The Guideline for Method Validation of DNA-Based Variety Identification Technology in plant

— Primarily by the SSR Analysis —

March 2008

Incorporated Administrative Agency National Center for Seeds and Seedlings

Table of Contents

1. Reference

Reference 1 Independence of markers

Reference 2 Demonstration of the complete or near-complete absence of null alleles in the marker region

Reference 3 Presentation format and deviation range of fragment size

Reference 4-1 Example of the use of presentation format in Reference 3

Reference 4-2 Example of inaccurate representation of fragment size in the use of presentation format in Reference 3

Reference 5 Method for examination of the reproducibility of marker performance

Reference 6-1 Fragment size by laboratory and marker

Reference 6-2 Example of deviations related to PCR and fragment analysis performed on different days

Reference 6-3 Deviations related to a sequencer, polymer, and a capillary

Reference 6-4 Example of mechanical deviations in fragment analysis

Reference 7 Method for fragment analysis in SSR

Reference 8 Guidelines for validation of qualitative analytical methods in the food industry

Reference 9 Example of developmental validation

2. Glossary

3. Bibliography

4. Working Committee

Introduction

Since the enactment of the Seeds and Seedlings Act in 1978 and subsequent full-scale introduction of the plant variety protection system, the number of registration applications of new varieties has steadily increased, and Japan is now one of the world's leading countries with respect to the annual number of registration applications of new varieties. Meanwhile, there is also an increase in the number of infringements of plant breeders' rights in which these new varieties are used without permission. The Policy about the Intellectual Property Strategy in the Ministry of Agriculture, Forestry, and Fisheries (Intellectual Property Strategy Headquarters in June 2006) stipulates that efforts should be made to promote the development of DNA-based variety identification technology and increase the number of applicable items as measures to prevent infringement of plant breeders' rights. As false labeling of food products becomes a significant social problem, people expect the early establishment and widespread use of DNA-based identification technologies that allow rapid identification of various varieties of plants.

DNA-based variety identification technologies have been developed for various plants i.e. testing research institutions, due to this the number of identifiable plant varieties is likely to increase. Validation of those technologies, which is important if these technologies are to be employed for actual identification of varieties, was performed for the identification of rice varieties with RAPD-STS markers by Otsubo *et al.* in the National Food Research Institute in 2004 and in a validation study for the identification of strawberry varieties with CAPS markers by Kunihisa, Matsumoto, *et al.* in the National Institute of Vegetable and Tea Science. However, other technologies have not adequately been investigated. Thus, as a policy of measures for DNA-based variety identification technology, the 'Working Group for Strengthening the Protection of New Plant Varieties and Promotion of Their Use' (Ministry of Agriculture, Forestry, and Fisheries on December 19, 2006) recommended that individually developed DNA-based variety identification technologies should be validated and that joint guidelines be established.

There are multiple guidelines for DNA analysis in human forensic medicine. In a food-related community, AOAC INTERNATIONAL (hereafter referred to as AOAC) and others have developed international harmonized guideline for analytical method validation; however, no guideline are now available for DNA-based plant variety identification technology.

In a FY 2007 support program for strengthening the competitiveness of the agricultural and food industry led by the Ministry of Agriculture, Forestry, and Fisheries, with the aim of establishing the infrastructure for appropriate practice of plant breeders' rights for new plant varieties and promoting the use of DNA-based variety identification technology, it was decided to review and prepare the guideline for method validation of DNA-based variety identification technology in plant and to develop guidelines for validation focusing on simple sequence repeats (SSR) that is the most representative method among DNA-based variety identification technologies.

The guidelines were developed by convening the 'Working Committee for Validation of DNA-Based Variety Identification Technology' involving members such as researchers engaged in validation and development of DNA-based variety identification technology (see Appendix 3-4. Working Committee)

and taking into consideration efforts for DNA identification in forensic medicine (such as the National Research Institute of Police Science).

Chapter 1: Definition and Scope of Validation

1. Definition of validation

As stated in ISO/IEC 17025, analytical method validation is defined as the confirmation by examination and the provision of objective evidence that the particular requirements for a specific intended use are fulfilled. In the context of DNA-based plant variety identification technology, the specific intended use means to use DNA-based plant variety identification technology for solution to infringement of plant breeders' rights. Particular requirements are for accurate detection of alleles (genotype) in those varieties when DNA-based variety identification technology is used. According to the guidelines for validation established by an FBI's working group on DNA analysis methods (Scientific Working Group on DNA Analysis Methods, or SWGDAM), the section "General Considerations for Validation of the DNA Analysis Procedure" stipulates that validation is the process by which the scientific community acquires the necessary information to:

(a) Assess the ability of a procedure to obtain reliable results.

(b) Determine the conditions under which such results can be obtained.

(c) Define the limitations of the procedure.

Thus, the validation process identifies aspects of a procedure that are unstable and must be carefully controlled and monitored.

In the development of this guideline, we believed that validation of similar technologies should be taken into consideration, because the details of validation depend on the type of technology to be validated. Thus, we consulted guidelines in the forensic field for detailed standards such as validation procedures and study items and also took into consideration the AOAC guidelines on which previous validation studies of rice and strawberry varieties were based. In parallel with the development of this guideline, we conducted a validation study of SSR for the identification of a sweet cherry variety developed in the Yamagata General Agricultural Research Center and incorporated the study results into these guidelines.

2. Scope of validation

In the SWGDAM guidelines, validation is classified into developmental validation and internal validation. The scope of validation in this guideline corresponds to that of developmental validation. Although, some confirmation of internal validation are included in the pilot study in this guideline, this guideline does not specifically address internal validation, because this guideline was prepared to provide guidance on validation of the procedures for variety identification technology. However, something comparable to internal validation may be necessary when analytical services involving DNA-based variety identification technology are provided. Thus, it should be appropriate to separately prepare guidelines together with accreditation of laboratories and verification of devices and reagents used.

This guideline focuses on the SSR method, but contains information commonly applicable to DNA-based variety identification technologies by RAPD-STS, and CAPS, etc. Thus, the titles of considerations common to SSR and others are labeled "(common)."

Note: **SSR method**

SSR is also called microsatellite or short tandem repeat (STR) . These indicate sequences of short repeat units on the genome. The difference in the length of alleles at a locus may be detected as polymorphism, because the number of repeats is polymorphic. Thus, the difference in the number of repeats may be detected by amplifying SSR regions by polymerase chain reaction (PCR) and analyzing the length. For variety identification, markers that allow specific amplification of SSR regions representing the genotype of each variety (combination of alleles) may be used to detect the genotype of each variety.

SSRs are abundant in eukaryotic genomes and highly polymorphic (because mutations due to replication errors of the number of repeats are likely to occur). Another advantage is that SSRs are amenable to PCR amplification, because most SSRs are 100 to 200 bp in length.[1] Moreover, SSR markers are highly reliable, because primers (SSR markers) for amplification of SSR regions are designed to target conserved regions, and SSR markers are co-dominant. On the basis of these features, the SSR method has recently been used as a means of determining parentage and identifying varieties in different species of animals and plants.

Chapter 2: Validation Procedures and Confirmation

Validation procedures include the characterization of Confirmation, which should be conducted by technology development institutions and others (developmental validation) and the subsequent conduct of an collaborative study to examine the reproducibility of the technology among different laboratories.

In addition, before the conduct of an collaborative study, a pilot study is conducted to ensure and assess the level of individual laboratories participating in the collaborative study.

In DNA-based variety identification technology, varieties are identified through the following four processes: (1) sample collection, (2) DNA extraction from samples, (3) DNA analysis, and (4) verification of analytical results against variety-specific data. Of these, processes (1) and (4) do not require demonstration of the reproducibility. Thus, the reproducibility of processes (2) and (3) should be demonstrated in the collaborative study.

The review and interpretation of the results of validation (Confirmation developed techniques and collaborative study) should be performed by the institution engaged in validation. On the basis of the validation results, the institution engaged in validation should document validated and unvalidated items in written procedures for the technology and clarify the range of validation.

1. Confirmation developed techniques

Confirmation developed techniques may include the following items, although it is not necessary to study all items. It depends on the type of DNA-based variety identification technology, but it is not necessary to study items that are unnecessary or have no influence on analytical results.

In contrast, some type or nature of technology may have confirmation for which a single-laboratory study is not enough to ensure the reproducibility and stability of analytical results in practical settings. Those confirmation should be examined in an collaborative study rather than in developmental validation.

A. Information on plants to be analyzed

a. Intra-variety polymorphism (common)

The frequency of intra-variety polymorphisms differs substantially depending on the type of plant propagation (seed propagation [autogamy or allogamy] or vegetative propagation). The presence of intra-variety polymorphisms makes it difficult to identify the variety by DNA-based variety identification technology. Vegetatively propagated plants have generally fewer intra-variety polymorphisms; however, it should be demonstrated that intra-variety polymorphisms are less likely to be present in the DNA region used for variety identification.

The frequency of intra-variety polymorphisms may be increased over time and by unintentional, artificial selection under different environments. To examine the presence of intra-variety polymorphisms, varieties that have been distributed extensively for years should be used, and temporally distant individuals (such as original strains and distributed strains) and geographically distant individuals should be collected to demonstrate that no inter-individual

polymorphism is detected with markers used for DNA-based variety identification.

b. Origin of plant variety to be analyzed (common)

When DNA-based variety identification technology is used to establish data on registered varieties and others, samples of the variety to be analyzed must be representative samples of that particular variety (authentic samples of the variety). Thus, it is necessary to accurately identify the origin of the variety analyzed.

For vegetatively propagated plants, samples should preferably be collected from original strains or trees (or their equivalents). If those samples are not available, samples free from contamination by other varieties should be used.

For seed-propagated plants, samples should preferably be seeds of the variety that was submitted to the Seeds and Seedlings Division of the Ministry of Agriculture, Forestry, and Fisheries at the time of application for variety registration.

c. Selection of variety of reference alleles

In SSR analysis, the size of fragments detected is slightly influenced by the reagents or analytical devices used. Thus, reagents and analytical devices should preferably be standardized as much as possible. For higher consistency of analytical data, it is desirable to select a variety of reference alleles for fragment size in advance and determine the size of fragments in the test variety on the basis of the distance from fragment peaks of the variety of reference alleles. When fragment size is determined with reference to a variety of reference alleles, the variety of reference alleles should always be tested in a DNA-based variety identification of the plant.

In the case of RAPD-STS or CAPS, in order to confirm that PCR amplification and treatment with restriction enzymes are appropriately performed, it is desirable to choose a variety that serves as a reference standard for each marker used and test that particular variety.

Thus, the variety of reference alleles should preferably be well documented with respect to the origin of its strain or tree (original strain or tree) and managed to avoid any mix-up with other varieties. Strains or trees must be amenable to sampling for DNA extraction, as DNA is used for each analysis.

The number of varieties of reference alleles used in SSR analysis should be determined with consideration of the number of alleles and the range of sizes detected with SSR markers used. When plant species with many alleles or a wide range of allele sizes are used, the peak-to-peak distance between variety of reference alleles fragments and test variety fragments is long, and the deviation in fragment size detected by fragment analysis is likely to be larger. Thus, it is desirable to add a variety of reference alleles to minimize the deviation .

B. Information on SSR markers

a. SSR marker performance

It is important to use markers that have high amplification efficiency, are highly polymorphic among varieties, and are easy to identify.

For SSR markers, it may be difficult to interpret the results of dinucleotide repeat SSRs, because of the influence of stutter bands (see Reference 7). Thus, the use of markers modified with pig-tailing should be considered.

Even when a technique is intended to interpret band patterns on a gel in a practical stage, it is desirable to perform fragment analysis with a sequencer in the development stage of markers and confirm that band patterns used for identification agree with the fragments.

b. Independence of markers

If the independence of markers is demonstrated, it is possible to calculate the probability that different varieties produce the same results with all markers used[2] and the variety identification rate (for example, the probability that two samples represent the same variety is 0.01% or less). The independence of markers may be evaluated by determining whether one marker and the other are known to be located in different linkage groups or determining whether two markers, if located in a single linkage group, are far distant from each other (see Reference 1).

c. Demonstration of the complete or near-complete absence of null alleles

When it is demonstrated that a marker used is not (or quite unlikely to be) associated with null alleles, the presence of a peak in fragment analysis (band on a gel) almost certainly indicates that primers anneal to the target locus. This condition eliminates the need to conduct a reproducibility study for each allele and makes the analysis efficient.

The complete or near-complete absence of null alleles may be demonstrated on the basis of the expected heterozygosity (He) in a population (about 100 varieties or more) and the observed heterozygosity (Ho) (see Reference 2).

C. DNA extraction (common)

a. DNA extraction method

To perform DNA-based variety identification, it is necessary to establish a method for extracting from samples adequate concentrations of DNA for PCR amplification.

For DNA extraction, it is desirable to use a commercially available kit. The use of self-prepared buffers and others may be associated with variable results, depending on the skill of the personnel, time from preparation to use, storage method, and the quality and manufacturer of reagents used for the preparation, and thus requires a study to ensure that the self-prepared reagents allow reproducible DNA extraction.

b. Sample site

One of advantages of DNA-based variety identification technology is that variety identification may be performed with only a part of the plant. To make use of this advantage, the portion (or expected portion) used for DNA extraction (variety identification) should preferably contain a part of the crop, seed, or seedling generally distributed.

It is necessary to determine the conditions for obtaining stable results, as the sample site may

influence DNA extraction efficiency, extraction procedures, or results of variety identification. In particular, tissues that develop or grow after fertilization should be carefully examined, which have some parts (seeds and fruits such as chestnuts) with a genotype (next generation) different from that of the variety analyzed.

c. Condition of samples

Some sample sites such as fruits continue to age and decompose day by day during the process of distribution. DNA is stable in living cells, but susceptible to degradation in dead cells, resulting in low extraction efficiency. Thus, it is important to determine the condition (limitations) of samples suitable for DNA extraction.

d. Sample storage

Appropriate storage of samples is important in quality control of studies.

Samples collected should be placed in a container and documented to avoid mix-up with other samples. To identify samples, the sample name, collection place, collection date, and the name of the person who collected the sample should be documented. Before storage of samples, it is desirable to take photographs showing the condition of the sample at the time of collection.

An appropriate storage method should be selected, depending on the sample site and the time required for DNA extraction.

When samples are frozen, attention should be paid to the following:

- Freeze samples quickly.
- Store at −30°C or below.
- Avoid placing too many samples in a freezer.
- Avoid thawing stored samples when opening and closing the freezer.

When liquid nitrogen is used, care should be taken to avoid contamination via damaged containers.

It is recommended that samples used for variety identification should be stored for a period during which re-testing etc. may possibly be needed. For the storage of PCR-amplified products, special care should be given to those points described in "D. Examination of the reproducibility of SSR marker performance" on page 10.

D. PCR conditions (common)

It is necessary to determine the appropriate PCR conditions for markers used (such as PCR cycle, concentration of DNA tested, and concentrations of reagents). It should be ensured that adequate PCR products are obtained without nonspecific reaction under the PCR conditions (particularly DNA concentrations) specified in written procedures.

When modifications to sample site and others result in different quality and quantity of DNA tested, appropriate PCR conditions should be determined accordingly.

E. SSR analysis

a. Dilution ratio of PCR products

It is desirable to determine the dilution ratio of PCR products suitable for fragment analysis, as the concentration of PCR products used in fragment analysis may influence the interpretation of results.

However, the dilution ratio should not be used as an absolute requirement, but as a guide, due to the amplification efficiency of PCR products may depend on the model of the thermal cycler, pipetting techniques, precision of pipettes, and type of Taq polymerase etc.

b. Determination of precision

Multiple data (measurements) obtained from DNA samples of a single variety should be used to establish the criteria for determination of the range of measurements within which measurements are considered the same. In SSR analysis, it is desirable to use a data presentation format on the basis of the distance from the fragment peak of the variety of reference alleles (difference in peak size, hereafter referred to as the distance) or establish the range of measurements that are considered to represent an identical fragment (allele) (see Reference 3 and Reference 4).

F. Considerations in processed goods (common)

When DNA-based variety identification technology is applied to processed agricultural products, attention should be paid to the following:

a. Quality and quantity of DNA extracted

DNA in processed goods is likely to be fragmented in the heating and manufacturing processes. Seasonings, preservatives, and others may interfere with DNA extraction.

Thus, it is important to select markers for short PCR amplified products and consider DNA extraction methods suitable for processed goods.

b. Mixture

Processed goods may contain materials other than the plant to be analyzed or other varieties of the plant to be analyzed.

It is thus necessary to determine the conditions for obtaining reliable results from mixed samples.

