



Mapping out the best root

Catholic University of Louvain (UCL) uses SAS for agronomy reasearch

Many of the world’s most important food and feed crops are members of the Poaceae family, including rice, wheat, maize, sorghum, sugar cane, barley, oat, rye and millet. Improving the performance of these crops can therefore make a huge difference to the world’s nutritional health.

The Faculty of Bio-engineering, Agronomy and Environment at the Catholic University of Louvain (UCL) is carrying out genetic research in order to gain a better understanding of the consequences of selective breeding and the evolutionary effects of natural selection. Such research supports efforts to improve the performance of a wide spectrum of animal and plant species.

Dr Xavier Draye, a researcher within the Faculty’s Laboratory of Crop Physiology and Plant Improvement has focused on sorghum (*Sorghum bicolor* L. Moench), a suitable species for gene positional cloning because of the relatively small size of its genome. Sorghum is a leading cereal in arid and semi-arid agriculture, ranking fifth in the world’s most heavily produced grain crops. Moreover, sorghum provides an important template for the study of closely related large-genome crops such as maize and sugar cane.

Draye explains why genetic research on sorghum and similar plants should allow improvements to crop qualities and yields, while reducing the environmental impact of growing these crops (for example by reducing the levels of fertilizer residues in the soil). “For example, we try to gather detailed knowledge of the root system size and development of the plant. We look in the genome to try and find the genes responsible for a specific growth or form of the roots that makes them more adapted to certain soils and environments.

“We analyse the root system in the soil and the plant’s genome, making a comparison between the two,” says Draye. “Our goal is to exploit the genetic variations between individual plants.”

Over the past two years direct sequencing of genomes (running to billions of pairs of bases for a crop plant such as maize) has become routine. The current challenge is to annotate the DNA sequence, in other words to find the genes (each a sequence of pairs of bases) and assign their functions. One promising strategy is to look at the association between the presence of certain sequences (also known as DNA markers) in the genome and desired characteristics of the roots.



Customer	Catholic University of Louvain (UCL)
Sector	Academic
Business Issue	Needed analytical solutions for agricultural science
Solution	SAS®
Benefits	Prepares students for successful careers in business

More than two decades of genomic research have yielded high-density maps of DNA markers for many major crops. These maps provide a valuable foundation for basic research in genome organization and evolution. However for this research to succeed it is absolutely essential to have access to libraries containing large inserts of genomic DNA, enabling the production of physical maps and the isolation of important genes. One such library, containing a large insert of Sorghum propinquum BAC (bacterial artificial chromosomes) has been constructed at Texas A&M University to analyse the physical organization of the sorghum genome and facilitate positional cloning of sorghum genes.

"We are speaking of really high numbers here," says Draye. "Just to give you an idea, with 1,000 to 10,000 markers, we try to map 30,000 to 150,000 clones. We use grids to spot the occurrence of certain clones, with more than 18,000 clones on a single grid." Keeping track of all this information is further complicated by other factors, according to Draye. "Grids can be re-used for examining (but get old and more imprecise to read), batches tend to overlap and research people come and go."

UCL therefore needed to develop a solution that could capture large amounts of data quickly and output intelligible results to researchers' queries. The solution would need to facilitate the archiving of the data

produced by screening macroarrays of ordered DNA clones, and it would also need to allow easy navigation of the data to detect any errors and uncertainties.

"Consequently we needed a monitoring system and a robust data warehouse, one that could integrate data from some 10 different databases." The solution needed to enable easy and accurate identification of the grids, while guiding the user through the various steps intuitively.

UCL chose SAS software for the complex statistical work involved in mapping DNA markers for genome sequencing. "The SAS applications development environment enabled us to develop the BAC-DMS solution very rapidly – in less than three months, working only part time," says Draye. "SAS has a very open systems design and allows easy links. SAS is also a highly developed language and a very robust and comprehensive statistical analysis tool, which means that an experienced SAS user can create a new application in less than half an hour. The fact that everything can be done in one single environment also accelerates the work."

"The robustness and power of SAS has enabled us to build a comprehensive data warehousing system for our fundamental and applied research in a very short period of time."

Dr Xavier Draye

Faculty of Bio-engineering, Agronomy and Environment, UCL

The solution is now well developed. It offers a versatile "click-a-clone" graphical interface for data entry, a data warehouse including a sophisticated exploration tool, a routine to export groups of overlapping clones for interactive clone ordering and a monitoring system and a monitoring system to support the management of a large number of reusable macroarrays.

"You do not need SAS skills to use the BAC-DMS solution," says Draye. "But you can access all of the native SAS windows during a BAC-DMS session, so virtually any database transaction can be performed by an experienced SAS user."

With this power and flexibility, UCL's bio-engineers and agronomists are better equipped to carry out fundamental and applied research that will help improve the characteristics and performance of food plants.



SAS World Headquarters
and SAS Americas
SAS Campus Drive
Cary, NC 27513 USA
Tel: (919) 677 8000
Fax: (919) 677 4444
U.S. & Canada sales:
(800) 727 0025

SAS Europe, Middle East & Africa
P.O. Box 10 53 40
Neuenheimer Landstr. 28-30
D-69043 Heidelberg, Germany
Tel: (49) 6221 416-0
Fax: (49) 6221 474850

www.sas.com