Computational approaches to standard-compliant biofilm data for reliable analysis and integration

Ana Margarida Sousa[1], Andreia Ferreira[1], Nuno F. Azevedo[2], Maria Olivia Pereira[1], Anália Lourenço[1*]

[1]IBB - Institute for Biotechnology and Bioengineering, Centre of Biological Engineering, University of Minho, Campus de Gualtar, 4710-057 Braga – PORTUGAL

[2]LEPAE, Department of Chemical Engineering, Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

Summary

The study of microorganism consortia, also known as biofilms, is associated to a number of applications in biotechnology, ecotechnology and clinical domains. Nowadays, biofilm studies are heterogeneous and data-intensive, encompassing different levels of analysis. Computational modelling of biofilm studies has become thus a requirement to make sense of these vast and ever-expanding biofilm data volumes.

The rationale of the present work is a machine-readable format for representing biofilm studies and supporting biofilm data interchange and data integration. This format is supported by the Biofilm Science Ontology (BSO), the first ontology on biofilms information. The ontology is decomposed into a number of areas of interest, namely: the Experimental Procedure Ontology (EPO) which describes biofilm experimental procedures; the Colony Morphology Ontology (CMO) which characterises morphologically microorganism colonies; and other modules concerning biofilm phenotype, antimicrobial susceptibility and virulence traits. The overall objective behind BSO is to develop semantic resources to capture, represent and share data on biofilms and related experiments in a regularized fashion manner. Furthermore, the present work also introduces a framework in assistance of biofilm data interchange and analysis – BiofOmics (http://biofomics.org) – and a public repository on colony morphology signatures – MorphoCol (http://stardust.deb.uminho.pt/morphocol).

1 Introduction

Microorganisms have evolved various strategies to survive and adapt to the ever changing environmental conditions. The formation of biofilms is an example of such adaptation strategies. Biofilms are structured and complex sessile communities of microorganisms that are able to survive virtually everywhere in Nature because of their ability to adhere to a surface and embed in a protecting, self-produced matrix of extracellular polymeric substances [1-2].

Due to their persistence and resistance to antimicrobial agents, biofilms cause a variety of problems in different areas of great importance to human development, such as the clinical, industrial and environmental settings. Biofilm-growing microorganisms affect, for instance, hygiene and food safety in the food industry [3], are responsible for nosocomial infections [4-7], acute and chronic infections [8-10], and clogging and contaminations in drinking water systems [11-12]. In turn, biofilms play an important role in the ecological balance and can be "engineered" to carry out beneficial tasks in several biotechnological and bioengineering

* To whom correspondence should be addressed. Email: analia@uvigo.es

processes, such as water and wastewater treatment, bioremediation and production of biocompounds in biofilm reactors [7,8].

Research in this field has a long trajectory, and similarly to other domains, biofilm research has benefited from the technological evolution occurred in the last decades [13-14]. The development of high-throughput biofilm-forming devices (e.g. the 96-well plate, the microtiter plate with coupons and the Calgary device) has enabled the simultaneous testing of large sets of conditions. In addition, biofilm studies involve technology and knowledge from multidisciplinary science fields. The implementation of automated spectrophotometry (e.g. microplate readers) and microscopy systems (e.g. scanning electron, atomic force, and confocal laser scanning microscopy) has empowered the large scale analysis of biofilm features, such as biofilm biomass, biofilm activity and microbial composition. The “omics” platforms also support the study of the transcriptome [15-16], proteome [17-19] and metabolome [20-21] of biofilms.

Biologists are increasingly recognising that computational modelling is crucial for data interchange and for making sense of the vast quantities of complex experimental biofilm data that are now being collected. The current inability to exchange biofilm data has its roots in the limited access to existing data and the lack of a common format for describing biofilm experiments and the results obtained [22]. This work is the first that addresses the standardisation of biofilm studies and proposes a machine-readable format for their representation.

After this introduction, the paper is organised as follows. Section 2 identifies the main challenges to be faced in biofilm data modelling. Section 3 presents the controlled vocabulary and the machine-readable format that have been created to meet modelling and processing requirements. Section 4 introduces the computational resources that have been developed to support biofilm data interchange and data integration. Section 5 resumes the rationale of the initiative and delineates future lines of work. Some conclusions and acknowledgments to researchers supporting the initiative can be found at the end of the paper.Heading for a subsection

2 Biofilm Data Modelling/Management

Even for a well-defined research domain, an universal language that fits all purposes and interests is very hard to be achieved. Nowadays, biofilm studies often encompass multidisciplinary approaches from biology, chemistry, medicine, material science and engineering, among others. Furthermore, biofilm experiments vary greatly in terms of tested conditions and methods of analysis used. Nevertheless, it is reasonable to assume that studies on similar subjects (e.g. stress response, antimicrobial susceptibility testing) should share elements of information - a set of minimum information. Consequently, a common intermediate format could be established to enable communication of the most essential aspects of the biofilm related studies.

