Report

of the

CORESTA Task Force

Genetically Modified Tobacco:

Detection Methods

CONTENTS

List of Contributing Authors 3
Summary 4
I. Introduction 6
II. Sampling Strategy and Methods 8
III. Sample Preparation and DNA Isolation Methods 16
IV. PCR-Based Screening Methods 22
V. Semi-Quantitative and Quantitative PCR Methods 42
VI. Immunological Detection Methods 59
VII. Selectable Genetic Marker-Based Detection Methods 65
VIII. Official and Validated Methods 73

APPENDICES

I. Internet Databases on GM Tobacco Field Trials 88
II. Tobacco DNA Extraction for PCR Analysis 89
III. Chinese Method for Screening GM Tobacco 93
IV. Glossary of Terms 114

LIST OF CONTRIBUTING AUTHORS

Mr. Gregor Bindler	*Philip Morris International, Switzerland*
Dr. François Dorlhac de Borne	*SEITA, France*
Dr. Ferruccio Gadani Task Force Coordinator and Editor	*Philip Morris International, Switzerland*
Dr. Evan Gregg	*British American Tobacco, UK*
Mr. Zhaokui Guo	*Heilongjiang Tobacco Research Inst., P.R.China*
Dr. Hubert Klus	*Austria Tabak AG, Austria*
Dr. Martin Maunders Co-Editor	*Advanced Technologies (Cambridge) Ltd, UK*
Dr. Lutz Mueller	*R. J. Reynolds Tobacco GmbH, Germany*
Dr. Hideaki Negishi	*Japan Tobacco Inc., Japan*
Mr. Hans Pijnenburg	*Philip Morris International, Switzerland*
Dr. Martin Ward Task Force Secretary and Editor	*Advanced Technologies (Cambridge) Ltd, UK*
Dr. Jacques Zuber	*Philip Morris International, Switzerland*

SUMMARY

The advent of modern biotechnology and the accelerated pace of scientific discovery in the areas of genetic modification (GM) and genomics are introducing fruitful developments in both the medical and agricultural sciences. However, the presence of genetically modified crops in food products has recently become a widely debated and contentious issue in Europe and other regions. Despite the substantial growth of the global GM crop area to 40 million hectares (as of 1999), public concern is being increasingly raised with regard to the potential impact of the technology. To respond to consumer preferences, Identity Preservation (IP) programs that can ensure traceability and segregation of GM crops are gradually being implemented for certain sectors of commodity crops (e.g., soybean and corn). Moreover, in many countries legislation has been introduced that mandates the labeling of food which contains or has been produced from a genetically modified organisms (GMO). Reliable tests are therefore being developed to identify transgenic materials in food by the detection of DNA (deoxyribonucleic acid) and proteins.

Although applications of biotechnology to tobacco and other *Nicotiana* species have been limited to the use of this plant as a model system and the evaluation of traits in field experiments, there are examples where new GM tobacco varieties have been developed and are awaiting approval before commercialization. Analytical methods that can detect genetic modification in tobacco are being developed in several laboratories.

During 1998, the CORESTA Scientific Commission has held discussions on this subject and recommended the creation of the Task Force Genetically Modified Tobacco - Detection Methods, which was officially inaugurated in December 1998 with the following mandate (quoted in the original wording):

- Review methods for detecting genetically modified tobacco (including cured leaf, tobacco products, plants and seed);
- Propose principles for sampling procedures.

Accordingly, the Task Force has thoroughly investigated these issues and has prepared a report, which reviews the recent scientific literature in the area of GM crop analysis. Eight Sections and three appendices provide details on the scientific and technical aspects of methodologies that can detect the presence of introduced DNA or novel proteins in a tobacco plant, and on sampling principles.

After the Introduction (Section I), the review starts with Section II on Sampling Strategies and Methods, which summarizes the main steps of a sampling plan and includes a brief glossary of terms that are frequently used in describing sampling procedures. ISO standards and practical approaches with potential relevance to GMO analysis are examined. Moreover, this section emphasizes the difficulties in devising a suitable sampling strategy until acceptance criteria (*de minimis* thresholds) have been defined and established for tobacco.

Section III, Sample Preparation and DNA Isolation Methods, summarizes the main strategies for the isolation and purification of plant genomic DNA suitable for the polymerase chain reaction (PCR), and highlights the importance of selecting the appropriate protocol for materials such as cured tobacco leaf.

Section IV, PCR-Based Screening Methods, reviews PCR methods to detect unique fragments of DNA that are very widely used in plant genetic modification. These fragments such as the 35S promoter from cauliflower mosaic virus, and the *nos* promoter and *nos* terminator from *Agrobacterium tumefaciens* are therefore diagnostic of genetic modification.

The advantages and drawbacks of confirmatory methods (e.g., nested PCR, nucleotide sequencing) are also highlighted.

The key issues of method validation (between different laboratories) and availability of GM tobacco reference materials, which are also discussed in Section VIII, are formulated here.

Section V on Semi-Quantitative and Quantitative PCR Methods includes a mathematical evaluation of PCR product formation, which shows the limits of conventional PCR in quantitative assays. The main approaches for semi-quantitative and quantitative determination of GMO levels, i.e. the quantitative competitive PCR (QC-PCR) and the real-time PCR, respectively, are reviewed and analyzed in detail.

Section VI, Immunological Detection Methods, provides insight into testing methods required to detect and quantify specific novel proteins in GM plant varieties, with a focus on adaptations of the widely used immunological assay ELISA (enzyme-linked immunosorbent assay), which is reported to be less sensitive than PCR.

Section VII, Selectable Genetic Marker-Based Detection Methods, discusses the use of GM crop analytical strategies based upon identifying the presence of selectable marker genes in transgenic plants. The activity of marker genes such as the neomycin phosphotransferase II (NptII) can be detected in seeds or plantlets directly (e.g. assessing the impact of a selective agent in seedling tests), or indirectly (e.g. enzymatic assay, bacterial complementation). However, at present these methods are less sensitive than PCR and are difficult to apply to cured leaf.

In Section VIII, Official and Validated Methods, the Task Force sets out to examine an additional dimension of the GMO analysis technical issues, that of international method validation and standardization. The first part of the section reviews the current status of official and validated methods in Switzerland, Germany and the European Union. The second part summarizes the activities of working groups within national and international bodies such as the French Standardization Association (AFNOR) and the European Committee for Standardization (CEN), that have recently acknowledged the need for harmonized protocols to validate GMO analysis methods. A critical issue appears to be the lack of a central repository with both information on GMOs (e.g., sequence data, primers, proteins) and samples (e.g., DNA, proteins, antibodies) of conventional and GMO reference materials.

The three Appendices include information from practical laboratory experiments conducted at laboratories of Task Force members, and a list of internet addresses of public databases featuring information on GM plant field trials.

Finally, a Glossary of Terms has been included to facilitate the reading of the Report by those readers who may not be familiar with genetic modification of plants and GMO analysis techniques.

I. INTRODUCTION

As modern crop biotechnology is coming of age, the commercial cultivation of genetically modified (GM) plant varieties has grown to 40 million hectares distributed in twelve countries [1], which represents approximately 4 % of the total world acreage. Most current GM varieties feature enhanced agronomic (or input) traits, such as built-in protection from pest and disease, that can reduce production costs and provide increased flexibility in crop management. More recently, genetic modification has been successful in enhancing quality (or output) traits of crop plants, e.g. modified fatty acid or amino acid profiles in vegetable oils and cereals, respectively.

However, the presence of genetically modified organisms (GMOs) in food products is becoming increasingly controversial in Europe because of public concern with regard to the potential impact of GMOs on the health and the environment. To respond to consumer preferences, the biotechnology, trade and food industry are gradually implementing Identity Preservation (IP) programs that can ensure traceability and segregation of GMOs in certain sectors of commodity crops, such as soybean and corn. Moreover, recent EU regulations have introduced GMO labeling provisions that require verification of the GM or non-GM status by the use of methods for the analysis of DNA (deoxyribonucleic acid) and proteins [2]. This has led most official government and several industry laboratories to develop analytical methods that can assist in the enforcement of the labeling provisions and in the implementation of IP programs by tracing GM agricultural materials throughout the entire supply chain, from seed to shelf.

Tobacco (*Nicotiana tabacum* L.) has been the most widely used model system for investigations in the areas of plant physiology, biochemistry, molecular biology and genetic engineering, and was the first plant to be genetically modified in 1983 [for a review, see 3, 4 and 5]. Potential commercial applications of biotechnology to tobacco as a crop have been limited to field trials conducted in several countries (see Appendix I), although there are examples where new GM tobacco varieties have been developed and are awaiting approval before commercialization.

