This document addresses automated, PCR-based, dideoxy-terminator, and primer extension sequencing done on gel- or capillary-based sequencers. Topics covered include specimen collection and handling; isolation of nucleic acid; amplification and sequencing of nucleic acids; interpretation and reporting of results; and quality control/assessment considerations as appropriate.

A guideline for global application developed through the NCCLS consensus process.
NCCLS...  
Global Consensus Standardization for Health Technologies

NCCLS is an international, interdisciplinary, nonprofit, standards-developing, and educational organization that promotes the development and use of voluntary consensus standards and guidelines within the healthcare community. It is recognized worldwide for the application of its unique consensus process in the development of standards and guidelines for patient testing and related healthcare issues. NCCLS is based on the principle that consensus is an effective and cost-effective way to improve patient testing and healthcare services.

In addition to developing and promoting the use of voluntary consensus standards and guidelines, NCCLS provides an open and unbiased forum to address critical issues affecting the quality of patient testing and health care.

PUBLICATIONS
An NCCLS document is published as a standard, guideline, or committee report.

Standard A document developed through the consensus process that clearly identifies specific, essential requirements for materials, methods, or practices for use in an unmodified form. A standard may, in addition, contain discretionary elements, which are clearly identified.

Guideline A document developed through the consensus process describing criteria for a general operating practice, procedure, or material for voluntary use. A guideline may be used as written or modified by the user to fit specific needs.

Report A document that has not been subjected to consensus review and is released by the Board of Directors.

CONSENSUS PROCESS
The NCCLS voluntary consensus process is a protocol establishing formal criteria for:

- the authorization of a project
- the development and open review of documents
- the revision of documents in response to comments by users
- the acceptance of a document as a consensus standard or guideline.

Most NCCLS documents are subject to two levels of consensus—“proposed” and “approved.” Depending on the need for field evaluation or data collection, documents may also be made available for review at an intermediate consensus level.

Proposed An NCCLS consensus document undergoes the first stage of review by the healthcare community as a proposed standard or guideline. The document should receive a wide and thorough technical review, including an overall review of its scope, approach, and utility, and a line-by-line review of its technical and editorial content.

Approved An approved standard or guideline has achieved consensus within the healthcare community. It should be reviewed to assess the utility of the final document, to ensure attainment of consensus (i.e., that comments on earlier versions have been satisfactorily addressed), and to identify the need for additional consensus documents.

NCCLS standards and guidelines represent a consensus opinion on good practices and reflect the substantial agreement by materially affected, competent, and interested parties obtained by following NCCLS’s established consensus procedures. Provisions in NCCLS standards and guidelines may be more or less stringent than applicable regulations. Consequently, conformance to this voluntary consensus document does not relieve the user of responsibility for compliance with applicable regulations.

COMMENTS
The comments of users are essential to the consensus process. Anyone may submit a comment, and all comments are addressed, according to the consensus process, by the NCCLS committee that wrote the document. All comments, including those that result in a change to the document when published at the next consensus level and those that do not result in a change, are responded to by the committee in an appendix to the document. Readers are strongly encouraged to comment in any form and at any time on any NCCLS document. Address comments to the NCCLS Executive Offices, 940 West Valley Road, Suite 1400, Wayne, PA 19087, USA.

VOLUNTEER PARTICIPATION
Healthcare professionals in all specialties are urged to volunteer for participation in NCCLS projects. Please contact the NCCLS Executive Offices for additional information on committee participation.
Abstract

Sequencing methods for genotyping have moved from the research laboratory into the clinical laboratory. Sequencing is an assay format of choice for very high complexity genotyping, especially when hundreds or thousands of bases of genetic sequence are analyzed. NCCLS document MM9-A—Nucleic Acid Sequencing Methods in Diagnostic Laboratory Medicine; Approved Guideline addresses automated, PCR-based, dideoxy-terminator, and primer extension sequencing done on gel- or capillary-based sequencers. Topics covered include specimen collection and handling, isolation of nucleic acid, amplification and sequencing of nucleic acids, and interpretation and reporting results. Quality control/assessment considerations are addressed for each step of the process as appropriate.

Committee Membership

Area Committee on Molecular Methods

Russel K. Enns, Ph.D.
Chairholder
Cepheid
Sunnyvale, California

Roberta M. Madej, M.S., M.T.
Vice-Chairholder
Roche Molecular Systems, Inc.
Pleasanton, California

Larry E. Bockstahler, Ph.D.
FDA Center for Devices and Radiological Health
Rockville, Maryland

Frederick S. Nolte, Ph.D.
Emory University Hospital
Atlanta, Georgia

Timothy J. O’Leary, M.D., Ph.D.
Armed Forces Institute of Pathology
Silver Spring, Maryland

Carolyn Sue Richards, Ph.D., FACMG
Oregon Health Sciences University
Portland, Oregon

Advisors
Dale H. Altmiller, Ph.D.
O.U. Medical Center
Edmond, Oklahoma

Lee Ann Baxter-Lowe, Ph.D.
University of California,
San Francisco

Mark Evans, Ph.D.
American Medical Association
Chicago, Illinois

Linda Ivor
Gen-Probe Inc.
San Diego, California

Robert B. Jenkins, M.D., Ph.D.
Mayo Clinic
Rochester, Minnesota

Alan L. Landay, Ph.D.
Rush-Presby.-St. Lukes Medical Center
Chicago, Illinois

Richard S. Schifreen, Ph.D., DABCC
Promega Corporation
Madison, Wisconsin

Janet L. Wood, M.T.(ASCP)
BD Diagnostic Systems
Sparks, Maryland

Subcommittee on Nucleic Acid Sequencing

Michael A. Zoccoli, Ph.D.
Chairholder
Celera Diagnostics
Alameda, California

Maria Chan, Ph.D.
FDA Center for Devices and Radiological Health
Rockville, Maryland

James C. Erker
Abbott Laboratories
Abbott Park, Illinois

Andrea Ferreira-Gonzalez, Ph.D.
Medical College of Virginia
Richmond, Virginia

Ira M. Lubin, Ph.D., FACMG
Geneticist
Centers for Disease Control and Prevention
Atlanta, Georgia

Staff
Lois M. Schmidt, D.A.
Staff Liaison
NCCLS
Wayne, Pennsylvania

Donna M. Wilhelm
Editor
NCCLS
Wayne, Pennsylvania

Melissa A. Lewis
Assistant Editor
NCCLS
Wayne, Pennsylvania
Contents

Abstract ............................................................................................................................................... i

Committee Membership ..................................................................................................................... iii

Foreword ............................................................................................................................................. vii

1 Scope .............................................................................................................................................. 1

2 Introduction ...................................................................................................................................... 1

3 Standard Precautions ...................................................................................................................... 1

4 Terminology ..................................................................................................................................... 2

4.1 Definitions .................................................................................................................................. 2

4.2 Abbreviations and Acronyms .................................................................................................... 4

5 Specimen Handling .......................................................................................................................... 5

5.1 Specimen Collection, Transport, and Storage ............................................................................ 5

5.2 Specimen and Nucleic Acid Retention ....................................................................................... 6

5.3 Specimen Quality Assurance .................................................................................................... 6

6 Isolation of Nucleic Acid .................................................................................................................. 6

6.1 Extraction of Nucleic Acid .......................................................................................................... 6

7 Amplification .................................................................................................................................... 8

7.1 Primer Design and Synthesis for PCR and Sequencing .............................................................. 8

7.2 Amplification Parameters ........................................................................................................... 9

7.3 Practices to Aid in Contamination Control ............................................................................... 10

7.4 Quality of Sequencing Template ............................................................................................... 11

7.5 PCR Product Purification ........................................................................................................... 11

7.6 Quantity and Quality of DNA Template .................................................................................... 13

8 Sequencing ...................................................................................................................................... 13

8.1 Sequencing Reactions ................................................................................................................. 13

8.2 Controls and Reference Standards ............................................................................................. 15

8.3 Assay Verification ....................................................................................................................... 15

8.4 Instrument Setup ......................................................................................................................... 16

8.5 Electrophoresis .......................................................................................................................... 17

8.6 Data Examination ....................................................................................................................... 17

9 Data Analysis .................................................................................................................................. 19

9.1 Sequence Review ......................................................................................................................... 19

9.2 Clinical Interpretation ................................................................................................................. 22

9.3 Reports ....................................................................................................................................... 24

References ....................................................................................................................................... 26

Summary of Consensus/Delegate Comments and Committee Responses ........................................ 28

The Quality System Approach .......................................................................................................... 30

Related NCCLS Publications ............................................................................................................ 31
Foreword

Nucleic acid sequencing is evolving rapidly as a diagnostic technique in the clinical laboratory. Applications of this analytic approach are seen in several areas, including cancer testing (genetic testing for cancer predisposition, assessment of gene mutations in cancer), genetics (carrier testing and diagnosis of genetically transmitted diseases), and microbiology (viral genotyping, sequences associated with drug resistance). Sequencing is an assay format of choice for very high complexity genotyping, especially when hundreds or thousands of bases of genetic sequence are analyzed. Recently, sequencing-based tests for HIV genotyping for assessing drug resistance have received FDA clearance for in vitro diagnostics use. Currently, some laboratories utilize manual or automated assays developed in-house (“home-brew”), while others use commercially available reagents and kits. There is a need for standardization of many aspects of the sequencing process. While the chemistry of the reactions may be driven by a particular supplier, sample handling, nucleic acid preparation and assessment, nucleic acid amplification, sequencing and sequence data assessment, and reporting of results are independent of any kit that might be used. This guideline is limited to automated, PCR, dideoxy-terminator, and primer extension sequencing done on gel- or capillary-based sequencers.

Obtaining accurate clinical results from sequencing-based methods requires control of the process from the acquisition of the sample to the interpretation of the sequence obtained. With the increase in the use of laboratory-developed (“home-brew”) sequencing-based genotyping assays and commercially available sequencing-based genotyping kits, a guideline for the development, verification, validation, and implementation of sequencing-based assays is required to guide laboratories and manufacturers. This guideline provides recommendations for all aspects of the sequencing process including specimen collection and handling, isolation of nucleic acid, amplification and sequencing of nucleic acids, and general interpretation and reporting of genotyping results.

It is hoped that this guideline will serve as a roadmap for laboratories to use in guiding themselves in implementing sequencing-based genotypic testing.

Key Words

Capillary electrophoresis, dideoxy-terminators, electrophoresis, gel electrophoresis, nucleic acid, primer, sequencing
Nucleic Acid Sequencing Methods in Diagnostic Laboratory Medicine; Approved Guideline

1 Scope

This guideline specifies recommendations for all aspects of the sequencing process including specimen collection and handling, isolation of nucleic acid, amplification and sequencing of nucleic acids, and general interpretation and reporting of genotyping results. It is the intent of this document to provide instruction for verifying that the sequence obtained is accurate and suitable for subsequent interpretation; to address general interpretation of the sequence; and to provide quality control/assessment considerations for each step of the process as appropriate.

The intended users of this guideline are manufacturers and laboratories involved in the development, verification, validation, and implementation of sequencing-based assays.

This guideline:

- does not address emerging methodologies such as array-based sequencing but is limited to automated, PCR, dideoxy-terminator, and primer extension sequencing done on gel- or capillary-based sequencers; and
- is not intended to provide very specific guidelines for interpreting every possible medical consequence that can be discerned from analyzing individual sequencing results.