2.1 Access to Biofilm Data

Most biomedical databases are populated by curating scientific literature and deriving data from secondary sources (e.g. sequence databases). Biofilms domain is different though, as biofilm publications only summarise the obtained results (mostly general statistics, such as mean and standard deviation values). Raw data is not provided in supplementary material or submitted to any public location. The source of biofilm data is the private archives of researchers. Furthermore, biofilm data files do not comply with a standard format. Data files vary widely from laboratory to laboratory, from researcher to researcher and even from a

researcher's experiment to the next (Figure 1). Data files are organised on an *ad hoc*, as-needed basis and lack comprehensive documentation on the experimental conditions evaluated in each biofilm study (often they are only mentioned by abbreviations or non-documented mnemonics) and, more important, on data quantification (e.g. a simple matter such as the units of measure used). Data curation is thus very hard to achieve without the help of researchers, even for experienced curators.

Fig. 1: Examples of original data files from already published studies.

So, for now, data access is dependent on the willingness of researchers to participate in interchange initiatives. It is necessary to raise awareness of the importance (and benefits) that data interchange can represent to individual studies while promoting collaborative studies and the harmonization of procedures across laboratories. Indeed, the authors of the present work truly believe that data interchange should be proposed as a community initiative and the discussion of such initiative should involve as many researchers as possible. Dissemination activities, for example in biofilm-related conferences, and discussion forums, are considered important purposes to bring forward researchers' participation.

2.2 Characterization of Biofilm Studies

As referred, biofilm studies vary greatly in terms of experimental methodology and analysis purposes. Nowadays, most studies take advantage of high-throughput technologies to test a large number of conditions simultaneously and acquire different data about biofilm "behaviour".

Many *in vitro* systems have been developed for developing and testing microbial biofilms. These systems include simple batch/static systems, batch systems with induced shear, flow cells and systems operated under continuous-flow conditions (e.g. rotating-disk reactor) that

generally provide a surface that can be removed and examined once it is colonised to assess biofilm formation. In recent years, microtiter-plate methods began to be used to grow and assess biofilms in large-scale [10,11]. Crystal violet staining and direct quantification by optical density were initially part of these new systems but, due to its larger data generation ability, the application of colorimetric assays has expanded in the meantime. Assays are now able to assess not only biofilm formation as a whole but also to quantify microbial numbers using, for instance, the Syto 9 assay; microbial physiological activity by the fluorescein diacetate assay; or even the extracellular biofilm matrix using the dimethyl methylene blue assay [12]. For all microtiter plate-based assays, the final results are based either on colour or fluorescence intensity at a certain wavelength, which means that rapid, quantitative analyses are obtained from a single equipment such as an automated multiscan reader.

To define a common intermediate format to exchange such heterogeneous biofilm data and account for multiple analysis goals (and implicitly, results interpretation), it is needed to define a set of minimum information – a signature – that reflects the most essential aspects of the biofilm studies. It is important to emphasise that results are fully comparable only for similar methods under identical conditions. Moreover, it is necessary to compile terminology about these elements of information such that descriptions are unambiguous and well-documented.

A biofilm signature should include information about the experiments but also “credits” information. The signature should identify the team that developed the study, a summary of its main findings (produced by the authors) and eventually other notes of interest, and the reference to any peer-reviewed publications that credit the quality and importance of the study. This information is important at different levels: it credits the authorship of the data made public, authors/submitters are accounted responsible for the quality of the overall signature, and researchers are provided first-hand notes (and other useful insights) of the experiments.

The information about the biofilm experiments depends heavily on the purpose of the study. Due to research interests of the authors of this work and their expertise in the field, this present work has focused on the documentation of studies about antimicrobial susceptibility, biofilm adhesion to abiotic surfaces and biofilm stress response. The set of minimum information that was considered essential to the documentation of these studies was as follows:

- the microorganisms composing the biofilm consortia and their sources (e.g. mutants, clinical isolates, reference strains);
- the experimental setup, including biofilm-forming devices, growth media and adhesion materials;
- the antimicrobial products tested, specifying the concentration and time of application;
- the analytical methods used, including data processing (e.g. the calculation of dilution rates or log reductions) and statistical validation (e.g. number of replicates and reproductions, negative and positive controls, etc.).