G. Need for publication in a peer-reviewed journal (common)

It is important for the developed DNA-based variety identification technology to be published in a peer-reviewed journal. This indicates that the technology has gained acceptance by two or more experts. The research paper describing the technology should preferably be cited in multiple papers. This is an important factor for which the results obtained by DNA-based variety identification technology are considered admissible evidence in court.

2. Collaborative study

An collaborative study is conducted primarily to demonstrate the reproducibility of the technology described in written procedures for DNA-based variety identification of individual plants. If poor results obtained in a study are followed by modifications to the written procedures, validation should be performed again to address the modified procedures.

A. Study design (common)

a. Number of laboratories

It is recommended that at least 10 laboratories should participate in the collaborative study from a statistical viewpoint. However, it may be appropriate for the sake of flexibly to determine the number of laboratories taking into consideration the quality of the participating laboratories and analysts (experience and past performance), prior use of markers in the technical field (such as previous use by multiple persons and previous use in the plant of interest published in peer-reviewed journals), costs for conduct of the study, and availability of participating laboratories (the number of laboratories is three at minimum and eight at maximum in the example of developmental validation shown in Reference 9).

b. Number of replicates

It is recommended that the number of replicates should be determined to fulfill the total number of samples (number of laboratories × number of replicates) required to evaluate the statistical reliability of the DNA-based variety identification technology. When fewer laboratories are available; however, it is not always necessary to impose too many replicates on a single laboratory taking into consideration the quality of analysts as well as the costs and labor required for the study. When errors (i.e. false positives and false negatives) occur in an collaborative study with a small total number of samples, the cause of error should be identified together with implementation of corrective measures. If the error appears to be confirmed to the technology, the study should be repeated at an equivalent or larger scale.

c. Variety of Reference alleles

All varieties of reference alleles should be provided for the collaborative study, as fragments derived from varieties of reference alleles are used as the reference standards for determination of fragments in the variety to be analyzed.

d. Others

Laboratories participating in the collaborative study should have capabilities (such as technical experience, past performance, and quality control) to produce stable results, except for potentially unstable factors associated with the DNA-based variety identification technology to be validated. When participating laboratories do not have quality assurance systems as evidenced by ISO 17025 accreditation(or its equivalent) or do not have adequate experience or prior performance with the technology, the institution responsible for the validation study should

implement a pilot study (Chapter 4) to assess and ensure the level of each participating laboratory.

B. Items examined for reproducibility (common)

Items examined for reproducibility in the collaborative study are "DNA extraction from samples" and "DNA amplification and fragment analysis." The reproducibility of these procedures should preferably be examined in a single process, because these are performed in a series of processes in actual analysis, but may be examined separately, as it is difficult to identify the cause when a study produces unsatisfactory results.

C. Examination of the reproducibility of the DNA extraction method (common)

DNA should be extracted from samples by the specified DNA extraction method and subjected to agarose gel electrophoresis or similar methods to examine whether the quality and quantity of DNA are satisfactory for PCR. When different sample sites require different methods for DNA extraction or modifications, the reproducibility should be examined for each sample site. In principle, a sample unit should consist of one leaf, one fruit, or equivalent so as to avoid contamination by different varieties.

The final decision should be made on the basis of the results of PCR amplification, as DNA may not be amplified by PCR even when an adequate quality and quantity of DNA is visible on agarose gel or by other procedures.

D. Examination of the reproducibility of the SSR marker performance

The examination of the reproducibility of the SSR marker performance aims to demonstrate that primers reliably anneal to a specific locus and that specific fragment size is stably detected.

For replicates in an intralaboratory study, fragment peak values may be compared directly. Comparison among laboratories with different analytical devices and reagents may be associated with larger deviations in peak values. Thus, it is desirable to use the presentation format for comparison as shown in References 3 and 4 (see References 6-1 and 6-3).

The reproducibility may be demonstrated by examining all fragment sizes detected with individual markers in test varieties (study method 1 in Reference 5). When the complete or near-complete absence of null alleles (alleles with mutations in a primer region that preclude allelic amplification) is demonstrated, investigation of the size of one fragment for each marker suffices (study method 2 in Reference 5).

To minimize data deviations in fragment analysis, each laboratory must pay attention to the conditions and procedures specified in "1. Confirmation developed techniques " on page 4 as well as the following considerations:

a. It is desirable to perform fragment analysis soon after PCR amplification.
   (This is because nonspecific fragments may be detected when PCR products are stored for a long time after PCR amplification.)

b. It is desirable to perform analysis of all varieties and replicates under identical conditions in

each laboratory, including the model of thermal cyclers, sequencers, and other equipment used for the examination of the reproducibility of marker performance, settings for devices used in PCR amplification, fragment analysis, and other equipment, and reagents.

c. Appropriate reagents for the specific model and settings must be used. Size standards suitable for fragment sizes detected should be used as much as possible.

d. Test varieties and varieties of reference alleles should preferably be compared in the same run on the same day (see Reference 6-2).

e. The same PCR products should be analyzed repeatedly, as mechanical deviations may occur during fragment analysis, (see Reference 6-4).

E. Documentation and submission of study data (common)

Institutions participating in the collaborative study must organize records on the following study data and submit them to the study coordinating institution within a specified period.

a. Validation study of the DNA extraction method

(1) Documents indicating the concentration and quality of DNA

Example: Photographs of agarose gel electrophoresis

(2) Documents evidencing PCR amplification of extracted DNA

Example: Images of fragments in SSR analysis, photographs of amplified products on a gel

b. Validation study of markers

(1) Documents for evaluation of whether the markers show their expected performance

Example in the SSR method

- Images of fragments in SSR analysis (for each marker and each DNA sample)
- Values of peaks that the analyst identified on the basis of images of fragments

Example in RAPD-STS and CAPS methods

(1) Photographs of gel showing band patterns of amplified products

c. Common considerations

(1) Name of device (models) and reagents used in the study (Table 1)

(2) Events noted during the study

Failure at work, differences in sample conditions and others, and other information on possible protocol deviations

Chapter 3: Validation of Modified Procedures for DNA Analysis

Modified procedures for validated technology must be evaluated for performance by analyzing identical or similar samples and comparing the modified procedures with the original procedures.

To demonstrate reproducibility, an collaborative study must be conducted again, depending on the type and level of modifications.

1. Modifications for which the reproducibility should be demonstrated at a single laboratory (common)

(1) Use of different models of devices

(2) Slight modifications to the DNA extraction method

(3) Use of different plants or sites for DNA extraction (with no modification to the extraction method)

(4) Addition of unconfirmed alleles in the use of markers with proven reproducibility (only for markers for which the complete or near-complete absence of null alleles is demonstrated)

(5) Slight modifications to markers (such as use of a different fluorescent dye)

2. Modifications for which the reproducibility should preferably be demonstrated additionally at multiple laboratories (common)

(1) In cases where the DNA extraction method is slightly modified and different plants or sites are used for DNA extraction

(2) In cases where a highly reliable marker (as evidenced by prior use in the plant of interest reported in a peer-reviewed journal or others) is newly added or modified

3. Modifications for which the reproducibility must be demonstrated from scratch (common)

(1) Major modifications to the DNA extraction method (modifications to the feature of the extraction method such as the use of a different kit)

(2) Addition of or change in markers

Chapter 4: Pilot Study

Laboratories participating in the collaborative study must have capabilities (such as technical experience, past performance, and quality control) to produce stable results, except for potentially unstable factors associated with the technology to be validated. At present, there seems to be only a few laboratories with established quality assurance systems (such as ISO 17025accreditation) for DNA-based variety identification technology and only a few laboratories with experience and proven track record on all DNA-based variety identification technologies established for individual plant species.

Thus, for the time being, institutions leading the collaborative validation study should play a central role in conducting an adequate pilot study before the conduct of collaborative study in order to improve the knowledge and experience of participating laboratories and their staff and establish systems for reliably producing valid results.

It may not always be necessary to conduct a pilot study if objective criteria to ensure the level of individual laboratories are established in the future.

1. Considerations in the pilot study (common)

The pilot study should preferably be designed to address the following considerations:

A. Review of the work processes in the collaborative study

In principle, the collaborative study is conducted according to the written procedures for DNA-based variety identification technology for each plant. In the pilot study, the processes should be reviewed to identify technical issues to be improved in each laboratory and written procedures requiring improvements (such as change in wording and amendments for the use of different devices) and establish the conditions for smooth conduct of the collaborative study.

B. Repeated analysis

In the pilot study, analysis should be repeated to examine the stability of results.

C. Examination of the ability to interpret results

In SSR analysis, it is necessary to examine the ability to determine the peak value for variety identification on the basis of images and multiple peak values obtained by fragment analysis. For this purpose, cases in which the interpretation requires some experience and knowledge should preferably be included in the pilot study (see Reference 7).

In RAPD-STS and CAPS methods, it is necessary to evaluate whether the analyst has the ability to correctly interpret the results on the basis of band patterns visualized on a gel. It should be ensured that the analyst well understands the method for interpretation specified in the written procedures, particularly interpretation of bands resulting from nonspecific amplification or incomplete digestion by restriction enzymes.

Chapter 5: Points to Consider

In the conduct of a validation study or the use of validated DNA-based variety identification technology, attention should be paid to the following points to increase the stability and reproducibility of results.

1. Work area (common)

For the prevention of contamination, the area for DNA extraction, area for the preparation of the PCR reaction solution, and area for manipulation of PCR products should preferably be physically separated from one another.

2. Sample grinding (common)

During sample grinding, care must be exercised to avoid contamination. Particular attention should be given to samples ground in a mortar, as scattered material will likely be found around the mortar. Thus, other samples should not be placed around the mortar, and gloves etc. should be changed before handling other samples.

Care should be exercised in avoiding thawing before grinding of the frozen samples (because thawing reduces the quality and extraction efficiency of DNA).

3. Determination of DNA concentration and purity (common)

Before extracted DNA is subjected to PCR, the concentration and purity (amount of contaminants) should be determined (to ensure that the concentration of DNA and contaminants such as protein do not influence the results of PCR).

Each laboratory may select appropriate procedures for the determination unless the coordinating institution designates a specific procedure. However, a spectrophotometer detects not only DNA but also RNA, and thus RNA should be adequately cleaned when a spectrophotometer is used.

4. PCR (common)

PCR must be performed with specific reagents under specific conditions as directed by the study coordinating institution.

It should be noted that, even under the same PCR conditions, the use of different models of thermal cyclers might lead to different PCR results (because of various conditions such as different temperature gradients).

5. Prevention of contamination (common)

It is recommended that sterile disposable plastic equipment should be used. Glass equipment, if used, should be sterilized by dry heat (at 180°C for 2 hours).

Gloves should be worn when handling DNA samples and PCR products.

Reagents used in other experiments should not be used.

Appendix 1: Reference Material and Discussion

1. Current status of the development of guidelines for validation

Guidelines for validation have been developed in different domains. AOAC guidelines for the collaborative study (2005)[3] provide the minimum criteria for the conduct of an collaborative study of qualitative analytical methods, although it is stated that the criteria are not part of the harmonized protocol (see Reference 8).

For human DNA analysis, SWGDAM guidelines have been developed.

In addition, for methods of qualitative analyses including DNA analysis, the International Union of Pure and Applied Chemistry (IUPAC) started drafting guidelines for collaborative studies of PCR-based methods for GMO detection in 2006.[4] The American National Standards Institute (ANSI) proposed the establishment of a new subcommittee "biomarkers," which is engaged in the technology for the detection of molecular biomarkers, in the ISO/TC 34 (food product committee). A technical working group of the International Union for the Protection of New Varieties of Plants (UPOV) has been working on the application of molecular technology to the variety registration review. However, no international guidelines for validation of DNA-based variety identification technology have been established.

2. Features of guidelines for validation in the forensic field

In Japan, no guidelines for validation have been established in the forensic field. In other countries, there are SWGDAM guidelines, European Network of Forensic Science Institutes (ENFSI) Guidance,[5] and DNA Advisory Board Quality Assurance Standards for Forensic DNA Testing Laboratories (DAB Standards, http://www.cstl.nist.gov/biotech/strbase/dabqas.htm), and the features are as follows:

In the SWGDAM guidelines, validation is divided into developmental validation performed by some laboratories in the development of a DNA analytical method and internal validation performed by each laboratory that plans to introduce a developed DNA analytical method.

Developmental validation is performed by manufacturers, public research institutions, etc. to demonstrate the accuracy, precision, and reproducibility of a procedure. Studies are conducted to characterize the details of markers such as DNA polymorphism and chromosomal localization, species specificity, sensitivity, stability, reproducibility, DNA polymorphism in a population, detectability, accuracy, and precision in a mixture, and basic procedures for PCR. The collaborative study is included in this type of validation. The number of laboratories is not stipulated.

Internal validation aims to examine the reliability and limitation of application and should include a study with known samples, study of reproducibility and precision, study to determine the size of DNA separated, study of the sensitivity and stochastic factors, detection study with a mixture, contamination study, and DNA analysis qualifying test. Through these studies, at least 50 samples should be analyzed. The DNA analysis-qualifying test may be performed in an internal study (for example, the use of internal quality control samples), external study (proficiency test), or collaborative study.

In validation of chemical analysis methods for food analysis in the AOAC guidelines , a process corresponding to internal validation is called method verification and distinguished from validation.

Method verification is performed by each laboratory that intends to introduce an analytical method to verify that the analytical method can be used within the range of validated performance characteristics, and thus the purpose is similar.

In the ENFSI guidance, technologies to be validated are classified into objective methods and subjective methods on the basis of their features. A subjective method is one where the experience and expertise of the analyst is the overriding determinant on whether or not analytical results are considered to match reference standards or fall within a certain range. Examples include handwriting and microscopic identifications.

Although, all SSR analyses are not based on subjective methods, valid interpretation of fragment analysis data requires experience and knowledge that allow the analyst to accurately read data, determine whether data on the test variety and data on the variety of reference alleles are matched or different, and determine whether data are appropriate (the presence or absence of problems with size standards and the appropriateness of the dilution ratio of PCR products).

Thus, SSR analyses also include processes in which experience and expertise of the analyst influence results. To address these issues, internal validation appears to be specified in the SWGDAM guidelines.

3. Number of laboratories in the collaborative study

The SWGDAM guidelines, ENFSI guidance, and DAB Standards indicate the need for collaborative study, but they do not specify the number of laboratories.

Reference 9 shows the details of developmental validation performed with DNA analysis kits (21-SNP multiplex and PowerPlex Y Validation).[6,7] For each of these kits, test items were divided into groups and validated to validate the entire kit. Of these, items to be examined for the reproducibility were examined in a collaborative study involving multiple laboratories. The number of laboratories participating in the collaborative study varied, with three at minimum and eight at maximum. This suggests that the numbers of samples and collaborating laboratories may be determined on the basis of the nature of technology (kit) to be validated.

In contrast, the AOAC guidelines and ISO 16140 (Microbiology of Food and Animal Feeding Stuffs — Protocol for the Validation of Alternative Methods) stipulate that a study of qualitative analytical methods should include at least 10 laboratories for statistical purposes regardless of the nature of technology.

The number of laboratories (at least 10) stipulated in the AOAC guidelines is based on the assumption that the true sensitivity (the proportion of positive samples correctly identified) or the true specificity (the proportion of negative samples correctly identified) would be 80%, and determined to mostly satisfy the criterion, that is inequality $362 \leq Lm^2$, where L is the number of laboratories and m is the number of repeat measurements required for the 90% confidence interval of the proportion calculated with normal distribution approximation (the proportion correctly identified) to fall within ±10% of the true value of 80% (72% to 88%) (see Reference 8).

The number of laboratories participating in the collaborative study of DNA-based plant variety identification technology should preferably be at least 10 to increase the statistical reliability but may be flexible taking into consideration the analytical technology (such as availability of kits and automation

of many work processes and devices including DNA extraction, PCR, and sequencers), diversity of plant species, and feasibility of collaborative studies (such as availability of relatively few laboratories knowledgeable about individual technologies).

It is labor intensive, costly, and inefficient to use a larger number of samples and participating laboratories to conduct a collaborative study above the standard required for a validation study. When the results of a preliminary validation study show stable analytical values within and between laboratories, efforts should be made to design an efficient study by reducing the numbers of replicates, laboratories, etc.

4. Admissibility of analytical results

It is desirable to establish the conditions so that study results are admissible in court, as DNA-based plant variety identification technology may be used as a countermeasure to the infringement of plant breeders' rights.

In the United States, Daubert Standards[8] are available as criteria for admissibility of scientific evidence in court and are adopted in many states. The Daubert Standards are based on a precedent in which decisions were made on admissible scientific evidence in court in 1993. This consists of the following four items:

1. Whether or not it could be tested and falsified
2. Whether it had been subject to peer review and publication
3. Whether there existed known or potential rate of error and standards controlling the technique's operation
4. Whether or not it was generally accepted within the scientific community

Scientific evidence provided by technology that meets these criteria is considered admissible in court unless the laboratory has quality assurance problems. The Japanese community of forensic medicine does not unconditionally accept the Daubert Standards, but reviews scientific evidence according to the standards. Thus, it is recommended that DNA-based plant variety identification technology to be validated should satisfy the Daubert Standards.

