

1-1-2015

## College of American Pathologists' Laboratory Standards for Next-Generation Sequencing Clinical Tests

Nazneed Aziz

Qin Zhao

Lynn Bry

Denise K. Driscoll

Birgit Funke

*See next page for additional authors*

Find similar works at: <https://stars.library.ucf.edu/facultybib2010>

University of Central Florida Libraries <http://library.ucf.edu>

This Article is brought to you for free and open access by the Faculty Bibliography at STARS. It has been accepted for inclusion in Faculty Bibliography 2010s by an authorized administrator of STARS. For more information, please contact [STARS@ucf.edu](mailto:STARS@ucf.edu).

---

### Recommended Citation

Aziz, Nazneed; Zhao, Qin; Bry, Lynn; Driscoll, Denise K.; Funke, Birgit; Gibson, Jane S.; Grody, Wayne W.; Hegde, Madhuri R.; Hoeltge, Gerald A.; Leonard, Debra G. B.; Merker, Jason D.; Nagarajan, Rakesh; Palicki, Linda A.; Robotorye, Ryan S.; Schrijver, Iris; Weck, Karen E.; and Voelkerding, Karl V., "College of American Pathologists' Laboratory Standards for Next-Generation Sequencing Clinical Tests" (2015). *Faculty Bibliography 2010s*. 6408.

<https://stars.library.ucf.edu/facultybib2010/6408>

---

**Authors**

Nazneen Aziz, Qin Zhao, Lynn Bry, Denise K. Driscoll, Birgit Funke, Jane S. Gibson, Wayne W. Grody, Madhuri R. Hegde, Gerald A. Hoeltge, Debra G. B. Leonard, Jason D. Merker, Rakesh Nagarajan, Linda A. Palicki, Ryan S. Robetorye, Iris Schrijver, Karen E. Weck, and Karl V. Voelkerding

# College of American Pathologists' Laboratory Standards for Next-Generation Sequencing Clinical Tests

Nazneen Aziz, PhD; Qin Zhao, PhD; Lynn Bry, MD, PhD; Denise K. Driscoll, MS, MT(ASCP)SBB; Birgit Funke, PhD; Jane S. Gibson, PhD; Wayne W. Grody, MD; Madhuri R. Hegde, PhD; Gerald A. Hoeltge, MD; Debra G. B. Leonard, MD, PhD; Jason D. Merker, MD, PhD; Rakesh Nagarajan, MD, PhD; Linda A. Palicki, MT(ASCP); Ryan S. Robetorye, MD; Iris Schrijver, MD; Karen E. Weck, MD; Karl V. Voelkerding, MD

• **Context.**—The higher throughput and lower per-base cost of next-generation sequencing (NGS) as compared to Sanger sequencing has led to its rapid adoption in clinical testing. The number of laboratories offering NGS-based tests has also grown considerably in the past few years, despite the fact that specific Clinical Laboratory Improvement Amendments of 1988/College of American Pathologists (CAP) laboratory standards had not yet been developed to regulate this technology.

**Objective.**—To develop a checklist for clinical testing using NGS technology that sets standards for the analytic

wet bench process and for bioinformatics or “dry bench” analyses. As NGS-based clinical tests are new to diagnostic testing and are of much greater complexity than traditional Sanger sequencing–based tests, there is an urgent need to develop new regulatory standards for laboratories offering these tests.

**Design.**—To develop the necessary regulatory framework for NGS and to facilitate appropriate adoption of this technology for clinical testing, CAP formed a committee in 2011, the NGS Work Group, to deliberate upon the contents to be included in the checklist.

**Results.**—A total of 18 laboratory accreditation checklist requirements for the analytic wet bench process and bioinformatics analysis processes have been included within CAP's molecular pathology checklist (MOL).