2.3 A Machine-readable Format for Biofilm Data

Standardizing biofilm data on a common format is essential for being able to move forward with large-scale collaborative analysis. It removes any impediment to sharing results and permits other researchers to reproduce the experiment, examine it carefully, propose extensions, and apply new techniques and analysis.

The definition of a common format to standardise biofilm data should take into account the costs associated to the curation of existing data. It is important to relieve researchers from the

need to acquire computational skills and/or adapt to new tools, and focus their attention on the curation of the large volume of existing data. Because many of the researchers working in the biofilm area use the familiar Microsoft Excel worksheets to store data from biofilm experiments, it has been decided to adopt this format as a first mean to implement the biofilm data interchange format. Experimental conditions are distributed into worksheets (Figure 2). Single value conditions (i.e. the conditions that remain constant in the experiment) are at one worksheet. Then, for each method of analysis used, one worksheet is created. In these worksheets, multi-value conditions (i.e. the conditions that have several tests) are organised hierarchically in descendent order. This hierarchical structure is allowed to grow vertically – as many sub-levels as condition tests and experiment reproductions there are – and no restriction is imposed to the length of the data series of each-level (i.e. the number of replicates for each condition test).

```
Input: E_cond: list of experimental conditions; M_a: list of
methods of analysis
Output: F_excel customised file

E_cond_single ← singleValueConditions(E_cond)
E_cond_multi ← E_cond − E_cond_single
w_s ← initialiseWorksheet('experiment constants')
for c ∈ sortByType(E_cond_single)
      w_s ← fillIn(c.name, c.value)
end
                              F_excel ← addWorksheet(w_s)

for m ∈ M_a
   w_m ← initialiseWorksheet(m.name)
   for c ∈ sortDescendNumValues(E_cond_multi)
     w_m ← fillInHierarchy(c.name, c.values)
   end
   F_excel ← addWorksheet(w_m)
end

return F_excel
```

Fig. 2: Pseudo-code of the systematic construction of biofilm structured data files.

3 Computational Resources in Support of Biofilm Studies

3.1 Controlled vocabularies and metadata annotations

To effectively search and analyse biofilm data, data should be well organized and semantically integrated. An important tool for searching, integrating and analysing data from databases is record metadata annotation, supported by controlled vocabulary. Hence, it was developed the first ontology on biofilms and their study – the Biofilm Science Ontology (BSO). BSO is a neutral species ontology, developed following the basic principles of the OBO Foundry [23]. BSO is designed to develop shared, structured and accurated vocabularies for the annotation of the general biofilm experimental workflow. BSO is intended for the broad research community, including bench microbiologists, clinical researchers, clinicians, curators and bioinformaticians. Structure and definitions are freely accessible at OBO foundry Web site (http://miabie.org/ontology). Definitions of BSO terms are provided from actual knowledge and consensual statements about biofilms obtained from reference works and discussed with well-known experts in the field.

BSO is modular, i.e., is composed by specialized ontologies covering all areas of biofilm studies. The Experimental Procedure Ontology (EPO) and the Colony Morphology Ontology (CMO) were two of the first modules that are being fully developed. EPO can be used to

annotate descriptions for biofilm experiments workflow, namely biofilm formation and all operating procedures involved (Figure 3).

CMO is one of the well-established modules of BSO due to the emergent significance of colony morphology observations in the understanding of antibiotic resistance, persistence and virulence [24]. Colony morphologies have been intensively studied due to its medical importance, without internationally accepted guidelines for colony description, classification or even designation. Besides, the main source of information about colony morphology variation is the peer-reviewed journal literature. Therefore, CMO was developed to consistently describe morphological traits of colony bacteria (Figure 4), as well as to enable data accessibility, use and understanding. CMO attempts to capture the precise meaning of colony morphology terms avoiding variability, heterogeneity and ambiguity. The retrieval, integration and analysis of colony morphology data will become undoubtedly easier using universal and unquivocal ontologies that have to be independent of colony morphotypes, laboratories, researchers or even the experimental protocols used.

