INTERNATIONAL STANDARD

ISO 16578

Second edition
2022-09

Molecular biomarker analysis — Requirements for microarray detection of specific nucleic acid sequences

Analyse moléculaire des biomarqueurs — Exigences relatives à la détection sur microréseaux de séquences d'acides nucléiques spécifiques

Reference number
ISO 16578:2022(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 34 *Food products*, Subcommittee SC 16, *Horizontal methods for molecular biomarker analysis*.

This second edition cancels and replaces the first edition (ISO 16578:2013), which has been technically revised.

The main changes are as follows:

— Annex A has been added to provide the practical determination of limit of detection for microarray platform (LODP).

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

Available methods for nucleic acid sequence discovery using nucleic acid containing samples are based on three technologies: the polymerase chain reaction (PCR), deoxyribonucleic acid (DNA) sequencing and oligonucleotide microarrays.[1]

DNA microarrays are used in biomarker identification and the measurement of gene expression. International harmonization efforts of experimental methods for microarray experiments began with the "Minimum Information about a Microarray Experiment (MIAME)—toward standards for microarray data"[2] and the US Food and Drug Administration's critical path project the Microarray Quality Control project (MAQC),[3][4] which began in 2005 and focused on technical aspects of gene expression measurements, robust technology platforms and the development of accurate and reproducible multivariate gene expression-based prediction models.

DNA microarrays are made either by chemically synthesizing DNA probes on a solid surface or by attaching pre-made DNA probes to a solid surface, e.g. a microplate or coated bead. Microarray assays can be designed to detect multiple single nucleotide polymorphisms (SNPs) simultaneously. High density oligonucleotide microarrays synthesized *in situ* using techniques such as photolithography, ink-jet deposition and robotic arraying of PCR products and pre-synthesized oligonucleotides are manufactured for detecting specific sequences over a wide range of variability.[5] This technology continues to develop along with PCR and next generation sequencing (NGS).

DNA microarrays are typically used to probe a solution of mixed labelled nucleic acids: hybridization of the labelled targets to the fixed probes on the array is detected, and their relative concentration to the remaining nucleic acid species in solution is measured. By generalizing to a very large number of spots of DNA, an array can be used to quantify an arbitrarily large number of different nucleic acid sequences in solution.[6]

Microarray technologies are used in food analysis for detection and identification of genetically modified organisms (GMO) analysis and other biomarkers.[7][8][9][10][11][12] The focus of this document is DNA microarray-based methodologies for food products and products of agriculture.

Molecular biomarker analysis — Requirements for microarray detection of specific nucleic acid sequences

1 Scope

This document specifies verification and validation parameters and processes for microarray detection and identification of specific nucleic acid sequences.

This document provides recommendations and protocols for:

— microarray design and manufacture;

— validation of hybridization specificity;

— interlaboratory validation of qualitative methods;

— determination of limits of detection for a microarray;

— determination of range of reliable signals;

— criteria for assessing technical performance of the microarray platform:

This document is applicable to all methods that use microarrays for detection of nucleic acids.

It does not apply to the following protocols:

— quantitative measurement;

— requirements for sample preparation prior to DNA microarray experiments.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/TS 16393, *Molecular biomarker analysis — Determination of the performance characteristics of qualitative measurement methods and validation of methods*

ISO 16577, *Molecular biomarker analysis — Vocabulary for molecular biomarker analytical methods in agriculture and food production*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 16577 and the following apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

3.1
limit of detection for microarray platform
LODP
lowest relative quantity of the *external measurement standard* (3.11) (or reference material) for which a positive identification can be achieved with reasonable or previously determined confidence or both in a defined matrix using a specific analytical method

Note 1 to entry: In qualitative testing, an estimate of the LOD is measured at the chosen probability of detection (POD).

3.2
range of reliable signal
concentrations of target sequence for which a method can provide results where the output signal is proportional to the concentration and/or copy number of the *external measurement standard* (3.11) (or reference material)

Note 1 to entry: Direct of derived proportionality is acceptable in this range.