**Conclusions.**—This report describes the important issues considered by the CAP committee during the development of the new checklist requirements, which address documentation, validation, quality assurance, confirmatory testing, exception logs, monitoring of upgrades, variant interpretation and reporting, incidental findings, data storage, version traceability, and data transfer confidentiality.

(*Arch Pathol Lab Med.* 2015;139:481–493; doi: 10.5858/arpa.2014-0250-CP)

Accepted for publication June 19, 2014.

Published as an Early Online Release August 25, 2014.

From Molecular Medicine (Dr Aziz), Laboratory Improvement Programs (Dr Zhao and Ms Palicki), and Laboratory Accreditation and Regulatory Affairs (Ms Driscoll), College of American Pathologists, Northfield, Illinois; the Department of Pathology, Brigham & Women's Hospital, Harvard Medical School, Boston, Massachusetts (Dr Bry); the Department of Pathology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts (Dr Funke); the Department of Clinical Sciences, University of Central Florida College of Medicine, Orlando (Dr Gibson); the Divisions of Medical Genetics and Molecular Diagnostics, Department of Pathology & Laboratory Medicine, Pediatrics, and Human Genetics, UCLA School of Medicine, UCLA Institute for Society and Genetics, Molecular Diagnostic Laboratories and Clinical Genomics Center, UCLA Medical Center, Los Angeles, California (Dr Grody); the Department of Human Genetics, Emory University School of Medicine, Decatur, Georgia (Dr Hegde); Robert J. Tomsich Pathology & Laboratory Medicine Institute, Cleveland Clinic, Cleveland, Ohio (Dr Hoeltge); the Department of Pathology, University of Vermont College of Medicine, Burlington (Dr Leonard); the Departments of Pathology (Drs Merker and Schrijver) and Pediatrics (Dr Schrijver), Stanford University School of Medicine, Stanford, California; the Department of Pathology & Immunology, Washington University School of Medicine, St Louis, Missouri (Dr Nagarajan); the Department of Laboratory Medicine & Pathology, Mayo Clinic in Arizona, Phoenix (Dr Robetorye); the Departments of Pathology & Laboratory Medicine and Genetics, University of North Carolina at Chapel Hill, Chapel Hill (Dr Weck); and ARUP Laboratories Institute for Clinical and Experimental Pathology, and Department of Pathology, University of Utah School of Medicine, Salt Lake City (Dr Voelkerding). Dr Aziz is now with Phoenix Children's Hospital, Phoenix, Arizona.

The authors have no relevant financial interest in the products or companies described in this article.

Reprints: Nazneen Aziz, PhD, Phoenix Children's Hospital, 1919 E Thomas Rd, Phoenix, AZ 85016 (e-mail: naziz@phoenixchildrens.com).

DNA sequencing has evolved from Maxam-Gilbert<sup>1</sup> and Sanger<sup>2,3</sup> methods in the 1970s to a set of technologies that are collectively referred to as *next-generation sequencing* (NGS).<sup>4–12</sup> The primary difference between NGS and first-generation technologies is that sequencing of millions of short fragments of DNA occurs in parallel instead of one DNA fragment at a time. Sequencing of DNA as a clinical test became routinely possible only after the automation of Sanger sequencing methods introduced in the mid-1990s, which used capillary gel electrophoresis with fluorescence-based detection.<sup>13,14</sup> The throughput of NGS far surpasses that of automated Sanger sequencing. The higher throughput and lower per-base cost of NGS have contributed to its rapid adoption in clinical testing,<sup>15</sup> despite the fact that several aspects of NGS analysis have much higher complexity. Examples include the acquisition and storage of data sets that far exceed those commonly generated in a Clinical Laboratory Improvement Amendments of 1988 (CLIA) laboratory and downstream challenges in computation and interpretation. Areas in which NGS testing is being

applied currently include inherited diseases, solid tumors, hematologic malignancies, infectious diseases, human leukocyte antigen analysis, and noninvasive prenatal screening to detect fetal chromosome defects.