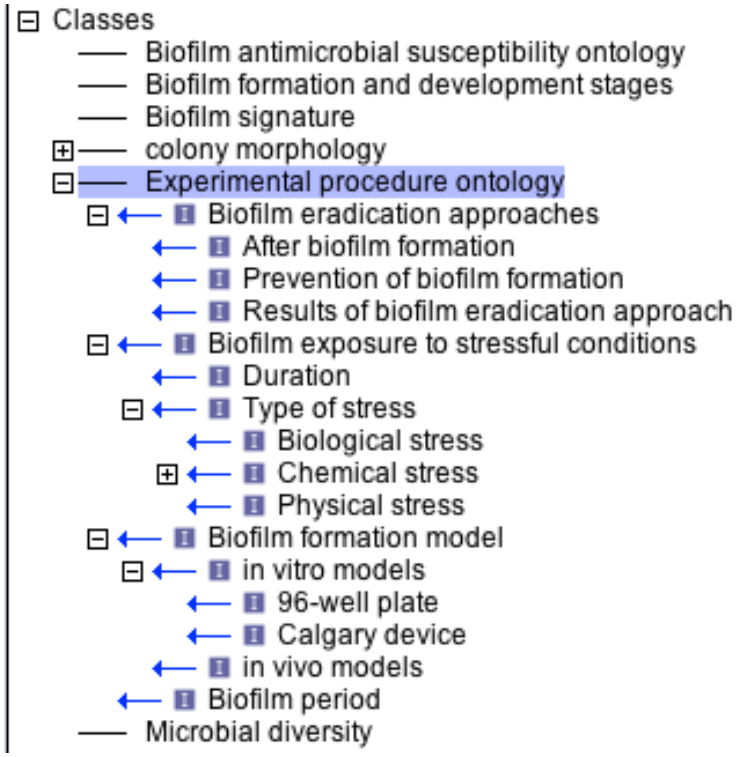

Fig. 3: General tree view of EPO within BSO structure.

Other modules of BSO are currently under development, namely the ontology of biofilm signature, antimicrobial resistance and virulence traits of biofilm cells. Such ontologies represent the knowledge underlying description of biofilm composition and architecture, biofilm-cell phenotype and biofilm matrix.

Through BSO using, biofilm data become interoperable and consistent because of its unambiguously description, search, integration and analysis. In addition, standardization can readily transform the biofilm research and applications and promote the integration of biofilms in other science areas.

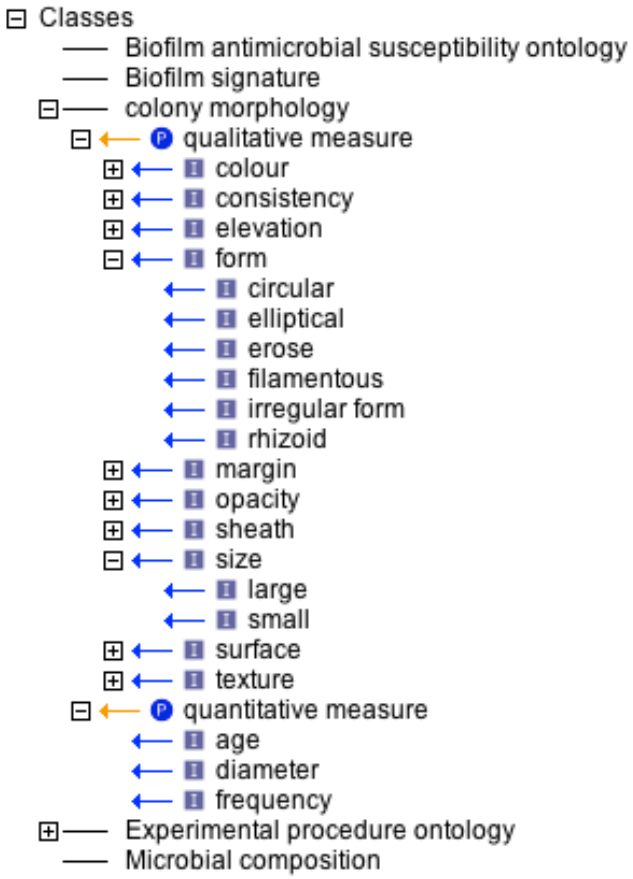

Fig. 4: General tree view of CMO insert in BSO structure.

3.2 BiofOmics Framework: Means to Compare and Interchange Data

BiofOmics is a novel Web framework that aims to make readily available biofilm studies and enable comparative and collaborative studies. It is meant to provide accommodation to existing data, but far more important, ensure data standardization.

The existence of a database compiling existing biofilm data in a computer-amenable way simplifies research in a number of ways: the search for similar experiments (Figure 5), the interchange of data between researchers and laboratories, the search for "open spots" (i.e. relevant but under-reported areas), and the comparative analysis of experiments (in particular, inter-laboratory collaborations). Besides the obvious value of widespread dissemination of biofilm research, researchers are also rewarded with the possibility to ameliorate the supplementary materials accompanying publications; a (major) step forward to endorse the transparency and high-quality of biofilm experimental data as well as the validity of the results and discussion being published.