Among the four criteria, particular emphasis is placed on criteria 2 and 4 in Japan, and technology that has been subject to peer-review and journal publication is considered the one accepted by experts. Technology that is cited in multiple papers is considered the one generally accepted by experts. It is therefore important for technology used in DNA-based variety identification to be peer-reviewed and published in a journal.

Appendix 2: Example of Application of the Guidelines (Sweet Cherry)

The results of validation of DNA-based identification of the variety sweet cherry according to this guideline are described below.

1. Pilot study

A. Considerations in the pilot study

a. Review of the process of the collaborative study

A pilot study was conducted before the conduct of an collaborative studyto demonstrate the reproducibility of the method for extraction of DNA from the leaf, peduncle, and epicarpand marker performance in accordance with written procedures for DNA-based variety identification of sweet cherries developed by the Yamagata General Agricultural Research Center, which is available as Reference 9 "DNA-based variety identification technology" on the homepage of the Plant Variety Protection (http://www.hinsyu.maff.go.jp/). In the pilot study of the DNA extraction method, leaves of the Taisho-nishiki, and peduncles and epicarpof the Satonishiki were used as samples. In the pilot study of marker performance, Satonishiki and Benishuho were used as samples together with three markers.

b. Repeat analysis

Studies of the DNA extraction method and marker performance were conducted in duplicate.

c. Examination of the ability to interpret results

To examine the ability to determine the peak value for variety identification on the basis of images and multiple peak values obtained by fragment analysis, three markers were used in the pilot study of marker performance that showed typical SSR waveforms and allowed detection of the homozygote and heterozygote in the test varieties.

When the pilot study designed as described above was conducted, participating laboratories offered opinions and suggestions on improvements in additions to the written procedures and asked the details of the study. We revised the written procedures for precautionary purposes in this study and explained to some laboratories about points to consider in the interpretation of the results of fragment analysis.

This allowed participating laboratories to gain some experience in SSR analysis and DNA-based variety identification technology for sweet cherries and allowed us to revise the written procedures.

2. Confirmation developed techniques

A. Information on plants to be analyzed

a. Intravariety polymorphism

As a variety for examination of the intravariety polymorphism, Satonishiki was used, as it is a variety that has been extensively distributed for many years. Twelve SSR markers were used to perform fragment analysis of the DNA extracted from the peduncle or epicarpof the Satonishiki

maintained and grown in the Department of Agro-Production Science of Yamagata General Agricultural Research Center as well as Satonishiki produced in Hokkaido, Fukushima, Yamanashi, Nagano, and Tasmania, which were purchased from the Ohta Market in Tokyo and some production districts in Japan. The results showed no difference in data among samples provided, suggesting a low frequency of intravariety polymorphisms.[9]

b. Source of the variety

All samples (leaf, peduncle, andepicarp) used were collected from the variety (trees) maintained in the Yamagata General Agricultural Research Center.

c. Selection of the variety of reference alleles

The appropriate number of varieties of reference alleles appeared to be about two, because study data at the Yamagata General Agricultural Research Center showed that the number of alleles for each marker was relatively small and that the range of allele sizes was relatively narrow.

Given this, the varieties of reference alleles selected were Benishuho, which was grown in the Yamagata General Agricultural Research Center with its original tree available, and Satonishiki whose trees (not original) were managed in the Yamagata General Agricultural Research Center. The latter was expected to have fewer intravariety polymorphisms and used for the development of many other varieties.

B. Information on SSR markers

a. Marker performance

Among SSR markers shown on an integrated linkage map of Prunus and others, 10 markers were selected that had higher amplification efficiency, exhibited higher polymorphisms among varieties, and were easy to identify.

b. Independence of markers

Markers for which chromosomal locations were revealed on an integrated linkage map of Prunus and others were selected to ensure the independence of markers. One marker on each of the eight chromosomes of the sweet cherry was selected, and two markers on G1 and G4 (both indicating linkage group numbers), located about 40 to 50 cM apart from each other, were added. A total of 10 markers were used in the validation study.

c. Demonstration of the complete or near-complete absence of null alleles

He and Ho values were reported in the literature in 8 of the 10 markers used in the validation study,[10,11] suggesting the complete or near-complete absence of null alleles. A similar study with the remaining two markers was conducted in the Yamagata General Agricultural Research Center, also suggesting that null alleles were unlikely to exist.

C. DNA extraction

a. DNA extraction method

DNA was extracted largely according to the written procedures for DNeasy Plant Mini Kit (QIAGEN). Modifications to the procedures are as follows.

(1) Add 8 μL of mercaptoethanol and 4 mg of insoluble polyvinylpolypyrrolidone to 400 μL of Buffer AP1.

(2) After addition of Buffer AE and centrifugation, transfer the eluate in a microtube to another DNeasy Mini spin column for centrifugation.

b. Sample site

Samples were taken from the peduncle and the epicarpof the fruits distributed in the market. Samples were also taken from the leaf, because analysis may be performed when the fruits are not available in the market.

c. Sample conditions

After the fruits were purchased and stored at either ordinary temperature or refrigerated (4°C) for 1 to 6 weeks, DNA was extracted from the peduncle and theepicarp, and fragment analysis was performed with one marker. The results showed that DNA was successfully extracted from the peduncle and epicarpat 6 weeks of cold storage. DNA was also extracted from the fruits at 5 weeks of storage at ordinary temperature; however, mold grew on some samples at 2 weeks of storage and on most samples at 5 weeks, making these samples inappropriate. Still, PCR amplification and fragment analysis were not affected.

d. Sample storage

Samples collected in the Yamagata General Agricultural Research Center were labeled and stored frozen at −30°C.

D. PCR conditions

PCR reaction solution was prepared as follows. For 9 of the 10 markers, PCR cycle **(1)** was used. For marker 6-1, PCR cycle **(2)** was used.

a. PCR reaction solution

×10 buffer	2.0 μL
dNTP	2.0 μL
H_2O	13.9 μL
Taq polymerase	0.1 μL
Marker	1.0 μL
DNA	1.0 μL
Total	20.0 μL

b. PCR cycle

(1)

94°C 3 min
94°C 30 sec
55°C 30 sec } 35 cycles
72°C 1 min
72°C 5 min
4°C ∞

(2)

94°C 3 min
94°C 30 sec
51°C 30 sec } 35 cycles
72°C 1 min
72°C 5 min
4°C ∞

E. SSR analysis

a. Dilution ratio of PCR products

Table 1. Dilution ratio of PCR products by marker used in the collaborative study (for reference).

Marker	1-1	1-2	2-1	3-1	4-1	4-2	5-1	6-1	7-1	8-1
Dilution ratio (fold)	10	20	10	60	30	80	20	40	20	20

b. Determination of precision

For measurements obtained from a single laboratory, direct comparison was made. In principle, fragments were considered to have the same size when the difference in size was less than 1 bp. For comparison of measurements between laboratories, measurements were converted into the data presentation format described in Reference 3.

F. Validation confirmation **in processed goods**

Validation was not performed in the present study, because processed goods are beyond the scope of this paper/guideline.

G. Peer review and publication

DNA-based variety identification technology for sweet cherries was peer-reviewed and published in the following journal.

Tadashi Takashina, Makoto Ishiguro, Koichi Nishimura, Toshiya Yamamoto. Identification of varieties of imported and domestic sweet cherry with SSR markers. DNA Polymorphism. 2007;15:101–104.

3. Collaborative study

The collaborative study was conducted in participating laboratories of the following 11 institutions.

- DNA profiling team (main lab), Incorporated Administrative Agency National Center for Seeds and Seedlings
- Citrus Genome Analysis Team, National Institute of Fruit Tree Science, Incorporated Administrative Agency National Agriculture and Food Research Organization

- Cereal Science and Utilization Laboratory, Food Resource Division, National Food Research Institute, Incorporated Administrative Agency National Agriculture and Food Research Organization
- Technical Research Division, Food Labeling Monitoring Department, Incorporated Administrative Agency Food and Agricultural Materials Inspection Center
- Biotechnology and Plant Breeding Division, Department of Agro-Production Science, Yamagata General Agricultural Research Center
- Cultivation Group, Fruit Tree Research Centre, Fukushima Agricultural Technology Centre
- Laboratory of Fruit and Ornamental Breeding, Plant Biotechnology Institute, Ibaraki Agricultural Center.
- Plant Biotechnology Laboratory, Biotechnology Division, Chiba Prefectural Agriculture and Forestry Research Center

- Laboratory of Japanese apricot , Fruit Tree Experiment Station, Wakayama Research Center of Agriculture, Forestry and Fisheries
- Department of Biotechnology, Fukuoka Agricultural Research Center
- Division of Biotechnology, Miyazaki Agricultural Research Institute

A. Items examined for reproducibility

DNA extraction method and marker performance

B. Examination of the reproducibility of the DNA extraction method

In each laboratory, DNA was extracted from a total of 18 samples, including 3 replicates each from the leaf, peduncle, and epicarpof the varieties of reference alleles Satonishiki and Benishuho. Extracted DNA was subjected to agarose gel electrophoresis to determine the concentration and impurity. Subsequently, SSR markers were used to perform PCR to confirm that appropriate PCR products were obtained.

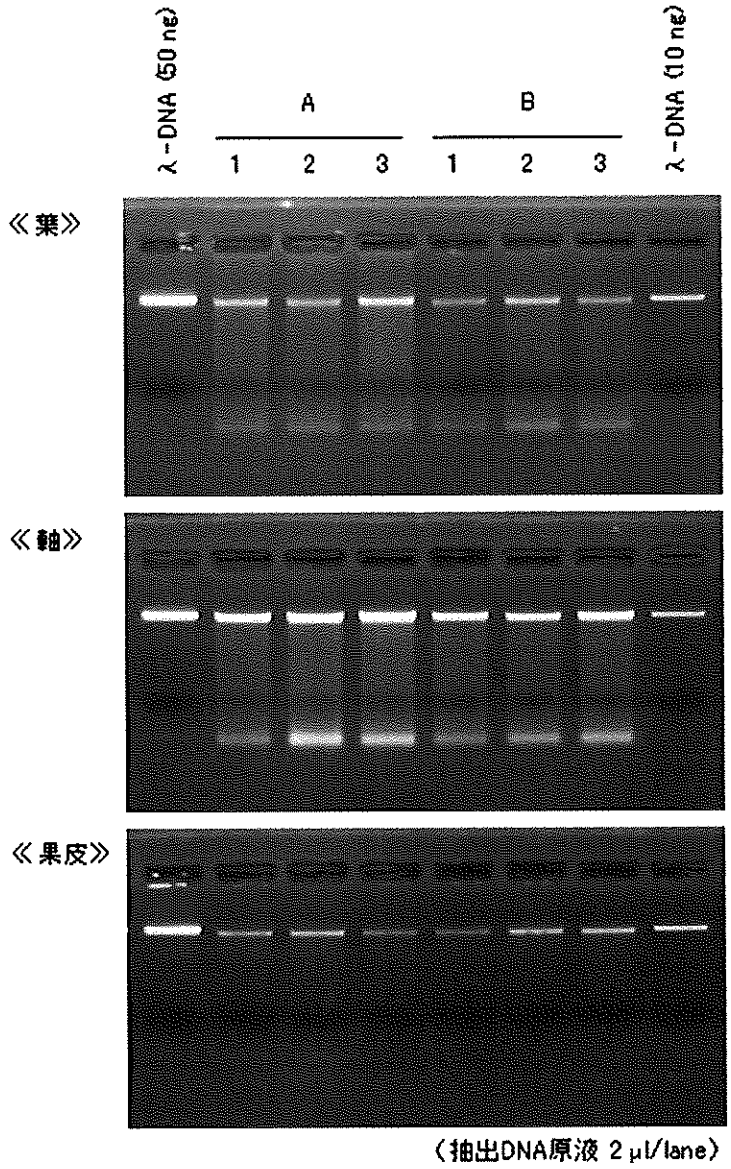
In agarose gel electrophoresis and examination of PCR products, genotyping of peduncle-derived DNA was feasible at all laboratories (Figure 1).

The efficiency of DNA extraction from the leaf was somewhat lower than that from the peduncle in some laboratories, because sample leaves were mostly mature leaves. Still, genotyping of extracted DNA was feasible in agarose gel electrophoresis and examination of PCR products at all laboratories.

The efficiency of DNA extraction from the epicarpwas lower in many laboratories, because it was rather difficult to sample the epicarpfrom frozen fruits and because more flesh on the epicarpwas associated with more impurities such as polysaccharides. This made visual examination of agarose gel difficult in some cases; however, PCR amplification produced good results in 10 laboratories. In one laboratory, PCR products were not detected with the specified protocol. Thus, modifications were made to the protocol such as reduction of DNA content added to the PCR reaction solution to one-tenth, a change in the number of PCR cycles from 35 to 40, and a change

in the dilution ratio of PCR products from 20-fold to 10-fold, which allowed the detection of PCR products. This suggests that a large amount of impurities such as polysaccharides in the extracted DNA was responsible for the failure with the specified protocol.

These findings demonstrate the reproducibility of the method for DNA extraction from the leaf and peduncle. For theepicarp, good results were not obtained with the specified protocol in one laboratory. However, analysis became feasible by reducing the amount of DNA for PCR to one-tenth and changing the number of PCR cycles and the dilution ratio of PCR products. In the process of collection of the epicarpfrom fruits, a large quantity of flesh adhered to theepicarp, and contaminants such as polysaccharides were likely to reduce the extraction efficiency. The written procedures for DNA-based variety identification technology for sweet cherries contains the wording "avoid flesh as much as possible." Taken together, (1) the DNA extraction method itself was not responsible and (2) good results were obtained in the other 10 laboratories, demonstrating reproducibility.



Leaf

Peduncle

Epicarp

(2 μL of undiluted DNA extract per lane)

Figure 1. Results of the reproducibility study of the DNA extraction method (example of the National Food Research Institute).

C. Examination of the reproducibility of marker performance

As the independence of markers and the frequency of null alleles were known, the reproducibility was examined according to study method 2 (see Reference 5).

DNA samples were collected from the varieties of reference alleles Satonishiki and Benishuho. In each laboratory, a total of 120 samples were analyzed (2 DNA samples × 10 SSR markers × 6 replicates).

Overall, replicates in each laboratory produced reproducible data. For eight markers, all DNA samples produced results that were considered to fall within the acceptable range (intralaboratory deviation less than 1 bp) in all laboratories (see Reference 6-1 and Figure 2).

For the remaining two markers, reproducible data were obtained in 10 laboratories. In one laboratory; however, some results did not fall within the acceptable range (intralaboratory deviations of 1.06 bp and 1.14 bp). The characteristics of data obtained with the two markers may be explained by mechanical deviations caused by analysis of DNA samples on different days or in different runs, because (1) both markers showed a slight difference in data between Satonishiki and Benishuho, (2) replicates of each variety produced reproducible data, (3) for the two markers, PCR and fragment analysis of DNA samples taken from Satonishiki and Benishuho were performed on different days (confirmed by the analyst), and (4) in the 10 laboratories reporting reproducible results, a similar trend was found with some markers when analyses of Satonishiki and Benishuho were performed on different days. For confirmation, PCR and fragment analysis of the two varieties were performed in the laboratory on the same day, which produced reproducible data.

As described above, for the eight markers, reproducible results were obtained from all DNA samples in 11 laboratories, demonstrating reproducibility.

For the remaining two markers, some results were not easily considered to fall within an acceptable range. However, (1) reproducible results were obtained in 10 laboratories and (2) the deviations were not confirmed to marker performance, demonstrating reproducibility.