The number of laboratories offering NGS testing has grown considerably in the past few years, despite the fact that specific CLIA/College of American Pathologists (CAP) laboratory standards had not yet been developed to regulate this technology. To address this need, the CAP formed an ad hoc committee, the NGS Work Group, to develop the first set of clinical laboratory standards for this nascent technology. Given that NGS-based testing represents an evolving technology with continued improvements in instrumentation, sequencing chemistries, and bioinformatic and computational analyses, the work group aimed to develop standards that provide a necessary regulatory framework for clinical NGS tests (which to date are laboratory-developed tests) without inhibiting further adoption of NGS-based testing technology.

Next-generation sequencing incorporates 2 processes: (1) the analytic wet bench process and (2) bioinformatics analysis of sequence data. The wet bench component generally includes any or all of the following processes: handling of patient samples, extraction of nucleic acids, fragmentation, barcoding (molecular indexing) of patient samples, enrichment of targets for exome or gene panels, adapter ligation, amplification, library preparation, flow cell loading, and generation of sequence reads. Sequence generation is almost entirely automated and the output consists of millions to billions of short sequence reads. The wet bench workflow is followed by intensive computational and bioinformatics analyses that use a variety of algorithms to map and align the short sequence reads to a linear reference human genome sequence. After mapping and alignment, variant calls are made at locations where nucleotides differ from the reference sequence. Separate processes develop content needed to analyze the clinical relevance of variants, either singly or in combination, relative to their contribution to a given clinical phenotype. For individual patient cases, identified variants are evaluated against annotated content to infer the potential for impairments to normal gene function (eg, premature transcript or protein truncation, impact of nonsynonymous amino acid changes to protein function, or alternative splicing). Interpretation requires integrating genomic findings with the patient's clinical phenotype in order to make an informed decision regarding causality and correlation of the deleterious mutation(s) with the patient's disease. The mapping, alignment, variant calling, and variant annotation steps, and, to some degree, clinical interpretation (if decision support tools are used), comprise the overall bioinformatics analysis workflow.

The CAP NGS Work Group approached the analytic wet bench process and the bioinformatics or "dry bench" analyses as 2 discrete processes requiring separate considerations for standards. This division was leveraged to support the fact that some laboratories use external facilities to conduct either portion of NGS-based testing. In a laboratory offering the entire process from wet bench through bioinformatics analysis, clinical validation of their test will incorporate the validation of both parts. A total of 18 laboratory accreditation checklist requirements for the analytic wet bench process and bioinformatics analysis processes have been included within CAP's molecular pathology checklist (MOL). The NGS checklist items

include new standards for documentation, validation, quality assurance, confirmatory testing, exception logs, monitoring of upgrades, variant interpretation and reporting, incidental findings, data storage, version traceability, and data transfer confidentiality. As described in this report, the work group's goal was to initially develop foundational accreditation requirements for NGS that could be applied across multiple testing areas including inherited disorders, molecular oncology, and infectious diseases. It was anticipated that once foundational requirements were in place, there would be the need to subsequently develop additional, discipline-specific (eg, molecular oncology) NGS checklist requirements, and this is further addressed in the "Comment" section. This report describes important issues considered by the NGS Work Group during the development of each of the new checklist requirements. In addition, this report serves as a supplement to the CAP NGS checklist requirements and therefore the contents are closely aligned to each requirement for the 2014 checklist.