The BiofOmics platform is already operative. Its facilities were validated through the introduction of existing experiments, representative of key research topics (and purposes of analysis) and involving latest high-throughput technologies and methodologies. As an example, experiments dealing with the resistance of biofilms to antimicrobial agents typically have data from biofilms before, during and after exposure to the agents. In this scenario, researchers will be most likely interested in comparing the values collected and for this purpose, values have to be adequately characterized in the database. Owing to the amount of information involved in high-throughput biofilm experiments, the number of data points involved in the tests exceeded 10000.

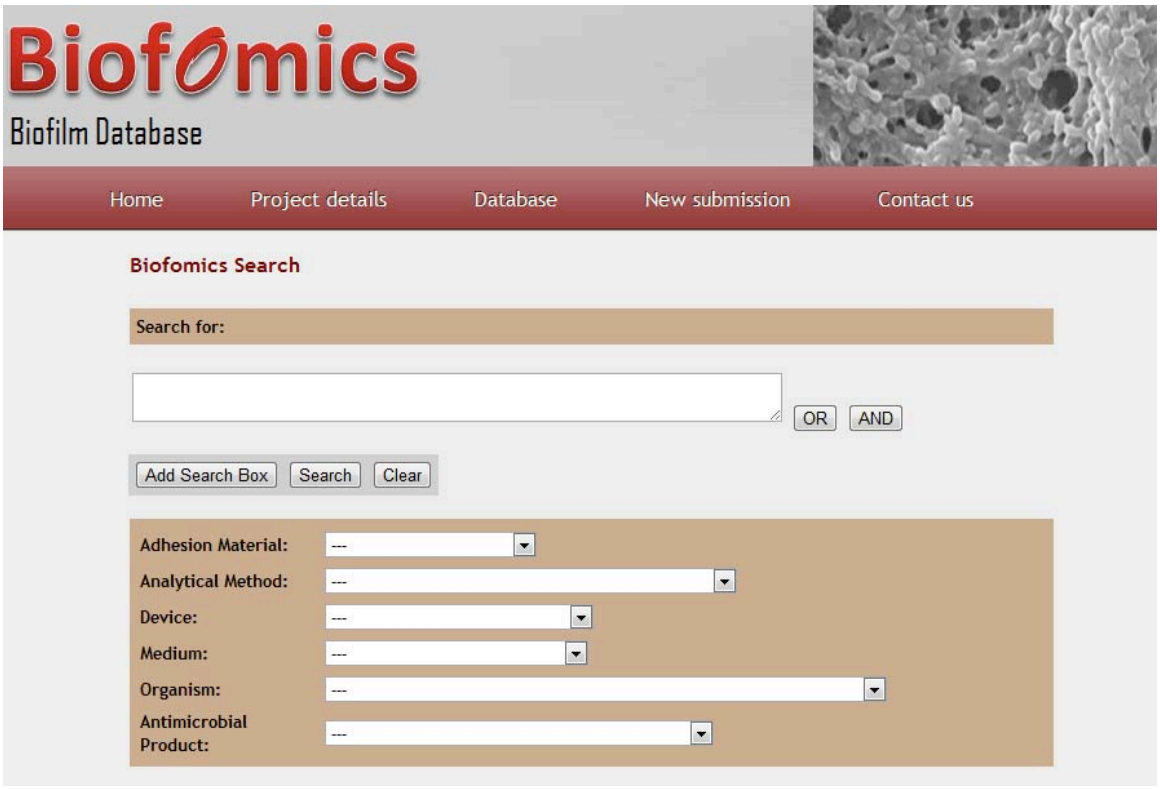

Fig. 5: BiofOmics facility for public biofilm data search.

3.3 MorphoColDB: a Database on Colony Morphotypes

MorphoColDB aims to be a comprehensible knowledge base on colony morphologies, i.e., the morphological characterisation of microorganism consortia under various environmental conditions and stages of their formation and development. Some morphological features of colonies are linked to important biological processes, such as antibiotic resistance [24] and microbial virulence. For instance, small colony variants of some bacteria are strongly associated to resistance to a broad range of antibiotics of clinical use. Due to its clinical importance, colony morphology studies are thus ever growing and MorphoCol was developed as an instrument to help the management and manipulation of this fast proliferation of information about colony morphology (Figure 6).

The main source of information about colony morphology variation is the scientific literature. Normally, manuscripts describe the setup of the experiments, as well as the colony morphologies observed. They also provide images about the colony morphologies, which are of obvious interests for observation and comparative purposes, but it is anticipated that most are bound to journal copyright constraints. Therefore, although the data curation workflow established in MorphoColDB is based in the curation of scientific literature (reached out mainly through PubMed), contact with authors is necessary as an additional source of data.