2 Introduction

With the sequencing of the human genome and an increasing number of other organisms, sequencing is fast becoming an important tool for genotyping in molecular diagnostics. Sequencing is routinely used in genotyping infectious disease organisms such as HIV and HCV. Additionally, in tissue typing for transplantation, high resolution HLA typing is done by sequencing. Emerging applications for use of sequencing in cancer biology, genetic susceptibility to disease, and understanding of complex genetic diseases are coming into practice. Genotypic testing has already demonstrated clinical utility in patient management and has resulted in genotyping moving rapidly out of the research laboratory and into the clinical laboratory setting. Sequencing is an assay format of choice for very high-complexity genotyping, especially when hundreds or thousands of bases of genetic sequence are analyzed. With the increase in the use of laboratory-developed (“home-brew”) sequencing-based genotyping assays and commercially available sequencing-based genotyping kits, a guideline for the development, verification, validation, and implementation of sequencing-based assays is required to guide laboratories and manufacturers. It is hoped that this guideline will serve as a roadmap for laboratories to use in guiding themselves in implementing sequencing-based genotypic testing.

3 Standard Precautions

Because it is often impossible to know what might be infectious, all patient and laboratory specimens are treated as infectious and handled according to “standard precautions.” Standard precautions are guidelines that combine the major features of “universal precautions and body substance isolation” practices. Standard precautions cover the transmission of all infectious agents and thus are more comprehensive than universal precautions which are intended to apply only to transmission of blood-borne pathogens. Standard and universal precaution guidelines are available from the U.S. Centers for Disease Control and Prevention (Guideline for Isolation Precautions in Hospitals. Infection Control and Hospital Epidemiology. CDC. 1996;17(1):53-80 and MMWR 1988;37:377-388). For specific precautions for
preventing the laboratory transmission of all infectious agents from laboratory instruments and materials and for recommendations for the management of exposure to all infectious disease, refer to the most current edition of NCCLS document M29—Protection of Laboratory Workers from Occupationally Acquired Infections.

4 Terminology

4.1 Definitions

Accuracy (of measurement) – Closeness of the agreement between the result of a measurement and a true value of the measurand (VIM93); See the definition of Measurand, below.

All base accuracy – Calculated by determining the percentage of the bases called that agree with the expected base call in the reference sequence.

Allele – In Genetics, any of several forms of a gene that is responsible for hereditary variation; NOTE: A pseudoallele is an almost-identical sequence to the allele found elsewhere in the genome.

Amplification – The enzymatic replication in vitro of a target nucleic acid; NOTE: The polymerase chain reaction (PCR) is a commonly used method of amplification.

Annealing – The hybridization of two complementary strands of nucleic acid, as in the hybridization of a probe or primer with the target DNA.

Antisense strand – Strand of DNA complementary to the sense strand.

Base call – Ability to distinguish presence of an adenosine (A), a thymidine (T), a cytidine (C), or a guanosine (G) at a given position within a sequence ladder or compilation of overlapping sequence ladders; NOTE: Positions may be ambiguous and represented by an S (G or C), W (A or T), K (G or T), Y (C or T), M (A or C), R (A or G), B (C, G, or T), D (A, G, or T), H (A, C, or T), V (A, C, or G), or N (any base).

Complementary – Describing the property of two strands of nucleic acid which can hybridize by specific-base pairing between the nucleotides.

Compression – Artifacts that are produced during electrophoresis due to the formation of stable secondary structures in the sequencing product DNA fragments; NOTES: a) The folded structures run faster through the matrix than equivalent unfolded DNA fragments resulting in odd spacing, or bands very close together followed by bands spaced further apart than normal; b) In some cases, bases can be lost entirely.

Consensus sequence – The final sequence generated from a compilation of overlapping sequence ladders following the completion of base calling at all positions; NOTE: The consensus sequence is believed to be representative of the source nucleic acid.

Denaturation – The conversion of double-stranded DNA or RNA to a single-stranded state with minimal secondary structure; NOTES: a) This is done by heating, increasing pH, or adding agents such as formamide or urea; b) Once denatured, nucleic acid molecules are available for hybridization with a primer or probe.

Diagnostic sensitivity – The proportion of patients with a well-defined clinical disorder whose test values are positive or exceed a defined decision limit (i.e., a positive result and identification of the patients who have a disease); NOTES: a) The clinical disorder must be defined by criteria independent of the test
under consideration; b) The term **Diagnostic sensitivity** (Europe) is equivalent to **Clinical sensitivity** (U.S.).

**Diagnostic/confirmatory testing** – 1) Testing generally performed to evaluate the genetic status of individuals at increased risk for a particular disorder due to a positive family history or symptoms; 2) A test that confirms the presence or absence of a substance by another methodology or procedure that is either more sensitive, more specific, or both; 3) A clinical condition as a follow up to testing previously performed that indicated the patient being at higher risk for having a clinical condition.

**Electropherogram** – The visualization of a sequence ladder characterizing the signal strength, noise, base spacing, and base calls as seen by the sequencing software; **NOTE:** This chart can be used as a tool to evaluate the quality and accuracy of the sequence.

**Electrophoresis** – see **Gel electrophoresis**.

**Extension/Elongation** – The 5′ to 3′ synthesis of DNA starting from an annealed primer to generate a complementary strand of DNA.

**Gel electrophoresis** – Separation of molecules in an electric field within a matrix of agarose or polyacrylamide according to size and charge.

**Genotype** – 1) The genetic makeup of an organism, or group of organisms, with reference to a single trait, set of traits, or an entire complex of traits; 2) The specific allelic composition of a gene, or set of genes, established at the DNA level.

**Hybridization** – Base pairing of complementary strands of nucleic acid by hydrogen bond formation, the binding of probe to specific nucleic acid sequences or amplification products; **NOTE:** Hybridization can be performed with both nucleic acid target and probe in solution, or with either one bound to a solid support such as a microtiter plate or membrane.

**Measurand** – Particular quantity subject to measurement (VIM93)\(^1\); **NOTE:** This term and its definition encompass all quantities, while the commonly used term “analyte” refers to a tangible entity subject to measurement. For example, “substance” concentration is a quantity that may be related to a particular analyte.

**Mismatch** – Hybridization of two DNA or RNA strands that are less than 100% complementary.

**Mutation** – An abnormal variation in DNA sequence which is found in association with, or which reflects a predisposition to disease; **NOTE:** Mutations must be distinguished from polymorphisms—DNA variants found in a population that neither harbors, nor is predisposed to disease.

**Noise** – Unwanted background signal which may adversely affect the ability to distinguish a positive signal.

**Oligonucleotide** – A small polynucleotide generally synthesized on a machine; **NOTE:** Oligonucleotides are used as amplification primers or as probes.

**Polymerase chain reaction (PCR)** – A common method of DNA amplification, utilizing pairs of oligonucleotide primers as start sites for repetitive rounds of DNA polymerase-catalyzed replication and alternating with denaturation in successive heating-cooling cycles.

**Polymorphism** – The occurrence together in a population of two or more alternative genotypes, each at a frequency greater than that which could be maintained by recurrent mutation alone; **NOTE:** A locus is
arbitrarily considered to be polymorphic if the rarer allele has a frequency of 0.01, so that the heterozygote frequency is at least 0.02.

**Positive predictive value** – The likelihood that an individual with a positive test result has a particular disease, or characteristic, that the test is designed to detect.

**Primer** – An oligonucleotide that hybridizes with one strand of a DNA or RNA template, providing a free 3′–OH end at which a polymerase starts the synthesis of a new, complementary strand.

**Resolution** – In DNA sequencing, the ability to clearly distinguish between adjacent positions; **NOTE:** Resolution is influenced by the parameters of electrophoresis, sequence ladder peak height (signal strength), and peak width.

**Ribonuclease** – An enzyme that catalyzes the breakdown of RNA molecules into smaller components.

**Sense strand** – The strand of a duplex DNA that has the same nucleotide sequence as mRNA except that T substitutes in DNA for U in RNA; **NOTE:** The sense strand is also called the coding strand. The other strand, which is the actual template for mRNA synthesis, is the anticoding or antisense strand.

**Sensitivity** – The change in response of a measuring system or instrument divided by the corresponding change in the stimulus (modified from VIM93)¹; **NOTES:** a) The sensitivity may depend on the value of the stimulus (VIM93)¹; b) The sensitivity depends on the imprecision of the measurements of the sample.

**Sequence** – The order of nucleotides (A, C, G, T, or U) in a strand of DNA or RNA.

**Signal strength** – In sequencing, the peak height associated with the termination of a sequence product leading to the generation of a sequence ladder; **NOTE:** Signal strength allows for the identification of a base above the noise.

**Specificity** – The ability of a measurement procedure to measure solely the measurand; **NOTE:** Specificity has no numerical value in this context; See **Measurand** above.

**Target sequence** – Area of nucleic acid to be detected and sequenced.

**Validation** – Confirmation through the provision of objective evidence, that requirements for a specific intended use or application have been fulfilled (ISO 9000)²; **NOTES:** a) WHO defines validation as the action {or process} of proving that a procedure, process, system, equipment, or method used works as expected and achieves the intended result (WHO-BS/95.1793)³; b) The components of validation are quality control, proficiency testing, validation of employee competency, instrument calibration, and correlation with clinical findings.

**Verification** – Confirmation through the provision of objective evidence that specified requirements have been fulfilled (ISO 9000)²; **NOTE:** A one-time process completed to determine or confirm test performance characteristics before the test system is used for patient testing.

### 4.2 Abbreviations and Acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>adenine</td>
</tr>
<tr>
<td>C</td>
<td>cytosine</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>dATP</td>
<td>deoxyadenosine 5′ triphosphate</td>
</tr>
<tr>
<td>dCTP</td>
<td>deoxycytidine 5′ triphosphate</td>
</tr>
<tr>
<td>ddATP</td>
<td>dideoxyadenosine 5′ triphosphate</td>
</tr>
</tbody>
</table>
Specimen Handling

The specimen container should be clearly labeled with a unique patient identifier. A requisition form and clinical and family history when applicable, should accompany every sample. It is recommended that each laboratory have written criteria for acceptance and rejection of specimens. For inherited disease testing, issues with regard to the need for informed consent are often raised. The decision as to whether informed consent is required before testing is carried out should take into account any applicable federal, state, or local requirements. NCCLS documents MM1—Molecular Diagnostic Methods for Genetic Diseases and MM5—Nucleic Acid Amplification Assays for Molecular Hematopathology provide detailed considerations for the selection of relevant criteria for the above items.

5.1 Specimen Collection, Transport, and Storage

From the safety standpoint, the collection and handling of specimens for transport and storage should proceed with the same precautions taken for any potential infectious materials, following current blood-borne pathogen, infectious disease, and appropriate regulatory guidelines (see Section 3).

Appropriate specimen collection, handling, and transport conditions are critical to ensure specimen integrity. Each laboratory should have written criteria for specimen collection, transport, and storage.

Specimen transport and storage conditions will depend upon the nucleic acid of interest, either DNA or RNA. NCCLS documents MM1—Molecular Diagnostic Methods for Genetic Diseases and MM5—Nucleic Acid Amplification Assays for Molecular Hematopathology provide detailed information regarding appropriate specimen transport and storage conditions for a variety of specimens that could be used for sequencing purposes.
5.2 Specimen and Nucleic Acid Retention

It is recommended that each laboratory have written criteria for specimen and nucleic acid retention. A primary concern with specimen and nucleic acid retention involves ethical and legal considerations such as informed consent and specimen reuse for the same and/or other testing. Some of these issues could be raised and documented during the informed consent process. These issues are currently being evaluated by different government agencies and professional societies, and recommendations can be expected. Until these recommendations and/or regulations are established, each laboratory should determine its own specimen and nucleic acid retention policy.