Given the importance of GMO analysis methods in the research, development and monitoring of GM tobacco varieties, CORESTA considered establishing a Task Force to review the scientific and technical aspects of methodologies that can detect the presence of newly introduced DNA or novel proteins in a tobacco plant. The Task Force Genetically Modified Tobacco - Detection Methods was therefore created within the Phytopathology Study Group of CORESTA in August 1998 and was officially inaugurated in December 1998 with 24 tobacco industry members from 13 countries.

This report is an overview of recent scientific literature in the area of GMO analysis and also includes information from practical laboratory experiments conducted at laboratories of Task Force members.

Note. The use of trade names and commercial sources throughout this report is for information purposes only and does not imply endorsement by CORESTA or CORESTA Task Force members. In addition, the views expressed in the document are those of the individual authors and may not necessarily reflect those of CORESTA.

References

1. James C. (1999) Global review of commercialized transgenic crops: 1999. ISAAA Brief No. 12, 1999, The International Service for the Acquisition of Agri-biotech Applications, Ithaca, NY.

2. Ward, M.R. (1999) A review of the regulations governing the production and use of genetically modified crops. In: Recent Advances in Tobacco Science, Proceedings of the 53rd Tobacco Science Research Conference (TSRC), pp. 91-91, 12-15 September 1999, Montreal, Canada.

3. Brandle, J.D., and Bai D. (1999) Biotechnology: uses and applications in tobacco improvement, in Davis, D. Layten and Nielsen M.T. (Eds.), Tobacco: production, chemistry and technology, Blackwell Science Ltd., Oxford, UK.

4. Gadani, F., Ayers D., Hempfling W. (1995) Tobacco: a tool for plant genetic engineering research and molecular farming, Part I. Agro-Food Industry Hi-Tech 6:19-24.

5. Koller D. (1999) Genetics and the future of tobacco: issues and research opportunities. In: Recent Advances in Tobacco Science, Proceedings of the 53rd Tobacco Science Research Conference (TSRC), pp. 37-60, 12-15 September 1999, Montreal, Canada.

II. SAMPLING STRATEGIES AND METHODS

Introduction 8
General Principles of Sampling 9
ISO Sampling Standards with Relevance for GMO Testing 10
Practical Approaches 11
A Few Notes of Advice and Caution 13
References 14

Introduction

There are many reasons why one would want to sample tobacco or tobacco products for analysis. The determination of essential leaf tobacco constituents, such as nicotine, sugars and proteins, and of agrochemical residues are routine procedures. Frequently, samples are taken from batches to make smoking articles for sensory testing. The measurement of the "tar" and nicotine yields in cigarette smoke is probably the most important analytical practice in the Industry. Last but not least, the examination of the crop or of finished product for transgenic tobacco requires appropriate sampling.

This is to say that sampling procedures for GMO testing are quite comparable and similar to sampling in general and can, in principle, often be done in combination. Special attention, however, must be paid to certain aspects which are specific for GMO testing. Here, it is important to differentiate between the two principal reasons of all analytical determinations.

On the one hand, the purpose may be the assessment of chemical or physical characteristics, on the basis of means, for a certain quantity of material or product. These means may then be taken as absolute numbers (e.g., nicotine and sugar content of raw tobacco) or compared to existing targets (e.g., weight and diameter of cigarettes) or evaluated in reference to existing ceilings (e.g., maximum residue levels for agrochemicals on raw tobacco or in finished product; yields for "tar" and nicotine in cigarette smoke).

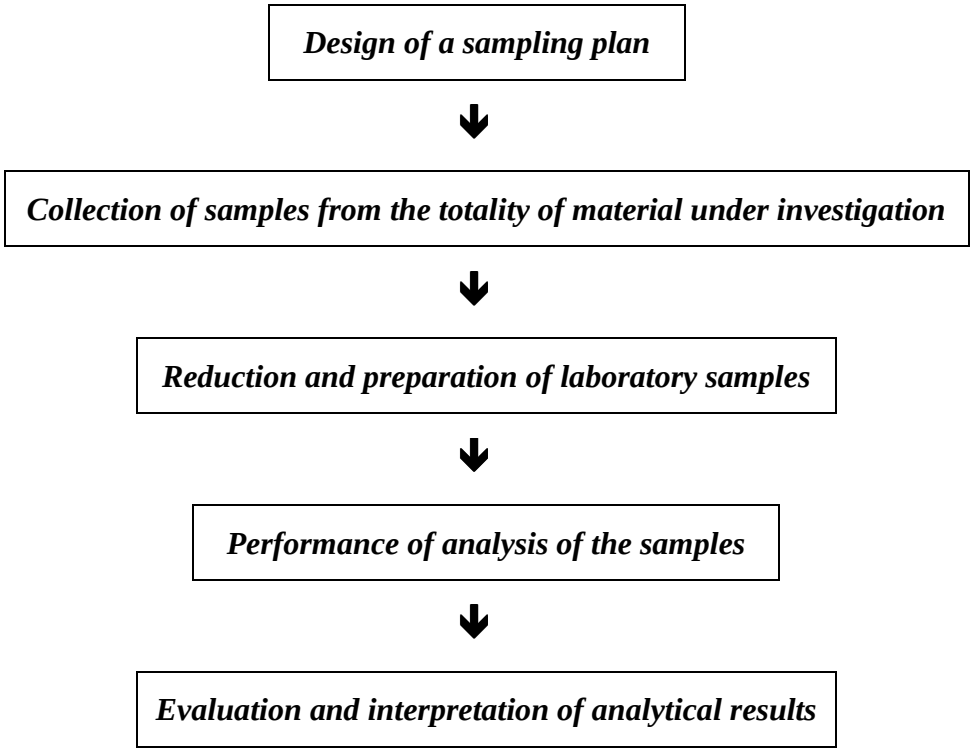
On the other hand, the determination of constituents or contaminations may be performed to demonstrate their absence (in practice, presence below the limit of quantitation) from the material or product under investigation. This is done in cases where, for instance, the maximum residue level for an agrochemical is set at zero. Similarly, under today's conditions, it is reasonable to assume that most analyses for transgenic tobacco have the purpose of demonstrating its absence from the material of interest.
However, it must be emphasized that no sampling strategy can be devised that would allow one to guarantee the total absence of transgenic material. Until acceptance criteria (thresholds) are defined and established also for tobacco (products), devising a suitable sampling strategy will always be difficult.

General Principles of Sampling

Sampling is the well-planned collection of representative parts from a larger amount of material for the purpose of analyzing certain physical or chemical properties. It may be required for two reasons: Either the whole amount of material under investigation is so large that it cannot be subjected to the analytical procedure in its entirety, or the method of testing is destructive (which is the case with most chemical methods) and no intact material is left after the analysis for commercial use or any other purposes.

Good sampling is an essential requirement for valid analytical measurements. Practical experience has shown that, not infrequently, too little attention is given to proper sampling techniques in relation to the analytical methodology subsequently applied to the samples. In these situations, the quality of the analytical procedure may be (far) superior to the quality of sampling – without true value being added to the whole investigation.

The principal steps in sampling and execution of analyses are the following:



The first three steps are discussed in this Chapter (the others being dealt with elsewhere). For this purpose, it may be helpful to explain a number of terms that are frequently used in describing sampling procedures. The following definitions are generally accepted and have been slightly modified and adapted for easier use in tobacco sampling:

Consignment	the quantity of tobacco dispatched or received at one time and covered by a particular contract or shipping document. It may consist of one or several lots
Lot (batch)	tobacco, presumed to be of uniform characteristics, taken from the consignment and available for analysis; a tobacco lot is generally from the same origin and variety, often specific for leaf position, color, ripeness, etc.
Sampling unit	a unit of the lot which is packaged separately (bale, case, carton, basket, sack); for bulk tobacco, a sampling unit should be 100 kg
Increment	a small quantity of tobacco taken from a single position in the sampling unit or the lot, generally by means of a sampling device; a series of increments is taken from different positions of the sampling unit or lot
Basic sample (aggregate sample)	the quantity of tobacco obtained by combining and mixing the increments taken from a specific sampling unit or lot; if the basic sample is further subdivided into identical portions, these are called *reduced samples*
Laboratory sample	a quantity of tobacco taken from the basic sample (or reduced sample) and intended for analysis; at this point, reference samples should routinely be set aside (e.g. for the seller, the buyer and for control analyses)
Test sample	a sample prepared from the laboratory sample according to the test procedure specified in the test method

ISO Sampling Standards with Relevance for GMO Testing

There are a number of official standards in place for the sampling of tobacco and tobacco products. It is worthwhile to examine them with respect to their utility in the present context.

The International Standard ***ISO 8243:1991 "Cigarettes - Sampling"*** was originally developed by ISO/TC 126 in 1988 and re-issued in 1991. It describes the sampling of finished product (cigarettes), essentially for the purpose of determining, and ranking, smoke yields for "tar" and nicotine. The standard contains plans for sampling both at the manufacturer's (importer's or wholesaler's) location, with samples being taken at a certain point of time or over a given time period, and from the market.