Another important step of this data curation workflow is the systematic and standardised labelling of the colony morphologies. Researchers use various terminologies, often even create new terms, to describe the morphological features observed in their experiments. MorphoColDB annotation process is supported by the CMO, which has gathered and processed these heterogeneous terminologies. Curators are responsible for the annotation of the submitted images. Furthermore, MorphoColDB enables the integration of multiple layers of information, namely colony morphology annotations (based on CMO) and data from transcriptomic, proteomic, and metabolomic analyses.

MorphoColDB offers a way to rationally inquire about bacteria more resistant or/and more virulent and may has a significant impact on medical decision support, namely in design of antibiotic therapy. To begin with, the knowledge base is gathering information on *Pseudomonas aeruginosa* and *Staphylococcus aureus* strains, two pathogenic bacteria commonly associated with biofilm-related nosocomial infections.

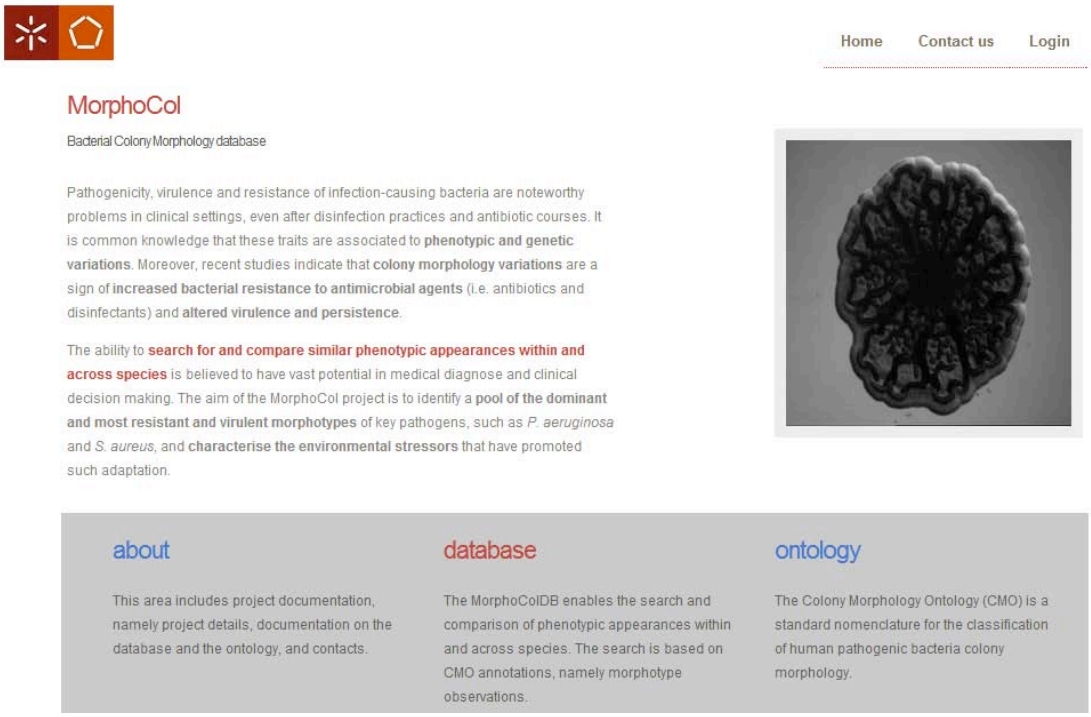

Fig. 6: MorphoCol site on colony morphology.

4 Rationale and On-going Work

In this work, some computational contributions to standardise biofilm data and enable data integration and large-scale decision making are presented.

The common intermediate format introduced with the present work is neutral with respect to programming languages and software encoding. By supporting this format for reading and writing biofilm signatures, different software tools (including programs for building and editing data files, analysis and simulation programs, databases, and other systems) can directly communicate and store the same computable representation of those files. This removes any barrier to sharing results and permits other researchers to start with an unambiguous representation of the data, examine it carefully, and apply new techniques and approaches (i.e. collaborative and “incremental” studies).

The frameworks, BiofOmics and MorphoColDB, are the first to deliver readily access to biofilm studies and encompass such a wide scope of areas of interest and various layers of information. They facilitate the gathering and annotation of the large volumes of biofilm data available. They also simplify and make more explicit the association between biofilm experiments and scientific statements, i.e., the connection between data and work findings, presenting them in a condensed and accessible form. Laboratories can rationalise their work and avoid losing information and know-how by storing data and protocols in homogeneous formats, objectively annotated and understandable; accessible results and experimental,

analysis and statistical protocols can avoid duplication of efforts; accessible results can also control the loss of competences by operator turnover.

The computational efforts developed in the scope of the present work are sustained by BSO-derived ontologies which facilitate the communication between researchers and scientific groups, unequivocally describing biofilm signatures. Indeed, the agreement about sharing semantic definitions is crucial to advance biofilm understanding.

