E

UPOV

BMT/13/32
ORIGINAL: English
DATE: November 21, 2011

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

GENEVA

WORKING GROUP ON BIOCHEMICAL AND MOLECULAR TECHNIQUES, AND DNA-PROFILING IN PARTICULAR

Thirteenth Session
Brasilia, November 22 to 24, 2011

DEVELOPMENTS CONCERNING THE VARIETY TRACER PROCEDURE

Document prepared by an expert from the Netherlands

Introduction to "Variety Tracer"

1. "Variety Tracer", the "Sherlock Holmes" concept in infringement matters, has been especially developed to answer questions about the identity of plant material, suspicion of repeated cropping, suspicion of infringement of Plant Breeders' Rights and patents and "Essential Derived Varieties" (EDVs). This product can provide evidence which can be used to settle (legal) discussions about identity. Within "Variety Tracer" crop-specific knowledge, morphological assessment and genetic research based on genotyping with molecular marker systems are combined. The independent status of Naktuinbouw as a self governing body (ZBO) regulated by the Ministry of Economic Affairs, Agriculture and Innovation (EL&I), is very important.

2. Variety Tracer projects consists of five steps:

1; *Problem analysis* in which the nature of the problem or question is discussed. Since every problem is unique, a research plan to solve the problem is different every time. Also, information on the crop is important (vegetatively propagated crops and cross-pollinated crops are investigated in different ways).

2; *Sampling.* Especially when legal steps are to be expected, a bailiff can be charged for purchases and seizures of plant material. Independent crop-specialist (e.g. Naktuinbouw employees) are then assisting the seizure.

3; *Morphological assessment* is carried out in accordance with the relevant UPOV/CPVO guidelines/protocols. Depending on the crop, this research is carried out by Naktuinbouw or in cooperation with the responsible EU examination office for the particular crop.

4; *Genetic conformity* is investigated by the use of DNA marker technologies. Which technology or technologies is/are exploited depends on the crop and the availability of DNA markers in the public domain for the particular crop. Naktuinbouw mainly uses AFLP® (Amplified Fragment Length Polymorphism). This technology was developed by Keygene N.V. in Wageningen, the Netherlands and Naktuinbouw is the only organization in the Netherlands with a service license. This technique is universally applicable for all crops in Naktuinbouw's working domain and even beyond. Not only plants, but all living material which contains DNA, therefore also plant pathogens, can be examined using this technique without prior sequence information. Therefore, AFLP® is a very flexible DNA marker system. For some major crops Naktuinbouw have SSR markers that can be used in genetic conformity studies.

5; *Reporting*. All results are statistically underpinned and described in a report. In the event of a dispute, this report may serve as evidence in a lawsuit.

Improved Variety Tracer Procedure

3. Experiences with the Variety Tracer Procedure have indicated the necessity to perform a so called 'AFLP-plus' procedure in the future. Although the technological developments rapidly improve and Next Generation Sequencing (NGS) makes sequence information more and more accessible for many small crops, there is not (yet) (whole genome) sequence information available for most species. Therefore, the AFLP® technology is an affordable option for performance of a genetic analysis in cases of suspected infringements. Although we cannot change the dominant and bi-allelic nature of AFLP® markers, we can improve our procedure to improve certain aspects. Naktuinbouw has improved the Variety Tracer procedure by a detailed research plan that describes the experimental set-up and selection of reference varieties that should be included in the analysis. The general requirements for a solid, reliable and unbiased research plan based on AFLP® are:

1. *Selection of samples to guarantee representative sample collection.*
    a. To cover the complete genetic width of the particular species, the reference collection to be used in the AFLP® analysis must represent the complete morphological diversity within the species currently available in the market. All relevant morphological characteristics should be present in comparative number of varieties.
    b. In addition to the genetic width, it is very important to zoom in on samples that are supposed to be genetic related. To fine tune the smallest genetic distance between related but not derived samples, we purpose to analyse varieties that originate from at least one common parent and/or progeny from a cross between two non related parents (if such a population is made available by an independent party).

c. The genetic analysis must include known and accepted mutant varieties and their original varieties. Although mutants are distinguishable from the original varieties based on morphological characteristics, their DNA fingerprints are expected to be nearly identical or completely identical.
d. The AFLP® analysis must contain varieties that are similar to the varieties of question on many morphological characteristics (based on the DUS characteristics) but not genetically related assuming their pedigree information.
e. From the varieties of question, several different origins (from different independent growers) must be included in the analysis to investigate the allowed genetic variation within a variety due to selection or genetic drift.
f. From all varieties included in the AFLP® analysis *duplo* samples must be analysed. This means that from the same leaf material two independent DNA extractions will be performed. The two DNA samples will proceed through the whole AFLP® procedure as independent samples and serve as a measure for the intrinsic error rate in the AFLP® protocol.