5.3 Specimen Quality Assurance

Samples should be maintained for an appropriate period of time to confirm sequencing results as required by regulation or recommended by laboratory practice standards. Patient samples or nucleic acids isolated can be used for quality control purposes once samples are anonymized, unless the person from which the sample came requests that the specimen not be used for this purpose. Regulations may require further measures such as informed consent from the patient prior to sample use for quality control purposes.

6 Isolation of Nucleic Acid

Numerous manual and several automated procedures exist for extraction of nucleic acids from clinical specimens to be used for sequencing. Several commercial nucleic acid isolation kits are available that may simplify and reduce the time of this procedure. Commercial sequencing kits may or may not include an extraction procedure as part of the kit. A key factor for the extraction of nucleic acids that are to be sequenced is that they should be as pure as possible so as to remove all inhibitors, competing nucleic acids, and other substances that may interfere with the amplification and cycle-sequencing reactions. Each laboratory should evaluate various DNA and/or RNA purification systems for yield efficiency, consistency, and nucleic acid purity before incorporation as a standard procedure.

6.1 Extraction of Nucleic Acid

The quality of sequencing results is directly related to the quality of the nucleic acid template. The template for sequencing should be free from protein, RNA, ethanol, salts, and any organic contaminants. The methods that are commonly used for extraction of nucleic acids that are to be sequenced are the same as those for amplification. The main categories include organic extraction and solid phase extraction methods.

6.1.1 Organic DNA Extraction Method

Organic extraction is a classical technique that uses organic solvents to separate the DNA from protein and other contaminants. Cells are lysed using a detergent, and then mixed with phenol, chloroform, and isoamyl alcohol. The correct salt concentration and pH are critical to ensure that contaminants are separated into the organic phase and that DNA remains in the aqueous phase. DNA is usually recovered from the aqueous phase by alcohol precipitation.

6.1.2 High Salt Method

A modification of the above method involves the use of high salt, detergent, and isopropanol to isolate and precipitate DNA from the solution. This method uses a detergent to release DNA and denature proteins from cells. RNA can be removed if necessary using RNase A. Proteins are precipitated by using a high salt solution that contains ammonium acetate, followed by isopropanol precipitation of DNA. The DNA precipitates out of solution in isopropanol because some of the salt from the protein precipitation solution will carry over into the alcohol and the salt and alcohol together will dehydrate the DNA,
bringing it out of solution. This system provides high purity DNA preparations without the need for organic solvents and can accommodate significant volume of starting specimen.

6.1.3 Solid Phase Extraction Methods

Solid phase extraction methods have almost replaced the organic extraction methods. Compared to the traditional liquid-liquid extraction, the major advantages are avoidance of dangerous organic solvents like chloroform and phenol, significant time savings, and the ultrahigh purity of the nucleic acid product.

6.1.3.1 Silica Technology

With the silica-based membrane method, nucleic acids are selectively adsorbed to the silica at high concentrations of chaotropic salt such as guanidine hydrochloride, guanidine isothiocyanate, sodium iodide, and sodium perchlorate. The bound DNA is washed, and then eluted with a low-salt buffer or with water. This system provides high purity DNA preparations with a choice of purification formats such as spin columns, strips, or 96-well plates.

6.1.3.2 Glass Fiber Technology

Cells are lysed during a short incubation with proteinase K in the presence of a chaotropic salt (guanidine HCl), which immediately inactivates all nucleases. Cellular nucleic acids bind selectively to glass fiber fleece in a special centrifuge tube. The nucleic acids remain bound while a series of rapid "wash-and-spin" steps remove contaminating small molecules. Finally, low-salt elution removes the nucleic acids from the glass fiber fleece. The process does not require precipitation, organic solvent extractions, or extensive handling of the nucleic acids.

6.1.3.3 Anion-Exchange Methods

Solid-phase anion-exchange chromatography is based on the interaction between the negatively charged phosphates of the nucleic acid and positively charged surface molecules on the substrate. DNA binds to the substrate under low-salt conditions; impurities such as RNA, cellular proteins, and metabolites are washed away using medium-salt buffers; and then ultrapure DNA is eluted using a high-salt buffer. The eluted DNA is recovered by alcohol precipitation. The matrix used is a special anion exchanger, packed into polypropylene columns. The sample binding takes place under low-salt concentration conditions. During the washing and elution steps, the salt concentration is increased stepwise. This method is suited for the rapid and efficient purification of DNA and RNA as well. Impurities are removed quantitatively because of their different binding behavior.

On occasion, it may be necessary to concentrate or enrich the specimen before it can be amplified for sequencing. Concentration of starting material is achieved by ways such as filtration or centrifugation. Enrichment can be achieved by binding to probes and/or antibodies.

6.1.4 RNA Extraction Procedures

Isolation of RNA is more challenging than isolation of DNA. A significant problem encountered during RNA extraction is the degradation of RNA by RNAses. RNAses are very difficult to inactivate and even resist boiling. In addition to their presence within the cells of the tissue being extracted, these enzymes can be easily introduced during laboratory manipulation from a variety of sources including sneezing, hair, and fingerprints. Extra care is required to ensure that all reagents, glassware, and plasticware are RNAses free. A number of commercial preparations are available that either inhibit or destroy this enzyme. The most commonly used procedure is treatment of water, nontemperature-sensitive reagents, nonvolatile reagents, glassware, and plasticware with diethylpyrocarbonate (DEPC) followed by autoclaving. Chaotropic agents such as guanidinium isothiocyanate are generally used in the purification
of RNA. These agents are capable of dissolving cellular membranes, causing nucleoproteins to dissociate from nucleic acid and inactivating ribonucleases. Purification of RNA can be achieved by organic or nonorganic procedures. The organic and nonorganic procedures for isolating RNA are similar to the ones already described for DNA with slight modifications. As a final step in purification, RNA should be stored in a solution containing an appropriate amount of RNase inhibitors.

6.1.5 Automated Nucleic Acid Extraction

Different manufacturers have introduced several stand-alone robotic automatic nucleic acid purification instruments with different success. A number of these automatic extractors use a modification of a manual procedure such as high-salt, glass fiber, and silica-based methods. Manufacturers of automatic extractors usually provide kits for DNA or RNA extraction from a variety of specimen sources. It is anticipated that fully automated instruments will be developed and implemented in the near future.

7 Amplification

Sequencing of a particular nucleic acid target requires, first of all, adequate quantity of the target sequence. In the majority of the cases, it is not possible to direct the sequence straight from the patient specimens. PCR has proven invaluable because it allows the amplification of a specific nucleic acid sequence without having to first clone in a microbial host.

There are two main elements that are crucial for successful sequencing of PCR products. Primer selection and amplification conditions should be optimized to yield sufficient quantities of the specific target without producing significant background. For additional information regarding amplification, refer to NCCLS documents MM3—Molecular Diagnostic Methods for Infectious Diseases and MM5—Nucleic Acid Amplification Assays for Molecular Hematopathology.

7.1 Primer Design and Synthesis for PCR and Sequencing

PCR uses repeated cycles of template denaturation, primer annealing, and polymerase extension to amplify a specific target sequence. The oligonucleotide primers determine the specific target sequence that is amplified.

Primers used for PCR must be designed using a known sequence for the target of interest. Sequencing primers can be designed as well using a known sequence from the target, or it is possible to add several noncomplimentary bases at the 5’-end of the PCR primer to generate a unique nontarget sequence for the sequencing primer. Sequencing can also be done using the PCR primers but this is not advisable. Several factors need to be taken into consideration when designing primers whether for PCR or sequencing. They are as follows:

- Primers should be at least 18 to 28 nucleotides in length to ensure good hybridization.
- Match Tm/Td of the primers to each other.
- The Tm selected should be optimized for the maximum sensitivity and specificity of the PCR and sequencing reactions.
- Avoid runs of an identical nucleotide, especially four or more runs of Gs.
- Keep C-G content between 40 to 70%, preferably 50 to 55%.
- Design primers that are not complementary to each other.
• Avoid complementarity at the 3’ ends.

• Avoid complementarity at the 5’ ends.

• Avoid palindromic sequences, particularly at the 3’ end, as they can form secondary structures.

• Avoid all known SNPs by checking primers against SNPs databases.

• Sequence conservation should be taken into consideration during primer design. If possible, primers should be designed in the most conserved region of the sequence of interest.

• Whenever possible and/or appropriate, primers should be designed at least 50 bases in the intron sequences in order to be able to interrogate intron/exon sequences with quality sequence data.

• Ambiguous positions are allowed, but they should be avoided in the 3’-most positions.

• For sequencing, the primers should not be closer than 40 nucleotides to the sequence of interest due to the difficulty of reading sequences immediately after the primer.

There are a number of different computer software programs available that can greatly assist with PCR and sequencing primer selection and design. Some of these programs are freely available on the Internet at: genome-www2.stanford.edu/cgi-bin/SGD/web-primer and frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi. There are also commercial sites on the Internet that provide access to computer software programs. It is advisable to perform homology checks against known sequences to identify possible homology of the primers with other known sequences that might impact primer specificity. Pseudoalleles or almost-identical sequences found elsewhere in the genome could sometimes confound sequencing reactions and interpretations. Sequence databases may reveal the presence of these, and efforts should be made for selective amplification of the desired sequence. While these rules and software can aid in primer design, practical optimization of amplification conditions to ensure sensitivity and specificity must be performed.

Sequence primer spacing is partially dependent upon the sequencing technology used. The optimal spacing between the primers is defined by the sequencing instrument and technology employed. Newer sequencing chemistries and methodologies will further increase the sequence length attainable. For many formats in use today, primers should be placed every 200 to 600 bases on both the sense and antisense strands for accurate base calls to be made.

The dyes for dye-labeled primers should be chosen to be compatible with the detection system of the sequencing platform used. The dyes and linkages should be stable to the temperature cycling between 40 and 92 °C, which occurs in cycle sequencing. The fluorescent dyes used for sequencing are all photolabile to some extent and should be protected from light during storage and use. Purified dye-labeled primers are readily available from commercial sources and should be qualified by sequencing a control or reference standard prior to use with clinical samples.

7.2 Amplification Parameters

While there are standard conditions that will amplify most target sequences, each new PCR application is likely to require optimization. Some problems that are often encountered are: lack of PCR product, low yield of the desired product, the presence of nonspecific background bands due to mispriming or misextension of the primers, and formation of primer-dimer that competes for amplification with the desired product.
The following PCR parameters need to be taken into consideration and, if necessary, optimized:

- amount of starting template;
- primer concentration;
- type of enzyme and enzyme concentration;
- use of chemically modified thermostable DNA polymerases that provide automatic hot start, allowing template DNA and primers to be mixed together and held at a temperature above the threshold of nonspecific binding of primer to template before nucleic acid amplification occurs;
- magnesium ion concentration;
- buffer composition and pH;
- annealing temperature; and
- number of cycles.

