

TerraGenome: a consortium for the sequencing of a soil metagenome

Vogel and colleagues invite the microbiology community to participate in an ambitious and extraordinary sequencing project to uncover the soil metagenome.

Timothy M. Vogel and Pascal Simonet are at the Environmental Microbial Genomics Group, Laboratoire AMPERE, Ecole Centrale de Lyon, Université de Lyon, 36 avenue Guy de Collongue, 69134 Ecully, France.

Janet K. Jansson is at the Lawrence Berkeley National Laboratory, Division of Earth Sciences, Berkeley, California 94720, USA.

Penny R. Hirsch is at Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK.

James M. Tiedje is at the Center for Microbial Ecology, Michigan State University, East Lansing, Michigan 48824, USA.

Jan Dirk van Elsas is at the Department of Microbial Ecology, Centre for Ecological and Evolutionary Studies, University of Groningen, PO BOX 149750 AA Haren, The Netherlands.

Mark J. Bailey is at the Centre for Ecology & Hydrology, CEH-Wallingford, Maclean Building, Crowmarsh Gifford, Wallingford, Oxon OX10 8BB, UK.

Renaud Nalin is at the LibraGen, 3 rue des Satellites, 31400 Toulouse, France.

Laurent Philippot is at INRA, Université de Bourgogne, UMR Microbiologie du Sol et de l'Environnement, CMSE, BP 86510, 21065 Dijon Cedex, France.

Correspondence to T.M.V.
e-mail: Timothy.vogel@ec-lyon.fr

The microorganisms in the 'living soil' are fundamental to all higher life on our planet and are responsible for terrestrial processes that determine our quality of life, including soil fertility, carbon cycling and nutrient cycling. Given the importance of soil functions to most aspects of our lives, surprisingly little is understood about the vulnerability of soil to perturbations or its functional resilience; for example, to changes in land use or climate. We do not fully understand soil biogeochemistry or spatial complexity, or how soil processes, such as carbon cycling, contribute to climate change. Although microorganisms are responsible for key functions in soils, only a small percentage (less than 0.5%) have been grown in the laboratory and genome sequences are only available for a select few.

Soil is the most biodiverse environment on the Earth: it is estimated to contain approximately 1,000 Gbp of microbial genome sequences per gram of soil! Compared with the Human Genome project (in which 3 Gbp were sequenced)¹ and sequencing projects that target microbial habitats, such as the Sargasso Sea (for which 6 Gbp were sequenced)², metagenomic sequencing of soil remains rudimentary and constitutes a new and ambitious challenge. We propose that soil should be our next global metagenomic sequencing initiative.

Sequencing the soil metagenome will bring considerable economic and environmental value. The soil microbial community is a gold mine for genes and pathways that encode novel biocatalysts for biosynthetic or biodegradation processes, including degradation of pollutants, synthesis of biofuels and production of novel drugs. Sequencing of the soil metagenome will also provide insights into the ecology of microorganisms that are beneficial to, or threaten, crop production, and that ensure the quality and provision of ecosystem services.

Recent developments in high-throughput sequencing methods have put the goal of deciphering the soil metagenome within our reach. Large-scale metagenomic sequencing efforts will be necessary to resolve the intricacies of the soil microbiome and to provide sufficient data to understand soil microbial community diversity and function. The success of soil metagenomics depends

on a combination of intelligent sample selection, efficient DNA extraction methods, cloning, screening strategies and sequencing approaches, together with open system data management and sharing. Owing to the magnitude of this task, we propose that a coordinated international effort should be established to combine the skills of the global scientific community to focus on sequencing and annotating the soil metagenome. To catalyse this process, we are seeking agreement and cooperation from the scientific community in reaching a primary objective: the complete sequencing of a 'reference' soil metagenome. The soil system chosen for investigation, Park Grass, is an internationally recognized agroecology field experiment that has been running for more than 150 years at the UK agricultural sciences institute, Rothamsted Research. This ambitious reference sequencing effort cannot be undertaken by a single laboratory or even by a single country. We therefore invite the international community to participate in this project, and hope to eventually expand the project to other soil sites. The information gleaned from this project will serve as a starting block or platform for other soil metagenomic sequencing efforts and will generate new hypotheses to test. This initiative will also spur complementary efforts in other 'omic approaches, such as transcriptomics, proteomics and metabolomics of soil to add more layers of information about gene expression, activity and function. In addition, we will need to develop and apply new approaches to cultivate the previously uncultivated and rare members of the soil community to assign functions to the vast number of unknown or hypothetical genes that will undoubtedly be found. The [TerraGenome international sequencing consortium](#), which is dedicated to soil metagenomics, has just been launched to coordinate these efforts.

1. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* **431**, 931–945 (2004).
2. Rusch, D. B. *et al.* The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol.* **5**, e77 (2007).

FURTHER INFORMATION

TerraGenome international sequencing consortium: <http://www.terragenome.org/>