ProFuelDB: an open-access database of physiological properties of biofuel-producing anaerobic prokaryotes

C.P. Lourenço[1], M.A. Pereira[1], A. Lourenço[1,2] and D.Z. Sousa[1]

[1] IBB-Institute for Biotechnology and Bioengineering, Centre of Biological Engineering, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal
(E-mail: *lourenco.carina@gmail.com; alcina@deb.uminho.pt; dianasousa@deb.uminho.pt*)
[2] Dpto. de Informática, ESEI - Escuela Superior de Ingeniería Informática, Universidade de Vigo, Edificio Politécnico, Campus Universitario As Lagoas s/n, 32004 Ourense, Spain (E-mail: *analia@uvigo.es*)

HIGHLIGHTS

Anaerobic microorganisms are attractive for the synthesis of fuels and chemicals, but information on the physiology of culturable anaerobes is dispersed in scientific literature. Herewith we present the ProFuelDB, a web-based publicly available database prototype compiling information on the physiology of anaerobic prokaryotes with relevance for biofuels production. It is foreseen that this prototype will evolve into a broader database of physiological properties of anaerobic prokaryotes with biotechnological relevance.

Keywords

Anaerobes; Physiology; Curated database

CONTEXT AND MOTIVATION

Anaerobic prokaryotes can be used for the cost-effective production of bulk chemicals and fuels (Weusthuis *et al.* 2011). The number of anaerobic isolates, with different physiological traits, is rapidly increasing and information on these microorganisms is spread over scientific literature.

Databases gathering phylogenetic information and cultivation conditions of prokaryotes are available (*e.g.* NCBI, EMBL-Bank, SILVA Database, and BacDive). However, more specific information on e.g. substrate utilization, electron acceptor usage, product formation and yields, or growth rates, is not easily retrieved. Therefore, a novel database for easy access to these data would be of interest to both scientist and technicians working in biotechnology and bioengineering fields.

Our goal was to bring together the expertise required to study the physiological properties of anaerobic prokaryotes, making use of our group's research experience in the field. As a prototype, we chose to develop the ProFuelDB database, with physiological information of anaerobic prokaryotes involved in the production of biofuels (Fig. 1).
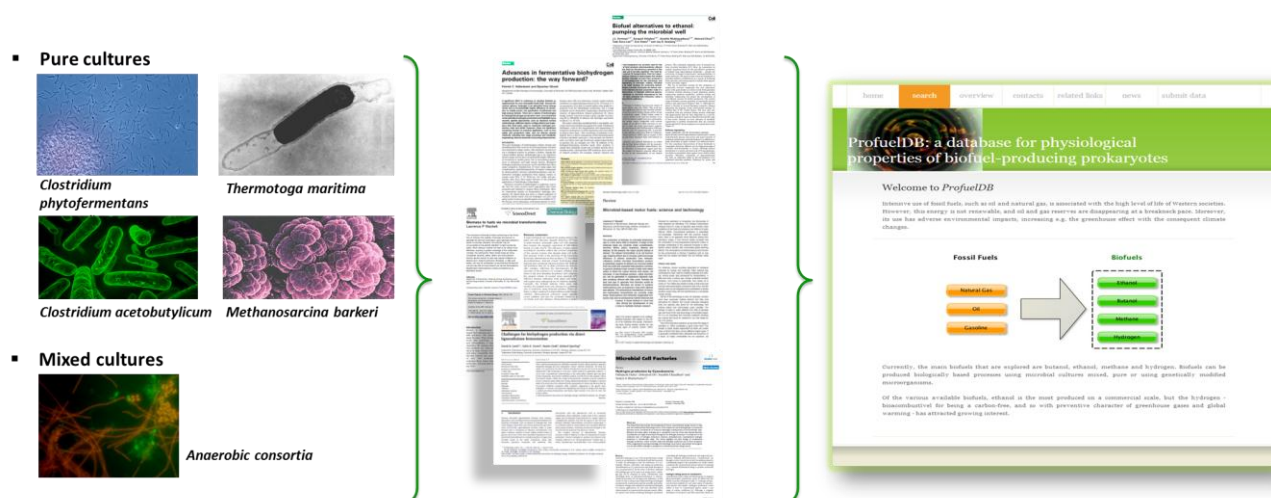

Figure 1. Conceptual idea of the ProFuelDB – gathering dispersed physiological information in a web-based searchable database.