Various additions or cosolvents such as dimethyl sulfoxide (DMSO), glycerol, and bovine serum albumin have been shown to improve amplification in many applications. Careful evaluation of these parameters will greatly increase the specificity of the amplification and decrease the amount of nonspecific products that can interfere with the sequencing reaction.4-17

7.3 Practices to Aid in Contamination Control

Due to the ability of PCR to amplify from minute amounts of target template or already amplified product (amplicon), laboratories are strongly encouraged to implement measures to avoid template and/or amplicon contamination. This can best be accomplished by developing a workflow process to minimize the potential for amplicon and target contamination.18-22 Physical separation of reagent preparation, specimen processing and nucleic acid extraction, and processing of the amplified product should be carefully designed. Unidirectional workflow must be strictly followed whereby personnel perform tasks in the clean areas first, and only after completing their tasks there, move into the postamplification area. In addition, supplies and instruments should be dedicated to each area. Protection of countertops with absorbent plastic-backed paper must be changed after each use as paper can collect dust and with it, amplified material. The use of bleach appears to be the most generally accepted method for decontaminating surfaces. Decontamination of countertops, laboratory equipment, and furniture should be performed by wiping with a freshly made 5% to 10% solution of bleach and then with a 75% ethanol solution or even water to dilute the bleach that could corrode the surfaces. All pipette tips should contain a hydrophobic filter in order to avoid any contamination of the tip of the pipettor with template or amplified material. Moreover, these tips should be discarded into a container which can be completely sealed before discarding them. Gloves should be changed periodically or as soon as any contamination is suspected.

A chemical means of carry-over prevention is through the incorporation of deoxyuridine triphosphate instead of deoxythymidine triphosphate into the amplicon during the amplification procedure. In this procedure, deoxyuridine triphosphate is incorporated into every amplicon, which makes them susceptible to destruction by uracil N-glycosylase (UNG) and heat. If uracil-containing amplicon is incorporated by accident into another sample, one can eliminate this carry-over contamination by treating the sample prior to amplification with UNG. Treatment with UNG will result in destruction of the uracil-containing amplicon. The uracil-containing amplicon is the substrate for UNG, while the template is not. The mechanism behind this chemical reaction is that UNG removes uracil residues from the amplicon and
leaves an intact phosphodiester backbone in the amplicon. During the first denaturation step of the amplification procedure, the phosphodiester bond breaks at the sites where the uracil residues were located. The fragment amplicon is no longer able to act as a template. Use of dUTP and UNG has been shown to be effective to control contamination if the uracil-containing amplicon does not exceed the amount between 10^6 and 10^7 copies per reaction. Uracil-containing amplicons can be sequenced with good success under the same conditions used for the usual thymidine-containing amplicons.

One of the shortcomings of this procedure is the need to optimize the amount of dTTP and dUTP. The dTTP and dUTP ratio is optimized to get a good yield of PCR product from the target. But the dUTP level needs to be high enough to ensure that UNG treatment of the amplicon will eliminate its ability to be amplified. Thus, it should be possible to incorporate this contamination control technology into almost all sequencing assays.

### 7.4 Quality of Sequencing Template

Evaluation of substances that could interfere with the ability of amplification and sequencing must be performed. Interfering substances are components present in the patient specimen that interfere with the amplification of the specific target. There are mainly two sources of interfering substances. Exogenous interfering substances are substances that are introduced to the specimen such as anticoagulant or substances that are given to the patient that may affect the test ability to detect or measure the target of interest. These could be drugs, parenteral nutrition, etc. On the other hand, endogenous interfering substances (e.g., lipids, bilirubin) could be a result of pathological conditions.

DNA template quality and quantity is critical for good quality sequencing results. Poor template quality is the most common cause of sequencing problems. Some of the problems derived from poor quality template are noisy data, peaks under peaks, weak signal, and even no usable sequence data. A common problem affecting template quality is the presence of contaminants such as proteins, chromosomal DNA, excess primers, dNTPs, enzymes and buffer components/salts from the PCR reaction, residual organic chemicals, and detergents. The quality of the DNA template for sequencing can be assessed by agarose gel electrophoresis where purified DNA should run as a single band on the gel.

### 7.5 PCR Product Purification

Purification of PCR product is essential for optimal sequencing results. Unused primers and nucleotides as well as salts should be eliminated prior to sequencing for optimal results. Dye-terminator sequencing is particularly sensitive to the presence of excess amplification primers. Both forward and reverse primers can act as sequencing primers, resulting in the generation of multiple labeled sequence ladders within a single reaction. These additional fragments cannot be distinguished from the intended target, making it impossible to interpret the sequence data. In addition, the excess primers and nucleotides carried over from the PCR reaction can alter the ratios of reagents in the sequencing reaction sufficiently to compromise both the fluorescent signal and the reaction fidelity.

A variety of approaches have been developed for the purification and sequencing of PCR products, including dilution for direct sequencing, organic precipitation, column purification, gel purification, and exonuclease I/shrimp alkaline phosphatase (ExoI/SAP) treatment. Each approach should be evaluated based on the requirements of the specific application.

#### 7.5.1 Dilution of PCR Product

This approach is useful in those cases where the PCR reaction has been rigorously optimized to use minimal concentrations of primers and nucleotides without interfering with the efficiency and fidelity of amplification. The PCR product can be diluted 1:5 or 1:10 in highly purified water (i.e., usually by a combination of techniques that can include carbon adsorption, distillation, deionization, reverse osmosis,
ultraviolet radiation, and other methods) and used directly as sequencing templates. Even though this approach will require no manipulation of the PCR product before sequencing, it will require extensive optimization to ensure that the sequencing reaction delivers sufficient quality sequence data. This requires that the PCR reaction is optimized regarding primer and nucleotide concentration, but also to avoid generation of nonspecific amplification products that can interfere with the quality of the sequencing data.

7.5.2 Gel Purification

Agarose gel electrophoresis separates PCR reaction products and byproducts by size. After gel electrophoresis, the desired PCR product can be excised from the gel and extracted by a variety of different methods, including commercial products such as columns.23,24 The primary advantage of gel purification is that the specific PCR product can be separated from the byproducts that are generated during the amplification process such as nucleotides, primers, and any other nonspecific amplified product. The major disadvantage of this process is that ethidium bromide staining and UV illumination are used to identify the desired reaction product on the gel and this process can nick the PCR product that will subsequently be used as sequencing template. Furthermore, this process is time consuming, produces low yield, is not amenable to automation, and most importantly, is not efficient in separating DNA fragments based on size, so some residual byproduct may remain. In addition, there are some agarose products that contain small molecules that can interfere with the sequencing enzymatic reactions.

7.5.3 Column Purification

There are a number of commercially available products for PCR product purification. The most widely used product uses a silica gel membrane that binds DNA preferentially and allows elimination of unused primers and salt-eluting DNA into very small volumes such as 50 µl.25,26 Most of these columns have different cutoff values based on the size of the DNA molecules. For example, some columns will separate oligonucleotides 40 bases or shorter from the PCR product, allowing nucleotides and primers to be eliminated. Ultrafiltration columns are other commercially available options. Ultracentrifugation columns allow separation based on molecular weight, for example, there are columns that have a 100-kDA cutoff and consequently can be used to separate the desired PCR products from single-stranded oligonucleotides of less than approximately 300 bases in length and double-stranded DNA less than 125 base pairs in length. Other commercially available columns use gel filtration chromatography to preferentially bind small DNA segments, allowing elution of larger product for sequencing. One of the major advantages of column purification is that the procedure is extremely rapid, simple, and amenable to automation. In addition, it has been shown that these products have a high degree of reproducibility. It is important to point out that byproducts generated during amplification with similar size to the intended amplicon will not be effectively removed using these procedures and these may interfere with subsequent sequencing. Careful optimization of the PCR reaction to avoid these byproducts should alleviate this problem.

7.5.4 ExoI/SAP

Another method for PCR product purification for sequencing is using an enzymatic reaction. Following PCR, treatment of PCR product with an ExoI degrades single-stranded oligonucleotide DNA such as unincorporated primers.27,28 Furthermore, treatment with shrimp alkaline phosphatase (SAP) dephosphorylates residual oligonucleotide primers and as a result, they become inactivated. Following treatment, the residual enzymes can be heat inactivated and the enzymatic products can be directly used for the sequencing reaction. The major advantage of this procedure is that it is simple, cost effective, and can be easily automated for high throughput applications. Furthermore, this method is highly reliable. The major disadvantages of this procedure are that this reaction also needs to be optimized for the best possible performance and that secondary byproducts of PCR reaction will not be removed. Consequently, the PCR must be optimized to prevent formation of byproducts.
7.6 Quantity and Quality of DNA Template

Another factor that has a major impact on the quality of the sequence data is the quantity and quality of DNA target used in the sequencing reaction. The quantity of DNA is typically determined by spectrophotometry at 260 nm, fluorometry, or electrophoresis.

The quantity of DNA used for the sequencing reaction will depend upon the sequencing chemistry used and the manufacturer of the reagents. It is advisable to review the manufacturer’s recommendations as to the amount of DNA required to obtain good sequencing data. In general, the amount of template will depend upon the PCR product size, double-stranded versus single-stranded template, etc. As a point of reference, PCR products between 100 to 200 bp will require 1 to 3 ng of template, and products between 1000 to 2000 bp will require 10 to 40 ng. If too much template is used, one can obtain very strong initial peak heights that fade rapidly as the sequencing reaction progresses. This can occur due to depletion of the dye label early in the reaction. On the other hand, if too little template is used, peak height and strength will be reduced, making data analysis difficult.

8 Sequencing

8.1 Sequencing Reactions

The technology employed in most sequencing is a combination of enzymatic chain termination with dideoxynucleotides and cycle sequencing with thermostable DNA polymerases. In general, sequencing reactions require the following components:

- sequencing primers;
- sequencing mix containing mixtures of deoxynucleotides and dideoxynucleotides and buffer;
- thermostable DNA polymerase;
- dye label (either the sequencing primers or the dideoxynucleotides can carry the label); and
- DNA template (see Sections 5 and 6 of this document).

8.1.1 Sequencing Primers

Refer to Section 7.1 for information on designing sequencing primers.

8.1.2 Sequencing Mixes

Sequencing mixes vary depending on whether dye-labeled primer or dye-labeled dideoxy-terminator chemistry is used.

When using dye-labeled primers, it is necessary to have four separate sequencing mixes for A, C, T, and G. There are multiple approaches to using dye-labeled primers. The simplest uses a single primer labeled with one dye. Each mix would contain dATP, dGTP, dCTP, and dTTP and one of the dideoxynucleotides ddATP, ddGTP, ddCTP, or ddTTP, formulated in ratios optimized to produce maximum read length and accuracy. This is because all the fragments will have the same dye label and it would be impossible to differentiate the sequencing DNA fragments terminated with each dideoxynucleotide if they were all included. Each sequencing mix should be analyzed separately on a lane in a gel or capillary, and the data for all four sequencing mixes should be combined by the sequencer software to give the complete sequence.
Another approach is labeling the same primer separately with four different fluorescent dyes. This results in four dye-labeled primers, sharing the same DNA sequence but having different dye labels. Sequencing ladders are generated in four separate base-specific reactions for A, C, T, or G. The products from these reactions can be combined and prepared for electrophoresis.

The forward and reverse sequencing primers can be labeled with different dye labels therefore allowing forward and reverse sequencing to be performed in the same reaction tubes. In this approach, a separate reaction is required for A, C, T, or G. Each reaction is analyzed in a separate gel lane or capillary.