In practice, however, the question of whether transgenic tobacco is contained in finished product at all or not should reasonably be examined, and controlled, at the level of the raw material, tobacco. Therefore, the sampling of cigarettes for GMO testing according to ISO 8243:1991 may be of limited practical importance.

The document which obviously has the highest practical relevance to the question under discussion is the International ***Standard ISO 4874:1981 "Tobacco – Sampling of batches of raw material – General principles".*** It is currently under revision in the form of a first working draft (ISO/TC 126 N 563 of June 1997). The standard "specifies the general principles to be applied when sampling batches of raw tobacco in order to assess either the mean value of one or more of its characteristics, or the heterogeneity of one or more of its characteristics." In general, heterogeneity describes the variability of a characteristic measured in different batches. Heterogeneity, of course, also exists if some analytical values are zero (indicating the absence of the analyte) and others are not zero (demonstrating its presence in certain amounts). Therefore, the standard does have relevance for sampling if the absence or presence of a certain characteristic (such as transgenic tobacco) is to be assessed.

The standard is applicable to the sampling of batches of various types of leaf tobacco as well as pre-treated raw tobacco. It provides a range of definitions, guidance for contractual arrangements, recommendations for sampling and sample storage equipment, and detailed rules for the sampling procedure (including the selection of sampling units, the number and kind of samples taken and the sampling report).

With GMO testing in mind, special attention should be given to the comment in Note 5 of ISO 4874:1981 on heterogeneity testing and the recommendations regarding the treatment of damaged sampling units (in Section 6.2 of ISO 4874:1981).

Turning the attention now to the other end (the beginning) of the production chain, there is no specific standard for the sampling of tobacco seeds. Practical guidance may be taken from a document developed by ISO/TC 34, namely the International Standard ***ISO 950:1979 "Cereals – Sampling (as grain)"*** if appropriate considerations of scale-down are being applied.

Finally, if a specific sampling standard were to be developed for tobacco GMO testing, it would be quite prudent to take guidance from the International Standard ***ISO 7002:1986 "Agricultural food products – Layout for a standard method of sampling from a lot"***.

Practical Approaches

For all practical purposes in today's business environment, the determination of transgenic tobacco has the objective to demonstrate its absence from, or under special circumstances its presence in, a given lot of commercial tobacco. This is of overriding influence on the development of sampling plans. Passing judgement based on available test results, the presence of transgenic varieties does not seem to be wide-spread in today's commercial tobaccos. Transgenic material is only expected in crops from certain areas (in certain varieties and from certain crop years). Consequently, in practice, sampling plans will focus strongly on those tobacco lots which are suspected to contain transgenic material.

On the other side, one should not be too reliant upon available information regarding the presence or absence of transgenic material in commercial tobacco. Therefore, duty of care requires also a screening-type approach to sampling, generally within the limits of resources and practicability.

Regarding sampling procedures for tobacco lots, there are not yet any binding standards or any generally accepted procedures. However, it is possible to draw conclusions, and derive recommendations, from the experience gathered to date in many laboratories. One has to recognize that generally the transgenic tobacco cannot be expected to be evenly distributed in a lot, frequently not even in a sampling unit, such as a carton or sack. Therefore, if sampling units of concern have been identified, based on information about their origin, it is recommended, as a matter of principle, **to take as many incremental samples as possible**. In practice, however, the number of increments will be limited by technical practicability and by the available testing capabilities. An economical way of examining a large number of incremental samples for the absence of transgenic material is pointed out below.

The examination of green tobacco offers an opportunity to assess the presence of transgenic material at an early stage in the supply chain. The testing of small green plants, following germination and before they are fully grown and subsequently mixed up with other product, looks like a very attractive option to ascertain whether a crop will be free from transgenic material or not because the work-up and analysis of this kind of green material is technically simpler than working with cured leaf or directly with seed tissue.

Turning next to raw (cured) tobacco, samples may have to be taken from packaged tobacco leaf, from tobacco ribs (quite important in practice), from tobacco dust (out of packages, from the factory floor or in the form of milled material) and from cut tobacco (during processing). Both tobacco ribs, tobacco dust and cut tobacco can be assumed to have been subject to some kind of mixing and blending which offers a certain advantage in terms of representativity but, at the same time, has probably "diluted" any transgenic material present. The collection of samples is easy in practice, and scoops, shovels and tweezers are the preferred collection devices.

From packaged (pressed) tobacco, samples are usually taken by means of long cylindrical samplers with a sharp round end which are driven through the packaged tobacco to produce so-called auger samples (Bohrmuster, échantillons). In case of pressed strips (which often come in cartons of up to 400 kg weight) it is advisable to take, at minimum, one incremental auger sample each in the vertical and the two horizontal directions. If more incremental samples are collected from a carton they should be taken in identical numbers in the three directions and in a surface pattern as widely spread-out as possible – the objective being always to make the sampling process as randomized as possible (9 or 12 increments are desirable). Another rather simple sampling approach is to remove the packaging material from the top surface of the carton and to cut, from the pressed strips, one or several rectangular incremental pieces.

If a larger number of cartons needs to be examined, it is advisable to take the auger incremental samples (or the cut-out pieces) from every fifth or tenth carton. Auger samples or cut-out pieces may also be taken from pressed oriental tobacco leafs which generally come in bales of much smaller size (typically around 30 kg).

Dried commercial tobacco intended for cigar wrapper manufacturing comes in the form of layered whole leafs (bundles). Here, the collection of auger samples or the cutting-out of pieces is only recommended if the intact leaf structure does not have to be preserved. Otherwise it much rather makes sense to untie the leafs and select individual ones for testing. In this case one must be aware of the fact that possible heterogeneity within the lot may lead to problems.

The combination of incremental samples (auger samples, cut-out pieces or particulate samples such as dust, ribs or cut tobacco) to produce the basic sample should be done according to established laboratory practices by cutting, mixing and milling. Care should be taken to ensure that no significant amounts of material are lost during these operations and that representativity is maintained during reduction and sub-sampling steps. Such steps should be kept to a minimum. In the end, most analytical laboratories prefer to receive laboratory samples having weights between 0.5 and 1 kg.

Presently, testing for genetic modifications is much less frequently done on the finished product, cigarettes, or on tobacco seeds although this may become more important in the future. For the sampling of cigarettes, the International Standard ISO 8243:1991 provides reliable guidance.

With respect to seeds, available collection and testing resources will determine sampling intensity. Ideally, the number of incremental samples taken should be as high as the number of individual seed production sites in terms of year, location and variety. Typically, sample size is around 100 mg. To economize the testing process, a relatively large number of seed samples may be combined into one laboratory sample if in fact only proof of absence of transgenic material is required (see below).

From both an economical and practical point of view, it is highly desirable that a standardized procedure for seed sampling should be developed and adopted on a global scale. It would enable analytical laboratories to conduct the analyses in a timely fashion, rather than with a high degree of urgency after the tobacco harvest when the grown crop is close to, or already in the process of, being marketed and commercial decisions have to be made rapidly.

A Few Notes of Advice and Caution

Sampling and analysis are linked inextricably. In spite of this, the task of taking and preparing samples is often assigned to people with insufficient training and experience while the analytical methods are being pushed to ever higher levels of accuracy, precision and sensitivity. But even the most sophisticated analytical results are of limited value if the preceding steps of sampling and sample preparation are based on inadequate strategy and sloppy execution. Therefore, it is absolutely necessary to ensure that those taking and preparing samples are properly trained and that sampling protocols are followed faithfully.

It is always worth emphasizing that good practices in handling samples need to be observed particularly in GMO testing because of the delicacy and sensitivity of the analytical methodology. It must be ensured that the test samples are not contaminated or transformed in any way during sampling, transport or storage. The use of suitable and well-sealing containers and the observation of storage conditions appropriate for tobacco samples is mandatory. Full documentation is required for every sample as well as the clear labeling of all sample containers by code numbers, bar codes or attached label with relevant information. Finally, any samples submitted for analytical examinations as involved as those in GMO testing deserve to be accompanied by a proper sampling report.

Because the analytical methodology of testing for transgenic tobacco is capable to demonstrate the presence of very small amounts, great care needs to be taken to avoid the accidental contamination of lots and – even worse – of samples. Such contamination may

occur under two conditions. If a consignment or lot contains damaged sampling units which potentially contain transgenic tobacco, extreme caution has to be taken that the other sampling units are not contaminated. The remedial action which needs to be taken in such cases depends on the specific situation.

The other "opportunity" for contamination is the whole sampling and sample preparation process itself. Any step in the procedure should be done with particular attention to the possibility of cross-contamination. It is good analytical practice to relegate sample processing (including the preparation of the laboratory samples), the work-up of the test sample and the execution of the analysis itself to separate laboratory rooms. Apparatus, devices and labware used in sampling, mixing and milling should not be interchanged and thoroughly cleaned after each use. The sample containers which necessarily have to be moved between the laboratories should equally be subjected to rigorous maintenance procedures (labeling, cleaning, examination for leaks, etc.)

