

Identification of plants that cause subsidence damage using genetic data analysis

by John Graham

The seemingly arcane field of identifying tree roots has in fact an important practical application. In assessing building damage caused by subsidence due to tree roots, it is important for insurance and compensation purposes to be able to identify the offending trees from analysis of the roots. An EU-funded project is developing a genetic database of common trees and shrubs. Efficient use of this database requires sophisticated software systems for the storage and comparison of DNA gels.

Accurate analysis of tree or shrub roots is often crucial in the outcome of subsidence claims where plants are implicated in damage to a building. In fact, insurers often make decisions on claims based on this information to obtain compensation from the owner of the plant in question.

Traditionally, root analysis starts with a site investigation company collecting samples at the site of the damaged property and posting them to a laboratory for analysis. The roots are sectioned and examined by microscope. The sections are compared by eye with pictures of known tree or shrub root cells and identification is based on finding a match. This crude method of identification relies heavily on the experience of the person analysing the samples and has the disadvantage of not being able to reliably identify individual species of tree or shrub. Difficulty arises when two roots from different plants appear to be the same. For example poplars and willows, which are from the same family, have root cells that can look very similar under the microscope, a problem which is exaggerated when the soil or weather conditions have dried the roots out.

Jonothan Cocking Associates have secured European Union funding to set up a laboratory to develop a root identification service based on genetic fingerprinting. The laboratory utilises DNA techniques such as PCR and gel electrophoresis, coupled with advanced image analysis and information management systems from Syngene (Cambridge, UK), to match roots with the trees or shrubs to which they belong. Microsatellite and RAPD (Random

Amplified Polymorphic DNAs) PCR methods are used to produce DNA from trees and shrubs. Microsatellites are tandem repeats of 2-10 base pairs which are often present as perfect or imperfect repeats in the DNA. They can be amplified using PCR primers that flank the microsatellite regions. RAPDs are another method of PCR where arbitrarily chosen 10 base primers are used to search for variations in the DNA. Following PCR, the amplified DNA is run on agarose gels, stained with ethidium bromide and the DNA is then visualised using Syngene's GeneGenius image analysis system.

Gel Analysis

These RAPD and microsatellite methods produced vast quantities of genetic fingerprint information, which is stored in the GeneGenius' GeneTools software package. The software allows the laboratory to store information associated with a gel in a way compliant with Good Laboratory Principles (GLP), perform band matching across a gel to a known standard and also export this information if necessary.

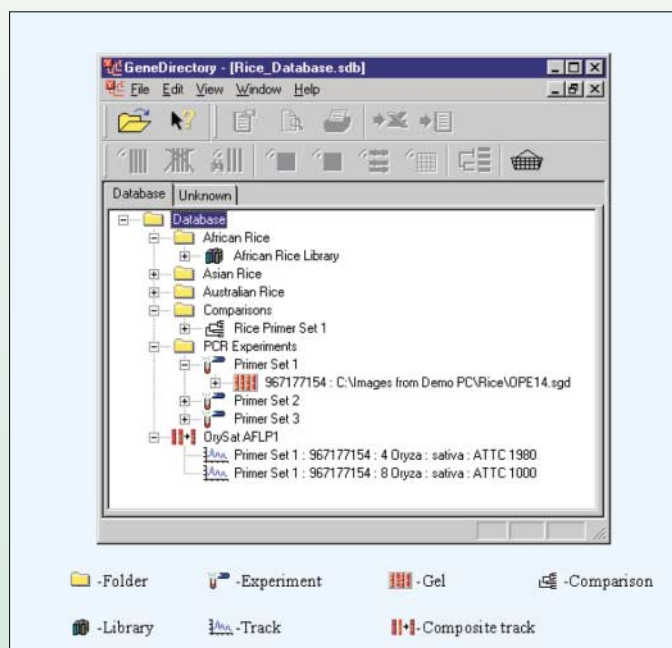


Figure 1. The hierarchical tree representation of the database. This is used to store gel images of plant root DNA.

Data Storage

It is necessary to compare gel band patterns between gels. To avoid doing the calculations by hand, the new GeneDirectory from Syngene was chosen. This is a gel data storage and analysis package that works in conjunction with GeneTools and enables complex genetic fingerprint studies to be performed across an unlimited number of gels. The software is simple and intuitive to use, making it particularly suitable for laboratory users who may have limited experience of this type of diagnostic software. The system's relational database and SQL queries enables large amounts of gel data to be stored and retrieved efficiently. Having a formal database structure makes it possible to lock records while others are accessing them. The system can thus be operated over a network for access by multiple users. In GeneDirectory the database is represented by a simple hierarchical tree view (Figure 1) which enables easy navigation around the data, a task which would be impossible if it were simply displayed as a list of tracks. The structures within the tree can be expanded and collapsed in a standard explorer-type manner.