3.3
DNA microarray
DNA chip
solid substrate where a collection of *probe DNA* (3.6) arranged in a specific design is attached in a high-density fashion, directly or indirectly, that assays large amounts of biological material using high-throughput screening methods

3.4
analytical power
power
probability that an analyte will not go undetected if it is present

3.5
sensitivity
smallest treatment response that will be detectable

3.6
probe DNA
single-strand nucleic acid defined by its property to target specific nucleic acid sequence by base complementarities, where the stringency of the binding is linked with the length and nucleic acid composition of the probes, along with reaction parameters

3.7
platform
device that supports a *DNA microarray* (3.3) technology

3.8
fluorescence detection
method of detecting hybridization using immobilized *probe DNA* (3.6) by measuring a fluorescence signal

3.9
colorimetric detection
method of detecting hybridization using immobilized *probe DNA* (3.6) by measuring a colorimetric signal

3.10
electrochemical detection
method of detecting hybridization by measuring electric currents of an electrode onto which *probe DNA* (3.6) are immobilized

3.11
external measurement standard
material or substrate prepared for testing the compatibility of the microarray-based methods of analysis, whose property value is derived as a consensus value based on collaborative experimental work under the auspices of a scientific or engineering group

3.12
cross-hybridization
non-specificity binding of *probe DNA* (3.6) to non-targeted nucleic acid

4 Principle

4.1 DNA microarray platform assay

A microarray platform assay consists at a minimum of the following steps:

— denaturation of the double- or single-stranded DNA or ribonucleic acid (RNA) analyte (DNA sample);

— hybridization of the target(s) to probe DNAs bound to a solid substrate;

— detection of each hybridized target(s) by an electrochemical, colorimetric or fluorescence signal;

— data analysis.

The laboratory shall verify the procedure used for each microarray assay step, using known measurement standards (or reference material) and appropriate controls. Requirements governing verification of DNA microarray-based methods shall be documented.

4.2 Microarray design and manufacture

4.2.1 General

The DNA microarray and device for analysis shall be validated as an integrated measurement system for specific nucleic acid analysis. The performance of a DNA microarray shall be specified in combination with its analytical device. The DNA microarray and the device that analyses it should specify the performance in an integrated state and evaluate its reliability.[13]

4.2.2 Control probe sequences and targeted probes

A DNA microarray analysis shall have the following types of probe DNAs incorporated:

— external measurement standards (or reference material);

— a positive control;

— a negative control;

— the nucleic acid sequence of interest.

It shall be designed to be verifiable.

The immobilized DNA probes for signal quality control, including but not limited to the positive and negative controls, shall be included in the microarray design as replicates located in different positions on the microarray.[14] Probe DNA design shall consider the Tm value, GC ratio and sequence specificity of the nucleic acid oligonucleotide. Oligonucleotide probe sequence information shall be provided according to the International Union of Pure and Applied Chemistry (IUPAC) nomenclature code for nucleic acids.[15] In order to improve legibility between “G” and “C”, lower-case “g” should be used in the description (i.e. C, g, A, and T shall be used to indicate bases). The quality of probe DNA shall be ensured by an appropriate method, e.g. spectroscopic analysis, mass-spectroscopy analysis.

4.2.3 Analytical power of microarray assay

Power determines the number of replicates that are needed to validate the performance requirement.[16] An estimate of expected variance (uncertainty) is required among replicates (see ISO/IEC Guide 98-3[17]). The probability of detection (POD), which is a measure of the variance of the limit of detection (LOD) for a qualitative (binary) analysis, should be used to ensure that method performance lies within the chosen confidence interval. ISO/TS 16393 provides guidance for determining the number of replicates required for validating a qualitative method.

Replication should be differentiated from repetition within an assay, i.e. repetition of a hybridization on the same immobilized DNA sample position. The power and detection limit of a method are not necessarily improved by doubling the number of hybridizations for each DNA sample and decreasing the number of samples by one half.

NOTE In practice, methods with lower power or detection limit or both can be used to detect the presumptive presence of a target. Secondary methods with higher power or detection limit or both can be used to confirm the presence of the target or to resolve presumptive results from the first screen. This is frequently the case when a screening method is used followed by a method to disclose a specific construct.