Some procedures during sampling may result in changes in the sample material and render it less amenable for testing. For example, milling may result in localized heating, which can in some cases affect the subsequent DNA extraction. Freezing, thawing, wetting and drying regimens can also damage the DNA to be extracted. It is wise that the procedures for sample preparation are developed in co-operation with the testing laboratory.

If the testing of commercial lots for transgenic tobacco is performed as a screening procedure, exercising due diligence, it is practical and may fully serve the purpose to do the GMO testing on (some of) the same samples that are prepared anyway for chemical, pesticide and sensory analysis.

In countries with closely watched tobacco variety programs or with stringent regulations regarding cultivation of transgenic crops, including tobacco, it may be justified to assume that no transgenic material is present in a consignment of different lots of raw tobacco or tobacco seeds. However, proof is required. Under these circumstances, it is economical to combine a larger number of incremental samples (20 or more) from the different tobacco or seed lots into one basic sample and test it. If the basic sample proves to be negative, all these tobaccos or seeds may be considered to contain no transgenic material, provided that careful consideration has been given to both the dilution due to sample combination and the limit of detection of the method in use. It must not happen that the concentration of transgenic material in the basic sample drops below the limit of detection as a result of excessive sample combination. However, if the first basic sample is positive, additional basic samples from fewer and fewer incremental samples will have to be prepared and tested until the exact source is identified.

Regulatory agencies have now begun to establish threshold levels for inadvertent commingling of transgenic raw material in food and feed (e.g., 1% being the threshold for food crops in Switzerland). A threshold, if exceeded, is the basis for requiring a GMO-related label on the packaging of a commercial product. On the other side, products which are shown to contain less than a fixed percentage of transgenic material may not require a label of this kind. While no such developments are currently under way for tobacco or tobacco products, they can be imagined to occur at some time in the future. If this in fact would happen, the question of the point of enforcement would become one of highest practical importance – with immediate effects on the choice of appropriate sampling procedures.

References

Although most of these references do not deal specifically with tobacco or tobacco products, the following brochures and papers may be quite worthwhile to read:

AFNOR (1999) Echantillonnage et contrôle des produits alimentaires (5th edition). AFNOR, Paris, France (http://www.afnor.fr).

Crosby, N.T. and Patel, I. (1995) General Principles of Good Sampling Practice. Published for the Laboratory of the Government Chemist by The Royal Society of Chemistry, Cambridge, U.K.

Stoeppler, M. (ed.) (1997) Sampling and Sampling Preparation - Practical Guide for Analytical Chemists. Springer Verlag.

Thompson, M. (1998) Uncertainty of Sampling in Chemical Analysis. Accred. Qual. Assur. 3: 117-121.

van der Veen, A.M.H. and Alink, A. (1998) A practitioner's approach to the assessment of the quality of sampling, sample preparation, and subsampling. Accred. Qual. Assur. 3: 20-26.

III. SAMPLE PREPARATION AND DNA ISOLATION METHODS

Introduction 16
Outline of Strategies 16
Overview of General Methods 16
  Alkali DNA Extraction 17
  Sample Grinding and DNA Purification 17
Overview of Tobacco Specific Methods 17
Summary 18
References 18
Detailed Methods 19

Introduction

Genomic DNA can be extracted easily from green leaf of tobacco (*N. tabacum)*, whereas the complex matrix of cured leaf can affect the suitability of DNA for GMO analysis This review describes some methods of sample preparation and DNA isolation applicable to tobacco, for the provision of DNA suitable for PCR reaction.

Outline of Strategies

Plant tissues contain significant quantities of polysaccharide and phenolic compounds, which can inhibit PCR. It is therefore important that DNA extracts should not be contaminated by these compounds. A number of methods have been used to isolate and purify plant for PCR reaction.

Overview of General Methods

Plant DNA isolation consists of two steps; (i) disruption of the sample tissue to release the DNA, and (ii) purification of the DNA from other compounds. Two basic methods for sample preparation are in common use. In one method, DNA is released by alkali treatment of the samples [1], and this DNA can then be used directly for PCR reaction. In the other method, the samples are ground in liquid nitrogen, and the DNA is subsequently purified from this ground material.

Alkali DNA Extraction

The merits of this method are its simplicity and that preparation of the DNA template is performed in a single tube. Plant tissue is briefly boiled in sodium hydroxide, neutralised with buffer, and re-boiled. The extract is then used directly for PCR. However, it has not been confirmed whether this method is applicable to cured tobacco leaf.

Sample Grinding and DNA Purification

The alternative approach is that the sample is frozen and ground to fine powder in liquid nitrogen. DNA can be extracted and purified from this powder by several possible methods, three of which are described below:

In one commonly-used method [2] the ground plant tissue is heated in a solution of the detergent CTAB, (cetyltrimethylammonium bromide) and then repeatedly extracted with chloroform. The DNA is then precipitated, washed and used for PCR.

A second common method of DNA purification involves phenol/chloroform extraction. This basic method has been modified successfully for isolation of tobacco DNA for PCR. The sample is heated in an extraction buffer before being repeatedly extracted with a phenol/chloroform mixture. DNA is precipitated from the extract, washed, and treated to remove any RNA present in the isolate.

Many kits are commercially available for DNA isolation from various plant materials. One such, the QIAGEN DNEasy Plant Kit has been used successfully in several laboratories. The method comprises incubation of the sample with extraction buffers followed by repeated column purification, using a set of pre-made buffers and columns supplied with the kit. These should ensure the reproducibility of the method and reduce the potential for sample cross-contamination.

Overview of Tobacco Specific Methods

As noted above, the Alkali Treatment preparation method has not been validated for application to cured tobacco leaf, whilst methods using liquid nitrogen-ground samples and subsequent purification can be applied to cured leaf. Modified CTAB, phenol/chloroform, and kit-based extraction methods have been developed for isolation of tobacco DNA for the purposes of PCR, and have been confirmed as applicable to cured tobacco leaf.

Summary

Numerous methods have been developed for the extraction of DNA from plant samples. These methods have been refined to enable the isolation of DNA of suitable quality for PCR analysis, and many examples are available both in the literature and as commercial kits. In order to accommodate cured tobacco leaf as the starting material, certain of these methods have been selected and modified, but their applicability should still be verified in individual laboratories.

References

1. Klimyuk, V.I., Carrol, B.J., Thomas, C.M. and Jones, J.D.G. (1993) Alkali treatment for rapid preparation of plant material for reliable PCR analysis. The Plant Journal. 3: 493-494.

2. Murray, M.G. and Thompson, W.F. (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acid Res. 8: 4321-4325.

Detailed Methods

Alkali DNA Extraction Method

1. Collect plant tissue into sterile microfuge tubes, containing 40µl of 0.25M NaOH, using the lid of the tube to sever the sample from the plant.
2. Incubate the samples in a boiling water bath for 30 sec and subsequently neutralize by addition of 40µl of 0.25M HCl, and 20µl of 0.5M Tris-HCl, pH8.0; 0.25% (v/v) Nonidet P-40 (Sigma), before boiling for a further 2 min.
3. Use the samples immediately for PCR.
4. Samples can be kept at 4º for several weeks. Incubate stored samples again for 2 min at 100º prior to PCR.

Note:

Quantities of sample materials for extraction by the following three methods are quoted for fresh green leaf samples. Suitable quantities of ground tobacco must be determined empirically.

Modified CTAB Extraction Method

Using 0.1-0.2 g (wet weight) of plant tissue, will yield approximately 50 µg of DNA.

1. Transfer the sample to a microfuge tube (1.5ml).
2. Add 300µl of 2 % CTAB solution (100mM Tris-HCl, pH 8.0, 20mM EDTA, pH 8.0, 1.4M NaCl, 2 % CTAB) and mix by inversion.
3. Incubate the mixture for 30 min at 65º.
4. Add equal volume of Chloroform / iso-amyl-alcohol (24:1, v/v) and shake gently for 5 min.
5. Centrifuge at 12,000 rpm for 15 min and transfer aqueous phase to a new tube.
6. Repeat steps 4 and 5.
7. Add 1-1.5 volume of 1 % CTAB solution (50mM Tris-HCl, pH 8.0, 10mM EDTA, pH 8.0, 1 % CTAB), mix by inversion, incubate at room temperature for 1 hour, and centrifuge at 8,000 rpm for 10 min.
8. Discard the supernatant and redissolve the precipitate by adding 400µl of 1M CsCl.
9. Precipitate and wash DNA by ethanol precipitation.

Modified Phenol/Chloroform Extraction Method

This method uses 1g (wet weight) of plant tissue (young leaves).