2. *Guarantee of anonymity*
All samples will be supplied to the lab technician in random order and under code and no further background information about the origin and background of the samples is distributed. Therefore, objectivity and independence is guaranteed.

3. *Determination of optimal reaction conditions*
To chose the optimal AFLP® reaction conditions a pre-screen is performed. Optimal reaction conditions are dependent on many different factors such as species, genome size and natural (botanic) genetic variation within the species.

4. *Separation of AFLP fragments, normalization and scoring*
It is of crucial importance to use a well defined normalization during separation of the DNA fragments (irrespective of the platform used). Naktuinbouw currently uses a gel-based platform (LICOR) with fluorescent dyes. With this system high quality gels are produced. DNA fragments in the range of 50-500 base pairs (bp) are analyzed. To normalize the gels an additional set of PCR fragments (Lambda DNA) called the “size-set” is added to every individual sample. This “size-set” contains 10 DNA fragments (app. every 50 bp) of known size and enables a reliable normalization of gels and makes it possible to combine different gels and data sets in the same analysis and database over time.

5. *Analysis and scoring of the DNA fingerprints*
Naktuinbouw uses the “BioNumerics” software (Applied Maths) to analyze and score the DNA fingerprints on normalized gels. The samples are genotyped based on the presence/absence of polymorphic DNA fragments (markers). In case of doubt during scoring, a sample will be scored as uncertain, which means that the marker is considered as neither present nor absent for this sample. Scoring is very strict; markers that are not clear are ignored in order to avoid false positive scores. Naktuinbouw’s philosophy is better to leave out not reliable information then to risk inaccurate scores. If too many uncertain scores will appear, the AFLP procedure will be repeated for the particular samples. Scoring is confirmed by another person.

6. *Genetic distance/similarity analysis*
In order to investigate the genetic relationship of the samples, a data set of absent/present marker scores is generated (scoring table). For a simplified

representation of these results, techniques such as clustering and ordination analyses are generally employed. The predecessor of these analyses is the construction of a similarity (or distance) matrix. The similarity matrices based on the 'Jaccard' and the 'Dice' coefficients are calculated (not only the Jaccard). "BioNumerics" software is used to produce all of these similarity matrices. To visualize the relationship between the samples three different charts or outputs will be presented: 1; the frequency distribution of pairwise comparisons based on the similarity matrices, 2; dimensioning techniques based on both the scoring table (Principle Component Analysis, PCA) and the similarity matrix (Multi Dimensional Scaling, MDS), 3; dendrograms using UPGMA parameters (Unweighted Pair-Group Method, Arithmetic average) based on different similarity matrices. Subsequently, the software allows to show the original DNA profiles (gel image) for every individual sample in the dendrogram. Since the samples that are most related are grouped together it is very useful to double-check their original DNA profile to verify if the differences in similarity directly correlate with polymorphic bands. This double-check will consequently reduce the human/technical error rate dramatically and improves the reliability and reproducibility of the analysis.

7. *Bias/non-bias of used markers*
   Since AFLP® is a random marker system the exact position of the markers is not known neither is the genome coverage. Since AFLP® has been used intensively in academia for the construction of genetic maps, we can assume that the genome coverage is not biased. To get an impression of the coverage of the markers used, a Principal Component Analysis on character coordinates is performed using BioNumerics. The maximum percentage of variation in two dimensions between the markers used is depicted in a graph. Alternatively, an additional genetic analysis can be preformed when crop specific markers and physical maps are available to prove the random distribution of the AFLP markers indirectly.

8. *Standard error for similarity estimates*
   Standard error for similarity estimates will be obtained as described by van Eeuwijk and Law, 2004.

9. *Cluster verification*
   Among the available cluster verification tools two will be applied: 1; cophenetic correlation to evaluate to what extent the dendrogram is a good representation of the similarity matrix, and 2; bootstrap analysis (resampling).

Naktuinbouw, PO BOX 40, 2370 AA Roelofarendsveen, The Netherlands:
Dr. Hedwich Teunissen, Daniël Deinum, Menno Hoekstra, Willem Wietsma, Dr. Raoul Haegens and Kees van Ettekoven. (contact: h.teunissen@naktuinbouw.nl)

Vondst Advocaten, P.O. Box 75781, 1070 AT Amsterdam, The Netherlands:
Mr. Tjeerd Overdijk, Mr. Hidde Koenraad

[End of document]