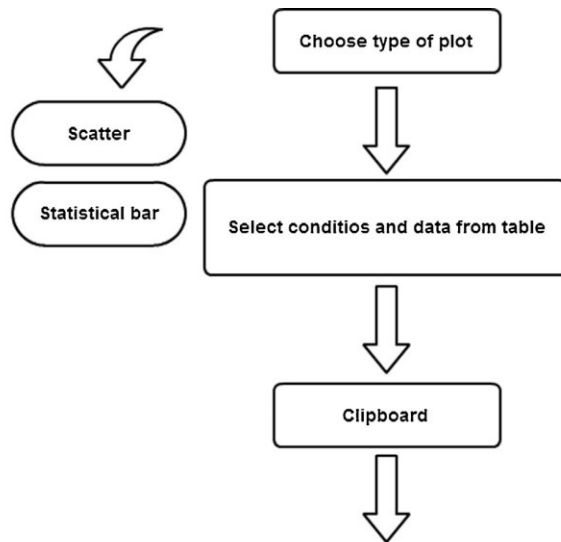
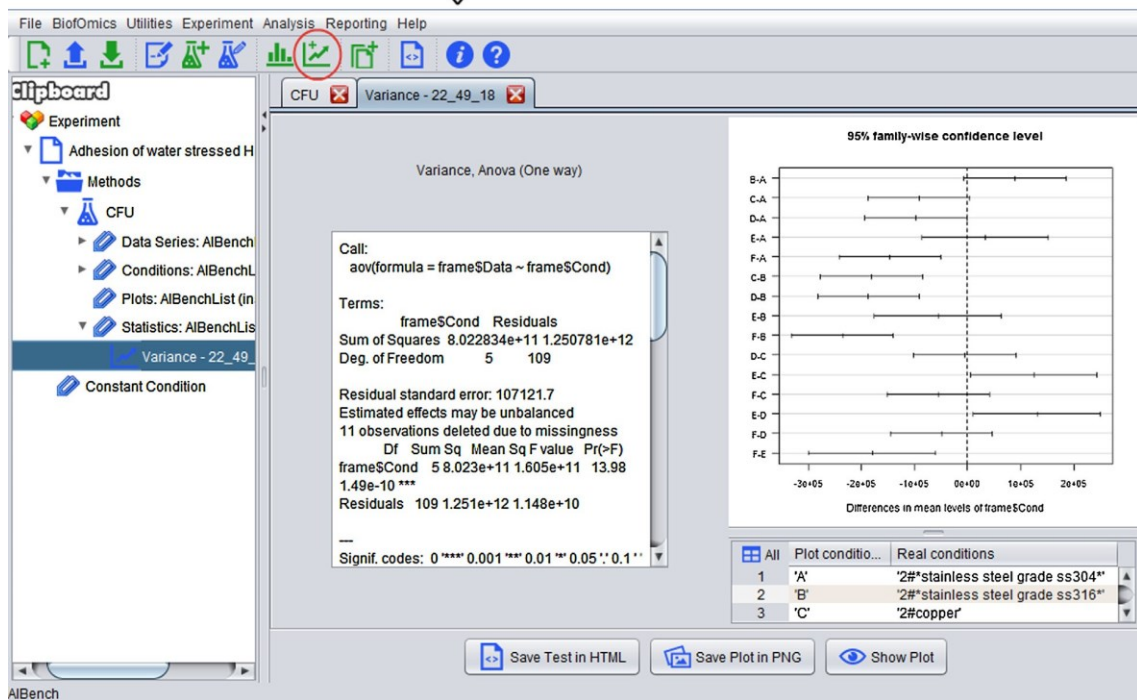*Fig. 6 – Example of the construction of 2D plots in BEW.*
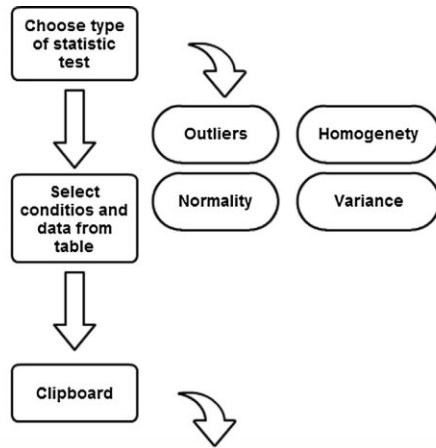
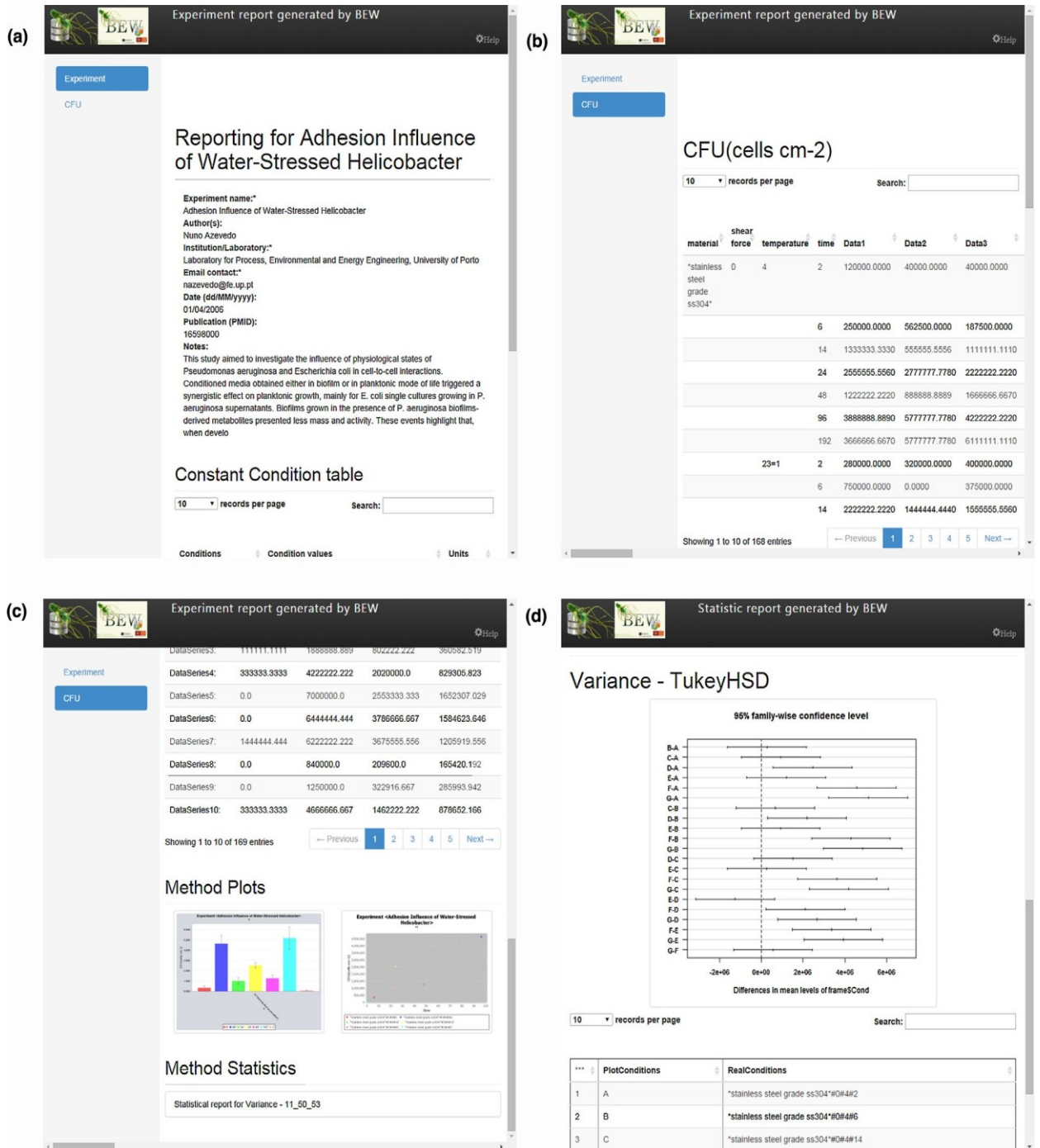*Fig. 7 – Example of the execution of statistical analyses in BEW.*