In dye-labeled dideoxy-terminator chemistry, ddATP, ddGTP, ddCTP, and ddTTP are all labeled with fluorescent dyes. The sequencing mix would contain dATP, dGTP, dCTP, and dTTP and limiting quantities of all the dye-labeled dideoxynucleotides ddATP, ddGTP, ddCTP, and ddTTP. The sequencing primers should be unlabeled and the sequencing DNA fragment should be labeled when a dye-labeled ddNTP is incorporated. This allows the sequencing reactions for A, G, C, and T to be performed in one reaction tube and analyzed in one lane on a gel or a single capillary. The sequencer software analyzes the data and generates a complete sequence from one lane. It is not possible to use multiple primers in a sequencing reaction, so forward (sense) and reverse (antisense) sequencing must be done in separate reaction tubes.

The advantages and disadvantages of dye primer and dye-labeled dideoxy-terminator chemistries for sequencing are summarized in Table 1.

<table>
<thead>
<tr>
<th>Chemistry</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dye Primer</td>
<td>• More even signal intensities</td>
<td>• Four reactions needed/less efficient</td>
</tr>
<tr>
<td></td>
<td>• Less or no clean-up needed before electrophoresis</td>
<td>• Compressions are possible</td>
</tr>
<tr>
<td></td>
<td>• Longer read lengths</td>
<td>• More pipetting steps to set up reactions</td>
</tr>
<tr>
<td></td>
<td>• Fewer structure-related stops</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Labeled primers are available for common vector binding sites</td>
<td></td>
</tr>
<tr>
<td>Dye-labeled dideoxy-terminator</td>
<td>• One reaction needed/efficient</td>
<td>• More structure-related stops</td>
</tr>
<tr>
<td></td>
<td>• Compressions are resolved</td>
<td>• Shorter read lengths</td>
</tr>
<tr>
<td></td>
<td>• Clean-up required before electrophoresis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Fewer pipetting steps to set up reactions</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Unlabeled primers used</td>
<td></td>
</tr>
</tbody>
</table>

8.1.3 Thermostable DNA Polymerase

There are many possible thermostable DNA polymerases to use in cycle sequencing. Special consideration is needed for selecting the enzyme for fluorescent dye-labeled dideoxy-terminator chemistry. Enzymes with mutations that decrease the discrimination between dNTPs and fluorescently labeled ddNTPs are preferred. Improved performance in dye-labeled primer sequencing can be obtained by including a thermal stable pyrophosphatase to eliminate problems with pyrophosphorolysis.
The choice of enzyme will determine the appropriate ratio of dNTPs to ddNTPs that would be needed in the sequencing mixes. Commercial mixes are available that are optimized for some of the DNA polymerases.

The key performance selection criteria for an enzyme should be the combination of equivalent efficiency of incorporation of the A, G, C, and T dideoxynucleotides leading to uniformity of band intensity for sequencing DNA fragments, the number of bases sequenced, and sequencing accuracy.

8.1.4 Dye Label

The key selection criteria for a dye label are that it is compatible with the sequencing chemistry and supported by the sequencing platform used. Commercial sequencing platforms have fluorescence detection systems that are optimized for certain excitation and emission wavelengths. These specifications will indicate which dyes can be detected and analyzed. Standards containing the dyes should be used to calibrate the sequencing platform. This is especially critical when using a fluorescent, labeled dye terminator chemistry, when four dyes are present in each lane or capillary.

8.2 Controls and Reference Standards

There are many steps between isolation of DNA or RNA and the final sequencing results. It is ideal to have both a control that is taken through the entire process from DNA isolation, through PCR and sequencing with a known sequence, and a reference standard that can be taken directly through the sequencing reaction.

The results for the control and the reference standard can be compared to a set of specifications derived during the verification of the sequencing assay. If the results of the control fail to meet specifications and the results of the directly sequenced reference sample pass specifications, then the troubleshooting can be directed to the steps prior to sequencing. If both fail specifications, initial troubleshooting should start at the sequencing step.

Controls should resemble as much as possible the sample to be sequenced. It is important that the control be stable and consistently prepared so that the results can be compared to the expected outcome. If the assay is designed to test for polymorphisms or mixtures, then the control should have a polymorphism or mixture as well.

Laboratories can obtain control materials from a variety of sources. Cell lines, clones, or reference samples are available from commercial suppliers such as Coriell Cell Repositories (http://locus.umdnj.edu/ccr/) or ATCC® (www.atcc.org). Controls can also be well characterized clinical specimens if properly consented from one’s own or from another laboratory. A sequencing standard is usually available from the sequencing system manufacturer, since this is needed for the installation and qualification of the system by the service technician. (Further information on how to use controls and sequencing standards for quality assurance of sequencing assays can be found in Sections 8.6 and 9.1.)

8.3 Assay Verification

Each new sequencing-based assay must be verified prior to using the assay to generate reportable clinical results. The assay should be subjected to analytical and clinical correlation studies to:

- characterize the locus/mutation(s) being detected (literature review);
- establish the performance properties of the assay to ensure the assay’s ability to provide consistent and reliable results;

\[\text{ATCC is a registered trademark of American Type Culture Collection.}\]

*An NCCLS global consensus guideline. ©NCCLS. All rights reserved.*
• establish the clinical utility of the assay (i.e., that the assay will measurably contribute to the
diagnosis of a disease or genetic status of the individual or family member tested);

• define aspects of the procedure which should be carefully regulated to maintain assay performance;

and

• define relevant limitations of the assay.

Such verification is necessary to ensure the safe and effective use of an assay for its intended use. Detailed protocols for assay verification are beyond the scope of this document. Each laboratory should develop its own verification protocols. Additional information covering assay verification is available in NCCLS document MM1—Molecular Diagnostic Methods for Genetic Diseases and in the Standards and Guidelines for Clinical Genetic Laboratories, 2003 Edition (American College of Medical Genetics [ACMG]; www.acmg.net).

Sequencing-based assays address either detection of a sequence variation where the mutation to be detected is known, or detection of a sequence variation from a reference or consensus sequence where the sequence variation might be anywhere in the region sequenced. In both cases, mixtures of sequences could be present (see Section 8.6.2 for additional information). The laboratory must identify and characterize samples that are positive for the known sequences to be detected or representative of the sequence variations to be expected. Depending on the rarity of the disorder, acquiring sufficient samples for verification studies may be difficult. Samples may have to be obtained from other clinical laboratories or researchers who study the disorder. Getting appropriate informed consent to use the samples in verification studies is recommended. These samples should be verified by testing by alternate methods such as allele or sequence-specific PCR amplification, or sequencing by other laboratories. The sequence of interest can be cloned from a clinical sample and multiple clones can be sequenced to verify the sample sequence. In addition, negative samples that have the expected reference or consensus sequence should be obtained in order to fully validate the assay for sensitivity and specificity.

It is also important that analytical parameters that might affect assay performance such as instrument run-to-run performance, different reagent lots, amplification and sequencing method variability, and operator-to-operator differences are addressed.

8.4 Instrument Setup

Both gel and capillary sequencers are complex instrument systems. For clinical applications, the instrument performance should be validated by installation and operation qualifications. The key aspects are to ensure that the instrument is properly installed and performing as specified by the manufacturer.

Critical systems to address in the installation are:

• all the components are installed properly;

• power requirements are adequately met including consideration of noninterruptible service;

• ventilation requirements are met; some sequencers generate significant waste heat that must be exhausted from the laboratory; and

• computer and software also require conditioned power and suitable back-up.

Operational qualification incorporates running a sequencing reference standard multiple times and assessing that operators of the sequencing system are properly trained to maintain the equipment. All
appropriate sequencing results must be obtained. Most sequencing systems are installed and tested by the manufacturer as part of the purchase. It is important that the sequencers receive required cleaning and preventive maintenance to function properly. Many systems in the sequencer are not accessible to the user and require calibration and service by the manufacturer. Proper alignment and function of the optical detection system is critical to high quality sequencing results. A service contract is essential for maintaining a sequencing instrument, and having access to a telephone at the instrument is valuable in troubleshooting with technical service. Most manufacturers offer training courses for operating their equipment and software. The cost of training of operators should be included as necessary in the purchase of the instrument. An operator should be qualified by assessing performance on controls and reference standards prior to running test samples.

8.5 Electrophoresis

Electrophoresis on gel and capillary sequencers is controlled by the instrument’s software. There are major differences between gel and capillary sequencers on the methods for preparing the separating medium and loading the samples. Capillary sequencers automatically load the separation medium in the capillary and load the sample. For gel-based sequencers, users have to cast a gel and load the samples manually.

8.5.1 Gel-Based Sequencers

The quality of the gel is a major contributor to the overall quality of the sequencing results on gel-based sequencing systems. Gels need to have uniform polymerization, no air bubbles, and the glass plates have to be free of lint and light scattering materials. It takes training and experience to consistently cast high quality gels. Visual examination of the gel can discern problems with air bubbles and/or uneven polymerization. Since multiple lanes are used on a gel, there is potential for some of the lanes to generate unusable data. This will result in the need to rerun the sequencing reactions that were affected on a new gel.

8.5.2 Capillary Array Sequencers

Capillary array sequencers use the capillaries over and over again until the maximum number of runs specified by the manufacturer is achieved. Each capillary is separately calibrated for the dyes used in the sequencing reactions, so that the software can perform the multicomponent analysis to identify each of the dye-labeled fragments. There can be variation among capillaries initially and over time, so recalibration should be done periodically. Changing of the capillary array requires optical recalibration and revalidation of sequencing performance. It is possible that a capillary will have a problem in a sequencing run that causes the loss of sequencing data. An air bubble in the separation media, sample loading problems, or other problems in the performance of a capillary could lead to a loss of data. These sequencing reactions would have to be reanalyzed. Unlike gels, which can be visibly inspected for obvious defects, the loading of the capillary is not easily observed. If your instrument system allows observation of the acquisition of the raw data, problematic capillaries can be identified. The choice of polymer is defined by the manufacturer of the sequencer and users should use the polymer as recommended by the manufacturer.

8.6 Data Examination

8.6.1 Sequence Quality Assessment

After electrophoresis is finished, the sequencing analysis software translates the electrophoretic data into sequencing data. The first assessment of the quality of the sequencing run is whether the software was able to generate a sequence for each sample. If multiple primers are used to sequence larger targets, the ability of the alignment software to generate the consensus sequence from each of the reactions is a good quality measure. Sequences should be cross-referenced to the electropherogram data by way of an output
file from an automated instrument. The important characteristics to review in the electropherogram include signal intensity (peak height), signal to noise ratio, clear spacing or resolution of the peaks, low background, and noise (minor peaks). Regularity of base spacing may be compromised by G- or C-rich regions. Background peaks should not exceed 5% of the peak height for a homozygous peak. When the electropherogram characteristics are acceptable, then the instrument software is usually able to generate high quality sequencing data. This is easily evaluated by looking at the number of bases called or not called by the software. In evaluating the sequence, the range of accurate base calling should be known or established. If the number of uncalled bases exceeds either the instrument manufacturer’s specifications or a laboratory’s specifications derived from verification studies, then the electropherograms should be visibly examined to determine whether the sample should be rerun. For automated instruments or purchased kits, the manufacturer should be consulted regarding the range of accurate base calling. The manufacturer’s manual for the sequencer and the sequence analysis software usually give examples of problematic sequencing data and what to do. Additional information can also be found on the Internet at http://seqcore.brcf.med.umich.edu/doc/dnaseq/interp.html.