After a number of international conferences, informal meetings, courses and workshops, the computational initiatives introduced by the present work have managed to catalyse a community of interested researchers, developers, and users who are now collaborating on evolving the above mentioned tools and creating new resources around them. This is a clear reflection of an urgent need in the community to address issues of harmonisation and interoperability. At the same time, it is believed that the challenges faced by these projects and the solutions that are arising have underlying components that would be faced by any effort to define a similar standard exchange format.

Indeed, the team is working on the definition of the Minimal Information About Biofilms Experiments – MIABiE (http://miabie.org) – which aims to provide guidelines to the adequate documentation of biofilm-related studies. MIABiE arose from the need to find an adequate and scientifically sound way to control and store the data from biofilm experiments, particularly those involving high-throughput devices. MIABiE is a member of MIBBI (http://mibbi.sourceforge.net/), an initiative that provides a common portal to minimum information checklists from all areas of biological and biomedical sciences.

5 Conclusions

Nowadays, biofilm studies are heterogeneous and data-intensive, encompassing different levels of analysis. The understanding of microorganism consortia would benefit greatly from the ability to integrate data from similar studies and confront distinct levels of analysis. Computational approaches to interchange studies and scale up biofilm analysis are desired. Besides limited access, no protocol exists on how to document biofilm studies, i.e., the minimum information required to guarantee self-contained and explanatory documentation. It does not take prescience to see that infrastructure such as SBML, databases, and more powerful analysis tools are needed to support continued progress in biofilm science.

The high-throughput technologies accelerated the rate of generating data on biofilm domains. As aforementioned, biofilm research relies increasingly on large collections of data sets. This “big data” dimension calls for the development of novel computational tools, specialised in biofilm data management, interchange and analysis. However, the authors of the present work are unaware of any efforts to standardise and disseminate biofilm data at large scale. Currently, sharing of biofilm data among researchers is poor at best, in detriment of research and community at large.

Access to primary high-throughput biofilm data alone is useless unless those data would be in form of accurate analysis and interpretation. Thus, data standardisation is crucial to make knowledge more explicit, help detect errors, ensure data reliability, and promote data interchange. Standardization augments the global value of results and leading to great advance of science knowledge. This advance is not limited to scientific purposes, as industrial interests might gain serious advantages.

The informatics approaches presented here to the standardised organisation of biofilm data demonstrated to be critically important for data sharing, unambiguous representation, validation and interpretation of data, semantic search and query, and data integration. In

addition, they ensure data quality, reliability and reproducibility. All approaches have been validated with a number of highly variable, already published experiments and it is already in practice, supporting the operation of the BiofOmics and MorphoCol databases. These databases facilitates data search and comparison as well as data interchange between laboratories (publicly accessible at http://biofomics.org and http://morphocol.org, respectively). All methods applied will certainly enable the successful application of biofilm researchers on clinical, medical and industrial field.

Acknowledgements

The authors thank, among others, Rosário Oliveira, Maria João Vieira, Idalina Machado, Nuno Cerca, Mariana Henriques, Pilar Teixeira, Douglas Monteiro, Melissa Negri, Susana Lopes, Carina Almeida, Alexandrina Rodrigues, Tom Coenye, Paul Stoodley and Hélder Lopes, for submitting their data and participating in the aforementioned standardisation initiatives. The financial support from the Institute of Biotechnology and Bioengineering - Center of Biological Engineering (IBB-CEB), the Fundação para a Ciência e Tecnologia (FCT) and European Community fund FEDER (Program COMPETE), projects PTDC/SAU-ESA/646091/2006/ and FCOMP-01-0124-FEDER-007480, and the PhD grant of Ana Margarida Sousa (SFRH/BD/31065/2006) are also gratefully acknowledged.

References

[1] McBain AJ. Chapter 4: In vitro biofilm models: an overview. *Adv Appl Microbiol.*, 69:99-132, 2009.

[2] Jain A, Gupta Y, Agrawal R, Khare P, Jain SK. Biofilms-a microbial life perspective: a critical review. *Crit Rev Ther Drug Carrier Syst.*, 24(5):393-443, 2007

[3] Van Houdt R, Michiels CW. Biofilm formation and the food industry, a focus on the bacterial outer surface. *J Appl Microbiol.*, 109(4):1117-31, 2010.

[4] Machado I, Lopes S P, Sousa A M, Pereira MO. Adaptive response of single and binary *Pseudomonas aeruginosa* and *Escherichia coli* biofilms to benzalkonium chloride. *Journal of Basic Microbiology*, 51: 1-10, 2011.

