Life under the Microscope

Four key partners from the **Microme** initiative discuss the key motivations and components of their new European project, and outline recent successes that the team have had so far

PK: Paul Kersey, European Bioinformatics Institute
CM: Claudine Médigue, CEA - Genoscope
CO: Christos Ouzounis, Centre for Research and Technology Hellas
VdL: Victor de Lorenzo, Centro Nacional de Biotecnología

Why is the study of microbial life so important to society?

VdL: Following the neolithic revolution, industrialisation, intensive agriculture, the green revolution and rampant globalisation, it is clear that our oil-based societies are destined to come to an end and that microorganisms and their catalytic power will be key players in whatever transition we enter into. We argue that such a transition will be eased by the onset of systems and synthetic biology, which allows a fast translation of biological data into useful information and then into valuable products and processes. Access to - and exploitation of - the virtually unlimited catalytic capacity of microorganisms, will endow the chemical and biotechnological sectors with a wealth of new opportunities that will foster a global shift.

Can you outline the aims and objectives of Microme, and what impact this project will have?

PK: Microme aims to extend the scope of microbial genome annotation from functional assignment at gene level to the systematic generation of pathways assemblies and genome-scale metabolic models, with an initial focus on bacteria. This is enormously exciting. New sequencing technologies are leading to the deciphering of dramatically increased numbers of microbial (meta-) genomes, and a key part of what differentiates these genomes is that they encode for different metabolisms that are active under different conditions and environments. By modelling these metabolisms, we can discover *in silico* properties of these systems not otherwise apparent, and apply these lessons to develop new biotechnological processes and understand microbial biology and evolution.

How is Microme bridging the gap between annotation and functional understanding?

CM: A solid reaction and enzymatic pathway annotation infrastructure is being developed in order to identify more easily new enzymatic functions, to fill identified gaps, and to collect and propagate that information once it has been acquired. In addition, integration of computational methods for model refinement, for comparative and evolutionary analyses, and for metabolic engineering and synthetic biology, will help validate and refine the resource. Microme will deliver to the scientific community a well-structured set of pathways and models across a significant range of microbial organisms.

In the second phase of the project, you plan to increase the quantity and quality of the data in the resource; how will this be performed?

CM: The quantity and quality of the data will increase through a continuous (and tightly coupled) process of pathway curation and projection. The curation process is performed in collaboration with microbiologists who have a strong knowledge of specific organisms (curation of complete pathway assemblies in reference genomes) and other specialists in metabolic processes (curation of reference pathways across a full spectrum of species). The reference pathways will be regularly projected to other sequenced genomes, and these projections will be analysed and fed back into the curation process.

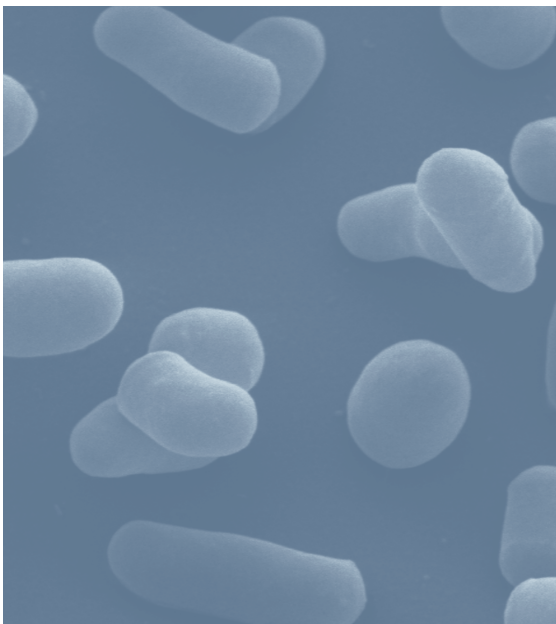
What will it mean for the progression of Microme to be supported by an established bioinformatics infrastructure?