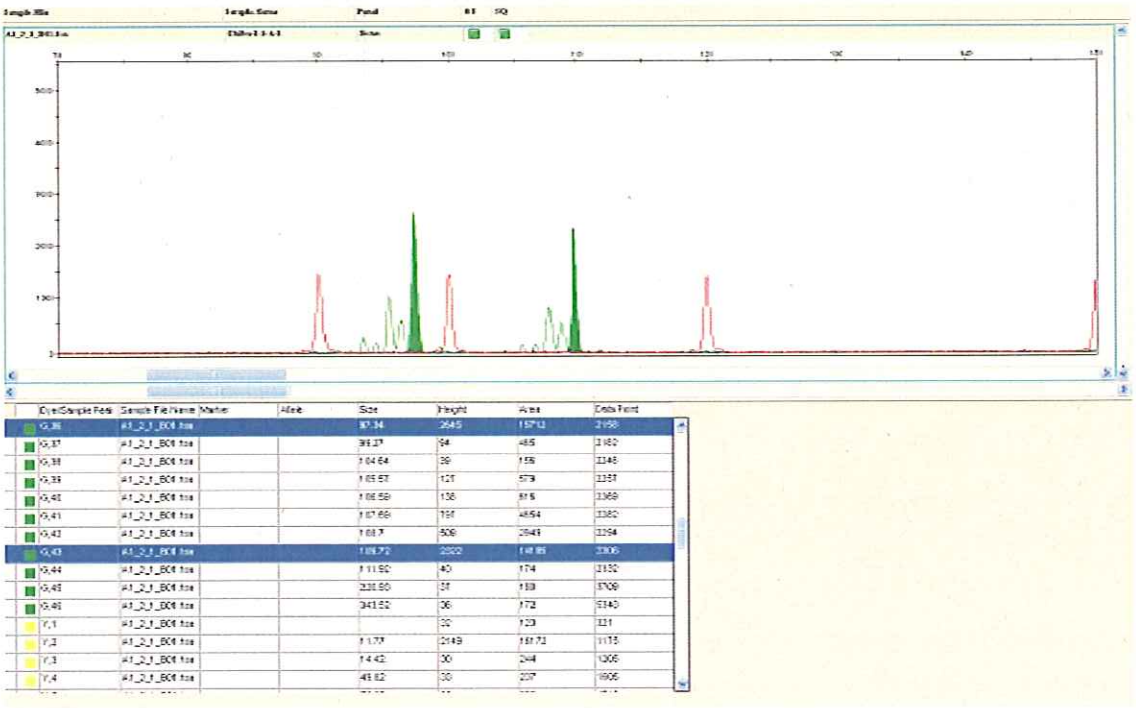


Figure 2. Results of the reproducibility study of marker performance (example in the Chiba Prefectural Agriculture and Forestry Research Center).

4. Data and reports submitted from each laboratory

A. Validation study of the DNA extraction method

As documents indicating the concentration and quality of DNA, photographs of agarose gel electrophoresis were submitted from the 11 laboratories.

As documents evidencing PCR amplification of extracted DNA, electropherograms of fragment analysis were submitted from the 11 laboratories.

B. Validation study of markers

Electropherograms of fragment analysis of DNA samples with each marker were submitted from the 11 laboratories. The charts were standardized to have a length of the horizontal axis of 80 bp, visible peaks of waveforms, and specified waveforms used for data interpretation.

Also submitted were lists of data on peaks that analysts identified on the electropherograms.

C. Common information

(1) Devices (models) and reagents used in the study

See Table 2.

(2) Issues raised during the study

- When DNA was extracted from the epicarpaccording to the written procedures, no DNA was identified on agarose gel or PCR. Analysis was feasible when the amount of epicarpwas halved together with modifications to the composition of PCR reaction solution, number of PCR cycles, and dilution ratio of PCR products.
- All parameters of the GeneMapper were manually set, because the sizing quality (SQ) of some primer peaks were always rated as low quality (displayed in red), indicating misalignment of size markers.

Table 2. List of devices used in individual laboratories.

	Thermal cycler	Sequencer	Number of capillaries	Capillary	Polymer	Size standard
National Center for Seeds and Seedlings	PE Applied Biosystems 9700	ABI 310	1	310 Capillaries 47 cm × 50 µm	POP-4	400HD ROX
National Institute of Fruit Tree Science	PE Applied Biosystems 9700	ABI 3100xl	16	3100/3130xl 36 cm Capillary Array	POP-4	400HD ROX
National Food Research Institute	BIO-RAD iCycler	ABI 310	1	310 Capillaries 47 cm × 50 µm	POP-4	400HD ROX
Food and Agricultural Materials Inspection Center	PE Applied Biosystems 9700	ABI 3130xl	16	3100/3130xl 50 cm Capillary Array	POP-7	400HD ROX
Yamagata General Agricultural Research Center	PE Applied Biosystems 9700	ABI 310	1	310 Capillaries 47 cm × 50 µm	POP-4	400HD ROX
Fukushima Agricultural Technology Centre	PE Applied Biosystems 9700	ABI 3130	4	3100-Avant/3130 36 cm Capillary Array	POP-7	400HD ROX
Ibaraki Agriculture Institute	PE Applied Biosystems 9700	ABI 3130xl	16	3100/3130xl 36 cm Capillary Array	POP-4	400HD ROX
Chiba Prefectural Agriculture and Forestry Research Center	PE Applied Biosystems 9700	ABI 3130xl	16	3100/3130xl 50 cm Capillary Array	POP-7	400HD ROX
Wakayama Research Center of Agriculture, Forestry and Fisheries	Gene Amp PCR System 2700	ABI 310	1	310 Capillaries 47 cm × 50 µm	POP-4	400HD ROX
Fukuoka Agricultural Research Center	ASTEC PC-808-02	ABI 310	1	310 Capillaries 47 cm × 50 µm	POP-4	400HD ROX
Miyazaki Agricultural Research Institute	PE Applied Biosystems 9700	ABI 3100	4	3100-Avant/3130 36 cm Capillary Array	POP-4	400HD ROX

Appendix 3: References and Others

1. Reference

Reference 1 Independence of markers

Two or more DNA markers may be used to identify many varieties or ensure the precision of variety identification. For these purposes, it should be verified that those markers are genetically independent. In this context, independence means that a combination of alleles at multiple genetic loci occurs randomly and follows the laws of probability. When markers are linked and not independent, the frequency of a specific combination is significantly high. Recombination between linked markers, which are located closely to each other, occurs rarely compared with markers located on other chromosomes. Thus, combinations of alleles do not occur randomly, and the frequency of a specific combination of alleles is disproportionally high or low. The use of linked markers does not improve the precision of variety identification, because analyses with multiple markers reveal no difference in marker genotypes.

An easy way of obtaining an independent marker set is to select a marker from each linkage group on the linkage map corresponding to the number of chromosomes. In cases where no linkage map corresponding to the number of chromosomes is available, the location of markers used on the linkage map is unknown, or the number of markers used should be larger than the number of chromosomes, you should examine the independence of markers used by yourself. When a segregating population is available for analysis of the marker set used, the observed and expected values of individual marker genotypes in segregants should be calculated, and chi-square test should be used to demonstrate that the observed values and the expected values are not significantly different. For example, in cases where two SSRloci A and B have alleles a_1, a_2, b_1, and b_2 with a_i and b_1 having frequencies of p and q, a_2 and b_2 should have frequencies of 1-p and 1-q, respectively. Thus, the frequencies of haplotypes (a_1b_1), (a_1b_2), (a_2b_1), and (a_2b_2) are pq, $p(1-q)$, $(1-p)q$, and $(1-p)(1-q)$, respectively. Substantial differences between observed and expected frequencies suggest a possible linkage, and the use of such makers should be avoided. When a segregating population is not available, the marker genotypes of varieties of the target crop should be determined. The frequencies of individual markers should be calculated on the basis of the data, and the observed values and expected values should be compared in a similar manner as in segregating population. In any case, a commonly used significance level of 5% is recommended.

In cases where marker independence is not demonstrated

In DNA-based variety identification with the use of markers whose independence is uncertain, a difference in analytical results between two samples indicates that the two samples represent different varieties (on the assumption that the markers used for the varieties are not associated with intravariety polymorphisms). When there is no difference in analytical results, the variety identification rate cannot be calculated (for instance, the probability of the same variety being 0.01%), because the probability that different varieties produce identical results with respect to all markers used cannot be calculated.

To identify the varieties with those markers on the basis of analytical results, it is necessary to meet the following conditions: (1) Samples are likely to represent a limited number of candidate varieties. (2) Information on all the candidate varieties is available in a database constructed with those markers. (3) On the database, the candidate varieties are distinguishable from all other varieties.

Table 3. Certainty of marker independence and variety identification.

	Marker independence	
	Certain	Uncertain
Any difference in analytical results between samples	The samples are considered to represent different varieties. (*)	The samples are considered to represent different varieties. (*)
No difference in analytical results between samples	The variety identification rate may be used to calculate the probability of the same variety.	The samples are considered to represent the same variety only when the varieties analyzed belong to known candidate varieties and when no pair of varieties in the candidate varieties produces the same analytical result.

* In cases where there are no intravariety polymorphisms or other problems.

Reference 2 Demonstration of the complete or near-complete absence of null alleles in the marker region

In forensic analysis on the basis of genotypes, no further analysis is necessary when the identity or kinship is ruled out; however, calculation of probability is needed when the relationship cannot be excluded (being identical). In this case, the population must be in Hardy-Weinberg Equilibrium (HWE) for loci analyzed. Appropriate selection of loci tested is essential for efficient individual identification. In general, highly polymorphic loci with adequate distribution of allele frequencies are desirable. However, extremely high polymorphisms may make interpretation difficult and reduce the reliability of the database. In practice, a microsatellite (SSR) with about 8 to 10 alleles is used for individual identification.

Representative measures for polymorphism are heterozygosity and polymorphic information content (PIC). Heterozygosity is the frequency of heterozygotes in a population and also the probability that two alleles randomly chosen from a population are different. This can be calculated by the following equation (1). The PIC is the probability of determining which parent each of two alleles at a gene locus of an offspring is derived from and calculated by the following equation (2).

$$H = 1 - \sum_{i=1}^{n} p_i^2 \qquad (1)$$

$$PIC = 1 - \sum p_i^2 - \sum \sum 2(p_i p_j)^2 \qquad (2)$$

High PIC values are advantageous in the analysis of disease-related genes in a family study. In forensic medicine, PIC is considered a measure of usefulness of the gene locus of interest in kinship analysis.

Null alleles are alleles that are undetectable at a SSR locus. Null alleles may be responsible for distortion of some SSR loci from Hardy-Weinberg Equilibrium. In SSR loci, null alleles often result from mutations in a primer-annealing site (or both primer annealing sites), interfering with the amplification of SSR alleles. This problem is likely to occur when SSR primers cloned from a certain species are used to perform genotyping in different species or genera. The presence of null alleles precludes genotype-based identification of individuals (varieties), and thus SSRloci without null alleles should preferably be selected as DNA markers. In the presence of null alleles, the observed heterozygosity (Ho) is reduced compared with the expected heterozygosity (He), and the null allele frequency in a population is estimated from the relationship. The null allele frequency is usually computed from population genotype data with analysis software (such as Genepop and CERVUS).

Reference 3 Presentation format and deviation range of fragment size

1. Presentation format for fragment size

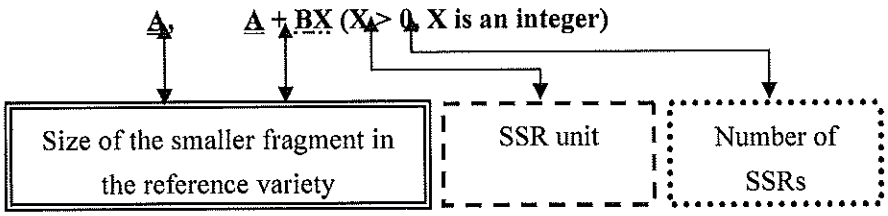
The fragment size scanned in SSR analysis is not based on the number of bases counted, but the figure calculated from the distance between peaks of size standards. Thus, slight deviations may occur, depending on the type of sequencers, size standards, capillaries, polymers, fluorescent dyes for primers, etc., which may result in the detection of different fragment sizes in a single variety.

This makes it difficult to use raw data on fragment size determined by fragment analysis as final data on varieties in database construction or interlaboratory comparison. To eliminate those deviations and share these data, we propose the following presentation format in which a variety of reference alleles for fragment size is used, and the fragment size in a test variety is based on the distance from the fragment peak of the variety of reference alleles.

Specifically, among fragments detected with a SSR marker in the variety of reference alleles, the smallest fragment is used as a reference fragment size for the SSR marker. The size of other fragments in the variety of reference alleles (when the variety of reference alleles shows multiple peaks) is indicated as the distance from the peak of the fragment. Next, among (multiple) fragments of the test variety, the fragment closest to (multiple) fragments of the variety of reference alleles is chosen (referred to as test fragment (1), and the distance between the two fragment peaks is calculated. The size of other fragments in the test variety is indicated as the distance from each test fragment (1). This takes advantage of the fact that deviations in fragment size are smaller when fragment size is indicated as the distance between fragment peaks.

Fragment sizes determined with SSR markers usually differ by a nucleotide repeat unit (in increments of two bases for dinucleotide SSRs and three bases for trinucleotide SSRs), and thus the presentation format is designed so that the fragment size increases in increments of repeat units.

As an example, the presentation format for fragment size in a variety of reference alleles is shown below. When alleles are heterozygous, two types of amplified products are detected. The smaller allele is expressed as A, and the larger one is expressed as A plus the product of SSR repeat unit (2 for dinucleotide repeat) and the number of SSRs ($X > 0$, X represents an integer). Fragment sizes in the fragment analysis are expressed to the second decimal place, but rounded off to give an integer X, because the actual number of bases is an integer.



In most cases, the values (A + BX) do not completely agree with the actual data obtained in fragment analysis (because data are expressed to the second decimal place). In the case of disagreement, X (integer) should be chosen so that the difference between the measurement and the value (A + BX) is minimized.

Method for presentation of fragment size in the variety of reference alleles (example)

* This example involves dinucleotide SSRs.

* The fragment sizes are 100.2 bp and 110.4 bp.

Under the conditions described above,

$A = 100.2$

$A + BX = 100.2 + 2 \times X$

Find X that satisfies the equation $100.2 + 2 \times X = 110.4$, and $X = 5.1$.

X must be an integer, and 5.1 rounded off to the nearest integer is 5.

Thus, 100.2 bp and 110.4 bp are presented as A, A + 10.

Method for presentation of the test variety on the basis of fragment size of the variety of reference alleles (example)

* This example involves dinucleotide SSRs.

* The fragment sizes in the variety of reference alleles are 100.2 bp and 110.4 bp and expressed as A, A + 10.

* The fragment sizes in the test variety are 106.4 bp and 112.6 bp.

Under the conditions described above,

Among fragment sizes in the variety of reference alleles and the test variety,

110.4 bp (variety of reference alleles) and 112.6 bp (test variety) are closest to each other.

$112.6 - 110.4 = 2.2$

Thus, 112.6 can be expressed as A + 10 + 2.2, that is, A + 12.2.

BX must be an integer, and A + 12.2 rounded off to the nearest integer is A + 12.

When 112.6 bp is presented as A+12,

$112.6 - 106.4 = 6.2$,

Thus, 106.4 bp can be expressed as A + 12 − 6.2, that is, A + 5.8.

BX must be an integer, and A + 5.8 rounded off to the nearest integer is A + 6.

Thus, on the basis of the fragment sizes of the variety of reference alleles, the fragment sizes of the test variety are presented as A + 6 and A + 12.

2. Example of inaccurate representation of fragment size in the presentation format

Reference 4-1 provides examples of the application of the presentation format for fragment size. In Reference 4-1, all measurements were converted into the presentation format that appropriately represented the original fragment size. Most data in the presentation format appear to represent the original fragment size, with some exceptions. Reference 4-2 provides examples of such exceptions. Although Laboratory Y and Laboratory I reported stable results as far as the distance between fragment peaks was concerned, slight deviations resulted in differences in final data presented after rounding, because the mean distance between peaks was 12.51 bp.

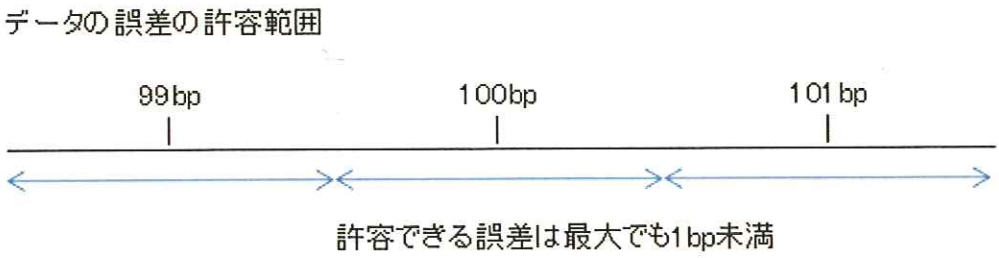
Data are not given in nucleotide bases, because calculation of fragment size in fragment analysis is not based on the number of bases counted, but based on the distance between size standard peaks. This type of calculation method is likely to produce large deviations (1) when nucleotide composition is biased (because different nucleotides have different sizes) and (2) when the distance between fragment peaks or the distance between the fragment peak and the size standard peak is long.

3. Acceptable deviation range

We investigated whether the analyst could determine that different final data represented the same fragment size when the variation among replicates was small and fell within a certain range as shown in Reference 4-2.