## WET BENCH ANALYTIC PROCESS

### NGS Wet Bench Process Documentation

#### **The Laboratory Uses a Standard Operating Procedure to Document the Analytic Wet Bench Process Used to Generate NGS Data.**

—The detailed documentation of the wet bench processes is a critical part of quality assessment in the clinical laboratory. All standard operating protocols of DNA/RNA sample preparation, fragmentation, library preparation, barcoding (molecular indexing), sample pooling, and sequence generation must be documented so that each step and subsequent manipulations can be traced. This includes documentation of all methods and reagents as well as instruments, instrument software, and versions used throughout the wet bench process. In addition, controls used need to be described. A few examples will be highlighted below. Targeted NGS assays (such as multigene panels or exome sequencing) allow selective capture of genomic regions of interest before sequencing, and detailed information regarding the captured region(s) (using genomic coordinates of capture probes and lists of genes) and target-enrichment protocols should be documented. Clinical laboratories that process different types of samples (eg, blood, formalin-fixed paraffin-embedded specimens) should develop standard operating procedures (SOPs) for each validated sample type. The reagents and protocols used for pooled analysis of patient specimens must be detailed and should include the sequence information of the barcodes used for each patient sample. Metrics and quality control parameters used to assess run performance must also be documented. Commonly used metrics include the percentage of reads mapping to the target region, the fraction of bases meeting specified quality and coverage thresholds, and average coverage/base and target region. The laboratory must define and document acceptance and rejection criteria for the wet bench process inclusive of sample preparation and sequencing. It is critical to determine and summarize regions that failed analysis (eg, due to inadequate coverage) if they are not covered by orthogonal technologies (such as Sanger sequencing).

Evidence of compliance for this requirement includes a written SOP that describes the analytic wet bench process and the ability to demonstrate that the laboratory follows its policies and procedures.

## NGS Wet Bench Process Validation

### **The Laboratory Validates the Analytic Wet Bench Process and Revalidates the Entire Process and/or Confirms the Performance of the Components of the Process as Satisfactory When Modifications Are Made. The Extent of Revalidation and/or Confirmation Is Modification Dependent.**

—Like all laboratory-developed tests in molecular diagnostics and other areas of the clinical laboratory, analytic performance of NGS procedures must be internally validated before clinical implementation. Next-generation sequencing analysis is a complex procedure with many steps within the wet bench workflow. Each step needs to be individually optimized to empirically determine optimal assay conditions and analysis settings. Once those are in place, an analytic validation must be performed for the whole test in a “beginning-to-end” fashion, including the entire wet bench process as well as the bioinformatic analyses. Essential performance characteristics that need to be determined during the validation are the analytic sensitivity and specificity, accuracy (the degree of closeness of measurements to the actual [true] value), precision (reproducibility and reliability), and limit of detection (if applicable). As for any molecular assay, validation should also be conducted independently for each accepted specimen type (blood, saliva, tissue, etc). Next-generation sequencing tests are typically designed to interrogate large and multiple regions of the genome, and its use can range from mutational hotspots for oncology applications to gene panels to exomes or genomes. As a consequence, NGS permits the detection of novel as well as known sequence variants, which necessitates a comprehensive approach to be able to determine test performance with adequate confidence. Because it is not possible to validate all theoretically possible variants that can occur, it is necessary to use a combination of a “methods-based”<sup>16</sup> and “analyte-specific” validation approach for determining a test’s analytic performance. Consulting the published literature for studies regarding the accuracy of the relevant NGS platform can be useful to inform the laboratory’s own validation work. In most cases, variants will have been identified via Sanger sequencing, considered (at least for now) the gold standard comparative technique. However, variant validation information may also be obtained from oligonucleotide microarray genotyping data in some cases. Several professional organizations have issued guidance regarding validation of molecular tests and, more recently, NGS tests in specific to which the reader is referred.<sup>17–21</sup>