PK: The expertise of the Microme partners in providing much of the existing European bioinformatics infrastructure provides a perfect platform for Microme. We have expertise in service delivery, and also in the source data from which Microme will be built, through our previous work on the main European genomics and biochemical resources. This frees us to concentrate on the missing elements, and deliver real results in pathway reconstruction. Crucially, our involvement in these other projects gives us the chance to feed back our results into these resources, and improve the annotation of microbial biology at every level.

Could you outline any recent successes that you have had?

CO: From a comparative analysis point of view, one area of strong interest is the reconstruction of the 'ancestral' core pathways and pathway variants. The most recent model proposes the partitioning of genomes into two classes, the 'paleome' (the ancient core genome) and the 'cenome' (the more recent additions). This notion greatly enhances our understanding of the essential components of a metabolic pathway (typically encoded in the paleome) and the additional components that have resulted by more recent steps (typically encoded in the cenome).

From a functional analysis point of view, great advances have been made by a number of Microme partners towards a much more precise assignment of unknown genome sequences (and the corresponding genes) to biochemical reactions, for example using probabilistic approaches.

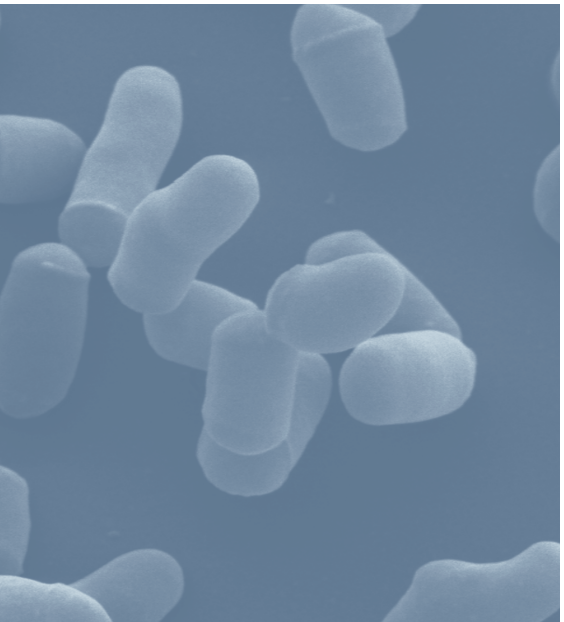


Harnessing Microbial Diversity

Microme is an ambitious new four year European project aiming to integrate approaches and resources in the identification and understanding of microbial metabolism. Researchers behind the project hope that this work will lead to exploitation of bacteria for new biotechnology applications in the future

MICROBES, OR MICROORGANISMS, are the most abundant and ancient organisms on Earth, having existed for almost 4 billion years. In this time, microbial life has evolved to form numerous different organisms, including bacteria, archaea, fungi and viruses. The huge diversity of microbes has allowed life to spread into the most extreme environments around the globe, from the dry heat of the Atacama Desert to the freezing conditions experienced at the South Pole. Studies have revealed that certain microbes can even survive intense radiation, extreme pressures, strong acids and total darkness.

Electron microscopic image of Xylanimonas cellulosilytica, a cellulose and xylane hydrolysing bacterium with bioenergetic potential



MICROBES FOR HUMANITY

Diverse and abundant, the integral role that microbes play in the Earth system is only now beginning to be understood. If science can tap into this resource, a new approach to combatting some of the most pressing problems facing humanity might be taken. Some of the potential uses for microbial biotechnology include the production of novel therapies, vaccines and treatments for diseases, creation of new renewable energy sources, management of environmental carbon dioxide, and the production of chemical catalysts and enzymes for industry.

Bacteria and other microbial life represent the largest reservoir of catalytic and enzymatic activities on the planet. Already some of the functions of microbes have been exploited in industry and medicine, but the majority of these diverse and potentially useful functions still remain unknown.

EXPLOITING DIVERSITY

The best way to gain an insight into the functions of microbes and thereby capitalise on them is to analyse their full DNA sequence, a fairly new science known as genomics. Microbes usually only have one chromosome, meaning that finding the sequence of their genome is relatively simple. However, whilst genomic DNA sequencing progresses at a fast pace, identifying and understanding the functions of sequences lag behind. This information is vital for tapping into the rich resources hidden in microorganisms.