In performing a sequence analysis to confirm a mutation, examination of sequence from both strands is recommended. Sequencing of both strands may not be needed for all applications, but it is very helpful when looking for heterozygous single nucleotide substitutions. When this is not practical, either the quality of the sequencing analysis should be reviewed (see above) or the mutation confirmed by a repeat analysis. Comparing the sequence of the sample to a reference sequence is a useful quality assessment tool. A high level of concordance indicates that a sample was sequenced properly and the sequence is from the target region. Since in most assays, only a subset of the sequence is important clinically, the rest of the sequence provides an internal standard for the sequence quality. In many applications, only a small segment of the sequence is expected to have variations, and most of the sequence can be compared to a reference or consensus sequence. The “all base accuracy” can be calculated and used to assess the quality of the sequence obtained. All base accuracy is calculated by determining for the total number of bases sequenced for a sample the percentage of the bases called that agree with the expected base call in the reference sequence. A subset of the sequence that is expected to be highly conserved between the patient sample and the reference sequence should be chosen for the analysis. Base-calling errors are often not evenly distributed throughout the sequencing run. The largest number of miscalls generally exists at the beginning of the sequence. Laboratories should adjust their sequencing reactions to account for this by beginning sequencing far enough upstream from the sequence of interest to minimize base miscalling for the DNA sequence of interest. Differences between the sample sequence and the reference sequence can be due to at least three causes: 1) The sample has a different sequence, either a polymorphism or a mutation, when compared to the reference or consensus sequence; 2) The infidelity of the enzymes used for PCR generated a base difference between the sample and the reference or consensus sequence; or 3) The automated sequence analysis software analyzed the electropherogram incorrectly, and has called an incorrect base in the sample. The expected level for all base accuracy depends on the amplification and sequencing chemistry used and the sequencing equipment chosen.

The role of the operator in getting high quality sequencing results should not be underestimated. Operators need to be thoroughly trained in executing the assay steps, operating the instrument, and interpreting the sequencing data. Operators should be qualified before performing the assay and proficiency testing should be employed.

8.6.2 Detection of Sequence Mixtures

When sequencing-based methods are designed to detect a mixture of sequences that may be in the sample, such as mutations in a background of wild type sequence, it is essential that the capability of the assay to detect mixtures is determined. The first step is the acquisition of reference samples or materials that can be used to generate samples that have known mixtures of sequences, for example, Sequence A or Sequence B. These samples must be taken through the entire method, including sample preparation, and analyzed with the methods to be used on patient specimens. Sufficient samples and replicates of defined
mixtures should be analyzed to determine the level of mixture (i.e., 30% of Sequence A with 70% of Sequence B) that can be reproducibly detected within a 95% confidence interval. It is important that all the variations in the method, such as instrument run-to-run performance, different reagent lots, amplification and sequencing method variability, and operator-to-operator differences, are addressed in the study.

Sequencing in both directions can add confidence to the base call of a mixture. However, if highest sensitivity is desired, a mixture may be evident in one direction and not confirmed in the other. The cause of this may be the differences that sequencing enzymes have in incorporating the different ddNTPs into a DNA strand, thus the base peak in one direction will be strong enough to call, but not strong enough in the opposite direction. Another possibility is that the noise in the sequence data is higher for one direction than the opposite direction, and the mixture is above the noise in one direction, but not in the other. If the highest possible sensitivity is important, then the verification study should incorporate single direction mixture determination in the analysis. The confidence level should be determined for both single and bi-directional sequencing sensitivity.

9 Data Analysis

9.1 Sequence Review

9.1.1 Quality Assurance and Quality Control

Sequence analysis for clinical purposes is performed to either confirm a previous finding or determine the sequence of a defined region of DNA. Methods for estimating error differ for each purpose. This can be appreciated in considering what an error rate of 1% means for each purpose cited. In confirming a previous finding, an error rate of 1% tells us that 99 out of 100 times we will expect the correct answer. An error rate of 1% in sequencing a 500 base pair fragment is expected to result in five incorrect base determinations.

Sequence controls should include a wild-type specimen having characteristics of actual test samples. Results from sequencing this control should be available for comparison to results from test samples and also useful for training operators. The wild-type sample should be run at intervals sufficient to detect changes in performance of the platform and consistent with current regulatory requirements.

Either reverse strand sequencing or a duplicate repeat of the sequencing reaction should be performed, even when a reference sequence is used. Sole use of a reference (control) sequence is not always a good quality assessment tool since it may be of higher quality than the clinical sample.

Quality assurance and control measures for sequence analysis can be broadly divided into two categories. The first covers efforts to ensure appropriate sample collection, preparation, and chemistries, and the operation of the technology platform according to specifications. The second includes efforts to ensure that the correct sequence is determined from the raw data generated. Interpretation of raw data is often performed through a variety of software tools. Operator review of this interpretation is critical to determine whether the data returned are reasonable or if problems may have occurred in the sequence analysis or software interpretation.

9.1.1.1 Software Interpretation

Raw data obtained from a sequencing analysis represent mobility shifts of nucleic acid fragments though a matrix (i.e., acrylamide). For acrylamide gels, in which a number of samples may be run simultaneously, the software must be capable of accurate lane tracking and assignment. Problems can occur if lanes are smeared, streaked, pinched, tilted, or wavy. Direct operator review of the raw data (the actual gel or computer recreation) is sometimes necessary to determine whether any of these conditions
exist. Further, the mobility of fragments depends on their size and specific dyes used in the detection process. Dyes can alter mobility shifts. Software uses mobility-shift files that contain data to compensate for changes in mobility due to the presence of dyes. Sequence platforms are often flexible and can accommodate multiple dye combinations. Use of the correct mobility-shift file is necessary for correct software interpretation of the results.

Laboratories performing diagnostic sequence analysis should be able to document the quality of sequence determinations. A number of software packages provide “quality scores” useful for this purpose. Many factors are often included in the calculation of a “quality score” and may include consideration of:

- **Signal strength** – This refers to the peak height and may be expressed in relative units or normalized against some standard (electronic or control sample) sequence.
- **Signal-to-noise ratio** – The ratio of an actual signal (peak) to other signals (peaks) in the same vicinity.
- **Overlap** – The extent to which adjacent peaks occupy the same or different positions.
- **Loss of signal intensity** – The extent to which signal strength is lost during the course of a sequence analysis.
- **Baseline variation** – The extent to which the baseline remains at a constant signal strength. The signal strength may be absolute or relative to normalization against some standard (electronic or control sample).
- **Compression** – Changes in the relative spacing between subsequent peaks.

Although quality scores are useful in documenting the overall quality of a sequence determination, the composite score may or may not reflect certain problems. For this reason, operator evaluation of both the raw and processed data is critical in assuring the quality of the procedure. The sequence should be visually evaluated to determine if any of the items listed above adversely affect the interpretation of the raw data.

Samples to be sequenced may contain mixtures of nucleic acids. These may or may not be expected but their presence must be evaluated through the software and visual inspection of the data by the operator. A mixture may result from:

- heterozygosity;
- mosaicism;
- fragments resulting from amplification of pseudoalleles;
- spiking with control materials;
- other multiple amplification products (intended or unintended); and
- contamination.

### 9.1.1.2 Examples of Problems in Interpretation of the Raw Data

It is important for the operator to understand and recognize the range of analytic outcomes that are clinically reasonable or may be attributable to problems with the technology platform or software. Problems and their causes can often be inferred through evaluation of the original or recreated mobility-shift pattern (i.e., the acrylamide gel), the raw unprocessed data, the processed data, and the quality score and indicators provided by the software used. It is also important to note agreement or no agreement between each strand when reverse strand sequencing is performed on duplicative samples. The quality of
sequence determined for the wild-type control may also be useful as a comparison in identifying problems.

Poor sequencing data can have several causes. The data may show:

- no recognizable signal;
- poor lane recognition and tracking;
- noise resulting in missed or incorrect base calls;
- signal loss after base-calling begins; and
- unexpected stops.

Potential causes for these problems may include:

- poor template (i.e., low concentration or poor preparation);
- inadequate purification of template prior to sequencing reaction;
- poor primer and template annealing;
- multiple annealing sites for primer;
- inappropriate concentration of template or primer;
- failed PCR amplification;
- template different from what is expected (i.e., due to unknown genomic sequence variation or unanticipated RNA processing);
- contaminating sequences present in reaction mix;
- poor primer design or purity (low Tm, primer-dimer formation);
- interfering secondary structure within the template;
- reagents not performing as expected or wrong concentrations used;
- wrong mobility file used; and
- low signal strength resulting in failure of software to interpret raw data.

Software interpreted sequence data may on occasion appear to indicate a mixture when in fact, a homogeneous sequence is present. This can occur if more than one primer is present in the sequencing reaction, a secondary hybridization site, or secondary structure within the template.

Changes in the operating characteristics of the nucleic acid sequencer while data are being collected can result in aberrant results. Use of a standard control during the sequence run provides one measure that the technology platform is operating as expected.

Troubleshooting low-quality sequencing runs can be complex and may involve investigating multiple possible causes. If unable to resolve the problem, the operator should consult the manufacturer or seller of the sequence reaction kits or technology platforms to help identify and fix difficulties.

Software tools exist that are both platform dependent and independent. PHRED is one prototypic software program able to estimate the likelihood of error based on peak mobility and shape analysis. This software package can be adapted to a number of sequence chemistries and technology platforms. PHRED scores are reported as 10x the log of likelihood of error. A PHRED score of 40 or higher is appropriate for sequencing a 1000 base pair fragment. This score means there is a 1/10 000 likelihood of error. PHRED was primarily designed as a research tool and is cumbersome to integrate into a clinical environment. One complexity of PHRED is its requirement for look-up tables that take into account gel formulation and sequencing chemistries. Commercial software packages have been developed more amenable to implementation in clinical laboratories. These programs are available from manufacturers of sequencing instrumentation that support their equipment as well as independent vendors that have developed programs useful in analyzing data from a number of sequencing platforms. Laboratories performing
diagnostic sequence analysis should use one of these programs to determine the sequence and as an aid in documenting the quality of the sequence run.

9.1.2 Sequence Editing

Any sequence editing that differs from what is reported by the software interpretation needs to be documented and justified. Included in the documentation should be data, a description of the data examination process, and a clear explanation for the decision made to edit the sequence. It is advisable to have the decision to edit the sequence reviewed by a second person familiar with the sequencing methodology used.

9.2 Clinical Interpretation

Results should be interpreted in the context of the patient’s clinical findings, other laboratory tests, and findings from the sequence analysis. Relevant information such as the prevalence and frequency of sequence variations found, family history, inheritance pattern, factors related to gene expression, and clinical significance (see Section 9.2.2) should be considered in deriving an interpretation. Sequence findings, normal variations, and mutations should be referenced to listings in GeneBank, other DNA sequence databases, or the published literature.

In developing an interpretation, confounding factors may impact on the accuracy of the sequence determination. These may include allele dropout (due to amplification artifact or primer-binding site mutation), unexpected deletions, insertion, other rearrangements, mosaicism, or other interfering DNA sequences. It is the responsibility of the laboratory performing the test not only to be aware of those genes and disease conditions in which these events are more likely to occur, but to recognize that they can occur in other situations as well. As appropriate, these factors should be noted and explained on the test report.