4.3 Validation of hybridization specificity

4.3.1 Theoretical assessment of specificity

A theoretical assessment of probe DNA specificity shall be performed consisting of *in silico* screening of the probe against a nucleic acid sequence database that is fit for purpose for the analysis that will be performed. Examples of DNA databases are:

— DNA Data Bank of Japan (National Institute of Genetics)[18];

— EMBL-EBI (European Bioinformatics Institute)[19];

— GenBank (National Center for Biotechnology Information)[20].

Examples of sequence similarity search applications are:

— BLAST[21];

— FASTA[22].

Specific sequences should be selected that are not likely to cross-hybridize and be tested experimentally.

4.3.2 Experimental assessment of specificity

The sequence specificity of the probe DNAs should be validated experimentally on the basis of exclusivity and inclusivity, i.e. on samples having nucleic acid sequences similar to, but not the same as, the target sequence, as well as on biomarkers identified through the *in silico* assessment (see 4.3.1) as presenting sequences homologies likely to cause cross-hybridizations: exclusivity and inclusivity.[23] The experimental conditions should be the same as those used routinely for the method

4.3.3 Experimental assessment of cross-hybridization

The validation process shall demonstrate that no cross-hybridizations occurs on a probe DNA that is capable of experimentally detecting an external measurement standard (or reference material) in the matrix. An experimental result is accepted only if the probe DNAs for detecting the external measurement standards (or reference material) are all positive and the probe DNAs for detecting negative controls are negative.

4.4 Interlaboratory validation of qualitative methods

4.4.1 General

Qualitative (binary) test results only provide data for detection or non-detection within a pre-determined performance range. Sensitivity is an indicator of signal strength but cannot serve directly as a measure of system performance. Detection limits are direct indicators of system performance. The performance range for the method should be determined in the validation study. The validation study should begin once the method is established or modified.

4.4.2 Detection limit

The detection limit is the true net concentration of target DNA in the sample. It will lead, with probability (1-β), to the conclusion that the amount of target in the sample is larger than that in the negative control material. It is defined as shown in Formula (1):

$$P_r(\hat{L} \le L_c \,|\, L = L_D) = \beta \tag{1}$$

where

$\hat{L}$ is the estimated value;

L_c is the critical value;

L is the expectation or true value;

L_D is the LOD.

NOTE 1 The limit of detection is estimated by:

$$L_D \approx 2t_{1-\alpha\nu}\sigma_o$$

where

L_D is the LOD;

$\alpha = \beta$;

$t_{1-\alpha\nu}$ is Student's t-distribution value, based on ν degrees of freedom for a one-sided confidence interval of 1-α;

σ_o is the standard deviation of the true value (expectation).

$L_D = 3{,}29\ \sigma_o$, when the uncertainty in the mean (expected) value of the blank is negligible, $\alpha = \beta = 0{,}05$ and L is normally distributed with known constant variance. However, L_D is not defined simply as a fixed coefficient (e.g. 3, 6) times the standard deviation of a pure solution background. To do so can be extremely misleading. The correct estimation of L_D can consider degrees of freedom, α and β, and the distribution of L as influenced by factors such as analyte concentration, matrix effects and interference.

This description of the LOD is suitable for nucleotide analyses such as those used for real-time PCR that present as non-normal distributions with heteroscedasticity (e.g. "counting" (Poisson) processes).[24] It is essential to specify the measurement process under consideration since distributions, standard deviations and blanks can be dramatically different for different measurement processes.

NOTE 2 An empirically derived determination based on the results of a collaborative trial is called the "practical LOD". It is defined as the lowest relative quantity of the target DNA that can be detected, given a known (determined/estimated) number of target taxon copies. The practical LOD is related to the test portion, and the quality/quantity of the template DNA, and $L_D = 3{,}29\ \sigma_o$ which has also been called the absolute LOD of the method with 95 % confidence.

4.4.3 Probability of detection

A qualitative (binary) analytical DNA microarray method with the purpose of demonstrating the presence or absence of a given target sequence in a sample shall provide objective evidence that it is adequate for its intended use. Method performance should be characterized with respect to the concentration of the target sequence. The POD permits comparison of probabilities across concentrations. Validation data can be graphically represented as a POD response curve by concentration with associated error bars of the mean POD value which characterizes the response probability curve as a function of measurand mass or concentration.