Genome annotation is the process of attaching biological information to DNA sequences. Whilst incredibly important to understanding genomic functions, it is not an exact science. Approximately 30 per cent of the genes in a newly sequenced genome will be annotated as 'hypothetical' or 'conserved hypothetical'. This essentially means that researchers do not know what function they perform.

Offering glimpses into the dynamics within cells, several technologies are now helping to determine cellular functions in terms of metabolic pathways – metabolic pathways show the interactions of genes which induce biochemical reactions. Whilst this has meant that a reasonably clear picture of metabolic processes can theoretically be found, in practice this has only been achieved for certain well-studied systems.

The main difficulty encountered when attempting to discover genomic functions across the full range of different microbes is that a certain catalytic activity will generally be specific to a certain species. This means that researchers do not have a reference from which to build an understanding of what is going on.

TOWARDS A FULLER PICTURE

Microme, 'A Knowledge-Based Bioinformatics Framework for Microbial Pathway Genomics', is a collaborative four year EC-funded project between 14 leading European organisations to develop a comprehensive infrastructure for microbial genomic investigations.

Primarily, Microme has been created to bridge the knowledge gap between genomic sequencing and understanding for all microbial life. Vítor Martins dos Santos is a leading figure in the Microme project. Dos Santos explains the motivation behind Microme: "A major goal of Microme is to move from the simple classification of 'genetic parts' (in essence, component lists of the genome of the microbe involved) to the building of multi-dimensional scaffolds quantitatively describing the metabolic and transport capabilities of the organism". To do this, functions encoded by different 'genetic parts' for each organism will be assembled in a model which accounts for interactions between the various functions. Dos Santos believes that this approach will yield rich results: "It reflects the 'ethos' of Systems Biology in that the whole is clearly more than the sum of the parts. Thus, that the microorganisms are more than a simple list of their genes and encoded functions," he points out. Essentially then, Microme is the first step towards the study of microorganisms as complete, biochemical systems.

However, a serious problem remains: biological functions of new genome sequences can only easily be found by comparing sections of the genome with reference genomes of other microorganisms. One way researchers hope

MICROME - A KNOWLEDGE-BASED BIOINFORMATICS FRAMEWORK FOR MICROBIAL PATHWAY GENOMICS

OBJECTIVES

The Microme project will assemble and develop a bioinformatics infrastructure for the high-throughput generation of whole-cell metabolic models from the sequence of bacterial genomes. Hypotheses generated by these activities will be tested experimentally using growth phenotype data; public access will be provided through a portal and fed back into genome annotation.

PARTNERS

- CEA - Genoscope
- European Bioinformatics Institute
- Centre for Research and Technology Hellas
- Centro Nacional de Biotecnología
- Spanish National Cancer Centre
- German Collection of Microorganisms and Cell Cultures
- Wageningen University
- Ceprodi & AMAbiotics
- Isthmus
- Molecular Networks GmbH
- Swiss Institute of Bioinformatics
- Université Libre de Bruxelles
- Wellcome Trust Sanger Institute
- Tel Aviv University