In SSR analysis, fragment sizes usually differ by a repeat unit, and thus a single repeat unit could be an acceptable range within which two fragment sizes were considered the same (±1 bp for dinucleotide repeats). However, fragment sizes may differ by something other than a single repeat unit because of nucleotide insertion, and deletion etc., (in this case, the size is presented as A + BX + α), and an acceptable deviation range of 1 bp (±0.5 bp) was deemed appropriate.

Acceptable deviation range



The acceptable deviation is less than 1 bp.

Figure 3. Schematic diagram of the acceptable deviation range.

Next, in the collaborative study of sweet cherries, intralaboratory deviations in fragment size (difference between the maximum and minimum) were compared, and the results are shown in Table 4.

Table 4. Intralaboratory deviations in alleles.

Laboratory	NC	NI	FN	FA	YA	FS	I	CH	WA	FO	MI	ALL
Mean deviation	0.21	0.27	0.24	0.17	0.20	0.33	0.17	0.16	0.29	0.29	0.31	1.78

Note 1: Mean deviations in 18 fragment sizes with 10 primers. Values are expressed in bp.

As shown in Table 4, the mean deviation in each laboratory is far below 1 bp, and it is reasonable to determine that the acceptable deviation range within which two fragment sizes are considered the same should be 1 bp (±0.5 bp).

The mean deviation in data among all laboratories is 1.78 bp. Direct comparison of fragment sizes between different laboratories should be avoided, because the mean exceeds the acceptable deviation range.

4. Interpretation by analysts

Taken together, when the presentation format is not simply applicable as shown in Reference 4-2, the analyst should examine whether the following condition is met and then determine whether fragment sizes are considered the same.

- The difference between fragment peaks measured in a single laboratory on the same day is small (less than 1 bp).

Before data interpretation, the analyst should review the considerations to minimize deviation (see "D. Examination of the reproducibility of SSR marker performance" on page 10).

Reference 4-1 Example of the use of presentation format in Reference 3

Marker 1-1			
		Fragment size (1)	Fragment size (2)
Y	S	243.61	243.61
		243.38	243.38
		243.33	243.33
		243.35	243.35
		243.37	243.37
		243.42	243.42
	B	237.42	243.27
		237.39	243.31
		237.40	243.20
		237.40	243.19
		237.28	243.19
		237.31	243.22
F	S	246.11	246.11
		246.04	246.04
		246.00	246.00
		246.08	246.08
		246.00	246.00
		245.93	245.93
	B	240.11	245.90
		240.21	246.00
		240.11	245.90
		240.11	245.93
		240.11	245.90
		240.21	245.97

All measurements of fragment size (1) are presented on the basis of the distance from fragment size (1) of variety B No. 1.

All measurements of fragment size (2) are presented on the basis of the distance from fragment size (1) of the same sample.

Marker			
		(1)	(2)
Y	S	6.19	0.00
		5.96	0.00
		5.91	0.00
		5.93	0.00
		5.95	0.00
		6.00	0.00
	B	0.00	5.85
		-0.03	5.92
		-0.02	5.80
		-0.02	5.79
		-0.14	5.91
		-0.11	5.91
F	S	6.00	0.00
		5.93	0.00
		5.89	0.00
		5.97	0.00
		5.89	0.00
		5.82	0.00
	B	0.00	5.79
		0.10	5.79
		0.00	5.79
		0.00	5.82
		0.00	5.79
		0.10	5.76

Rounded to the nearest integer

Marker			
		(1)	(2)
Y	S	6	0
		6	0
		6	0
		6	0
		6	0
		6	0
	B	0	6
		0	6
		0	6
		0	6
		0	6
		0	6
F	S	6	0
		6	0
		6	0
		6	0
		6	0
		6	0
	B	0	6
		0	6
		0	6
		0	6
		0	6
		0	6

Fragment size (1) of variety B No. 1 is presented as A.

Marker			
		(1)	(2)
Y	S	A+6	A+6
		A+6	A+6
		A+6	A+6
		A+6	A+6
		A+6	A+6
		A+6	A+6
	B	A	A+6
		A	A+6
		A	A+6
		A	A+6
		A	A+6
		A	A+6
F	S	A+6	A+6
		A+6	A+6
		A+6	A+6
		A+6	A+6
		A+6	A+6
		A+6	A+6
	B	A	A+6
		A	A+6
		A	A+6
		A	A+6
		A	A+6
		A	A+6

Reference 4-2 Example of inaccurate representation of fragment size in the use of presentation format in Reference 3

Marker 1-2			
		Fragment size (1)	Fragment size (2)
Y	S	97.78	110.20
		97.77	110.24
		97.79	110.18
		97.85	110.20
		97.84	110.28
		97.71	110.23
	B	97.77	110.16
		97.78	110.27
		97.81	110.26
		97.75	110.17
		97.79	110.23
		97.80	110.24
I	S	97.32	109.81
		97.25	109.83
		97.21	109.79
		97.24	109.83
		97.22	109.83
		97.26	109.82
	B	97.26	109.84
		97.24	109.84
		97.20	109.79
		97.21	109.88
		97.19	109.83
		97.28	109.83

All measurements of fragment size (1) are presented on the basis of the distance from fragment size (1) of variety B No. 1.

All measurements of fragment size (2) are presented on the basis of the distance from fragment size (1) of the same sample.

Marker			
		(1)	(2)
Y	S	0.01	12.42
		0.00	12.47
		0.02	12.39
		0.08	12.35
		0.07	12.44
		-0.06	12.52
	B	0.00	12.39
		0.01	12.49
		0.04	12.45
		-0.02	12.42
		0.02	12.44
		0.03	12.44
I	S	0.06	12.49
		-0.01	12.58
		-0.05	12.58
		-0.02	12.59
		-0.04	12.61
		0.00	12.56
	B	0.00	12.58
		-0.02	12.60
		-0.06	12.59
		-0.05	12.67
		-0.07	12.64
		0.02	12.55
		Mean	12.51
		Maximum	12.67
		Minimum	12.35
		Difference	0.32

Despite stable data, the mean is about 12.5; thus,

Rounded to the nearest integer

Marker			
		(1)	(2)
Y	S	0	12
		0	12
		0	12
		0	12
		0	12
		0	13
	B	0	12
		0	12
		0	12
		0	12
		0	12
		0	12
I	S	0	12
		0	13
		0	13
		0	13
		0	13
		0	13
	B	0	13
		0	13
		0	13
		0	13
		0	13
		0	13

After rounding, data are presented as if these values represent different fragment sizes.

Fragment size (1) of variety B No. 1 is presented as A.

Considering data errors (difference between the maximum and minimum) etc., the analyst should determine whether these data represent the same fragment size

Marker			
		(1)	(2)
Y	S	A	A+12
		A	A+12
		A	A+12
		A	A+12
		A	A+12
		A	A+13
	B	A	A+12
		A	A+12
		A	A+12
		A	A+12
		A	A+12
		A	A+12
I	S	A	A+12
		A	A+13
		A	A+13
		A	A+13
		A	A+13
		A	A+13
	B	A	A+13
		A	A+13
		A	A+13
		A	A+13
		A	A+13
		A	A+13

Reference 5 Method for examination of the reproducibility of marker performance

1. Study method 1

This method is used to examine the reproducibility of the size of fragments (alleles) detected in SSR analysis and validate the SSR markers. When PCR amplification of DNA of varieties c and d with marker A allows detection of fragments 138 bp and 150 bp in size in variety c and fragments 150 bp and 173 bp in size in variety d, this method is used to investigate whether the results are reproducible at other institutions (in a strict sense, the reproducibility of the distance between fragment peaks rather than fragment sizes).

As study method 1 is designed to investigate the reproducibility of the size of fragments (alleles) in varieties, the scale of the study is larger as the number of alleles (number of different fragment sizes) increases. Fragment (allele) sizes other than those proven reproducible are not considered validated. However, study method 1 allows validation of SSR markers associated with null alleles.

There are two methods for examining the reproducibility of fragment size. One method examines a combination of fragment sizes. For example, the use of marker A allows detection of fragments 150 bp and 173 bp in size in variety d, variety e, etc. The other method addresses individual fragment sizes separately. For example, the use of marker A allows detection of a 138-bp fragment in variety a, variety b, etc. and a 150-bp fragment in variety b, variety c, etc. The latter method appears to be suitable when the number of alleles per SSR marker is large.

Study method 1 may be used when marker independence is not demonstrated. It should be noted; however, that the variety identification rate cannot be calculated under these circumstances (Reference 1).

Table 5. Example of study method 1.
The reproducibility of five alleles is investigated with marker A, and that of three alleles is investigated with marker B.

Variety	Marker A	Marker B
品種名	マーカーA	マーカーB
品種a	138／173	216／224
品種b	138／138	216／224
品種c	138／150	216／224
品種d	150／173	218／224
品種e	150／173	224／224
品種f	140／152	216／218

2. Study method 2

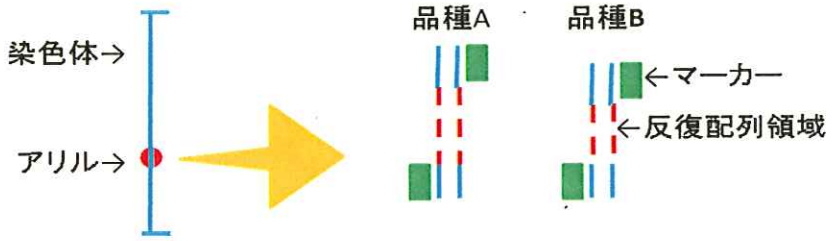
Study method 2 is based on the notion that the reproducibility of SSR marker performance to be demonstrated means that primers reliably anneal to the target locus. As long as SSR markers anneal to the target locus (under appropriate conditions for PCR etc.), alleles of interest are eventually amplified,

and detected fragment sizes should almost exactly represent alleles (genotype) of the tested variety. Thus, it is not necessary to demonstrate the reproducibility of each allele.

To demonstrate the annealing of primers to the target locus in different varieties, evidence of the complete or near-complete absence of null alleles at the locus is required (see Reference 2).

In study method 2, reproducibility is investigated by testing one or more varieties of reference alleles to see if expected sizes of fragments are detected. The scale of the study is small, because it is not necessary to investigate other alleles.

Another advantage of this method is that it does not require demonstration of the reproducibility of SSR markers, which were previously validated in an collaborative study, on unknown alleles detected in new varieties.



Chromosome
Allele
Variety A
Variety B
Marker
Repeat sequence region

Figure 4. Schematic diagram of study method 2.

As long as reliable annealing of primers to the target locus is demonstrated, the allelic difference between variety A and variety B does not influence marker performance.

Table 7. Comparison of study methods 1 and 2.

	Study method 1	Study method 2
Subject of reproducibility study	All alleles for each marker	Locus for each marker
Study scale	Number of markers × number of alleles in test variety × number of replicates × number of institutions (large scale of study)	Number of markers × number of varieties (1 or a few) × number of replicates × number of institutions (small scale of study)
Addition of new markers	Validation should be performed again: Number of new markers × number of alleles in test variety ×	Validation should be performed again: Number of new markers × number of varieties (1 or a few) ×

	number of replicates × number of institutions	number of replicates × number of institutions
Addition of alleles for existing markers	Validation should be performed again: Number of alleles added × number of varieties added × number of replicates × number of institutions	No additional validation is needed.
Investigation of the reproducibility of alleles	All alleles in test varieties should be directly investigated in an collaborative study.	Specific alleles in test varieties should be directly investigated in an collaborative study.

Reference 6-1 Fragment size by laboratory and marker in the interlaboratory collaborative study of sweet cherries

Laboratory	Fragment size	Marker																	
		1-1		1-2		2-1		3-1	4-1		4-2	5-1		6-1		7-1		8-1	
		(1)	(2)	(1)	(2)	(1)	(2)	(1)	(1)	(2)	(1)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
A	Maximum	238.17	244.25	97.21	109.89	118.63	142.97	138.11	168.94	203.34	165.81	147.11	151.00	115.99	122.24	160.22	164.35	155.72	157.75
	Minimum	238.05	244.00	97.07	109.71	118.30	142.84	137.97	168.57	203.00	165.53	146.71	150.79	115.79	122.12	160.11	164.17	155.62	157.62
	Difference	0.12	0.25	0.14	0.18	0.33	0.13	0.14	0.37	0.34	0.28	0.40	0.21	0.20	0.12	0.11	0.18	0.10	0.13
B	Maximum	237.64	243.57	97.18	109.79	117.96	142.28	137.89	168.61	203.06	165.69	147.16	151.18	115.70	122.25	160.25	164.24	155.43	157.50
	Minimum	237.20	243.31	97.09	109.65	117.86	142.19	137.42	168.41	202.84	165.56	146.99	151.07	115.58	121.64	159.92	164.02	154.92	156.87
	Difference	0.44	0.26	0.09	0.14	0.10	0.09	0.47	0.20	0.22	0.13	0.17	0.11	0.12	0.61	0.33	0.22	0.51	0.63
C	Maximum	238.32	244.19	97.37	109.93	118.46	142.97	137.90	168.70	202.93	165.90	147.31	151.37	116.01	122.36	160.47	164.60	156.07	158.15
	Minimum	238.12	244.02	97.08	109.69	118.23	142.69	137.64	168.38	202.69	165.65	146.98	151.14	115.99	122.22	160.23	164.23	155.84	157.89
	Difference	0.20	0.17	0.29	0.24	0.23	0.28	0.26	0.32	0.24	0.25	0.33	0.23	0.02	0.14	0.24	0.37	0.23	0.26
D	Maximum	239.76	245.47	97.54	109.89	119.64	144.37	138.56	169.36	203.07	166.82	147.93	152.02	116.32	122.77	161.51	165.48	157.36	159.43
	Minimum	239.52	245.29	97.37	109.75	119.46	144.28	138.47	169.24	202.96	166.65	147.82	151.89	116.24	122.40	161.34	165.28	157.08	159.17
	Difference	0.24	0.18	0.17	0.14	0.18	0.09	0.09	0.12	0.11	0.17	0.11	0.13	0.08	0.37	0.17	0.20	0.28	0.26
E	Maximum	237.42	243.61	97.85	110.28	118.32	142.82	138.10	168.57	202.31	165.94	147.38	151.47	116.22	122.36	160.54	164.54	156.51	158.58
	Minimum	237.28	243.19	97.71	110.16	118.20	142.71	137.85	168.42	202.15	165.79	147.24	151.38	115.99	122.03	160.43	164.39	156.10	158.13
	Difference	0.14	0.42	0.14	0.12	0.12	0.11	0.25	0.15	0.16	0.15	0.14	0.09	0.23	0.33	0.11	0.15	0.41	0.45
F	Maximum	240.21	246.11	97.03	109.48	119.45	144.54	138.71	169.97	203.95	167.20	147.61	151.70	116.20	122.78	161.86	166.09	157.39	159.46
	Minimum	240.11	245.90	96.89	109.35	119.22	144.38	138.52	169.30	202.81	166.93	147.48	151.62	116.10	121.72	161.59	165.50	157.21	159.26
	Difference	0.10	0.21	0.14	0.13	0.23	0.16	0.19	0.67	1.14	0.27	0.13	0.08	0.10	1.06	0.27	0.59	0.18	0.20
G	Maximum	238.17	244.11	97.32	109.88	118.27	142.65	137.98	168.87	203.20	165.95	147.23	151.29	116.17	122.41	160.52	164.63	155.97	158.02
	Minimum	238.04	243.92	97.19	109.79	118.03	142.51	137.73	168.70	203.08	165.71	147.11	151.24	116.08	122.30	160.32	164.36	155.73	157.79
	Difference	0.13	0.19	0.13	0.09	0.24	0.14	0.25	0.17	0.12	0.24	0.12	0.05	0.09	0.11	0.20	0.27	0.24	0.23
H	Maximum	239.54	245.36	97.40	109.81	119.50	144.22	138.55	169.27	203.07	166.65	147.88	151.92	116.22	122.69	161.36	165.41	157.14	159.13
	Minimum	239.46	245.24	97.32	109.70	119.23	144.11	138.36	169.11	202.91	166.41	147.78	151.82	116.12	122.36	161.27	165.18	156.83	159.04
	Difference	0.08	0.12	0.08	0.11	0.27	0.11	0.19	0.16	0.16	0.24	0.10	0.10	0.10	0.33	0.09	0.23	0.31	0.09
I	Maximum	237.90	244.24	97.39	109.86	118.41	142.76	137.90	168.65	202.85	165.95	147.34	151.24	116.10	122.36	160.36	164.53	155.93	158.01
	Minimum	237.71	243.68	97.17	109.62	118.03	142.61	137.32	168.37	202.67	165.37	146.91	151.15	116.01	122.12	160.12	164.12	155.76	157.81
	Difference	0.19	0.56	0.22	0.24	0.38	0.15	0.58	0.28	0.18	0.58	0.43	0.09	0.09	0.24	0.24	0.41	0.17	0.20
J	Maximum	238.67	245.19	97.50	110.04	119.13	143.80	138.31	168.88	203.08	166.56	147.27	151.27	116.07	122.39	160.35	164.38	156.75	158.74
	Minimum	238.34	244.29	97.37	109.92	118.52	143.50	138.09	168.50	202.93	166.09	147.01	151.27	115.97	122.25	160.12	164.16	156.43	158.48
	Difference	0.33	0.90	0.13	0.12	0.61	0.30	0.22	0.38	0.15	0.47	0.26	0.00	0.10	0.14	0.23	0.22	0.32	0.26
K	Maximum	238.12	244.02	97.21	109.79	118.32	142.66	137.86	168.76	203.11	165.84	147.13	151.20	115.99	122.26	160.47	164.56	155.90	157.95
	Minimum	237.71	243.65	97.06	109.69	117.80	142.50	137.48	168.36	202.90	165.46	146.77	151.13	115.83	122.01	160.17	164.23	155.40	157.42
	Difference	0.41	0.37	0.15	0.10	0.52	0.16	0.38	0.40	0.21	0.38	0.36	0.07	0.16	0.25	0.30	0.33	0.50	0.53
Overall	Maximum	240.21	246.11	97.85	110.28	119.64	144.54	138.71	169.97	203.95	167.20	147.93	152.02	116.32	122.78	161.86	166.09	157.39	159.46
	Minimum	237.20	243.19	96.89	109.35	117.80	142.19	137.32	168.36	202.15	165.37	146.71	150.79	115.58	121.64	159.92	164.02	154.92	156.87
	Difference	3.01	2.92	0.96	0.93	1.84	2.35	1.39	1.61	1.80	1.83	1.22	1.23	0.74	1.14	1.94	2.07	2.47	2.59

The table above shows the results of fragment analysis of 6 replicates each of 2 varieties of sweet cherry (Sato-nishiki and Benishuho) with 10 markers in 11 laboratories. The maximum and the minimum represent the maximum and minimum fragment size of a single allele detected by fragment analysis in each laboratory. Symbols (1) and (2) represent the types of alleles detected with individual markers. For markers 3-1 and 4-2, two varieties analyzed were homozygous and showed one type of fragment size.