1. Prepare extraction buffer by adding 20µl of 2-mercaptoethanol to 1000µl of 2 x stock extract solution (100mM Tris-HCl, pH 8.0; 20mM EDTA; 1 % SDS; 1.5M NaCl), and preheat to (50º). Add 1 ml to sample and mix.
2. Transfer 600µl of the mixture to a 2ml tube and incubate for 20 min at 50º.
3. Add 500µl of phenol/ chloroform (1:1) to the lysate, mix by vigorously shaking and shake for at least 60 min, at 37-45º.
4. Centrifuge for 20 min at 15000 rpm at 25º and transfer 400µl of the supernatant to a new tube.
5. Add 400µl of 1x stock extract solution (50mM Tris-HCl, pH8.0; 10mM EDTA; 0.5 % SDS; 1.5M NaCl) and 500-µl of phenol/ chloroform to the supernatant and shake for 30 min at 37º.
6. Centrifuge for 30 min at 15000 rpm at 25º and transfer 600µl of the supernatant to a new tube.
7. Add 500µl of chloroform and shake for at least 10 min at room temperature.
8. Centrifuge for 30 min at 15000 rpm at 25º and transfer 450µl of the supernatant to a new tube.
9. Add 150µl of 10M ammonium acetate and 300µl of iso-propanol and mix by inversion.
10. Centrifuge for 30 min at 15000 rpm at 25º and discard the supernatant carefully.
11. Wash DNA with 70 % ethanol twice and treat with RNase.

QIAGEN DNeasy Plant Mini Kit Method

The maximum amount of sample material using this kit is 0.1 g (wet weight, young tobacco leaves) which should yield 20-25 µg DNA. Up to 1g (wet weight) of plant tissue, can be extracted using the QIAGEN DNeasy Maxi Kit which should yield 120 µg DNA.

All buffers and columns described below are supplied with the Kit.

1. Transfer the tissue powder and liquid nitrogen to an appropriately sized tube and allow the liquid nitrogen to evaporate. Do not allow the sample to thaw.
2. Add 400µl of Buffer AP1 and 4µl of RNase A solution (100 mg/ml), and vortex vigorously.
3. Incubate the mixture for 10 min at 65º. Mix 2-3 times during incubation by inversion.
4. Add 130µl of Buffer AP2 to the lysate, mix, and incubate for 5 min on ice.
5. Apply the lysate to the QIAshredder spin column in a 2 ml collection tube and centrifuge in a micro centrifuge for 2 min at maxium speed.
6. Transfer flow-through fraction from step 5 to a new tube with out disturbing the cell debris pellet.
7. Add 0.5 volume of Buffer AP3 and 1 volume of ethanol (96-100 %) to the cleared lysate and mix by pipetting.
8. Apply 650-µl of the mixture from step 7, including any precipitate, which may have formed, onto DNeasy mini spin column sitting in a 2ml collection tube. Centrifuge for 1 min at >6000 x g and discard flow-through.
9. Repeat step 8 with remaining sample. Discard flow-through and collection tube.
10. Place DNeasy column in a new 2 ml collection tube, add 500µl Buffer AW onto the DNeasy column and cenrifuge for 1 min at >6000 x g. Discard flow-through and reuse the collection tube in Step 11.
11. Add 500µl Buffer AW to DNeasy column and centrifuge for 2 min at maximum speed to dry the column membrane.
12. Transfer DNeasy column to a 1.5 ml or 2 ml microfuge tube and pipet 100-µl of preheated (65º) Buffer AE directly onto the DNeasy column membrane. Incubate for 5 min at room temperature and then centrifuge for 1 min at 6000x g to elute.
13. Repeat elution (Step 12) as described.

IV. PCR-BASED SCREENING METHODS

Introduction 23
Outline of Detection Strategies 23
Strengths of PCR 24
Weaknesses and Limitations 24
Overview of General PCR Methods 24
General 24
DNA Quality 24
Primer Design 25
PCR Reaction 26
Nested PCR 26
Multiplex PCR 27
Gene Targets 27
Generic Transgenes 27
Specific Transgenes 29
Recombinant Transgene Assays 29
Confirmatory tests 30
Nested PCR 30
Restriction Analysis 31
Probe Hybridisation 31
Sequencing 31
Overview of Tobacco Specific Methods 32
DNA Quality and PCR Reaction 32
Gene Targets 33
Generic Transgenes 33
Specific Transgenes 34
Recombinant Transgene Assays 34
Validation of Results 34
Interpretation 35
Confirmation 35
Nested PCR 36
Restriction Analysis 36
Probe Hybridisation 36
Sequencing 36
Summary 36
References 40
Primer Sequences 40

Introduction

Genetic modification of tobacco has been an established laboratory procedure for many years. The process of producing GM plants, through the direct introduction of transgenes, is largely independent of the genes involved, and there are common features in the structures of the gene constructs employed. Hence, despite there being a variety of transgenes in use, generic strategies can be devised for the detection of most genetic modifications.

Detection of transgenes, which may not be being expressed at that time, can only be achieved by analysis of plant DNA. By their very nature, transgenes are novel, and can be distinguished from the surrounding tobacco genome, but at the practical level this requires either some knowledge of the inserted DNA, or a guess at the likely candidate sequences.

Transgenes conferring resistance to viral and bacterial diseases of tobacco are known to have been employed and are therefore possible candidates for screening. Two or more transgenes may be introduced simultaneously, and one of these may be a "selectable marker", necessary only during the laboratory stage, but usually still present in the final plant, such as the *nptII* gene. In addition, other DNA features found in most modifications include the 35S promoter from Cauliflower Mosaic Virus, and the *nos* promoter and *nos* terminator from *Agrobacterium tumefaciens*.

Methods for detection of transgenes in tobacco have been developed in a number of laboratories and this review summarises the methods used. Where possible, practical details have been attributed to their originators. In other cases where information is proprietary or apocryphal, a general description has been given.

Outline of Detection Strategies

The most common methodology employed for GM screening is PCR-based detection of transgenes followed by gel electrophoresis and comparison with standard samples. The process uses the enzyme Taq DNA Polymerase to amplify minute quantities of transgene DNA from plant material to a detectable level.

Most methods for transgene detection in plant products have been devised for the analysis of food materials [1]. In general, two strategies are adopted, based either on the detection of one or more "generic" transgene sequence (usually the 35S promoter), or on screening for a range of specific sequences.

As a complication, the transgenes being sought may comprise segments of DNA derived from viruses or bacteria that occur in the natural environment. Detection of these sequences in a sample may then be misinterpreted as indicating GM plants. Even RNA viruses (e.g. TMV, CMV), may pose a problem, since under certain circumstances Taq DNA Polymerase is capable of converting RNA to DNA [2,3], which could be detected in the subsequent assay. However, this potential ambiguity can be avoided by judicious selection of the target regions of transgenes being detected (See "Recombinant Transgenes" below), and a reliable diagnosis can be made.

Strengths of PCR

A major advantage of a PCR-based detection-strategy is that it is extremely sensitive. It can be highly specific, whilst the basic test process is similar whichever genetic modification is being screened for. PCR can detect transgenes whether or not they are being expressed and it can even detect partial gene sequences and non-coding regions.

Weaknesses and Limitations

The sensitivity of PCR means that the process is sensitive to contamination, and its high specificity can also lead to a failure to detect slight variants of the target sequences. Quantitation of GM material present in samples requires detailed procedures and will be addressed elsewhere [4]. A threshold of detection must be determined, and zero tolerance is scientifically unverifiable at present.

Overview of General PCR Methods

General

PCR is a common procedure in most analytical laboratories and basic operational procedures are well established, and will not be described here. These procedures are designed to eliminate the possibility of cross-contamination between test samples and controls. In general PCR screening should be conducted in a separate location to sample preparation, and any other work with transgenes. Some laboratories decontaminate working surfaces and apparatus between runs using ultraviolet light [5], or chemical reagents such as DNA Zap [6].

DNA Quality

As described in Section III, a critical aspect of the testing procedure is the isolation of DNA of suitable quality for PCR. The process of harvesting, storing, and processing plant materials introduces many abuses to the DNA which ideally the molecular biologist would avoid, including exposure to intracellular enzymes and chemicals during natural senescence, heating, freezing and thawing, wetting and drying. These lead to DNA fragmentation, and so all DNA preparations should be assessed for their fragment size profile, usually by visual appearance on gel electrophoresis. One advantage of PCR, however, is that it can utilise relatively short DNA fragments.

The PCR process is also prone to inhibition by other materials co-extracted from the matrix, including metals, carbohydrates, phenolics, and salts. This is particularly a problem with food analysis. Most DNA extraction methods address these problems, often using silica-based purification stages, whilst use of strong denaturants in the process may also be advantageous.