As the NGS Work Group debated NGS validation requirements, the concept of requiring a minimum number of samples for inclusion in a validation was extensively discussed. It was concluded that adding a minimum sample number requirement was premature given the ongoing evolution of NGS technology and the diversity of applications being implemented in diagnostic laboratories. Further, the concern existed that establishing a minimum sample number requirement may result in laboratories conducting an insufficient validation for a given NGS diagnostic application. The work group noted that NGS validations reported in the literature have varied considerably in sample number size (eg, ~20–80 plus samples),<sup>22–32</sup> reflecting that individual laboratories are on a validation “learning curve.” The total number of samples that needs to be run to appropriately validate an NGS test is driven partly by the size of the test (larger assayed regions will have more

variants available for deriving their technical performance), by the number of specific analytes (variants) that need to be assessed, by the possible requirement to determine limit of detection across a range of allele frequencies, and by the number of runs and samples needed to determine precision. At this juncture in time, the NGS Work Group concluded that statistical considerations with regard to the number of samples cannot be universally or comprehensively applied across the numerous assays that are possible when using NGS (eg, amplicon versus targeted capture; small numbers of genes versus exome or genome; inherited disease versus oncology versus infectious disease) as the sequencing methodology. Therefore, we have described different scenarios (eg, samples needed for methods-based approach, samples needed to assess reproducibility and reliability, and clinical samples used to assess diagnostic specificity and sensitivity), each of which will necessitate samples whose numbers will vary with the context of each assay. We emphasized the principles of validation in the requirements and several analytic performance parameters as highlighted below.

Analytic sensitivity can be assessed by using a methods-based approach that aims at maximizing the number of sequence variants that are compared to a gold standard method to increase confidence of analytic performance. These values may then be extrapolated to all bases. For this methods-based approach, pathogenicity of analyzed variants does not matter as this has no bearing on their technical detectability. However, it is important to determine this “baseline” performance by using as many different genomic regions as possible, as sequence context can be an important influence. In addition, laboratories should determine analytic performance separately for all variant types that are relevant for the test (eg, single nucleotide variants, indels, copy number variants, structural variants, homopolymers). Approaches to maximize the number of appropriately identified variants may include cumulative analysis of different in-house-developed tests (eg, different gene panels), provided that they rely on identical protocols. In addition, several publicly available databases provide exome/genome-wide variant calls that can be used in the clinical validation efforts (eg, HapMap or 1000 Genomes). In addition, the Centers for Disease Control and Prevention and National Center for Biotechnology Information have collaborated to establish a Web browser to facilitate access to 2 well-sequenced genomes (NA12878 and NA19240)<sup>33</sup> and to provide access to clinical-grade targeted data sets (gene panels) and exome/genome-wide data sets created by various laboratories.<sup>34</sup> These databases provide access to large sets of variants that can aid in deriving technical performance specifications. However, an analyte-specific validation may be necessary in addition to the more global methods-based approach when the NGS test includes genes that are known to harbor well-known, disease-causing variants. In such cases, it is important to include traditional positive controls with patient samples, including relevant variants (eg, p.F508del in *CFTR*) to demonstrate adequate detection by that NGS test. Analytic specificity is often calculated by using “negative” samples (ie, samples that have no pathogenic variant) to determine the fraction that is correctly identified as negative. However, this concept does not work well for NGS-based tests. Once again, a methods-based approach can be leveraged to calculate analytic specificity across the assayed region, for example, by determining the false-positive rate (fraction of

variants detected that are incorrect calls). It is also useful to determine the average number of false-positive calls for the regions tested in a clinical sample. Note that the analytic specificity accounts for numerous sources of type I error, including base-calling error, errors due to misalignment, and variant-calling errors. Determining the limit of detection is important for assays that interrogate samples with heterogeneous genotypes (eg, tumor specimens, maternal blood used for noninvasive prenatal testing for fetal aneuploidies, and mosaic specimens). This can be challenging given that Sanger sequencing, which is often used as an alternate gold standard technology during validation, is not as sensitive as NGS. Sample “mixing” experiments (eg, dilution of samples with known allele frequencies) may be used to assess the limit of detection of each variant type. Precision (interrun and intrarun variability) should be determined by using at least 3 samples. For tests that are performed with single-lane sequencers, intrarun variability may be determined by using bar-coded replicates of the same sample.