At the LOD, a positive identification of specific probe can be achieved with reasonable or previously determined confidence or both in a defined matrix using a specific analytical method. In qualitative testing, an estimate of the LOD is measured at the chosen POD.

When LOD for each target probe is evaluated, method validation shall include determining the POD for the performance range of the method as described in ISO/TS 16393.

4.4.4 Limit of detection for microarray platform

4.4.4.1 General

When the number of probes is too large to evaluate the LODs for all target probes, the LODP can be used as a performance characteristic representing the LOD applicable to the whole microarray platform.

4.4.4.2 Practical LODP

An empirically derived determination based on the results of an experimental trial by use of an external reference material called the "practical LODP" has been developed for DNA microarrays. Like the practical LOD, it is defined as the lowest relative quantity of the selected external reference material that can be detected with 95 % confidence (1 in 20 false negative rate), given a known (determined/estimated) copy number of a reference material. The practical LODP is related to the test portion, and the quality/quantity of the template DNA, and $L_{DP} = 3{,}29\ \sigma_0$ which has also been called the "absolute LODP" of the method with 95 % confidence. The LODP is commonly rounded to one significant figure; therefore, LODP = 3 σ_0.[25]

NOTE The LODP determination method is described in Annex A.

4.4.5 Range of reliable signal

The range of reliable signal shall be determined for a DNA microarray method. The values should be confirmed via an interlaboratory trial using appropriate certified reference materials or reference materials. Information may also be derived from single laboratory studies.[26][27]

4.4.6 Test sample

A solution or an extract containing DNA/RNA molecules appropriate to the field of application is prepared so that there is no demonstrated hybridization inhibition or interference with the electrochemical, colorimetric and/or fluorescence detection.

4.4.7 Measuring system

The criteria for choosing instrument settings (e.g. background setting, normalization setting) shall be determined and documented. The measurement accuracy and uncertainty of instruments and equipment used for microarray detection of specific nucleic acid sequences can affect the validity of reported results.[28] The following instruments shall be calibrated:

— thermal cyclers;

— hybridization ovens, or other hybridization apparatus;

— DNA microarray scanner;

— apparatus or equipment for measuring DNA/RNA integrity and concentration.

Any calculations or models used to derive analytical result shall be validated and documented.

4.4.8 Estimation of measurement uncertainty

The appropriate measurement uncertainty shall be determined for each component of the DNA microarray assay. An overall measurement uncertainty shall be determined for the integrated method. Samples shall be representative. When evaluating the uncertainty of measurement, all components which are of significance in the given situation shall be taken into account using appropriate methods of analysis.

NOTE For further information, see the ISO 5725 series[29] and ISO/IEC Guide 98-3[17].

4.4.9 Microarray reagents

The characteristics and quality of reagents (e.g. fluorescent dye(s), reverse transcription enzyme, buffers), and the amount of an external measurement standard (or reference material) that is added to the reaction mixture should be validated.

5 Expression of results

5.1 General

Qualitative determinations will return a binary result: positive or negative. The criteria used to choose a method for a particular application, i.e. in a specific matrix where the analyte falls within a specific range of concentration can be based on the POD for the minimum and maximum levels of target and the measurement uncertainty associated with the method (see ISO/TS 16393).

5.2 Expression of a negative result

A negative result shall never be expressed as zero or "target not present".

The following sentence shall appear in the test report:

> "For target analyte X, the target sequence Y was not detected. The LOD of the method is x %."

If it cannot be demonstrated that the amount of target DNA included in the PCR is sufficient for a positive identification to be obtained with reasonable or previously determined confidence or both in a defined matrix using a specific analytical method, then the following sentence shall be added:

> "However, the amount of the target DNA extracted from species X may not have been sufficient for a positive identification to be obtained for this sample."

In addition, if applicable:

> "The LODP of the method was X."

5.3 Expression of a positive result

The following text shall appear in the test report:

> "For sample X, target sequence Y was detected."

In addition, if applicable:

> "The probability of detection is x %".