The principles, methodology and implementation of ProFuelDB are flexible and extensible, namely to the information requirements of a broader open-access database of physiological properties of anaerobic prokaryotes. Hopefully, ProFuelDB will be a step forward to the concretisation of a multi-centre collaboration of a community-centred knowledge base.

PROFUELDB DESIGN AND IMPLEMENTATION

Much of the initial effort of the project was focused in data gathering, curation workflow definition and database development (Fig. 2). We identified the information of interest in the literature (i.e. how it is presented) as well as issues regarding data ambiguity, inconsistency and insufficiency. Curation workflow was designed to collect as much information as possible, but information contextualisation is a key premise. The database accounts for anaerobes profiling and characterisation of their physiological properties. Data records identify the source of every piece of information at all times, enabling the characterisation of multiple experimental scenarios (e.g. the use of various substrates or the presentation of different product yields depending on the strain and environmental conditions tested).
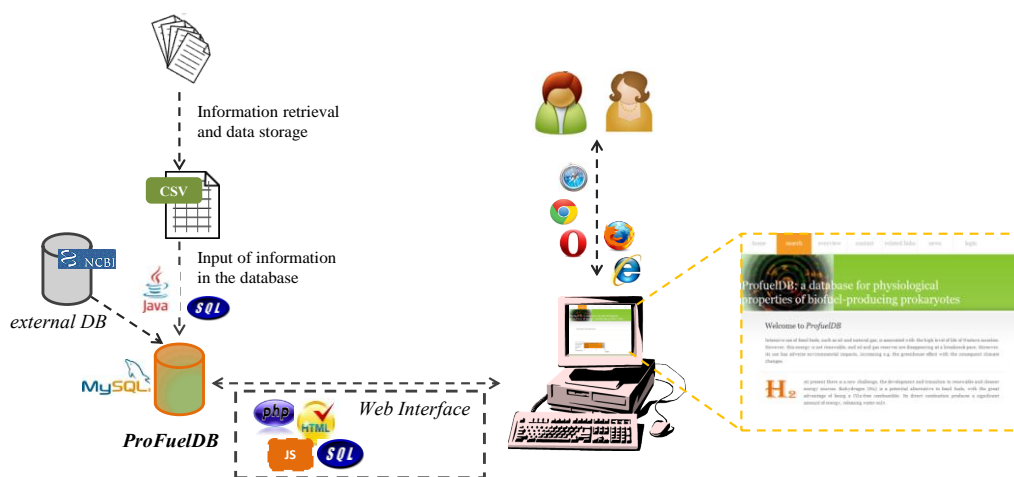

Figure 2. Workflow defined for the design and implementation of the ProFuelDB.

The Web server and all parts of the database are hosted at the Centre of Biological Engineering of the University of Minho, Portugal. ProFuelDB runs on a CentOS platform (version 5.6) with Apache HTTP server (version 2.0), MySQL server (version 14.14, distribution 5.1.52) and PHP 5.1.6. Apache, MySQL and PHP technology were preferred as they are open-source software and platform-independent. Moreover, MySQL supports multi-threading and multi-user environments and thus, it is well-suited to support extended database usage.

PROFUELDB CONTENTS

Currently, ProFuelDB keeps record of physiological properties of 77 anaerobes reported in 82 scientific publications as potential interveners in bio-hydrogen production. The microrganisms included are predominantly neutrophils (73%) and mesophilic (59%), but there is information on others as well. A total of 78 compounds were identified as substrate or product of the reactions played by these organisms in bio-hydrogen production.

REFERENCES
Weusthuis, R.A., Lamot, I., van der Oost, J., Sanders, J.P.M. 2011 Microbial production of bulk chemicals: development of anaerobic processes. Trends in Biotechnology **29**(4),153-158.