Homologous sequences such as pseudogenes can interfere with accurate variant calling and therefore pose significant challenges for correctly analyzing affected genes. An upfront bioinformatics homology analysis is useful to determine possible interference by homologous sequences. In addition, read-mapping quality can be used to identify problematic regions. If such genes are included in the NGS test, the laboratory must devise a method to ensure that identified variants are not due to pseudogene sequence and must document the accuracy of the method. When pooled sequencing of bar-coded samples is performed, the laboratory must document that individual sample identity is maintained throughout the wet bench process.

The extent of revalidation and confirmation is dependent on the magnitude of the introduced changes and their potential consequences. For example, minor changes, such as the introduction of a new lot of capture reagent that has already undergone comprehensive validation, can be addressed by confirming adequate performance. In this example, it would be deemed acceptable if the laboratory sequences a previously tested sample and documents that the main run metrics (eg, coverage, read quality) are unchanged and that the same results are obtained. Conversely, a major change, such as the introduction of a new sequencing platform or different target enrichment method, would require a more extensive revalidation.

#### **NGS Wet Bench Process—Quality Management Program**

**The Laboratory Follows a Documented Quality Management Program for the NGS Analytic Wet Bench Process.**—CAP-accredited laboratories must develop and follow a quality management plan. The CAP All Common Checklist (COM) applies to every part of a multispecialty laboratory and includes entire sections on Quality Management and Test Method Performance. However, NGS Wet Bench Process—Quality Management Program was added to the NGS portion of the checklist to highlight the particular needs of laboratories performing NGS. No two quality management programs are alike. Each is shaped by the laboratory’s scope, clinical market, and expertise, and the laboratory director is given wide latitude in the design of the quality assurance program. The design of the program must be written, and compliance with that design documented. A good quality assurance program for laboratories performing NGS will include the following attributes<sup>35,36</sup>:

1. The quality assurance program follows the path of workflow. The programs should assess preanalytic steps occurring before NGS, analytic testing, and postanalytic processes used in sequence analysis through reporting.
2. The NGS quality program should be integrated within the institution’s overall quality assurance program. If it is part of a larger institution, such as a hospital or medical center, the NGS quality program should fit well within its overall context.
3. The program should address common problems that arise in the course of testing. “Problems” include events that can affect the test result or its clinical use as well as nonconformance with the laboratory’s own policies and procedures. Documentation includes both review of the effectiveness of corrective actions taken and the revision of policies and procedures intended to prevent recurrence.
4. The overall goal of the quality program aims to ensure that testing is clinically relevant. This is particularly important for tests such as NGS, for which no comparative analytic result of greater sensitivity may exist. The appropriateness of test orders and analytic decisions must be grounded in medical science and evidence.
5. The program should also encourage laboratory employees to communicate concerns about the quality of laboratory testing. The investigation of employee complaints and suggestions must be a part of the quality assurance program.

#### **NGS Confirmatory Testing**

##### **The Laboratory Has a Policy That Documents Indications for Confirmatory Testing of Reported Variants.**—

While the accuracy of NGS technologies is continuing to improve, it is widely accepted that most NGS-based sequencing assays will yield false-positive and false-negative results. CAP preferred to give laboratories performing NGS-based assays flexibility in determining when confirmatory testing should be performed, how this testing is performed, and whether to recommend confirmatory studies for follow-up testing for additional family members, which may or may not be NGS based. For example, some laboratories might determine during validation studies that confirmatory testing of identified variants was not necessary owing to the very high coverage achieved by their assay (ie, 1000× coverage of a single-gene NGS-based assay) and/or very high confidence in the identified variants.<sup>37,38</sup> However, others may find that they need confirmatory testing by an alternative method to achieve the desired confidence in the variants that are reported. Some laboratories might decide that they will perform confirmatory testing on variants for a predetermined trial period and then reevaluate this decision at a later date. Each laboratory performing NGS must have a policy in place that clearly documents indications for confirmatory testing and/or documents how their assay validation determined that such testing was not required. Laboratories must be able to document compliance with their confirmatory testing policy and show evidence of ongoing monitoring of their NGS assay(s) to ensure that the benchmarks achieved during the validation process are maintained during the routine performance of NGS-based clinical testing and variant reporting. CAP also desired to give laboratories flexibility in deciding the methods used to perform any needed confirmatory testing. Although Sanger sequencing is likely to be the method most commonly