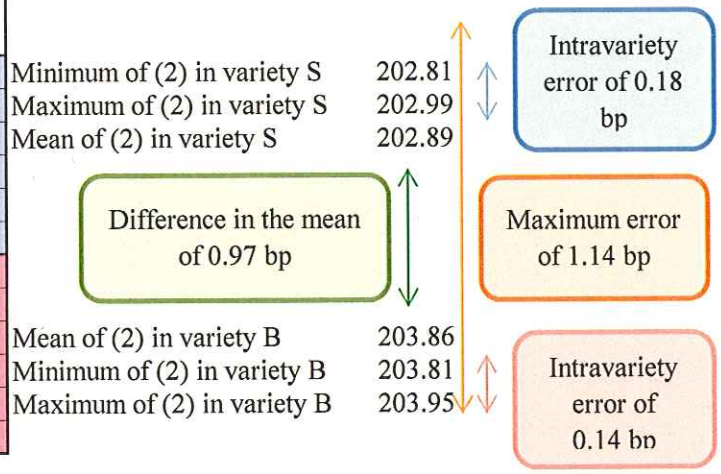
The table above shows stable results, and all intralaboratory errors in fragment size were less than 1 bp (with a mean of 0.24 bp), except for markers 4-1 (2) and 6-1 (2) in laboratory F.

However, interlaboratory errors often exceeded 1 bp, suggesting that it may be difficult to directly compare fragment size measurements. For interlaboratory comparison of fragment sizes, it is deemed necessary to use reference varieties that serve as reference standards for fragment sizes as shown in References 3 and 4 and convert measurements into the data presentation format in which data are presented on the basis of the distance between fragment sizes.

The results of markers 4-1 (2) and 6-1 (2) in laboratory F were explained by increased errors that resulted from PCR and fragment analysis of the two varieties on different days.

Reference 6-2 Example of errors related to PCR and fragment analysis performed on different days

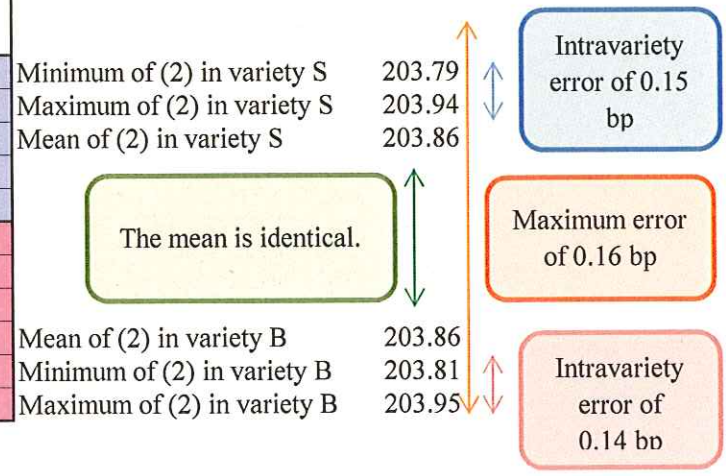
		Marker 4-1	
		Fragment size (1)	Fragment size (2)
F	S	169.37	202.91
		169.37	202.99
		169.30	202.88
		169.43	202.89
		169.30	202.81
		169.30	202.83
	B	169.85	203.83
		169.94	203.85
		169.89	203.85
		169.83	203.87
		169.97	203.81
		169.96	203.95



* PCR and fragment analyses of variety S and variety B were performed on different days. There were differences in fragment size (2) between varieties S and B, with a mean difference of 0.97 bp and a maximum difference of 1.14 bp.
However, replicates of each variety analyzed on the same day showed stable data.
Similar cases reported from other laboratories suggest that the large difference in data between varieties S and B may be explained by mechanical errors associated with PCR and fragment analyses performed on different days.

Results of PCR and fragment analysis performed on the same day

		Marker 4-1	
		Fragment size (1)	Fragment size (2)
F	S	169.78	203.79
		169.90	203.85
		169.83	203.86
		169.90	203.94
		169.96	203.87
	B	169.85	203.83
		169.94	203.85
		169.89	203.85
		169.83	203.87
		169.97	203.81
		169.96	203.95



Reference 6-3 Errors related to a sequencer, polymer, and capillary

Data were compared by type of sequencers, polymers, and capillaries used in laboratories participating in the interlaboratory collaborative study of sweet cherries. The results indicate that it is difficult to directly compare data on fragment size between different laboratories even if they use the same devices and reagents.

For direct comparison of data on fragment size, results of PCR and fragment analysis performed in a single laboratory must be used.

For interlaboratory comparison of data, it is deemed necessary to use reference varieties that serve as reference standards for fragment sizes as shown in References 3 and 4 and convert measurements into the data presentation format in which data are presented on the basis of the distance between fragment size peaks.

Although, relatively stable data were obtained with a 3100/3130xl 50cm Capillary Array, only two laboratories used the capillaries. Given this, the result should not lead to the conclusion that data obtained with a 3100/3130xl 50cm Capillary Array may be compared directly between laboratories.

Type of sequencer	Number of laboratories	Fragment size	Marker																	
			1-1		1-2		2-1		3-1	4-1		4-2	5-1		6-1		7-1		8-1	
			(1)	(2)	(1)	(2)	(1)	(2)	(1)	(1)	(2)	(1)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
ABI 310	5	Maximum	238.67	245.19	97.85	110.28	119.13	143.80	138.31	168.94	203.34	166.56	147.38	151.47	116.22	122.39	160.54	164.60	156.75	158.74
		Minimum	237.28	243.19	97.07	109.62	118.03	142.61	137.32	168.37	202.15	165.37	146.71	150.79	115.79	122.03	160.11	164.12	155.62	157.62
		Difference	1.39	2.00	0.78	0.66	1.10	1.19	0.99	0.57	1.19	1.19	0.67	0.68	0.43	0.36	0.43	0.48	1.13	1.12
ABI 3100	2	Maximum	238.12	244.02	97.21	109.79	118.32	142.66	137.89	168.76	203.11	165.84	147.16	151.20	115.99	122.26	160.47	164.56	155.90	157.95
		Minimum	237.20	243.31	97.06	109.65	117.80	142.19	137.42	168.36	202.84	165.46	146.77	151.07	115.58	121.64	159.92	164.02	154.92	156.87
		Difference	0.92	0.71	0.15	0.14	0.52	0.47	0.47	0.40	0.27	0.38	0.39	0.13	0.41	0.62	0.55	0.54	0.98	1.08
ABI 3130	4	Maximum	240.21	246.11	97.54	109.89	119.64	144.54	138.71	169.97	203.95	167.20	147.93	152.02	116.32	122.78	161.86	166.09	157.39	159.46
		Minimum	238.04	243.92	96.89	109.35	118.03	142.51	137.73	168.70	202.81	165.71	147.11	151.24	116.08	121.72	160.32	164.36	155.73	157.79
		Difference	2.17	2.19	0.65	0.54	1.61	2.03	0.98	1.27	1.14	1.49	0.82	0.78	0.24	1.06	1.54	1.73	1.66	1.67

Polymer	Number of laboratories	Fragment size	Marker																	
			1-1		1-2		2-1		3-1	4-1		4-2	5-1		6-1		7-1		8-1	
			(1)	(2)	(1)	(2)	(1)	(2)	(1)	(1)	(2)	(1)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
POP-4	8	Maximum	239.54	245.36	97.85	110.28	119.50	144.22	138.55	169.27	203.34	166.65	147.88	151.92	116.22	122.69	161.36	165.41	157.14	159.13
		Minimum	237.20	243.19	97.06	109.62	117.80	142.19	137.32	168.36	202.15	165.37	146.71	150.79	115.58	121.64	159.92	164.02	154.92	156.87
		Difference	2.34	2.17	0.79	0.66	1.70	2.03	1.23	0.91	1.19	1.28	1.17	1.13	0.64	1.05	1.44	1.39	2.22	2.26
POP-7	3	Maximum	240.21	246.11	97.54	109.89	119.64	144.54	138.71	169.97	203.95	167.20	147.93	152.02	116.32	122.78	161.86	166.09	157.39	159.46
		Minimum	239.52	245.29	96.89	109.35	119.22	144.28	138.47	169.24	202.81	166.65	147.48	151.62	116.10	121.72	161.34	165.28	157.08	159.17
		Difference	0.69	0.82	0.65	0.54	0.42	0.26	0.24	0.73	1.14	0.55	0.45	0.40	0.22	1.06	0.52	0.81	0.31	0.29

Capillary	Number of laboratories	Fragment size	Marker																	
			1-1		1-2		2-1		3-1	4-1		4-2	5-1		6-1		7-1		8-1	
			(1)	(2)	(1)	(2)	(1)	(2)	(1)	(1)	(2)	(1)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
310 Capillaries 47 cm ×	5	Maximum	238.67	245.19	97.85	110.28	119.13	143.80	138.31	168.94	203.34	166.56	147.38	151.47	116.22	122.39	160.54	164.60	156.75	158.74
		Minimum	237.28	243.19	97.07	109.62	118.03	142.61	137.32	168.37	202.15	165.37	146.71	150.79	115.79	122.03	160.11	164.12	155.62	157.62
		Difference	1.39	2.00	0.78	0.66	1.10	1.19	0.99	0.57	1.19	1.19	0.67	0.68	0.43	0.36	0.43	0.48	1.13	1.12
3100-Avant/3130	2	Maximum	240.21	246.11	97.21	109.79	119.45	144.54	138.71	169.97	203.95	167.20	147.61	151.70	116.20	122.78	161.86	166.09	157.39	159.46
		Minimum	237.71	243.65	96.89	109.35	117.80	142.50	137.48	168.36	202.81	165.46	146.77	151.13	115.83	121.72	160.17	164.23	155.40	157.42
		Difference	2.50	2.46	0.32	0.44	1.65	2.04	1.23	1.61	1.14	1.74	0.84	0.57	0.37	1.06	1.69	1.86	1.99	2.04
3100/3130 xl 36cm	2	Maximum	238.17	244.11	97.32	109.88	118.27	142.65	137.98	168.87	203.20	165.95	147.23	151.29	116.17	122.41	160.52	164.63	155.97	158.02
		Minimum	237.20	243.31	97.09	109.65	117.86	142.19	137.42	168.41	202.84	165.56	146.99	151.07	115.58	121.64	159.92	164.02	154.92	156.87
		Difference	0.97	0.80	0.23	0.23	0.41	0.46	0.56	0.46	0.36	0.39	0.24	0.22	0.59	0.77	0.60	0.61	1.05	1.15
3100/3130 xl 50cm	2	Maximum	239.76	245.47	97.54	109.89	119.64	144.37	138.56	169.36	203.07	166.82	147.93	152.02	116.32	122.77	161.51	165.48	157.36	159.43
		Minimum	239.46	245.24	97.32	109.70	119.23	144.11	138.36	169.11	202.91	166.41	147.78	151.82	116.12	122.36	161.27	165.18	156.83	159.04
		Difference	0.30	0.23	0.22	0.19	0.41	0.26	0.20	0.25	0.16	0.41	0.15	0.20	0.20	0.41	0.24	0.30	0.53	0.39

Note: Symbols (1) and (2) indicate the type of alleles.

Reference 6-4 Example of mechanical errors in fragment analysis

The table on the left shows the results of fragment analysis of varieties S and B with marker 1-1 in laboratory F.

For fragment size (2), the first measurement of variety S is slightly larger.

Visual inspection of fragment images revealed no sign of abnormalities in fragment analysis.

As indicated in the following cases, among a set of samples subjected to fragment analysis, the first sample applied to a capillary sequencer may produce a somewhat different result.

- The laboratory used a single-capillary sequencer.
- A laboratory using a four-capillary sequencer reported a case in which the first four data were smaller than the subsequent data.

Marker 1-1				
		Order of fragment analysis	Fragment size (1) (variety B	Fragment size (2)
F	S	1	245.19	/ 245.19
		2	244.87	/ 244.87
		3	244.71	/ 244.71
		4	244.53	/ 244.53
		5	244.50	/ 244.50
		6	244.37	/ 244.37
	B	7	238.67	/ 244.63
		8	238.38	/ 244.29
		9	238.53	/ 244.51
		10	238.66	/ 244.55
		11	238.34	/ 244.30
		12	238.67	/ 244.63

Reference 7 Method for fragment analysis in SSR

Alleles amplified with SSR markers are detected as fragments (waveforms). These fragments have some features. The understanding of these features allows accurate analysis of detectable fragments.

1. Characteristic appearance

In fragment analysis of PCR products prepared with SSR markers, smaller stepwise fragments may be detected in the vicinity of the target fragment. These fragments usually consist of nucleotide repeats of SSR and show a characteristic appearance as shown in Figure 5, with three or four fragments detected before the target fragment (fewer bp) and one or two fragments detected after the target fragment (more bp).

At one base apart from those fragments, other fragments called stutter bands may be detected (Figure 6). The presence of stutter bands makes fragment analysis difficult. Stutter bands may preclude accurate interpretation particularly in a DNA sample that possibly contains DNA of multiple varieties or in high-polypoidy varieties. The use of 3 repeat units or longer SSRs reduces the influence of stutter bands on analysis; however, those SSRs tend to be less polymorphic than 2 repeat units SSRs. Thus, it is difficult to develop effective markers for variety identification. It should be noted that 3 repeat units or longer SSRs are patent-protected until 2008.

The influence of stutter bands is reduced by adding the pig-tailing (GTTTCTT) to a reverse primer. Thus, it is recommended that pig-tailing be added to a reverse primer when 2 repeat units SSRs are used.