FUNDING

EU 7th Framework Programme

ARTICLE CONTRIBUTORS

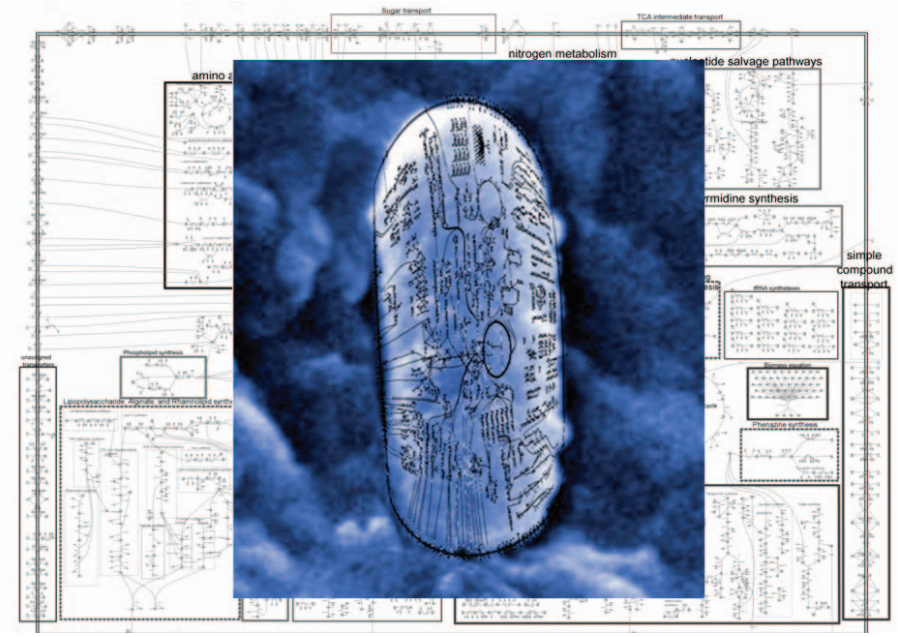
Paul Kersey, European Bioinformatics Institute
Claudine Médigue, CEA - Genoscope
Vítor Martins dos Santos, Wageningen University
Christos Ouzounis, Centre for Research and Technology Hellas
Victor de Lorenzo, Centro Nacional de Biotecnología

CONTACT

Dr Paul Kersey
Team Leader, PANDA Dataflow
EMBL-European Bioinformatics Institute
Wellcome Trust Genome Campus
Hinxton, Cambridge CB10 1SD, UK

T + 44 (0)1223 494601
F + 44 (0)1223 494468
E pkersey@ebi.ac.uk

www.microme.eu



The metabolic network of *Pseudomonas putida*, a soil bacterium with remarkable ability for the degradation of organic solvents. Puchałka *et. al* PLoS Comput Biol 4(10) 2008.

to get over this difficulty is to conduct massive searches of the Microme database to find correlations between DNA sequence types and various other physical and chemical factors, such as the nutrients available to the organism or its biogeography. Microme partners believe that altogether new biological functions might be found using this computational method.

AMBITIOUS AIMS

To integrate resources and approaches into a single workable framework, Microme partners have identified specific objectives which they hope to achieve in the four year project:

- The development of a software infrastructure capable of supporting the curation, projection and display of microbial pathway data
- The delivery of a set of genome-scale metabolic networks, each for a given bacterial species
- The delivery of genome-scale, metabolic models derived for a subset of the species included in Microme
- The development of a suite of software tools exploiting pathways and models from the Microme resource for comparative analyses and metabolic engineering purposes
- The development of a web portal integrating curated pathways, curated and projected pathway assemblies, and models for bacterial data. The portal will incorporate additional data and novel pathway analysis tools developed throughout the Microme project

COLLABORATION

Arching over these objectives is the aim of creating an infrastructure for genome annotators, modellers and other researchers to collaborate. The hope is that this will lead to greater efficiency in the knowledge-building process and may even lead to the development of viable biotechnology solutions to real-world problems.

To achieve this, it is not surprising that researchers across the scientific disciplines are involved in the project. Dr Paul Kersey is from the European Bioinformatics Institute which coordinates Microme. He emphasises the extent to which Microme is reliant on collaboration: "Collaboration really is key. No single group in the consortium has end-to-end expertise; but together, we command the wide repertoire of knowledge necessary to deliver the project".

PUBLIC IMPORTANCE

Microme is not the first metabolic pathway resource. But the existing resources are usually closed to the public, whereas Microme follows the open source model. One of the most interesting outcomes of the project will be that it will be used to engage the public in issues surrounding microbiology, as Kersey explains: "Microme is an infrastructure project – what is important is what it enables. There are always potential risks associated with new technologies, and it is important the public understand these and are empowered to make their own assessments of how they compete with the risks of not finding biotechnological solutions".

He continues by arguing that eventually, the public will have to take notice of the microbial world: "Bacteria may not be loved by most people, but they provide the most promising resource of new chemical activity potentially capable of supporting high standards of living". Clearly, this project is a timely one, and there are high hopes for the far reaching impact of this research in the years to come.