The identity of the molecular biomarker can be included, if available.

5.4 Expression of inconclusive or ambiguous results

Results from all test portions shall be consistent. When at least one test portion gives a positive result and at least one gives a negative result, the analysis shall be repeated. If at least one repetition of the procedure, beginning with the nucleic acid extraction, gives ambiguous results such as a positive and a negative result, the report should state that the sample is negative at the LOD. Results within the same test portion shall be consistent. In case of +/− results for the two replicates, repeat the two analyses for the respective test portion. If the two novel replicates are tested +/− or −/−, the test portion is considered as negative.

6 Test report

Reporting should be carried out as specified in the applicable standards (e.g. ISO/IEC 17025[28]).

The test report shall include at least the following information:

a) all information needed to identify the sample;

b) the date of receipt;

c) any particular information relating to the laboratory samples and any concomitant restrictions that can apply;

d) all information related to the test sample (the type, number of the sample);

e) any conditions regarding sample transport, as well as all storage, if applicable;

f) statement about the date and the type of sampling procedure(s) used, if applicable;

g) a description of the nucleic acid extraction method used;

h) identification of the analysis method and a general description of the microarray process that forms the basis of the analytical method(s);

i) both the positive and negative controls;

j) the type of external measurement standards (or reference materials);

k) LOD or practical LODP and the range of reliable signal for detection analyses;

l) any calculations or models used to derive the analytical result;

m) any outstanding points observed during testing.

Annex A
(informative)

Practical determination of limit of detection for microarray platform (LODP)

A.1 General background

LODP is defined in this document for experimental determination of the limit of a series of representative targets on a given platform. It is quite useful for qualitative analysis with a high number of probes on microarray platform, when more practical determination method of LODP is provided. In this annex, the information of method used for practical determination of LODP[25] is described, considering the principles of widely used definition of LOD including relative LOD.

A.2 Principle

Signal strength of the background signal plus three standard deviations (SD) is widely used for the LOD and criteria for differentiating positive and negative results in qualitative analysis. Whereas the negative signal has been defined differently for each microarray platform, signals obtained from non-probe spots (NPS) installed on the microarrays is defined as the "background" of microarrays in this annex. LODP is determined as the lowest concentration at which the average signal exceeded the average of the NPS + 3 SD and the signal is significantly different from those of the lower and higher adjacent concentration points measured with a diluted series of reference materials. For reliable qualitative analysis, the positive results can be defined as the signals higher than those corresponding to LODP and negative results as lower signals, without determining the LOD for all target probes. The use of LODP also enables comparisons of platform performances without checking sequence dependencies. It assists to select reliable and fitting platforms for experimental purposes.

A.3 Practical method for determining the LODP of a microarray

A.3.1 General

This clause provides the method to determine the LODP of a microarray. The following protocol to determine a LODP of a DNA microarray is shown visually in Figure A.1.

a) Prepare platforms on which spots containing probes of five reference materials and NPS are located.

b) Adjust the reference material concentrations to meet those of the log distribution and conduct measurements on the target platform. Measurement should be conducted on at least three platforms.

c) The lowest concentration that shows average signal values exceeding the threshold placed at average of NPS signals + 3 SD is identified as concentration A.

d) Investigate if a significant difference is observed between the concentration A and the lower adjacent concentration (concentration B).

e) If there is no significant difference between the signal of concentration A and that of B, redefine the concentration A to the adjacent higher concentration, and perform step d) again. Repeat this process to find the pair of concentrations A and B with a significant signal difference. Since the LODP should be determined between A and B, a linearity test should be conducted in range A to B.

f) Adjust the reference materials to meet the concentration of the linear distribution and conduct measurements on the target platform. Measurement should be conducted on at least three platforms.

g) The lowest concentration that shows average signal values exceeding the threshold placed at average + 3 SD of NPS signals is identified as concentration X.

h) Investigate if a significant difference is observed between the signal of concentration X and that of the lower adjacent concentration (concentration Y). In addition, investigate whether a significant difference is observed between the signal of concentration X and that of the higher adjacent concentration (concentration Z).

i) If there are significant differences both between the signal of concentration X and that of Y, and between concentration X and Z, concentration X is determined as the LODP.

j) If there is no significant difference in either concentration pairs, redefine concentration X to the higher adjacent concentration, and perform step h) again.

k) Repeat this process until concentration X is determined as the LODP.

