ESF Science Meeting – Final Report

MOLTER database workshop, May 19-20, Oslo, Norway

"A prototype for a MOLTER (MOLecular structures as drivers and tracers of TERrestrial C fluxes) database"

Funded by the

- ESF Networking Programme MOLTER: Molecular Structures as drivers and tracers of terrestrial C fluxes

And the

- Norwegian Research Council

Scientific Report
Program
List of participants

The aim of the workshop was to conceive suggestions for the design and functioning of a database containing quantitative and qualitative data of soil molecules. Such a database will help us to understand the effect of environmental conditions and soil management on the abundance and behavior of individual soil molecules. In the light of a warming climate and rising atmospheric CO_2 concentrations such information will be useful to better predict the response of soil organic matter to higher temperatures and changing climatic conditions.

The global carbon cycle: uncertainties for future climate projections

The properties of the global carbon cycle and its behavior as a response to climate change are still largely uncertain. Huntingford et al (2009) suggest that the unknown behavior of the carbon cycle is responsible for 40% of the uncertainty related to global warming projections. Knowledge of the mechanisms and factors controlling the stabilization or breakdown of soil organic matter is essential to improve our understanding of the dynamics of the carbon cycle at regional to global scales. Carbon pools in terrestrial environments are composed of contrasting molecular structures; these display contrasting properties, which ultimately govern their fate in the environment. The stability of soil organic carbon is thought to be due to the chemical recalcitrance of specific compounds. Structural complexity and chemical

bonding seem to be greatly responsible for a compound's fate and longevity in the soil environment (Kiem and Kögel-Knabner, 2003), and therefore directly influence the climate-carbon cycle feedback. The geographical location might play a significant role as well: certain compounds might be more abundant under certain climatic conditions, decompose faster or be more soluble.
Knowledge of the abundance and behavior of individual molecular structures will enable us to better predict the climate sensitivity of soil carbon stores worldwide and, hence, the magnitude of future CO_2 emissions. Therefore, an improved understanding of the mechanisms that control terrestrial carbon fluxes is both scientifically and politically important (Chapin et al., 2009).

Why a molecular database? What are the advantages and what requirements should such a database fulfill?
It is necessary to compile information in order to be able to get a comprehensive picture of the abundance and the dynamics of the most important individual soil molecules.
The database will not only serve as a tool for the collection and storage of larger amounts of data, but, more importantly, enable us to compare and extract information on interactions between molecular transformations and soil management as well as geographical location (differing soil type and climatic conditions). In the light of a warming climate such information will help us to better predict the response of soil organic matter to higher temperatures and changing climatic conditions.

The database workshop: Towards a database and on-line analysis tool for soil molecular research.
Molecular databases have been identified as an important tool by the ESF MOLTER networking programme (www.molter.no) to foster synergies in scientific activities at the European level on molecular research in terrestrial environments in general and soils in particular. The aim is to generate a widely-used online tool, which will generate the scientific synergies and advance European research in the field of soil molecular research, in agreement with the mandate of ESF networking programmes (www.esf.org). To reach this general goal, multiple possibilities exist in terms of scope, format, and analysis tools. In this context, MOLTER sponsored an exploratory workshop in Oslo in May 2010 on the development of soil molecular databases.

The overall framework of the database is to help our understanding of the complexity and composition of soil organic matter. Archiving is the first function of the database. However, this function is not sufficient in itself for ensuring the success at creating wide scientific synergies in the research community. Indeed, an archiving database is at risk of both being underused and receiving insufficient data inputs. Processing and re-processing of archived raw data is a second function of a computational database. A third function consists in performing statistical analyses on the processed data. Combining these first three functions, i.e. archiving, processing and statistical analyses, appears to be the primary objective of modern soil molecular database efforts developed at the level of the individual research laboratory or institute. Examples of such efforts under development are the soil NIR/MIR spectral database of the CEMAGREF in Grenoble – France, and the soil lipidomics database at Kansas State University. In other words, modern soil molecular database are mostly developed at the institute level for optimizing data analysis for one particular research

method. The transition towards an international computational database for soil molecular data will require additional functions.

Exploring the impact of environmental drivers on SOM molecular composition in a spatially explicit way represents the fourth desired function of the computational database. Within the MOLTER network, a prototype database for soil lignin composition across European environmental conditions has been developed. A first evaluation of the prototype suggests that this type of database is best geared towards one defined international collaborative project. A modern international soil database would need, in addition, to accommodate for multiple data types (from individual data to entire spectras) and integrate online analyses tools. These analyses should include multivariate statistics (PLS, …) and data fusion techniques in order to cross-compare and cross-calibrate multiple molecular techniques, with the aim of better understanding the complexity and composition of soil organic matter.

Our vision is that of a soil molecular database that would be of service to a large scientific community through online freeware analysis tools, both for raw data processing and statistical analyses. As the database grows, the associated statistical tools will help us infer series of molecular properties from a single type of measurement, e.g. infra red spectras, thermogravimetry. Such a database would need to be originally seeded with multiple sets of analyses conducted on standard samples. At last, the computational database should have a long-term goal of integrating expert-system functions, which would aim at estimating soil ecological parameters (e.g. SOM stabilility, soil structure, potential N_2O emission) based on soil spectral and molecular measurements.