chosen for confirmatory testing of NGS-identified variants, CAP did not want to mandate such testing in order to provide clinical laboratories with the flexibility to use other appropriate confirmatory testing methods consistent with the existing expertise of the laboratory and the type and frequency of variants requiring confirmation (ie, allele-specific polymerase chain reaction, melting curve analysis, other NGS-based method).

### Laboratory Records

**Methods, Instrument(s), and Reagents Used for Processing and Analyzing a Sample (or Batch of Samples) Can Be Identified and Traced in the Laboratory's Records.**—Comprehensive records of laboratory assay “runs” are essential to document the conditions and events associated with the complex processes and algorithms involved in the performance and interpretation of clinical NGS-based analyses. Accordingly, such archived information must be maintained within an overarching framework where all reagents, primers, sequencing chemistries, and platforms used for the analysis of each patient sample are traceable. Such records must contain a description of the test performed including the nature of the targeted sequence (eg, genome, exome, specific genes for targeted panels, transcriptome, or methylome) and depth of coverage (eg, range and average). It is also necessary to cite details of the analysis, including any publications or Web sites (with dates accessed) describing the pertinent parameters or other information and/or notations relative to the testing and reporting processes. While all details of the analysis need not be included in the patient report, it is critical that the laboratory maintain a documentation system from which detailed information regarding the analysis of individual patient specimens can be obtained.

### Exception Log

**The Laboratory Maintains an Exception Log for Patient Samples Where Steps Used in the NGS Analytic Wet Bench Process Deviate From Standard Operating Procedures.**—The laboratory must document any deviation from the SOP along with an explanation for the deviation, and the resulting outcome. Examples of anticipated deviations may include altered processing upon receipt of a suboptimal specimen, changes to the library preparation, and sequencing of libraries with suboptimal concentrations.

Exceptions may pertain to specimen quality and to the analytic process. At the time of specimen accessioning, an assessment is made as to whether or not a sample is in optimal condition for testing. If there is a concern, this can be documented on the worksheet or on a pending log and communicated to a supervisor or laboratory director. The director may decide to proceed with the testing, but should communicate the issue to the ordering physician and document this communication electronically or on the worksheet. One example of such a scenario is a sample that was not transported under optimal conditions. A decision may be made to process the sample and to proceed with subsequent testing only if the DNA specimen is found to be adequate.

Issues related to specific steps of the wet bench procedure should be reported to the laboratory supervisor or the director of the laboratory. It can then be assessed whether or not the testing was compromised and if the testing can be completed. If, after troubleshooting, the testing is assessed as satisfactory, the results can be interpreted by the

laboratory director, provided that the quality controls of the run and the sample results are deemed adequate. All aspects of the testing issue(s) should be thoroughly documented in an “exception log,” including the troubleshooting, the resolution, and the pertinent communications (especially regarding who was involved and who was informed by whom and on what date), and may also be incorporated into the monthly quality assurance report.

On occasion, the laboratory SOP itself may have to be revised to improve phrasing, to make process steps more clear, or to remove small inaccuracies in order to optimize the protocol. In such cases, the proposed correction should ideally be supported by at least 2 additional individuals, including the laboratory supervisor and either the technologist who developed the assay or a reference technologist. Any such corrections must be approved, signed, and dated by the director of the laboratory. This is not an exception log issue per se but rather a correction in the manner the assay is described.