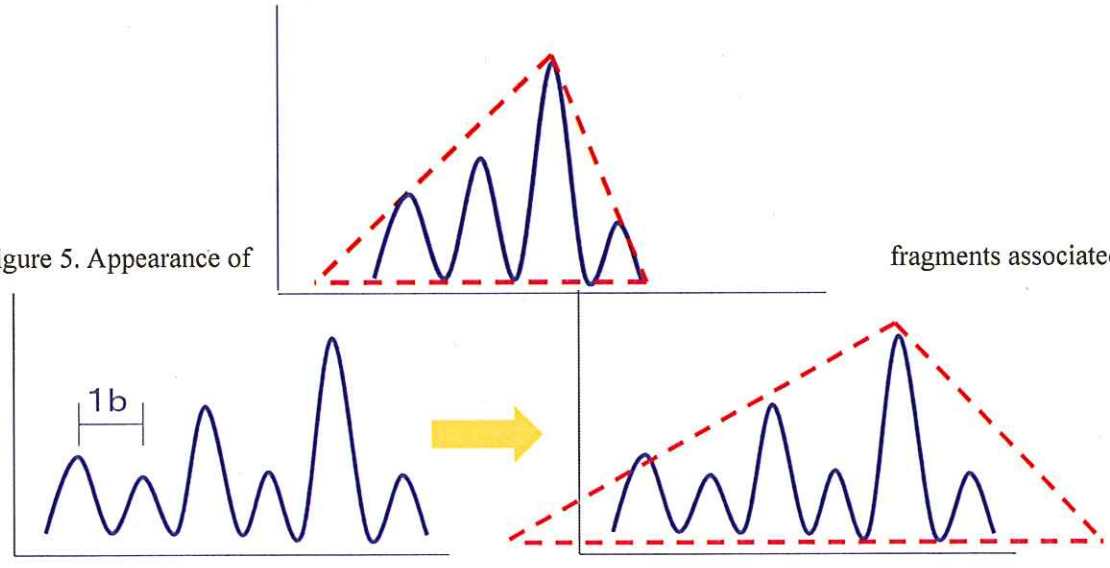


Figure 5. Appearance of fragments associated with SSR markers.

Figure 6. Appearance of fragments and stutter bands associated with 2 repeat units SSR.

2. Shorter fragments have higher amplification efficiency.

For SSR markers, shorter fragments tend to have higher amplification efficiency. For example, when 100-bp and 120-bp fragments are amplified, the peak of the 100-bp fragment is higher than that of the 120-bp fragment. The greater the difference in length between two fragments, the greater the difference in height (Figure 7).

Thus, caution should be exercised when PCR products are diluted e.g. with sterile water or inadequately amplified PCR products are used for fragment analysis. When those PCR products are used and the peak of a detectable fragment is less than half that of the size standard, other fragments may be overlooked (if fragments do not show an SSR-specific appearance or if the height of the fragment peaks are similar to that of background noise). To avoid this, the highest peak among detected fragment peaks should preferably be higher than that of the size standard. When fragment peaks are low, fragment analysis should be performed again with a different dilution ratio of PCR products.

It also should be noted that use of the same marker might result in different heights of fragment peaks, depending on whether alleles of the test variety are homozygous or heterozygous (Figures 8 and 9).

Understanding Figures 5 to 9 allows accurate analysis of fragments that are seemingly difficult to analyze.

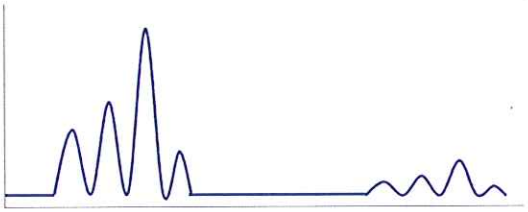


Figure 7. Relationship between fragment peaks when alleles are distant from each other (schematic diagram).

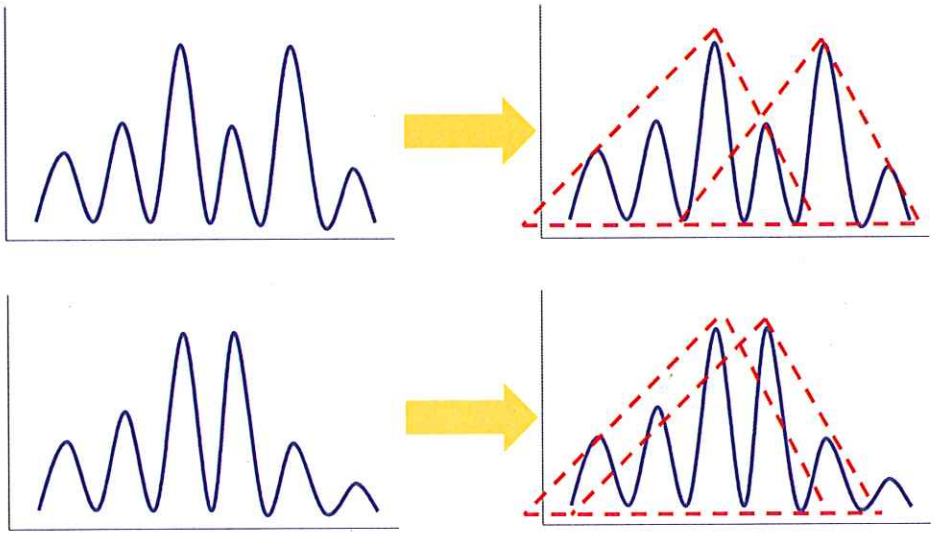


Figure 8. Examples of fragments suggesting the presence of two peaks.

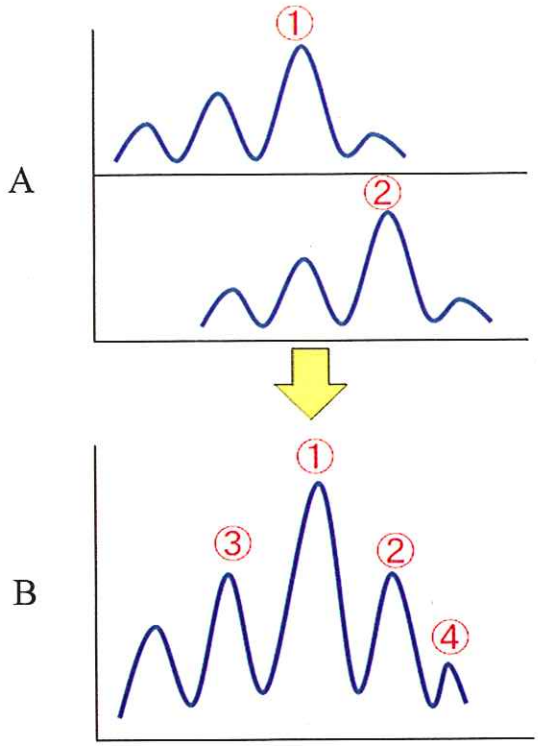


Figure 9. Example of analysis of waveforms.

Fragments as shown in Figure 9B are likely to result from the overlap of two types of fragments as shown in Figure 9A when the following two criteria are met. In this case, the analyst should conclude that there are two peaks, that is, peak 1 and peak 2.

- Peak 2 is similar to peak 3.
- Peak 2 is smaller than peak 1.

In addition, the presence of small peak 4 is a supplementary criterion.

3. Others considerations

When the concentration of PCR products used in fragment analysis are low, fragment peaks are low. When the concentration is too high, fragment peaks may exceed the upper limit of analysis, or an accurate size may not be displayed (Figure 10). To determine whether the concentration of the PCR products are adequate, it is important to simultaneously analyze a fluorescent dye other than that used for markers and size standard in fragment analysis. If the concentration of the PCR product is too high, the fluorescent dye that should not be detected may be detected as shown in Figure 11. This phenomenon occurs when the intensity of a fluorescent dye attached to the PCR product is too high, influencing another fluorescent dye of similar color. In this case, the PCR product should be diluted and analyzed again.

It is also important to ensure that the size standard fragments are appropriately displayed. As fragment size is calculated from the distance between size standard peaks, skewed peaks of size standard preclude accurate analysis of fragment sizes. When size standard peaks are skewed, re-analysis is needed.

Fragments that do not show an SSR-specific pattern as shown in Figure 5 (fragments without

accompanying low fragments in the vicinity and those with a fragment width greater than 1 bp) are not derived from alleles amplified with SSR markers, and those fragments should not be regarded as SSR fragments (Figure 12).

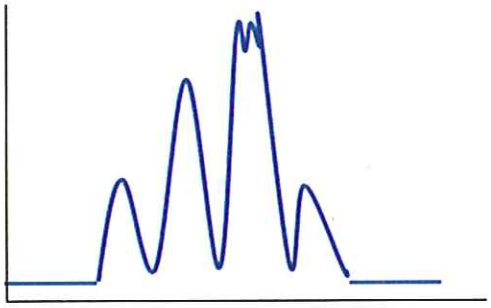


Figure 10. Irregular fragments confirmed to high concentrations of PCR product (1).
An irregular peak suggests that the concentration of the PCR product is too high.

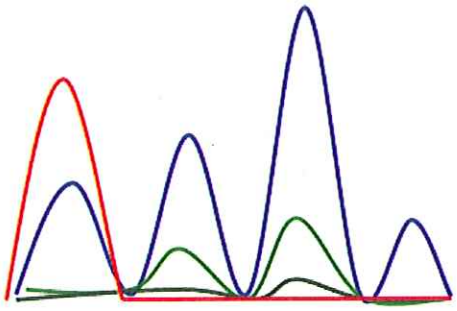
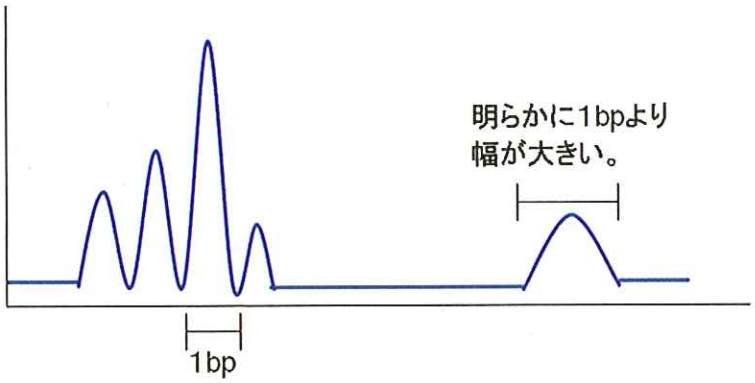


Figure 11. Irregular fragments confirmed to high concentrations of PCR products (2).
Blue, target fragments; red, size standard; green/yellow, noise.
There are blue fluorescent peaks in the figure above, and other fluorescent dyes are influenced by the high concentration of PCR product.



The width is substantially greater than 1 bp.
Figure 12. Example of a fragment unrelated to the products amplified with the SSR markers.
The peak on the right of the figure above has no accompanying, SSR-specific, small waveforms.
The peak with a width greater than 1 bp is unrelated to DNA.

Reference 8 Guidelines for validation of qualitative analytical methods in the food industry

As of March 2008, except for ISO 16140, there are no international guidelines for collaborative study of qualitative analytical methods comparable to the harmonized protocol for quantitative chemical analysis developed by IUPAC/ISO/AOAC International or the harmonized protocol for analysis of medicinal products developed by the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH).

In February 2006, the IUPAC started drafting guidelines for collaborative study of qualitative analytical methods and screening (such as detection of allergens by ELISA methods, GMO detection by PCR methods, and rapid screening with lateral flow devices).[4]

The guidelines for the collaborative study (2005)[3] developed by the AOAC international (AOAC) are those for quantitative analytical methods, but provide minimum criteria for the conduct of an collaborative study of qualitative analytical methods, although it is stated that the criteria are not part of the harmonized protocol.

AOAC minimum criteria for the conduct of an collaborative study of qualitative analytical methods

Ten laboratories should report the results of two levels of a positive sample per matrix and six samples per level as well as six negative samples per matrix.

In addition, there are AOAC methods committee guidelines for validation of qualitative and quantitative food microbiological official methods of analysis.[12] The guidelines provide guidelines for designing a method comparison study of qualitative analysis of food microorganisms conducted in a single laboratory (comparison of the performance between the alternative method and the reference method) and for designing an collaborative study. The requirements regarding the number of laboratories in an collaborative study and the number of replicates are the same as those in the minimum criteria described above.

The statistical background of these minimum criteria is detailed later. On the assumption that the true sensitivity (the proportion of positive samples correctly identified) or the true specificity (the proportion of negative samples correctly identified) would be 80%, the number of laboratories (10) and the number of replicates (6) are determined so as to mostly satisfy inequality (5), which gives the requirements for the number of laboratory L and the number of replicates m required for a 90% confidence interval for the estimate of the proportion of those correctly identified to fall within 80% ± 10% of the value (72% to 88%).

ISO 16140:2003 (Microbiology of Food and Animal Feeding Stuffs — Protocol for the Validation of Alternative Methods) stipulates the following protocol for collaborative study of qualitative analytical methods.[13]

Minimum requirements for the conduct of the collaborative study of qualitative analytical methods stipulated in ISO 16140:2003

- To report the results without outliers from at least 10 laboratories.
- To use at least three levels of samples including negative samples, samples with levels slightly higher than the detection limit for the alternative method, and samples with levels

about 10 times higher than the detection limit.

- To measure at least eight replicates of blind samples per level by the alternative and reference methods in each laboratory.
- To report at least 480 measurements per sample (2 methods × 10 laboratories × 3 levels × 8 replicates).
- To conduct a study of an alternative method with the use of at least one type of product belonging to each food category selected from the list in Annex B, which outlines sample types to be tested for each microorganism.

Statistical background of the numbers of laboratories and samples required for the collaborative study of qualitative analysis

Described below is the result of a study conducted by McClure (1990)[14] that is incorporated into the AOAC minimum criteria for the conduct of collaborative study of methods for qualitative analysis.

1) The basic idea is that the numbers of laboratories and samples are determined so that sensitivity p_+ obtained from analytical results (sample data) should fall within a certain range (acceptable range) $P_+ \pm d$ of the sensitivity in a population (true sensitivity P_+).
2) The sensitivity obtained from sample data is associated with an estimated error, and thus the confidence interval (a range within which the sensitivity in the population is estimated to fall) is calculated. It should be ensured that the confidence interval should fall within the acceptable range.
3) On the assumption that sensitivity values in the population follow a normal distribution, the confidence interval for sensitivity values obtained from sample data is:

$$p_+ - \frac{z_{1-\alpha} S_a}{L^{1/2} m} \le P_+ \le p_+ + \frac{z_{1-\alpha} S_a}{L^{1/2} m} \qquad (1)$$

where L is the number of laboratories, m is the number of samples, and $z_{1-\alpha}$ is the z score with the upper probability of standard normal distribution being α. The standard deviation S_a of correct number $a_i (i = 1, \cdots, L)$ obtained from each laboratory is:

$$S_a = \sqrt{\frac{\sum_{i=1}^{L} (a_i - \bar{a})^2}{L-1}} \qquad (2)$$

where $\bar{a}$ is the mean of a_i ($\bar{a} = \sum_{i=1}^{L} a_i \Big/ L$).

$S_a / (L^{1/2} m)$ in inequality (1) is the standard error of sensitivity p_+ obtained from sample data (see term A4 in **Annex 1**).

4) In order for the confidence interval described in inequality (1) to fall within the acceptable range $P_+ \pm d$, the following condition should be met:

$$\frac{z_{1-\alpha} S_a}{L^{1/2} m} \le d \qquad (3)$$

Square both sides of inequality (3), and solve for Lm^2 to give the following.

$$\left(\frac{z_{1-\alpha}S_a}{d}\right)^2 \leq Lm^2 \qquad (4)$$

Inequality (4) is used to determine the numbers of laboratories and samples.

5) In a report by McClure (1990), the minimally acceptable criteria for qualitative analytical methods was to accurately identify at least 80% of positives and negatives (the sensitivity and specificity being 80% or more) and be able to estimate the detection rate at a two-sided significance level of 10% (one-sided α = 0.05) or within ± 10% of 80% (± 0.08) (d = 0.08), and the numbers of laboratories and samples required to meet the criteria were investigated.

6) When S_a^2 in inequality (4) is determined, the left side value is given, and the conditions that Lm^2 should satisfy are determined. In a computer simulation (see **Annex 1**), McClure (1990) generated simulation data to explore S_a^2 and obtained an estimate (S_a^2 = 0.8558).

7) Substitution of $z_{0.95}$ = 1.645, d = 0.08, and S_a^2 = 0.8558 into inequality (4) gives the following:

$$362 \leq Lm^2 \qquad (5)$$

8) To satisfy inequality (5), McClure (1990) recommended the number of laboratories (L = 15) and the number of samples (m = 5) (the number of known positive samples $n_1 = m = 5$, the number of known negative samples $n_2 = m = 5$).

$$362 \leq 15 \times 5^2 = 375 \qquad (6)$$

9) McClure (1990) reported that the number of laboratories L may be reduced, but should be at least 10 and that n_1 and n_2 should be determined so that the number of samples should satisfy the following: $Ln_1 = Ln_2 \approx 75$. McClure argued that the conditions ensure that sensitivity bias does not exceed 10%, because it is proportional to the inverse of the number of laboratories, and that the distribution of the test statistic of outliers (Cochran's Q test) can be approximated by a chi-square distribution. However, Cochran's description[15], which was cited as grounds to believe that involvement of at least 10 laboratories reduces sensitivity bias below 10%, is applicable to the bias in the ratio estimate with auxiliary variables, but not applicable to the proportion correctly identified in qualitative analytical methods. According to Tate and Brown,[16] Cochran's Q test statistic can be well approximated by a chi-square distribution when at least four test samples (including positives and negatives) are used and when at least six laboratories are available after exclusion of those with outliers. Thus, the numbers of laboratories and replicates are determined by inequality (5) in order to estimate the proportion of correctly identified (80%) with a precision of ±10% at a significance level of 10%.