Building a Database Library

The Jonathan Cocking's Laboratory have developed a library of gel images of plant root DNA using Syngene's Gene Directory. This was done by classifying tracks of DNA bands according to the species of tree or shrub that the root DNA originated from. Once classified, the track is added to a root library data set, which is continuously being developed into an increasingly comprehensive data collection of DNA root fingerprints.

Data Comparison

Data from the DNA root library database can be compared in several ways. The simplest of these is when an unknown track is introduced into the database. A search routine in the software identifies the nearest match in terms of band positions and band-quantities. More complex ways to view data are cluster analysis and genotyp-

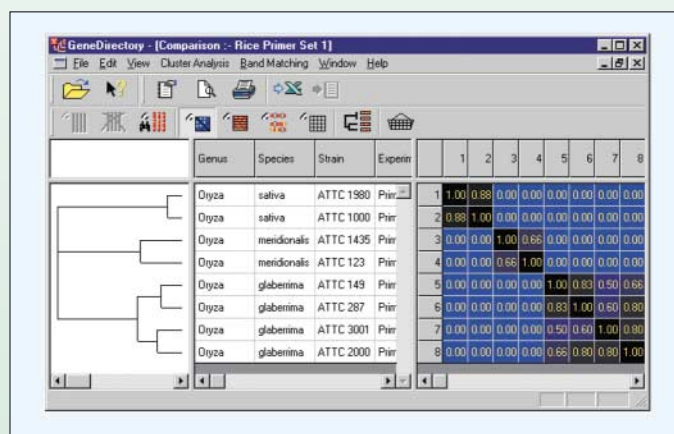


Figure 2. Example of a dendrogram and similarity matrix. This portrays the calculated similarity between, say, an unknown sample and the rest of the database and enables the unknown sample to be clustered, or grouped.

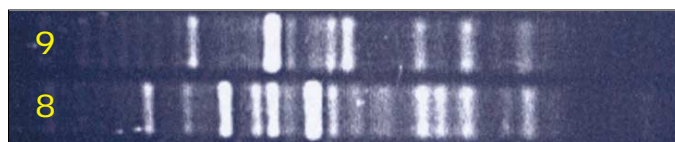


Figure 3. Example of gel patterns of RAPDs from different species. Lane 8 White Willow, Lane 9 Weeping Willow. Correct classification of such patterns requires sophisticated software packages.

ing, both of which offer different ways of comparing band patterns. The cluster analysis viewing method is particularly suitable for relating the DNA root library to unknown sample roots and involves comparing the migration-distance values of DNA bands on a gel track. A similarity measure is calculated between an individual track and every other track in the database, eventually generating a similarity matrix. The similarity measure runs from zero (no similarity) to one (identical). Once a similarity matrix is generated a dendrogram (Figure 2) can be drawn which shows how the root DNA should be grouped, or clustered.

Results

So far it has been shown that both RAPDs and microsatellite PCR products on agarose gels can produce unique DNA band patterns of trees and shrubs at the species level (Figure 3). In practice, however, it has been found that the RAPD method produces DNA that is easier to visualise on agarose gels. This technique has therefore been chosen for further expansion of the library database. Already a library of one hundred unique DNA profiles of trees and shrubs that are most commonly associated with causing building subsidence has been constructed.

Discussion

The use of GeneDirectory software for storage and comparison of agarose gel profiles of RAPD and microsatellite plant root DNA means that it is possible to compare automatically DNA tracks of known root samples with unidentified root DNA, therefore rapidly identifying a match. The manual alternative of comparing gel pictures by eye is extremely tedious, time-consuming and extremely expensive. In fact manual gel comparison is so cost-ineffective that this approach is not economically viable compared with the less accurate microscopy approach.

Using the software comparison of the data, it is envisaged that between 10 and 20 root damage cases per week will be analysed using the DNA fingerprinting technique. Initially all the DNA results will be compared with microscopy analysis to check the accuracy of the DNA based technique.

The Author

John Graham, Research Associate, Jonathan Cocking Associates, Halifax, West Yorkshire, HX4 0AD, UK. Tel +44 1422-376335. Email:jcocking1@aol.com
www.jonathancockingassociates.co.uk