Software packages are available that provide clinical treatment recommendations based upon the sequence determined. For instance, several packages are available that provide recommendations for HIV antiviral therapy. While recommendations to assess the validity of such software are beyond the scope of this guideline, discordance in the interpretation of sequence results has been reported among a number of these software packages. Therefore, users (laboratory personnel and clinicians) must take steps to understand the usefulness, reliability, and limitations of clinical interpretations generated by such software.

9.2.1 Nomenclature

As sequence analysis becomes increasingly important in making clinical decisions, it is critical to have a clearly defined nomenclature to avoid misinterpretation of the laboratory results. Nomenclature is necessary to describe the region of the genome or other nucleic acids being investigated (i.e., the gene(s) or relevant noncoding regions, episomes, rRNA) and a common language for describing the sequence variations therein.

In 1996, the Ad Hoc Committee on Mutation Nomenclature recommended a standard system for describing sequence variations for human DNA. This effort led to the development of The Human Gene Nomenclature Committee under the auspices of the Human Genome Variation Society. This nomenclature continues to evolve and changes are made public through publication. We recommend that this system be the standard by which sequence variations are described. Some of these recommendations have been described in NCCLS and other publications and salient points are outlined below.
9.2.1.1 Reference Sequence

A single reference sequence should be established or referenced. A genomic reference sequence is preferred to overcome difficult cases. However, if this is not available, a cDNA reference sequence should be used instead. The laboratory report should indicate the nomenclature used in describing the reference sequence. The standard for protein coding sequences is to have the nucleotide corresponding to the A of the ATG start codon referred to as position +1. There is no nucleotide 0. The position preceding the 5′ A should be designated as position -1. For mtDNA sequences, a standard reference sequence has been published and is commonly used. Accordingly, the origin of heavy strand replication is designated as the first base.

9.2.2 Clinical Significance

Clinical significance of a test should be based upon evidence that the sequence or variation thereof is associated with disease, a predisposition to disease, or relevant clinical condition. In describing the clinical significance of a human sequence variation, consideration must be given to prevalence rates in the patient population and other epidemiological data.

In reporting the clinical significance of human sequence variations, it is important to describe its relevance to the condition for which the test was referred. We recommend using the following categories in describing the relevance of sequence variations. Sequence variations can occur in protein coding and noncoding regions as well as regions that regulate RNA synthesis and other cellular activities. These are adapted from guidelines adopted by the American College of Medical Genetics:

- **Sequence variation is previously reported and is a recognized cause of the disorder.** Reported when there exists credible documentation of a sequence variation and a clinical outcome.

- **Sequence variation is previously unreported and is of the type which is expected to cause the disorder.** Insertions, deletions, and frame shifts can disrupt normal protein synthesis or regulation of cellular processes such as transcription and translation.

- **Sequence variation is previously unreported and is of unknown significance.** These variations represent base changes that do not change the coding sequence or effect known processing or regulatory pathways. However, such sequence variations may still produce cryptic splice sites, affect regulatory processes, interrupt exonic splicing enhancers and suppressor sites, or participate in other mechanisms that may be associated with disease.

- **Sequence variation is previously reported and is a recognized neutral variant.** Reported when evidence is available that the sequence variation has been consistently observed in a normal population without association to disease or predisposition.

In cases where the outcomes of the finding of a sequence variation are not understood, family studies may help clarify the clinical situation.

9.2.3 Limitations

Gaining useful information from sequence analysis is predicated on having an appropriate sample, interrogating the appropriate region of the nucleic acid, and understanding the technical limitations.

Interpretation of sequence analysis may be limited by the following:

- Appropriate interpretation of sequencing results depends, in part, on knowledge of its relevance to the medical condition referred. A clinically valid correlation between the presence of a particular
sequence and a medical condition can be based on clinical, molecular, biochemical, and/or epidemiological data. Low penetrance and/or variable expressivity may complicate the interpretation. Other genes, sequence variations, and environmental factors may contribute to the medical condition and should be considered in the interpretation.

- Some deletions, insertions, rearrangements, and other alterations may not be detectable by sequence analysis. For instance, in humans, a deletion of the sequence under investigation in one chromosome will not be obvious if only sequence analysis is used as the diagnostic tool. In such cases, a heterozygous deletion may appear as a homozygous normal individual. Likewise, a rearrangement may yield a correct sequence determination but not indicate that a recombination has taken place.

- Tissue-specific expression may impact the sequencing results and interpretation when RNA is the starting material.

- Limited sensitivity may preclude detection of sequences present.

- Low titer, mosaicism, mitochondrial heteroplasmy, presence of other DNA sequences, or other factors can limit the detectability of a particular sequence and should be considered in interpretation of the results.

- Recent (whole) blood transfusions may confound the sequencing results and interpretation.

### 9.3 Reports

#### 9.3.1 Critical Elements

Reporting of test results derived from molecular analysis has been described in NCCLS and other documents and is applicable to DNA sequence analysis. Salient features include the name of the individual, date of birth, specimen collection date, and other data. In addition to these, critical to the reporting of results derived from sequence analysis, the following should also be included:

- The gene and/or chromosomal region interrogated should be clearly identified (CFTR gene, etc.).

- That part of the gene and/or chromosomal region interrogated should be specified (exons, splice sites, etc.).

- The sequence and/or variation determined and position assignments should be specified and assigned according to accepted nomenclature. Both the nucleotide and the codon where a sequence change occurs should be indicated and frame shifts should be noted, including the codon position that results in a stop.

- When appropriate, the corresponding change in amino acid sequence should be specified according to accepted nomenclature.

- For human sequence determinations, results should indicate findings relevant to each allele.

- If available, other supportive data should be indicated that had been used in deriving the interpretation (i.e., family history, other clinical findings, homology to related sequences).

- The clinical relevance of the findings (see Section 9.2.2).

- Analytic limitations (see above).
• Interpretive limitations of the analysis (i.e., a negative result does not rule out contributory mutations present elsewhere in the genome).

• All databases used in the analysis should be referenced in the report. If applicable, a website reference for identifying gene-specific databases should be included.

9.3.2 Confidentiality

All patient results should remain confidential and adhere to regulatory requirements. Generally, results should only be made available to the referring healthcare professional. In some cases, results may be communicated directly to the patient. In these situations, specific policies should be established to help ensure that the patient understands the results, the limitations of the test, and recommendations for follow-up.

9.3.3 Preservation of Records

The laboratory should retain records for at least ten years and/or at least as long as mandated by regulatory authority. In some cases, such as tests performed on newborns, it is advisable to maintain records until the age of maturity.
References


45 American College of Medical Genetics; Standards and Guidelines for Clinical Genetics Laboratories; 2003. (Available at: http://www.acmg.net/Pages/ACMG_Activities/standards-2002/standardsMenu.n.htm.)

Summary of Consensus/Delegate Comments and Committee Responses

MM9-P: *Nucleic Acid Sequencing Methods in Diagnostic Laboratory Medicine; Proposed Guideline*

**Section 5.3, Specimen Quality Assurance**

1. Quality Assurance seems incredibly brief and does not address as a guideline how to set up QA related to sequence diagnostics.

   • The subcommittee has reviewed Section 5.3 and has found that specimen quality assurance is adequately covered. Quality assurance is addressed throughout the document and in significant detail in Section 8.2, Controls and Reference Standards, and in Section 9.1.1, Quality Assurance and Quality Control.

**Section 8.2, Controls and Reference Standards**

2. Positive and negative controls are not needed for every sequencing run, since the sequence itself is an internal control.

   • The subcommittee agrees with the commenter and it is not stated in the document that positive and negative controls are needed for every sequencing run. However, it has been reported that the FDA has required a company to specify running a control with every run for its FDA-cleared genotyping product.

**Section 8.6.1, Sequence Quality Assessment**

3. Sequencing of both strands may not be needed for all applications, but can be helpful when looking for heterozygous single nucleotide substitutions.

   • The second paragraph of Section 8.6.1 has been revised with the addition of the commenter’s statement as the second sentence.

**Section 9.2, Clinical Interpretation**

4. The issue of software generated clinical interpretations which interface with wet sequence software are not addressed.

   • Recommendations to assessing the validity of such software is beyond the scope of this guideline; however, for informational purposes the following paragraph has been added in Section 9.2:

   Software packages are available that provide clinical treatment recommendations based upon the sequence determined. For instance, several packages are available that provide recommendations for HIV antiviral therapy. While recommendations to assess the validity of such software are beyond the scope of this guideline, discordance in the interpretation of sequence results has been reported among a number of these software packages. Therefore, users (laboratory personnel and clinicians) must take steps to understand the usefulness, reliability, and limitations of clinical interpretations generated by such software.
The Quality System Approach

NCCLS subscribes to a quality system approach in the development of standards and guidelines, which facilitates project management; defines a document structure via a template; and provides a process to identify needed documents through a gap analysis. The approach is based on the model presented in the most current edition of NCCLS document HS1—A Quality Management System Model for Health Care. The quality system approach applies a core set of “quality system essentials” (QSEs), basic to any organization, to all operations in any healthcare service’s path of workflow (i.e., operational aspects that define how a particular product or service is provided). The QSEs provide the framework for delivery of any type of product or service, serving as a manager’s guide. The quality system essentials (QSEs) are:

- Documents & Records
- Organization
- Personnel
- Equipment
- Purchasing & Inventory
- Information Management
- Occurrence Management
- Process Improvement
- Service & Satisfaction
- Process Control
- Assessment
- Facilities & Safety

MM9-A addresses the quality system essentials (QSEs) indicated by an “X.” For a description of the other NCCLS documents listed in the grid, please refer to the Related NCCLS Publications section on the following page.

Path of Workflow

A path of workflow is the description of the necessary steps to deliver the particular product or service that the organization or entity provides. For example, NCCLS document GP26—Application of a Quality Management System Model for Laboratory Services defines a clinical laboratory path of workflow which consists of three sequential processes: preanalytic, analytic, and postanalytic. All clinical laboratories follow these processes to deliver the laboratory’s services, namely quality laboratory information.

MM9-A addresses the clinical laboratory path of workflow steps indicated by an “X.” For a description of the other NCCLS documents listed in the grid, please refer to the Related NCCLS Publications section on the following page.

<table>
<thead>
<tr>
<th>Preanalytic</th>
<th>Analytic</th>
<th>Postanalytic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient Assessment</td>
<td>Test Request</td>
<td>Specimen Collection</td>
</tr>
<tr>
<td>Specimen Transport</td>
<td>Specimen Receipt</td>
<td>Testing &amp; Review</td>
</tr>
<tr>
<td>Laboratory Interpretation</td>
<td>Results Report</td>
<td>Post-test Specimen Management</td>
</tr>
</tbody>
</table>

MM6

Adapted from NCCLS document HS1—A Quality Management System Model for Health Care.
Related NCCLS Publications*

MM1-A  Molecular Diagnostic Methods for Genetic Diseases; Approved Guideline (2000). This document provides guidance for the use of molecular biologic techniques for clinical detection of heritable mutations associated with genetic disease.

MM5-A  Nucleic Acid Amplification Assays for Molecular Hematopathology; Approved Guideline (2003). This document addresses guidelines for a variety of amplification-based laboratory tests, including polymerase chain reaction (PCR), transcription-based amplification system (TAS), strand displacement amplification (SDA), ligase chain reaction (LCR), and other methods now widely used in diagnostic hematopathology. This guideline provides a basis for laboratory implementation and quality assurance in this important area of diagnostic molecular medicine.