DNA quality is normally assessed by a PCR test for an endogenous DNA sequence, which is present in the plant genome whether GM or not. This screen verifies the "PCR quality" status of the extracted DNA and guards against false negatives. The endogenous target should be a longer sequence than the transgene targets to ensure that genomic DNA fragments of a sufficient length are present. In some laboratories the PCR quality test is conducted simultaneously with the transgene test using multiplex PCR (Described below). Most laboratories test for the PCR of a chloroplast sequence, although some labs use a nuclear DNA sequence. This sequence could be specific for a particular crop plant, such as a Brassica seed protein gene or a soybean lectin gene, or alternatively a general "plant" sequence.

The chloroplast sequence test is well established [1,5,7,8] though a nuclear DNA assay has some advantages in that it specifically evaluates DNA from the cellular location where putative transgenes would be located. Nuclear and chloroplast DNA may differ in their ease of extraction, and may degrade to different extents during curing. Also, unlike chloroplast sequences, the nuclear sequences are present at a low copy number per genome, similar to that expected for transgenes, and the assay is therefore conducted at the appropriate level of sensitivity to detect any transgenes. The respective strengths and weaknesses of these two tests are illustrated in Table 1.

Table 1. Comparative advantages and disadvantages of two types of PCR ability tests using different endogenous genes.

Test	Advantages	Disadvantages
Chloroplast Gene	Established Process	Not representative of integrity of (nucleus-located) transgenes
		Different sensitivity to transgene detection
Nuclear Gene	Representative of integrity of (nucleus-located) transgenes	Not standardized
	Similar sensitivity to transgene detection	

Primer Design

Primer design is a crucial aspect of a successful PCR, but this has not always been the first consideration in the analysis of processed plant material, and primers have often been selected empirically. If the luxury of sufficient sequence information is available, some key elements in primer design are:

- Maximise the primer specificity
- Use primers of 18 - 30 nucleotides in length
- Avoid GC rich, repetitive, or self-complementary sequences
- Use a stable 5' end and a less stable, very specific 3' end to primer
- Select a high annealing temperature if possible
- Avoid primer pairs capable of binding to each other (primer-dimers)
- Keep the expected PCR fragment as short as possible

PCR Reaction

The basic PCR process consists of multiple components and reaction steps. Independent development in different labs has resulted in widely varying parameters, some of which are outlined below.

PCR is usually conducted in microtubes or microtitre plates, and reaction volumes vary from 10 to 100µl. The quantity of template DNA used also varies considerably. PCR reaction schemes differ with respect to times, temperatures, and numbers of amplification cycles, often for the same assay in different laboratories. For some assays conditions can be standardised which enables multiple assays to be conducted simultaneously on the same thermocycler, and also permits multiplex PCR. This standardisation however may mean compromise which can lead to sub-optimal reaction in some cases.

Most PCR tests are assessed by agarose gel electrophoresis, and results are scored visually as the presence or absence of a DNA fragment of the appropriate size. Certain PCR assays which yield a single, well-characterised product can be scored without the need for resolution by electrophoresis.

Nested PCR

A second round of PCR may be performed to add specificity to the assay (see Confirmation), or to increase the sensitivity. This "nested PCR" utilises a second pair of primers directed to the same target sequence, but internal to the first pair. These may or may not overlap the first pair of primers.

Nested PCR can be conducted by

(i) sampling a small aliquot from the first PCR into a second reaction mix,
(ii) adding new primers part way through the cycle, or
(iii) decreasing the annealing temperature part way through the cycle to enable the second pair of primers to come into play. This method precludes the need to re-open the reaction tube (potentially allowing cross-contamination), but does require careful primer design.

Table 2. Advantages and disadvantages of Nested PCR compared to basic PCR

Advantages	Disadvantages
Increased sensitivity	Increased opportunity for contamination
Increased specificity	Increase may not be significant
	More complex and lengthy

Multiplex PCR

Multiplex PCR entails the simultaneous PCR of two or more different target sequences in the same reaction tube. This can be used to save time and resources when screening for several transgene sequences, and also to verify the effectiveness of the PCR reaction by combining the "PCR-quality" and transgene screens in a single reaction.

A drawback of multiplex PCR is the potential competition for reagents between the two reactions, especially if the two targets are present in significantly different concentrations. Hence a GM target at low concentration may be "starved" of reagents and fail to amplify to a detectable level. This may be exacerbated by differing efficiencies of the two sets of primers. The optimal annealing temperature for each primer pair may differ and a multiplex reaction is of necessity conducted under compromise conditions. The two reactions may not be equally effective under such conditions.

When the copy numbers of the two targets differ significantly, as is the case when using chloroplast and nuclear sequences, empirically adjusting the ratio of primer concentrations can be used to compensate, but primer ratios established in this way may not be universally useful for different sample materials (leaf, stem, different tobacco varieties, and different degrees of degradation).

Comparisons of PCR efficiency are dependent upon the PCR products being of similar size. However, multiplex PCR demands that products differ in order to be resolved from each other on electrophoresis gels following reaction. Multiplex PCR also complicates any non-gel-based PCR detection method, since post-reaction separation of the products is required.
Table 3. Advantages and disadvantages of Multiplex PCR compared to basic PCR

Advantages	Disadvantages
Offers internal PCR control	Templates may compete
Saves resources	PCR may be non-optimised
	Requires gel analysis

Gene Targets

Generic Transgenes

The ideal PCR target is a generic sequence, found in all GM plants but not in the native plant nor in associated microbes. Ideally it should not occur in any processing additives either. Current generic targets include the 35S-promoter (most used), the *nos*-promoter, the *nos*-terminator, and the npt*II* gene. In addition, transfer DNA (T-DNA) sequences from *A. tumefaciens*-derived transformation vectors may be useful for transgenic plant screening. However, a survey of such "generic" targets in a selection of existing transgenic food materials indicates that even using combinations of targets does not guarantee infallible detection of genetic modification [See Table 4.].

Analyses will soon need to include second generation targets, including novel promoters for controlled expression. It is also likely that in future, vector sequences and especially antibiotic markers will be removed from the final product.

Table 4. Presence of "generic" transgenes in 28 different transgenic plants currently approved or seeking approval for release in Europe. [Data summarized from Reference 8.]

Transgenic sequence	Plants containing one or more	GM plants undetectable using this screen
35S	22	6
nos-terminator	16	12
35S, and nos-terminator	26	2
NPTII	17	11
35S, and NPTII	25	3
35S, nos-terminator, and NPTII	26	2
nos-promoter	7	21
35S, and nos-promoter	25	3

35S Promoter

This test is used in most laboratories screening GM plants and relies on the 35S promoter being the most commonly used sequence in GM plants. There are natural variants of this sequence and also variations in the detailed assembly of transgene constructs, so it must be confirmed that any primers used are both effective and equally efficient in all cases.

Neomycin phosphotransferase II (*nptII*, Kanamycin resistance)

The *nptII* gene is found in most early transformation vectors, and is often used in GM verification during the process of producing transgenic plants. However, since the gene is of bacterial origin, it is may be unreliable because of potential bacterial contamination [5].

Hygromycin phosphotransferase

The hygromycin phosphotransferase gene is used in some transformation vectors as an alternative selectable marker to *nptII*. A screen for this marker has been developed for food analysis [7].

nos-promoter and *nos*-terminator (*nopaline synthase* promoter and terminator)

These sequences are also common features of many transformation vectors. They are derived from *Agrobacterium tumefaciens*, a soil bacterium, and therefore may not be ideal for assays on roots [8] or other soil-carrying material. However, strains of *Agrobacterium tumefaciens* found naturally in soil are often avirulent, i.e. they lack the Ti-plasmid that carries the T-DNA and its oncogenes (e.g. *nos*).

Specific Transgenes

In addition to the screening for generic sequences, a number of laboratories have developed tests for specific transgenes, based upon well-characterised gene modifications. Such target genes include;

- *cry*1A gene (Bt - insect resistance)
- EPSPS gene (Glyphosate/RoundUp resistance)
- *bar* gene (BASTA resistance)
- polygalacturonase gene (delayed fruit ripening)

Recombinant Transgene Assays

When the transgenes under scrutiny are derived from micro-organisms present in the environment, false positives are a potential hazard. However, transgenes comprise microbe-derived sequences "spliced" to other DNA elements (promoters and terminators) in combinations which do not occur in the native micro-organism. Hence specific DNA junction regions occur, and by detecting these junctions, a reliable diagnosis can be made. Presence of the junction sequence confirms GM status of the plant for the particular transgene construct, since neither homologous sequences from the unmodified plant genome, nor genes from contaminating bacteria or virus will give a positive signals in these tests.

Table 5. Advantages and disadvantages of different transgene targets.