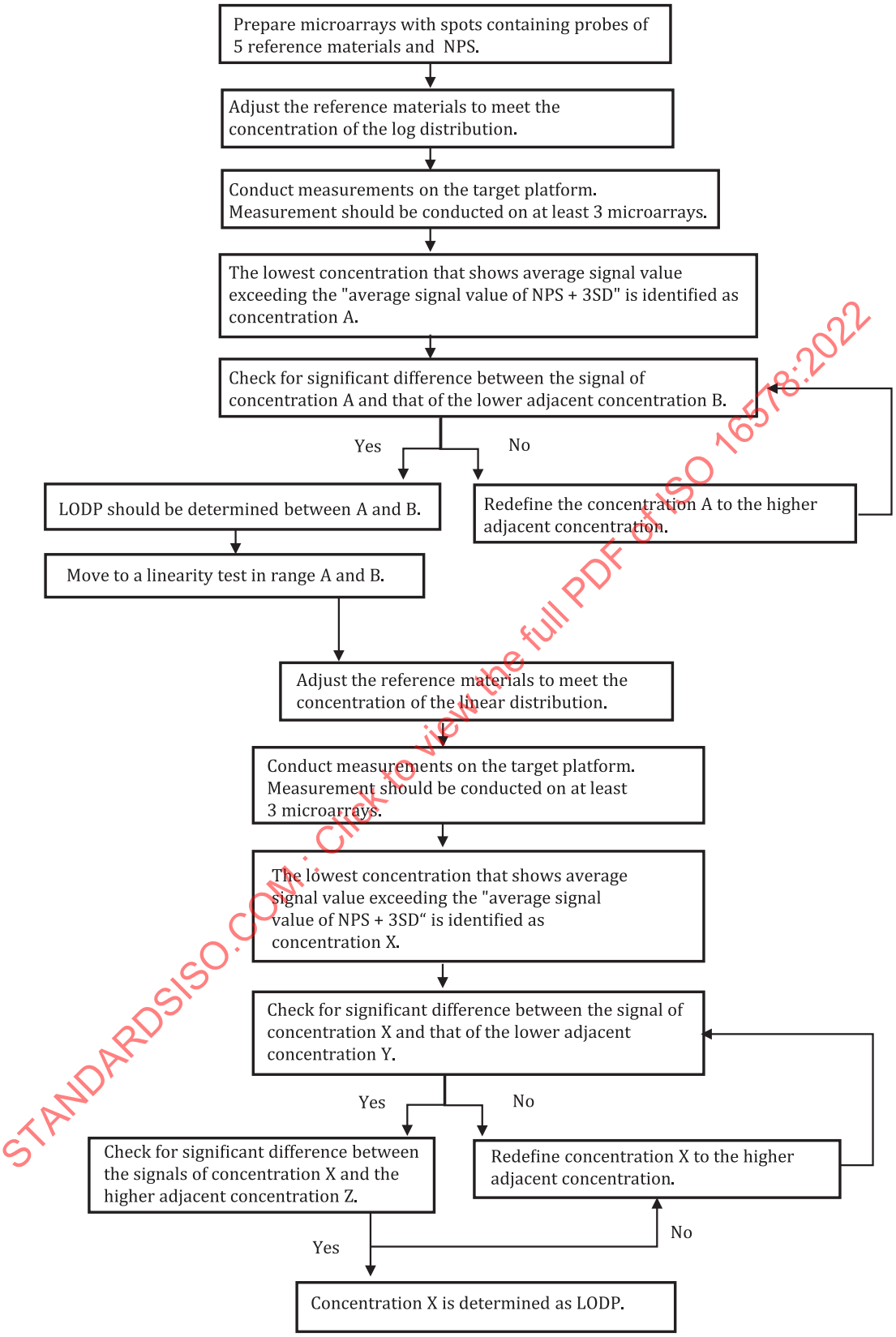

Figure A.1 — Flowchart for LODP determination