10) AOAC guidelines for the collaborative study in 1995[17] and in 2000[18] adopted L = 15 and m = 5, which were changed in 2002[19] (L = 10 and m = 6).

$$362 \approx 10 \times 6^2 = 360 \qquad (7)$$

Annex 1 Simulation method by McClure (1990)

SAS uniform procedure was used to generate uniform random numbers (0 to 1). The uniform procedure output was considered 1 (positive) when the output was greater than 0.2, and the output was considered 0 (negative) otherwise (corresponding to a detection rate of 80%). A data set (L = 10 laboratories, m = 5 replicates per laboratory) was prepared and analyzed with the following randomized block, two-way analysis of variance model.

$$X_{ij} = \mu + \alpha_i + \pi_j + \varepsilon_{ij} \quad \text{(A1)}$$

where X_{ij} is the response value (0 or 1), m is the true sensitivity, α_i is the variation of response values related to laboratory i, π_j is the variation of response values related to analysis sample j, and e_{ij} is the random error. Specifically, with the use of SAS Varcomp procedure (nested design analysis of variance), the data set was analyzed by equation (A1) to compute the degree of variance of α_i, π_j, and e_{ij}. When the analysis was repeated 500 times, the variance of π_j was not significantly different from the random error in all the 500 analyses (p = 0.25), and thus equation (A1) was simplified into the following one-way analysis of variance model:

$$X_{ij} = \mu + \alpha_i + \varepsilon_{ij} \quad \text{(A2)}$$

Given equation (A2), the unbiased variance of sensitivity is:

$$\sum_{i=1}^{L}(a_i/m - \bar{a}/m)^2 \Big/ (L-1) = \frac{1}{m^2}\sum_{i=1}^{L}(a_i - \bar{a})^2/(L-1) = \frac{S_a^2}{m^2} \quad \text{(A3)}$$

The standard error of sensitivity is the square root of the unbiased variance in equation (A3) divided by L:

$$\frac{S_a}{L^{1/2}m} \quad \text{(A4)}$$

S_a^2 values are calculated from individual data sets, and the mean of S_a^2 values in 500 data sets is 0.8558.

Reference 9 Example of developmental validation

PowerPlex Y Validation

Reference: Krenke *et al.* (2005) *Forensic Sci. Int.* 148(1):1-14

Study Completed	Description of Samples Tested (performed in 7 labs and Promega)	# Run
Single Source (Concordance)	5 samples x 8 labs	40
Mixture Ratio (male:female)	6 labs x 2 M/F mixture series x 11 ratios (1:0,1:1,1:10,1:100,1:300,1:1000,0.5:300, 0.25:300,0.125:300, 0.0625:300, 0.03:300 ng M:F)	132
Mixture Ratio (male:male)	6 labs x 2 M/M mixtures series x 11 ratios (1:0,19:1,9:1,5:1,2:1,1:1,1:2,1:5,1:9,1:19,0:1)	132
Sensitivity	7 labs x 2 series x 6 amounts (1/0.5/0.25/0.125/0.06/0.03)	84
Non-Human	24 animals	24
NIST SRM	6 components of SRM 2395	6
Precision (ABI 3100 and ABI 377)	10 ladder replicates + 10 sample replicates + [8 ladders + 8 samples for 377]	36
Non-Probative Cases	65 cases with 102 samples	102
Stutter	412 males used	412
Peak Height Ratio	N/A (except for DYS385, but no studies were noted)	
Cycling Parameters	5 cycles (28/27/26/25/24) x 8 punch sizes x 2 samples	80
Annealing Temperature	5 labs x 5 temperatures (54/58/60/62/64) x 1 sample	25
Reaction volume	5 volumes (50/25/15/12.5/6.25) x [5 amounts + 5 concentrations]	50
Thermal cycler test	4 models (480/2400/9600/9700) x 1 sample + [3 models x 3 sets x 12 samples]	76
Male-specificity	2 females x 1 titration series (0-500 ng female DNA) x 5 amounts	10
TaqGold polymerase titration	5 amounts (1.38/2.06/2.75/3.44/4.13 U) x 4 quantities (1/0.5/0.25/0.13 ng DNA)	20
Primer pair titration	5 amounts (0.5x/0.75x/1x/1.5x/2x) x 4 quantities (1/0.5/0.25/0.13 ng DNA)	20
Magnesium titration	5 amounts (1/1.25/1.5/1.75/2 mM Mg) x 4 quantities (1/0.5/0.25/0.13 ng DNA)	20
	TOTAL SAMPLES EXAMINED	**1269**

Validation of 21-SNP multiplex

Study Conducted	Description of Samples Tested	# Run	If redone now
Single Source (Concordance)	32 samples for 3 individuals	96	24
Mixture Samples	not performed		10
Mixture Ratio (male:female)	not performed		
Mixture Ratio (male:male)	not performed		
Sensitivity	1ng, 500pg, 250pg, 125pg, 63pg, 31pg, 16pg, 0pg x 12 dilution series	96	16
Sensitivity	Artificially degraded samples	18	16
Non-Human	dog, cat, guinea pig, ferret, horse, chicken, wolf, toad, rat, bull, cow, deer, badger, pigeon, spider monkey, otter, gorilla, chimpanzee and orangutan, *Micrococcus luteus*, *Escherichia coli* and *Clostridium perfringens*	22	8
Precision (ABI 3100 and ABI 377)	used samples from single source and dilution series		
Non-Probative Cases	6 cases, various crime scene samples plus reference buccals	18	18
Stutter	n/a		
Peak Height Ratio	used samples from single source and dilution series		
Cycling Parameters	not part of validation, part of previous work		
Annealing Temperature	not part of validation, part of previous work		
Reaction volume	not part of validation, part of previous work		
Thermal cycler test	not part of validation, part of previous work		
Male-specificity	n/a		
Other: Population database samples	201 White Caucasian samples, 71 Indian Sub-continent and 86 Afro-Caribbean	358	300
Other: Reproducibility	Different sample types - blood, hair, semen, saliva, trace swabs	96	32
Other: CE injection parameters	Dilution series	8	8
	TOTAL SAMPLES EXAMINED	**712**	432

2. Glossary

Locus

A specific region of genomic DNA is called a locus. A polymorphic locus may be a gene, a part of a gene, or a region other than genes.

Allele

Individual nucleotide sequences with polymorphisms at a locus are called alleles. Alleles may result from substitution, insertion, or deletion of a part of nucleotide sequence in a DNA region, or different numbers of repeat units (number of repeats) in a tandem repeat sequence.

Primer

A short single-stranded DNA that is complementary to a template DNA and used to synthesize double-stranded DNA by DNA synthetase (DNA polymerase) is called a primer. Primers are required for polymerase chain reaction (PCR).

Marker

A symbol indicating a specific polymorphism in a specific DNA region with the use of primers, probes, restriction enzymes, and others used for polymorphic analysis is called a marker.

Fragment analysis

Analysis of the length of PCR products prepared by PCR amplification of a genomic template with fluorescent dye-labeled primers and subsequent capillary electrophoresis is called fragment analysis.

Sequencer

A sequence is an automated device used to read the nucleotide sequence recorded on DNA. It may be called a DNA sequencer. Technology for DNA nucleotide sequencing is used for investigation gene structure and function, research on revolutionary lineage, individual identification and analysis, and many other purposes.

RAPD

RAPD is an abbreviation for random amplified polymorphic DNA. When genomic DNA templates are amplified by PCR with random primers, variations in the genomic DNA sequence may be associated with a difference in primer binding, which leads to different sizes and numbers of DNA fragments amplified with random primers. In the RAPD method, these differences are detected by agarose or polyacrylamide gel electrophoresis. RAPD markers are generally dominant markers.

When sequence-tagged site (STS) primers are used, unnecessary bands common across varieties disappear and only specific bands of interest are visualized, which solves RAPD-related problems of poor reproducibility and misidentification.

CAPS

CAPS is an abbreviation for cleaved amplified polymorphic sequence. The CAPS method is also called the PCR-RFLP method. This technique identifies RFLP by digestion of PCR products with specific restriction enzymes. The CAPS method may be used when the nucleotide sequence of amplified products is not known; however, the nucleotide sequence information of amplified products is helpful. When comparison of nucleotide sequences of DNA fragments in a specific region reveals any difference in the nucleotide sequence at the restriction site, the corresponding restriction enzyme should be used.

When there is no difference in an appropriate restriction site, derived CAPS (dCAPS) is often used in which specially designed primers are used to introduce a restriction site into PCR products.
CAPS markers and dCAPS markers are co-dominant markers and may detect single nucleotide polymorphisms.

SWGDAM

SWGDAM stands for Scientific Working Group on DNA Analysis Methods. Its predecessor Technical Working Group on DNA Analysis Methods (TWGDAM) was a working group organized by the Federal Bureau of Investigation (FBI) in response to a growing interest in the quality assurance in forensic medicine. The TWGDAM issued and revised TWGDAM guidelines in 1989, 1991, and 1995, and then it was reorganized into the SWGDAM that leads scientific working groups and FBI's scientific working groups.

AOAC International

The Association of Official Agricultural Chemists established in the United States in 1884 was renamed the Association of Official Analytical Chemists and then the AOAC International in 1992. The organization works to provide or promote the development and use of validated analytical methods, harmonization, and quality assurance at laboratories on an international basis.

IUPAC

The International Union of Pure and Applied Chemistry, an association of chemists established in 1919 for the international advancement of chemistry, standardizes chemical nomenclature, terminology, analytical methods etc.

ISO

The International Organization for Standardization (ISO), which was established in 1947, is a private nongovernmental organization that works to develop international standards. ISO consists of one member in each country, and each member is the organization for the development of standards in the country.

Harmonized protocol

A harmonized protocol is one for which multiple organizations for the development of standards reach a consensus. The collaborative study protocol for quantitative methods for chemical analysis is based on a consensus reached by organizations for the development of standards in individual countries that participated in the workshop organized by the IUPAC, ISO, and AOAC International. The harmonized protocol for the use of standardized analytical methods and presentation of their performance characteristics is based on a consensus reached by organizations for the development of standards in individual countries that participated in the workshop organized by the IUPAC. Harmonized protocols for proficiency testing, internal quality control, single laboratory validation, and recovery rate were developed collaboratively by the three bodies of the IUPAC, ISO, and AOAC International. ICH guidelines are also harmonized protocols.

ICH

ICH, an abbreviation of the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use, is an organization that consists of regulatory authorities for medicinal products and experts in the pharmaceutical industry in Japan, the United States,

and Europe for promotion of international standardization of technology for application of medicinal products etc. The ICH has published guidelines for validation of analytical methods.

ENFSI

FNFSI, an abbreviation of European Network of Forensic Science Institutes, is an expert group in forensic science in Europe. It was established in 1995 to share knowledge and experience in the forensic field. Its members include 53 laboratories in 31 countries (as of 2005). Its standing committees include the Expert Working Group Committee (EWGC), the Quality & Competence Committee (QCC), and the European Academy of Forensic Science (EAFS).

DAB Standards

DAB Standards stands for DNA Advisory Board Quality Assurance Standards for Forensic DNA Testing Laboratories. DAB Standards provide quality assurance standards for forensic laboratories performing DNA analysis. Under the DNA Identification Act enacted in the United States in 1994, the DNA Advisory Board was established by the Director of the Federal Bureau of Investigation (FBI) for the development, revision, and recommendation of quality assurance standards.

Haplotype

Haplotype is the contraction of haploid genotype and represents a genetic composition (specifically, DNA sequence) on a single chromosome of living organisms. In diploid organisms, a haplotype is a set of alleles at individual genetic loci on one of a pair of chromosomes. A haplotype may refer to the entire genome (across multiple chromosomes). In this meaning; however, the term specifically indicates a set of genes derived from either parent. At present, in a narrow sense, a haplotype often indicates a set of polymorphisms (such as single nucleotide polymorphisms [SNPs]) on a single chromosome that are statistically associated or genetically linked.

3. Bibliography

(1) Maki H. (1997) PCR Tips. Shujunsha Co., Ltd. p.95.

(2) DNA-Based Variety Identification Technology Committee. (2003) Theory of variety identification in plant variety identification: basic considerations in DNA-based plant variety identification.

(3) AOAC Int. (2005) Appendix D: Guidelines for Collaborative Study Procedures to Validate Characteristics of a Method of Analysis. In Official Methods of Analysis of AOAC Int. 18th ed., Gaithersburg, MD, USA. http://eoma.aoac.org/app_d.pdf

(4) IUPAC (2006) Establishment of guidelines for the validation of qualitative and semi-quantitative (screening) methods by collaborative trial: a harmonized protocol. http://www.iupac.org/projects/2005/2005-024-2-600.html

(5) ENFSI (2006) Standing committee for quality and competence (qcc): validation and implementation of (new) methods.

(6) Dixon LA, Murray CM, Archer EJ, Dobbins AE, Koumi P, Gill P. (2005) Validation of a 21-locus autosomal SNP multiplex for forensic identification purposes. *Forensic Sci Int.* 154: 62–77. http://www.cstl.nist.gov/biotech/strbase/validation.htm

(7) Krenke *et al.* (2005) *Forensic Sci. Int.* 148(1):1–14. http://www.cstl.nist.gov/biotech/strbase/validation.htm

(8) The National Academies Press (2006) Discussions of the Committee on Daubert Standards: Summary of Meetings.

http://books.nap.edu/catalog.php/record_id=11696#toc
(9) Takashina T, Ishiguro M, Nishimura K. (2007) Identification of varieties of imported and domestic sweet cherry with SSR markers. DNA Polymorphism. 15:101–104.
(10) Takashina T, Matsuda N, Yamamoto T, Hayashi T, Kimura T, Nishimura K. Variety identification and parentage analysis of sweet cherry with SSR markers. Information on Research Results at the National Agricultural Research Center for Tohoku Region in FY 2003. http://www.affrc.go.jp/ja/research/seika/data_tohoku/h15/seibutu/h15seibutu006
(11) Takashina T, Matsuda N, Kimura T, Yamamoto T, Nishimura K. (2004) Use of SSR markers in sweet cherry variety identification. Breeding Research. 6 (supple 1) 181.
(12) Feldsine P, Abeyta C, Andrews WH. (2002) AOAC INTERNATIONAL Methods Committee Guidelines for Validation of Qualitative and Quantitative Food Microbiological Official Methods of Analysis. J. AOAC Int. 85(5):1187–1200. http://www.aoac.org/vmeth/MICGUIDE.pdf
(13) ISO 16140 (2003) Microbiology of food and animal feeding stuffs — Protocol for the validation of alternative methods.
(14) McClure FD. (1990) Design and analysis of qualitative collaborative studies: minimum collaborative program. J. AOAC. 73:953–960.
(15) Cochran WG. (1977) Sampling Techniques. 3rd ed. John Wiley & sons, New York, p.160.
(16) Tate MW, Brown SM. (1970) J. Am. Statist. Assoc. 65:155–160.
(17) AOAC Int. (1995) AOAC INTERNATIONAL Guidelines for Collaborative Study Procedures To Validate Characteristics of a Method of Analysis. J. AOAC Int. 78(5): 143A–160A.
(18) AOAC Int. (2000) Appendix D: Guidelines for Collaborative Study Procedures To Validate Characteristics of a Method of Analysis. In Official Methods of Analysis of AOAC Int. 17th ed., Gaithersburg, MD, USA.
(19) AOAC Int. (2002) Appendix D: Guidelines for Collaborative Study Procedures To Validate Characteristics of a Method of Analysis. In Official Methods of Analysis of AOAC Int. 17th ed. rev.1, Gaithersburg, MD, USA. http://www.aoac.org/vmeth/Manual_Part_6.pdf

4. Working Committee

Chair: Akemi Yasui	Director of the Analytical Science Division, National Food Research Institute, Incorporated Administrative Agency National Agriculture and Food Research Organization
Member: Shigehiro Naito	Leading researcher, the Foodmetrics Laboratory, Analytical Science Division, National Food Research Institute, Incorporated Administrative Agency National Agriculture and Food Research Organization
Member: Satoru Matsumoto	Senior researcher, the Molecular Genetics and Physiology Research Team, National Institute of Vegetable and Tea Science, Incorporated Administrative Agency National Agriculture and Food Research Organization
Member: Toshiya Yamamoto	Leader of the Fruit Genome Research Team, National Institute of Fruit Tree Science, Incorporated Administrative Agency National Agriculture and Food Research Organization
Member: Tadashi Takashina	Research Specialist, the Biotechnology and Plant Breeding Division, Department of Agro-Production Science, Yamagata General Agricultural Research Center