MM6-A  Quantitative Molecular Methods for Infectious Diseases; Approved Guideline (2003). This document provides guidance for the development and use of quantitative molecular methods, such as nucleic acid probes and nucleic acid amplification techniques of the target sequences specific to particular microorganisms. It also presents recommendations for quality assurance, proficiency testing, and interpretation of results.

* Proposed- and tentative-level documents are being advanced through the NCCLS consensus process; therefore, readers should refer to the most recent editions.
Active Membership (as of 1 October 2004)

**Sustaining Members**
- Abbott Laboratories
- American Association for Clinical Chemistry
- Bayer Corporation
- BD
- Beckman Coulter, Inc.
- bioMérieux, Inc.
- CLMA
- College of American Pathologists
- GlaxoSmithKline
- Ortho-Clinical Diagnostics, Inc.
- Pfizer Inc.
- Roche Diagnostics, Inc.

**Professional Members**
- American Academy of Family Physicians
- American Association for Clinical Chemistry
- American Association for Respiratory Care
- American Chemical Society
- American Medical Technologists
- American Society for Clinical Laboratory Science
- American Society for Microbiology
- American Society of Hematology
- American Type Culture Collection, Inc.
- Asociacion Mexicana de Bioquimica Clinica A.C.
- Assn. of Public Health Laboratories
- British Society for Antimicrobial Chemotherapy
- Canadian Society for Medical Laboratory Science - Société Canadienne de Science Laboratoire de Médical
- Canadian Standards Association
- Clinical Laboratory Management Association
- COLA
- College of American Pathologists
- College of Medical Laboratory Technologists of Ontario
- College of Physicians and Surgeons of Saskatchewan
- ESMID
- International Council for Standardization in Haematology
- International Federation of Biomedical Laboratory Science
- International Federation of Clinical Chemistry
- Italian Society of Clinical Biochemistry and Clinical Molecular Biology
- Japan Society of Clinical Chemistry
- Japanese Committee for Clinical Laboratory Standards
- Joint Commission on Accreditation of Healthcare Organizations
- National Academy of Clinical Biochemistry
- National Association of Testing Authorities - Australia
- National Society for Histochemistry, Inc.
- New Zealand Association of Phlebotomy
- Ontario Medical Association Quality Management Program-Laboratory Service
- RTPA Quality Assurance Programs
- PTY Limited
- Sociedad Española de Bioquimica Clínica y Patología Molecular
- Sociedade Brasileira de Análises Clínicas
- Taiwanese Committee for Clinical Laboratory Standards (TCCLS)
- Turkish Society of Microbiology

**Government Members**
- Armed Forces Institute of Pathology
- Association of Public Health Laboratories
- BC Centre for Disease Control
- Caribbean Epidemiology Centre
- Centers for Disease Control and Prevention
- Centers for Medicare & Medicaid Services
- Centers for Medicare & Medicaid Services/CLIA Program
- Chinese Committee for Clinical Laboratory Standards
- Commonwealth of Pennsylvania Bureau of Laboratories
- Department of Veteran Affairs
- Deutsches Institut für Normung (DIN)
- FDA Center for Devices and Radiological Health
- FDA Center for Veterinary Medicine
- FDA Division of Anti-Infective Drug Products
- Iowa State Hygienic Laboratory
- Massachusetts Department of Public Health Laboratories
- National Center of Infectious and Parasitic Diseases (Bulgaria)
- National Health Laboratory Service (South Africa)
- National Institute of Standards and Technology
- National Pathology Accreditation Advisory Council (Australia)
- New York State Department of Health
- Ontario Ministry of Health
- Pennsylvania Dept. of Health
- Saskatchewan Health-Provincial Laboratory
- Scientific Institute of Public Health, Belgium Ministry of Social Affairs, Public Health and the Environment
- Swedish Institute for Infectious Disease Control

**Industry Members**
- AB Biodisk
- Abbott Laboratories
- Abbott Diabetes Care
- Acrometrix Corporation
- Advaxis Corporation
- Affymetrix, Inc.
- Ammirati Regulatory Consulting
- Anna Longwell, PC
- A/S ROSCO
- AstraZeneca Pharmaceuticals
- Avantis
- Axis-Shield POC AS
- Bayer Corporation - Elkhart, IN
- Bayer Corporation - Tarrytown, NY
- Bayer Corporation - West Haven, CT
- BD
- BD Consumer Products
- BD Diagnostic Systems
- BD Thailand Ltd.
- BD VACUTAINER SYSTEM
- Beckman Coulter, Inc.
- Beckman Coulter (K.K.), Japan
- Bio-Development SRL
- Bio-Inova Life Sciences
- Biomatik Inc.
- Bio-Rad Laboratories, Inc.
- Bio-Rad Laboratories, Inc. – France
- Bio-Rad Laboratories, Inc. – Plato, TX
- Blaine Healthcare Associates, Inc.
- Bristol-Myers Squibb Company
- Canadian External Quality Assessment Laboratory
- Cepheid
- Chen & Chen, LLC
- Chiron Corporation
- ChromaVision Medical Systems, Inc.
- Clinical Micro Sensors
- The Clinical Microbiology Institute
- Conical
- Copan Diagnostics, Inc.
- Cosmetic Ingredient Review
- Cubist Pharmaceuticals
- Dade Behring Inc. - Caputipo, CA
- Dade Behring Inc. - Deerfield, IL
- Dade Behring Inc. - Glasgow, DE
- Dade Behring Inc. - Marburg, Germany
- Dade Behring Inc. - Sacramento, CA
- David G. Rhodos Associates, Inc.
- Diagnostic Products Corporation
- Dipgen Corporation
- Eiken Chemical Company, Ltd.
- Elanco Animal Health
- Eletca Lab s.r.l.
- Enterprise Analysis Corporation
- F. Hoffman-La Roche AG
- Fort Dodge Animal Health
- Gen-Probe
- GlaxoSmithKline
- Greiner Bio-One Inc.
- Immunicon Corporation
- Immunix, Inc.
- Instrumentation Laboratory
- International Technidyne Corporation
- I-STAT Corporation
- Johnson and Johnson Pharmaceutical Research and Development, L.L.C.
- K.C.J. Enterprises
- LillieScan, Inc. (a Johnson & Johnson Company)
- Macionis Diagnostics
- Medical Device Consultants, Inc.
- Merck & Company, Inc.
- Minigrip/Zip-Pak
- Nansphere, Inc.
- National Pathology Accreditation Advisory Council (Australia)
- Nippon Becton Dickinson Co., Ltd.
- Nissui Pharmaceutical Co., Ltd.
- Norfolk Associates, Inc.
- Novartis Pharmaceuticals Corporation
- Olympus America, Inc.
- Optimer Pharmaceuticals
- Pfizer Animal Health
- Pfizer Italy Srl
- Powers Consulting Services
- Premier Inc.
- Procter & Gamble Pharmaceuticals, Inc.
- QSI Consulting
- Radiometer America, Inc.
- Radiometer Medical A/S
- Replidyne
- Roche Diagnostics GmbH
- Roche Diagnostics, Inc.
- Roche Laboratories (Div. Hoffmann-La Roche Inc.)
- Sartorius, Inc.
- Schering Corporation
- Schleicher & Schuell, Inc.
- Second Opinion
- SFBC Anapharm
- Streck Laboratories, Inc.
- SYN X-Pharma Inc.
- Synex Corporation (Japan)
- Synex Corporation (Long Grove, IL)
- Theradoc
- Theravance Inc.
- Thrombocyte, Inc.
- THYMED GmbH
- Transasia Engineers
- Trek Diagnostic Systems, Inc.
- Vetquenoil S.A.
- Vicuron Pharmaceuticals Inc.
- Vyrix, Inc.
- Wyeth Research
- XDIX, Inc.
- YD Consultant
- YD Diagnostics (Seoul, Korea)

**Trade Associations**
- Advamed
- Japan Association of Clinical Reagents Industries (Tokyo, Japan)

**Associate Active Members**
- Academisch Ziekenhuis - VUB (Belgium)
- Alfred L. duPont Hospital for Children (DE)
- All Children’s Hospital (FL)
- Allis Health System (Herstil, IL)
- American University of Beirut Medical Center (NY)
- Anne Arundel Medical Center (MD)
- Antwerp University Hospital (Belgium)
- Arkansas Department of Health
- ARUP at University Hospital (UT)
- Associated Regional & University Pathologists (UT)
- Atlantic Health System (NJ)
- Aurora Consolidated Laboratories (WI)

**AZ Sint-Jan (Belgium)**
- Azienda Ospedaliero Di Leco (Italy)
- Barreiros-Jewish Hospital (MO)
- Baxter Regional Medical Center (AR)
- Baystate Medical Center (MA)
- Bhagwan Duren Laboratories (Turkey)
- BC Biomedical Laboratories (Surrey, BC, Canada)
- Bermuda Hospitals Board
- Braintree Memorial Hospital (TX)
- Brownell Medical Center (FL)
- Cadham Provincial Laboratory (Winnipeg, MB, Canada)
- Calypso Laboratory Services (Calgary, AB, Canada)
- California Pacific Medical Center
- Canterbury Health Laboratories (New Zealand)
- Cape Breton Healthcare Complex (Nova Scotia, Canada)
- Carillon Consolidated Laboratory (VA)
- Carolina Medical Center (NC)
- Cathay General Hospital (Taiwan)
- Central Texas Veterans Health Care system
- Centro Diagnostico Italiano (Milano, Italy)
- Champlain Valley Physicians (NY)
- Chang Gung Memorial Hospital (Taiwan)
- Changi General Hospital (Singapore)
- Children’s Hospital (NE)
- Children’s Hospital & Clinic (MN)
- Children’s Hospital Medical Center (Akron, OH)
- Children’s Hospital of Central Texas (TX)

**Chinese Association of Advanced**
- Blood Bankers (Beijing)
- CHR St. Joseph Warquarques (Belgium)
- Clarian Health - Methodist Hospital (IN)
- CLSI Laboratories (PA)
- Community Hospital of Lancaster (PA)
- Community Hospital of the Monterey Peninsula (CA)
- CompuNet Clinical Laboratories (OH)
- Cook Children’s Medical Center (TX)
- Cook County Hospital (IL)
- Covance Central Laboratory Services (IN)
- Croynton University Medical Center (NE)
- Danish Veterinary Laboratory (Denmark)
- Detroit Health Department (MI)
- DFS/CLIA Certification (NC)
- Diagnostics da America S/A (Brazil)
- Dr. Everett Chalmers Hospital (New Brunswick, Canada)
- Duke University Medical Center (NC)

**Dwight David Eisenhower Army Medical Center (GA)**
- Eastern Health Pathology (Australia)
- Emory University Hospital (GA)
- Enzo Clinical Labs (NY)
- Evangelical Community Hospital (PA)
- Fairview-University Medical Center (MN)
- Florida Hospital East Orlando
- Focus Technologies (CA)
- Focus Technologies (VA)
- Foothills Hospital (Calgary, AB, Canada)
- Franciscan Shared Laboratory (WI)
- Fresno Community Hospital and Medical Center
- Gamna Dynacare Medical Laboratories (Ontario, Canada)
- Geisinger Medical Center (PA)
- Guthrie Clinic Laboratories (PA)
- Hagerstown Medical Laboratory (MD)
- Harris Methodist Fort Worth (TX)
- Hartford Hospital (CT)
- Headwaters Health Authority (Alberta, Canada)
- Health Network Lab (PA)
- Health Partners Laboratories (VA)