	Advantages	Disadvantages
35S Promoter	Most common transgene	Not 100% generic
	Established process	May become obsolete soon
	Viral source limited to certain crops	Some endogenous homologies
nptII	Common marker transgene	Not 100% generic
		May become obsolete soon
		Occurs in associated bacteria
		Some endogenous homologies
nos Terminator	Common transgene	Not 100% generic
	Established process	Occurs in associated bacteria
Specific transgenes	Specific	Limited use
		Requires information
		Can occur in "environment"
Recombinant transgenes	Highly specific	Very limited use
	Cannot occur in "environment"	Requires information

Confirmatory tests

Test samples which have been found positive for transgene sequences by PCR assay can be further analysed to provide information on the exact nature of the genetic modification. Several approaches are possible:

Nested PCR

Any of the PCR assays used in the basic analysis can be enhanced by further rounds of PCR to improve the specificity of detection. The method can employ either one (semi-nested PCR) or two (nested PCR) additional primers which will identify further regions of the transgene sequence within the PCR fragments already produced. A positive result is indicated by the production of a new PCR fragment of a defined length. Depending on the nature of the nested reaction, either two or one (the smaller) PCR products may be detected following the reaction. The method thereby confirms the presence and identity of the transgene by showing that the DNA fragment detected in the initial assay possesses the expected transgene sequence at various internal regions.

Restriction Analysis

This approach also bases the confirmation of transgene identity upon the detection of specific internal sequences. In many cases the full DNA sequence of the candidate transgene is known, which will include the identification of the presence and position of restriction endonuclease sites within it. By challenging the PCR fragment obtained in the initial transgene screen with one or more restriction endonucleases, the fragment is cleaved into smaller pieces of defined length. The presence of these cleavage sites within the fragment can then be used to confirm the identity of the transgene sequence. When using this confirmation assay, it is important that the initial PCR product is purified to remove reagents which may inhibit the action of the restriction endonucleases and so lead to false negative results [5].

Probe Hybridisation

Cloned copies of the transgenes or synthetic oligonucleotides can be used to confirm the identity of PCR fragments by Southern hybridisation. This method has advantages over the Nested PCR and Restriction Analysis approaches in that it confirms identity (or at least homology) along the full length of the transgene sequence rather than at short isolated sections. The hybridisation procedure can also be adjusted to tolerate certain variations in the transgene sequence and so can accommodate variants, without requiring foreknowledge of the variant sequence. It is important that the appropriate controls are included in this assay to ensure that the hybridisation is conducted under appropriate levels of stringency.

Sequencing

Whilst the previous methods depend upon the detected fragment possessing a known sequence, DNA sequencing could identify the sequence of the target analyte *de novo*, based solely upon homology at one or both ends of the fragment, and can therefore resolve doubts on uncertain results. Primer sequences used for the PCR assay may also be employed for sequencing of either the resultant PCR product, or even the genomic DNA itself. Since the transgene target sequences are relatively short, sequencing is not a problem if suitable quality template DNA is available.

Sequencing provides entire and absolute information on the transgene at the nucleotide level, and therefore can determine the precise variant of the transgene employed, including any artefacts introduced during vector construction, etc. The full identity and therefore possibly even the (laboratory) origin of the transgene can be determined. Sequencing can also provide information on the site of insertion of the transgene into the plant genome, and whether there are several insertions in the same plant. This information can act as a "fingerprint" for the transgenic plant and can discriminate between plants which carry the same transgenes, but which are derived from individual transformation events. Again this knowledge can lead to the identification of "pedigrees" and histories for transgenic plant material.

Table 6. Advantages and Disadvantages of different Confirmatory tests

Test	Advantages	Disadvantages
Nested PCR	Increased specificity	Increase may not be significant
	Established Process	Not independent of first test
Restriction Analysis	Increased specificity	Increase may not be significant
	Established Process	Requires information
		Prone to inhibition
Probe Hybridisation	Assesses whole of transgene	Homology, not identity
	Tolerant of minor variation	More complex technology
		Requires DNA probe
Sequencing	Absolute identity	More complex technology
	Can determine variants	Not routinely employed
	Can distinguish transformation events	
	Requires no additional info	

Overview of Tobacco Specific Methods

General PCR methods for detection of GM material have been developed in by several laboratories for application to cured tobacco, and although most protocols remain proprietary, many are largely similar to the method of Pietsch et al [1].

DNA Quality and PCR Reaction

The process of curing tobacco leaf causes much degradation of the genomic DNA and produces many chemical compounds which may inhibit the PCR reaction. Removal of these compounds during the DNA preparation stage is important for successful PCR.

The quantity of template DNA (tobacco) used in the PCR ranges from 5 - 100ng. Apocryphal data suggests that more template DNA yields more numerous positive results, though this may not be comparable between laboratories.

Primer concentrations used vary from 2 - 10.0 µM. In practice, primer lengths are normally around 20 - 25 nucleotides (the shortest reported being 16 [1]), and generally primers pairs are of approximately equal length (± 2 nucleotides). More variable is the GC content of the primers, which although commonly 50-55%, can range from 45 to 65% (with the lowest being 36%). In many instances pairs of primers are significantly mismatched with regard to their GC content.

PCR reaction schemes are broadly similar, reaction times varying with the thermocycler used, and ramping rates being set as fast as possible. Primer annealing temperatures for most assays are standardised at approximately 55°, irrespective of the primer characteristics. This enables multiple assays to be conducted simultaneously, and permits multiplex PCR. In some labs, individual PCR tests are optimised with respect to annealing temperature.

Most assays use 30-35 amplification cycles, although some labs use particular assays of 45-50 cycles. This may increase the sensitivity of the test, but care is necessary in these extended runs as the effect of minor contamination or PCR artefacts is significantly amplified.

PCR results are assessed by gel electrophoresis (1.4 - 4.0% agarose). Ideally a negative result is indicated by total absence of PCR product, though some tests require only the absence of a product of a particular diagnostic size. In some labs an arbitrary threshold is set before a result is called positive (e.g. greater than 2% of the intensity of the positive control).

Gene Targets

Generic Transgenes

35S Promoter

This test is used in all laboratories screening tobacco. Homologous sequences have also been detected in non-transgenic tobacco (or possibly from accompanying bacterial, fungal, and viral genomes present on the harvested leaf).

Most laboratories use the primers described by Pietsch et al [1] (see Appendix), which give rise to a 195bp PCR product. Other primers are used in some screening laboratories, often resulting in larger PCR products (300-400bp).

Neomycin phosphotransferase II (*nptII*, Kanamycin resistance)

This test is used in some laboratories, but it may be unreliable for cured leaf due to bacterial contamination [5]. Presence of *nptII*-homologous sequences have also been found in the unmodified tobacco genome. The test is used however for screening green leaf and seedlings [5].

Several different sets of primers have been used in NPTII screening, giving rise to products ranging in size from 400 - 800 bp. Those used by Pietsch et al [1] produce a much smaller PCR product of 173bp.

The *hygromycin phosphotransferase* gene is not commonly used in screening tobacco samples as yet.

nos-promoter and *nos*-terminator (*nopaline synthase* promoter and terminator)

These sequences are commonly used in screening tobacco. Since they are derived from a soil bacterium they may not be ideal, but although tobacco can indeed be a host to *A. tumefaciens*, the bacterium is rarely found on tobacco crops. This suggests that the likelihood of finding *A. tumefaciens nos* sequences on non-GM tobacco leaves is low.

Specific Transgenes

Tests have been developed for several specific transgenes, based upon intelligence that such modifications have been used in tobacco. The following genes have been targeted:

- TMV coat protein gene
- CMV coat protein gene
- PVY coat protein gene
- TMV 54kD gene
- Bt gene (*cry*1A)

Recombinant Transgene Assays

The risk of non-cruciferous plants being infected with CaMV is extremely low, although related viruses do infect tobacco. There is however a problem when screening for transgene sequences derived from soil bacteria, and also natural pathogens such as TMV and CMV. Even if a native virus does not contain DNA, the viral RNA genome may be converted to a form detectable by the PCR (as described above).

Assays have been developed to detect promoter-coding region junctions, including 35S-TMV and *nos-nptII*, which are the most common constructs using these genes, and 35S-chloroplast transit sequence [9]. Presence of the junction sequence confirms GM status of the tobacco for that particular transgene construct.

Validation of Results

A positive PCR result only means that a product has been successfully amplified, but the tobacco sample template may not necessarily be the source. Likewise, a negative PCR result only indicates that a product has not been amplified. It does not necessarily imply that the transgene is not present. These problems are addressed by the use of duplicate samples and appropriate controls. Each PCR run performed includes the following controls:

- Verified positive control.
- Verified negative cured leaf sample extract.
- No-DNA blank controls

In all assays, it is important that the best available controls are employed to validate the result. Ideally the positive control comprises a sample of DNA isolated from GM tobacco leaf containing the transgene(s) being assayed, and which has been cured, treated, and extracted in a comparable manner to the test samples. However, verified cured GM tobacco is not always available for some of the transgenes being screened, and the exact molecular identity of the potential GM sample may not be known. If the GM material is available in another form (e.g. seed), this can be used to produce a positive control DNA sample prepared from field grown air-dried leaf or even fresh seedling material. As a final option, DNA from the original transformation vector used to produce the GM plants may be used. However, these latter options are much less desirable as they are hardly comparable to the test samples.