Fig. 8 – Snapshot of a Web report automatically generated by BEW. The first page describes the experiment in terms of general purpose and authors, and the methods of analysis employed (a). For each method of analysis, data is presented (b), and may be complemented by associated plots (c) and statistical tests (d).

Table 1 – Some of the cases studies that supported BEW development. Data statistics are merely illustrative of the number of methods of analysis (# of methods) involved in the study, the number of conditions evaluated (# of tested conditions) and the average number of data points collected (avg # data points).

PubMed identifier	Short description of the experiment	Data statistics
16907822	Adhesion of water stressed *Helicobacter pylori* to abiotic surfaces	# of methods: 1 # of tested conditions: 2 Avg # data points: 42
16598000	Adhesion influence of water-stressed *Helicobacter pylori* to SS304 and polypropylene	# of methods: 1 # of tested conditions: 4 Avg # data points: 168
22770521	Antibiotic resistance of mixed biofilms in cystic fibrosis the impact of emerging microorganisms on infection treatment	# of methods: 6 # of tested conditions: 3 Avg # data points: 45
21193844	Casbane diterpene as a promising natural antimicrobial agent against biofilm-associated infections	# of methods: 1 # of tested conditions: 2 Avg # data points: 70
21107231	Farnesol as antibiotics adjuvant in *Staphylococcus epidermidis* control in vitro	# of methods: 1 # of tested conditions: 2 Avg # data points: 14
22313289	Silver colloidal nanoparticles: antifungal effect against adhered cells and biofilms of *Candida albicans* and *Candida glabrata*	# of methods: 2 # of tested conditions: 2 Avg # data points: 108
20401483	In vitro biofilm activity of non-*Candida albicans* species	# of methods: 2 # of tested conditions: 1 Avg # data points: 21
21479268	Discriminating multi-species populations in biofilms with peptide nucleic acid fluorescence in situ hybridisation (PNA FISH)	# of methods: 4 # of tested conditions: 3 Avg # data points: 219
22313289	Silver nanoparticles: influence of stabilising agent and diameter on antifungal activity against *Candida albicans* and *Candida glabrata* biofilms	# of methods: 2 # of tested conditions: 2 Avg # data points: 44
22354293	The involvement of sphingolipid biosynthetic intermediates in governing the miconazole resistance of sessile *Saccharomyces cerevisiae* cells	# of methods: 2 # of tested conditions: 3 Avg # data points: 90
18642340	Relationship between mass transfer coefficient and liquid flow velocity in heterogenous biofilms using microelectrodes and confocal microscopy	# of methods: 2 # of tested conditions: 3 Avg # data points: 1036
–	Emergent bacteria in cystic fibrosis: in vitro biofilm formation and resilience under variable oxygen conditions	# of methods: 7 # of tested conditions: 3 Avg # data points: 30
21388333	*Listeria monocytogenes* and *Salmonella enterica enteritidis* biofilms susceptibility to different disinfectants and stress-response and virulence gene expression of surviving cells	# of methods: 2 # of tested conditions: 3 Avg # data points: 166