[5] Frei E, Hodgkiss-Harlow K, Rossi PJ, Edmiston CE, Bandyk DF. Microbial Pathogenesis of Bacterial Biofilms: A Causative Factor of Vascular Surgical Site Infection. *Vasc Endovascular Surg.*, 45(8):688-696, 2011.

[6] Rodrigues LR. Inhibition of bacterial adhesion on medical devices. *Adv Exp Med Biol.*, 715:351-67, 2011.

[7] Høiby N, Bjarnsholt T, Givskov M, Molin S, Ciofu O. Antibiotic resistance of bacterial biofilms. *Int J Antimicrob Agents*, 35(4): 322-32, 2010.

[8] Bowler PG, Duerden BI, Armstrong DG. Wound microbiology and associated approaches to wound management. *Clin Microbiol Rev.*, 14(2):244-69, 2001.

[9] Simões M. Antimicrobial strategies effective against infectious bacterial biofilms. *Curr Med Chem.*, 18(14):2129-45, 2011.

[10] Smith A, Buchinsky FJ, Post JC. Eradicating chronic ear, nose, and throat infections: a systematically conducted literature review of advances in biofilm treatment. *Otolaryngol Head Neck Surg.*, 144(3):338-47, 2011.

[11] Gião MS, Azevedo NF, Wilks SA, Vieira MJ and Keevil CW. Persistence of *Helicobacter pylori* in Heterotrophic Drinking-Water Biofilms. *Applied and Environmental Microbiology*, 74(19): 5898-5904, 2008.

[12] Jang HJ, Choi YJ, Ka JO. Effects of diverse water pipe materials on bacterial communities and water quality in the annular reactor. *J Microbiol Biotechnol.*, 21(2):115-23, 2011.

[13] Ramage G, Culshaw S, Jones B, Williams C. Are we any closer to beating the biofilm: novel methods of biofilm control. *Curr Opin Infect Dis.*, 23(6):560-6, 2010.

[14] Cos P, Toté K, Horemans T, Maes L. Biofilms: an extra hurdle for effective antimicrobial therapy. *Curr Pharm Des.*, 16(20):2279-95, 2010.

[15] Wood TK. Insights on *Escherichia coli* biofilm formation and inhibition from whole-transcriptome profiling. *Environ Microbiol.*,11(1):1-15, 2009.

[16] Coenye T. Response of sessile cells to stress: from changes in gene expression to phenotypic adaptation. *FEMS Immunol Med Microbiol*., 59(3):239-52, 2010.

[17] Silva MS, De Souza AA, Takita MA, Labate CA, Machado MA. Analysis of the biofilm proteome of *Xylella fastidiosa*. *Proteome Sci.*, 9:58, 2011.

[18] Jiao Y, D'haeseleer P, Dill BD, Shah M, Verberkmoes NC, Hettich RL, Banfield JF, Thelen MP. Identification of biofilm matrix-associated proteins from an acid mine drainage microbial community. *Appl Environ Microbiol.*, 77(15):5230-7, 2011.

[19] Cabral MP, Soares NC, Aranda J, Parreira JR, Rumbo C, Poza M, Valle J, Calamia V, Lasa I, Bou G. Proteomic and functional analyses reveal a unique lifestyle for *Acinetobacter baumannii* biofilms and a key role for histidine metabolism. *J Proteome Res.*, 10(8):3399-417, 2011.

[20] Xu H, Lee HY, Ahn J. Growth and virulence properties of biofilm-forming *Salmonella enterica* serovar typhimurium under different acidic conditions. *Appl Environ Microbiol*., 76(24):7910-7, 2010.

[21] Workentine ML, Harrison JJ, Weljie AM, Tran VA, Stenroos PU, Tremaroli V, Vogel HJ, Ceri H, Turner RJ. Phenotypic and metabolic profiling of colony morphology variants evolved from *Pseudomonas fluorescens* biofilms. *Environ Microbiol*., 12(6):1565-77, 2010.

[22] Azevedo NF, Lopes SP, Keevil CW, Pereira MO, Vieira MJ. Time to "go large" on biofilm research: advantages of an omics approach. *Biotechnol Lett.*, 31(4):477-85, 2009.

[23] Smith B, Ashburner M, Rosse C, Bard J, Bug W, Ceusters W, Goldberg LJ, Eilbeck K et al. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nature Biotechnology*, 25:1251–1255, 2007.

[24] Sousa AM, Machado I, Pereira MO. Phenotypic switching: an opportunity to bacteria thrive. *In* Méndez-Vilas (Ed.). Science against microbial pathogens: communicating current research and technological advances, pages 252-262, Formatex Research Center, 2011.