Negative controls comprise non-transgenic cured tobacco DNA, and some laboratories also include "extraction blanks", i.e. dummy samples passed through the extraction process, but containing no starting sample material.

All control samples are prepared and evaluated by the same procedures as used for test samples. In the PCR tests, blank buffer mixes (No DNA) are also included.

Results from the PCR tests are only accepted if all control samples perform as expected. Results are usually rejected and the analyses repeated if PCR fragments are weak, or if several bands are seen. However, Pietsch et al [1] found several spurious PCR products from 35S and *nptII* assays of plant material, and these were ignored. Negative results ideally require the total absence of PCR product, although in some cases a very faint product is observed but discounted. This is presumably indicative of a low level of contamination in the laboratory.

Interpretation

Interpretation of the analysis can depend upon which PCR tests are applied. When only a single test (e.g. 35S) is performed, the result depends solely on the outcome of this test. If several sequences are screened, it is possible that conflicting results may be observed, yet there is no common course of action when this situation arises. Some laboratories retest the samples, some reject the samples if any test is positive, and others form a consensus based on the distribution of positive and negative results. The congruency of results depends on the exact construct present, and may be affected by there having been some re-arrangement of the transgenes in vivo. Results must be coherent and internally consistent, and compatible with any intelligence on the sampling regime.

Confirmation

Test samples which have tested positive are usually subjected to further analysis. In practice only Nested PCR and Restriction Analysis are used routinely.

Nested PCR

This is conducted as described above. The process adds little to the original specificity, and in practice when a PCR product is detected, its identity is usually "confirmed" by nested PCR. A more meaningful question to be resolved may be the source of the transgene, i.e. did it originate in the sample or has contamination occurred.

Restriction Analysis

This approach has been used to verify the 35S and nos-terminator assays [8]. Digestion of the 35S fragment with the restriction enzyme Asp700 (XmnI), and the nos-terminator fragment with NsiI yields diagnostic bands, which can be resolved on agarose gels. Other restriction sites are present in these fragments and have also been used in verification assays.

Probe Hybridisation

Confirmation by hybridisation is offered by some testing laboratories, and is recommended in food analysis [7,9]. A commercial kit is available for the detection of PVY infection using PCR and Oligonucleotide Hybridisation [10], and this might be applicable as a transgene screen.

Sequencing

Transgene sequencing is not being conducted routinely by any laboratory analysing tobacco. An unresolved problem with all confirmation assays is the action to be taken following an unconfirmed result. There is no common response to this situation. It must be borne in mind that a single nucleotide change due to a PCR-introduced mutation, or to a variant form of the transgene sequence would prevent reaction in both Nested PCR and Restriction Analysis.

Summary

Effective methods for PCR-based screening of transgenes in tobacco have been developed, using where appropriate similar technology introduced for food analysis. There is a common basic methodology in most laboratories, although details of the process are not standardised. However, differing strategies have been adopted as regards the choice of transgene targets, the criteria of validation, and the interpretation of results. This is in part due to the limited shared intelligence on potential targets and the dearth of control samples, but also to the fact

that the consequential action following a screening run may dictate precisely what technical information is required. A number of tobacco-specific issues remain, in addition to the technical considerations, not least of these being the definition of tolerance limits for the levels of GM material in tobacco.

Table 7. A summary of the features of a PCR-based screening process, identifying the merits and limitations of each step.

Basic PCR Process	**Advantages**	**Disadvantages**
	DNA relatively robust to curing process	DNA quality can vary with source
	Degraded DNA can be used	Matrix affects the process
	High sensitivity	Sensitive to contamination
	Single basic process for all tests	Can detect "environmental" DNA
	Can be automated	Not quantitative (Basic process)
	Permits physical isolation of transgene	Sensitive to parameter changes - cycle no, concentrations
	Independent of expression	May miss minor variants
	Can use non-coding sequences	Detection threshold must be defined
	Can use partial genes	

PCR Refinements	**Advantages**	**Disadvantages**
Nested PCR	Increased sensitivity	Increased opportunity for contamination
	Increased specificity	May not be significant increase
		More complex/lengthy
Multiplex PCR	Internal PCR control	Templates may compete
	Saves resources	PCR may be non-optimised
		Requires gel analysis

PCR Ability Test	**Advantages**	**Disadvantages**
Chloroplast gene	Established process	Not representative of integrity of (nucleus-located) transgenes
		Different sensitivity to transgene detection
Nuclear gene	Representative of integrity of (nucleus-located) transgenes	Not standardised
	Similar sensitivity to transgene detection	

Transgene Screens	Advantages	Disadvantages
35S Promoter	Most common transgene	Not 100% generic
	Established process	May become obsolete soon
	No "environmental" source (when screening tobacco)	Some endogenous homologies
nptII	Common marker transgene	Not 100% generic
		May become obsolete soon
		Occurs in associated bacteria
		Some endogenous homologies
***nos* Terminator**	Common transgene	Not 100% generic
	Established process	Occurs in associated bacteria
Specific transgenes	Specific	Limited use
		Requires information
		Can occur in "environment"
Recombinant transgenes	Highly specific	Very limited use
	Cannot occur in "environment"	Requires information

Confirmatory Tests	Advantages	Disadvantages
Nested PCR	Increased specificity	May not be significant increase
	Established process	Not independent of first test
Restriction Analysis	Increased specificity	May not be significant increase
	Established process	Requires information
		Prone to inhibition
Probe Hybridisation	Assesses whole of transgene	Homology not identity
	Tolerant of minor variation	More complex technology
		Requires DNA probe
Sequencing	Absolute identity	More complex technology
	Can determine variants	Not routinely employed
	Can distinguish transformation events	
	Requires no additional info	

References

1. Pietsch, K., Waiblinger, H.U., Brodmann, P. and Wurz, A. (1997) Screeningvaerfahren zur Identifizierung "genetechnisch veränderter" pflanzlicher Lebensmittel. Deutsche Lebensmittel-Rundschau 93: Heft 2, 35-38

2. Jones, M.D. and Foulkes, N.S. (1989) Reverse transcription of mRNA *by Thermus aquaticus* DNA polymerase. Nucleic Acids Res. 17: 8387-8388

3. Ruttimann, C., Cotoras, M., Zaldivar, J. and Vicuna, R. (1985) DNA polymerases from the extremely thermophilic bacterium *Thermus thermophilus* HB-8. Eur. J. Biochem. 149: 41-46

4. This report, Section V, Semi-quantitative and quantitative PCR methods.

5. This report, Appendix III, Chinese Method for Screening GM Tobacco.

6. Ambion Inc., Austin, TX, USA

7. Official Collection of Investigation Methods - 35 LMBG. (1996) Examination of foodstuffs. Evidence of genetic modification of potatoes by PCR (Polymerase Chain Reaction). Amplification of the modified DNA sequence and identification by hybridization with a DNA probe.

8. Joint Research Centre, Environmental Institute, Consumer Protection and Food Unit. (1998) Screening method for the identification of genetically modified organisms (GMO) in food. Detection of the CaMV 35S promoter and NOS terminator by means of polymerase chain reaction (PCR).

9. Official Collection of Investigation Methods - 35 LMBG. (1998) Examination of foodstuffs. Detection of genetically engineered soy beans by amplification of the altered DNA sequence using the PCR (Polymerase Chain Reaction) and hybridizing the PCR product with a DNA probe.

10. PE Applied Biosystems, Foster City, CA, USA: http://www.pebio.com/

11. Swiss Food Compendium 52B (1997). Methods in Molecular Biology.

Primer Sequences

Chloroplast [1,7,8]

5' CGA AAT CGG TAG ACG CTA CG
5' GGG GAT AGA GGG ACT TGA AC

Yield a 550bp product, i.e. longer than most transgene targets.

35S [7,8,11]

5' GCT CCT ACA AAT GCC ATC A
5' GAT AGT GGG ATT GTG CGT CA

Yield a 195bp product.

nptII [1]

5' GGA TCT CCT GTC ATC T
5' GAT CAT CCT GAT CGA C

Yield a 173bp product

***nos*-terminator** [1,8,11]

5' GAA TCC TGT TGC CGG TCT TG
5' TTA TCC TAG TTT GCG CGC TA
Yield a 180bp product.

Hygromycin phosphotransferase [7]

5' CGC CGA TGG TTT CTA CAA
5' GGC GTC GGT TTC CAC TAT

Yield a 839 bp product