Future actions:
These concepts were presented at the ESF MOLTER- sponsored SOM 2010 conference in September 2010 in France together with a first prototype of computational database for soil molecular research:

Cecillon, L., D.P. Rasse, K. Knoth de Zarruk, A. Budai, R. Sørheim, R. Jeannotte, G. Ragan, E. Ancelet, F. Bray, J.-J. Brun, S. De Danieli, C.E. Thompson, A.F. Plante, G. Hudson. Towards a database and on-line analysis tool for soil molecular research. Organic matter stabilization and ecosystem functions. SOM 2010. Presqu'île de Giens, France, Spet 19-23 2010.

In addition, we are now working on an article with the same title for a special issue of Biogeochemistry, associated with the SOM 2010 conference.

The outcome of the Oslo workshop will be presented to the MOLTER steering committee at the November 2010 meeting in Amsterdam. Future support of MOLTER to the database will be discussed.

New analyses have been conducted as a direct result of the workshop (in Univ. Pennsylvania in collaboration with CEMAGREF), in order to start testing the database prototype. Several participants to the workshop (in Norway, France, Switzerland, USA) have been informally discussing venues for larger proposals to support the development of the database. These efforts are expected to continue.

Visibility and acknowledgement of ESF funding

All participants were informed beforehand that the workshop was funded by ESF and that ESF would cover all costs for European as well as part of the non-European participants. Additionally, ESF and MOLTER logos were printed on the first page of the workshop program. The chair of MOLTER, Dr. Daniel Rasse, opened the workshop with a presentation on MOLTER and ESF Networking Programmes in general. ESF funding was also acknowledged in the SOM 2010 presentation. .

Synergies

The MOLTER database workshop was suppoted by an additional grant of the Norwegian research Council of 12 500 euros. Workshop participants from the USA (University Pennsylvania, and Kansas State University) were supported through this parallel funding.

Agenda for the MOLTER database workshop

May 19-20, 2010 in Oslo

Wednesday 19.05.2010:

10:00-10:20: Coffee/Welcome/Introduction

10:20-10:35: Roald Sørheim, Bioforsk: *The Soil and Environment Group at Bioforsk*

10:35-10:55: Daniel Rasse, Bioforsk: *Why a soil molecular* d*atabase?*

10:55-11:15: Katrin Knoth de Zarruk, Bioforsk: *The "MOLTER database" – first attempts.*

11:15-11:30: Questions/Discussion

11:30-11:45: Coffee break

11:45-12:15: Lauric Cecillon, CETIOM: *"Innovative data processing for MOLTER DB".*

12:15-12:30: Questions/Discussion

12:30-13:30: Lunch

13:30-14:30: Sebastien de Danieli, Frederic Bray and Estelle Ancelet, Cemagref: "*The Cemagref Model."*

14:30-15:00: Questions/Discussion

15:00-15:15: *Coffee break*

15:15-15:30: Alain Plante, University of Pennsylvania: "*Project plans"*

15:30-17:00: Discussion/Sum up/ideas

17:30: Walk to restaurant

18:00: Joint dinner

Thursday 20.05.2010:

9:00-09:10: Coffee

9:10-10:10: Richard Jeannotte and Gail Ragan, Kansas Lipidomics Research Center: *"Lipidome DB: database of molecular lipid species applicable to soil systems".*

10:10-10:30: Questions/Discussion

10:30-10:45: Coffee break

10:45-11:00: Cathy Thompson, GNS: "*Fully Automated Thermochemolysis for High-Throughput Monitoring of Soil Biomarker Abundance".*

11:00-11:15: Questions/Discussion

11:15-11:45: Gordon Hudson, Macaulay Land Use Research Institute: "*Soil database".*

11:45-12:00: Questions/Discussion

12:00-13:00: Lunch

13:00-15:00: Final Discussion/Workshop sum-up

15:00: End

List of participants:

A total of 13 scientists participated in the database workshop. Additional funding from the Norwegian Research Council was obtained in order to invite two additional participants from outside Europe and to cover working hours for the organization of the workshop (* non-ESF funding).

1. Ancelet, Estelle, Cemagref (France)
2. Bray, Frédéric, Cemagref (France)
3. Budai, Alice, Bioforsk (Norway)
4. Cecillon, Lauric, CETIOM (France)
5. de Danieli, Sébastien, Cemagref (France)
6. Hudson, Gordon, Macaulay Land Use Research Institute (UK)
7. Jeannotte, Richard, Kansas Lipidomics Research Center (USA)
8. Knoth de Zarruk, Katrin, Bioforsk (Norway)
9. Plante, Alain, University of Pennsylvania (USA)*
10. Ragan, Gail, Kansas Lipidomics Research Center (USA)*
11. Rasse, Daniel, Bioforsk (Norway)
12. Sørheim, Roald, Bioforsk (Norway)
13. Thompson, Cathy, GNS (NZ)