### Monitoring of Upgrades

**The Laboratory Has a Policy for Monitoring, Implementing, and Documenting Upgrades to Instruments, Sequencing Chemistries, and Reagents or Kits Used to Generate NGS Data.**—Laboratories must be aware of upgrades to ensure that they are not using obsolete methods. The laboratory must implement a policy to monitor and implement upgrades to instruments, sequencing chemistries, and reagents or kits used to generate NGS data. The policy should address how laboratories performing NGS-based testing can ensure that they are using the most up-to-date sample library preparation as appropriate for that assay, clonal fragment amplification, and sequencing methods in this rapidly evolving environment provided that these newer methods have been validated by the laboratory to improve the quality, reproducibility, and accuracy of the assay. The policy should also address the methods used to monitor upgrades and when a relevant upgrade(s) will be implemented and further validated before productive clinical use. For example, the laboratory's policy may be to monitor and implement upgrades at specified intervals (such as quarterly, biannually, or annually), depending on the relevance of the new upgrade for enhancing assay performance. Additionally, since the implementation of upgrades may require revalidation of the entire wet bench process, or at least the relevant steps, it may be convenient to set time intervals accordingly.

### BIOINFORMATICS PROCESS

A variety of open-source and commercial bioinformatics algorithms and software is available for analyzing NGS data.<sup>39</sup> While these tools continue to improve, they each have strengths and weaknesses with respect to their performance in diagnostic applications. Operationally, the bioinformatics processes applied to NGS data can be conceptualized into 3 major steps. First, is the generation of a sequence read file consisting of a linear nucleotide sequence (eg, ACTGGCA), with each nucleotide assigned a numerical value (termed its *base quality score*) that correlates to its predicted accuracy. The generation of sequence read files uses instrument-specific software that analyzes several physical parameters, such as signal to noise ratios, during the sequencing run. Sequence read files are usually configured in the FASTQ file format, which contains the compilation of individual sequence reads, each with its own

identifier, and an associated base quality score for each nucleotide. FASTQ files have become a dominant form of information exchange in the field of NGS. The next step consists of aligning the sequence reads to a reference sequence, typically a human genome reference sequence, to identify differences between the patient sequence reads and the reference. Identified variants may include single nucleotide variants, insertions and deletions, copy number variants, and other structural variations (translocations, inversions, etc). Identified variants are then annotated to provide information regarding their impact on gene and protein function. Separate processes within the laboratory implement, or otherwise develop, curated content for assessing the clinical relevance of particular variants to a given disease or condition. Lastly, annotated variants are interpreted within the context of the patient's phenotype to render a clinical report. For gene panels and exome or genome sequences, the large list of annotated variants is typically reduced by excluding variants with a higher population frequency and by focusing on rare variants that are of greatest predicted deleterious impact that correlate with patient phenotype.<sup>40,41</sup> When analyzing exome or genome sequences within a family unit, variant prioritization typically takes into account variant cosegregation within the family, based on affected versus unaffected family members. Variant prioritization during the tertiary step uses previous knowledge of association of variants and disease within public or private databases of human mutation, such as the Human Gene Mutation Database (HGMD),<sup>42</sup> Online Mendelian Inheritance in Man (OMIM),<sup>43</sup> and/or other disease/locus-specific databases.<sup>44</sup>

Developing a cohesive diagnostic pipeline that incorporates bioinformatics steps, and content development for variant annotation, usually requires the integration of multiple algorithms and software applications. As such, laboratories must empirically determine which algorithms and associated bioinformatics tools to apply to each diagnostic application. An iterative pilot process commonly uses known patient samples and training data sets, which may be synthetic or from prior cases, to test algorithms and software parameters. Having established a working set of bioinformatics tools and parameters, the laboratory performs a bioinformatics validation with a larger set of samples to determine analytic sensitivity and specificity for the types of variants assayed (eg, single nucleotide variants, insertions and deletions, homopolymer or repetitive sequences, or copy number variants) and reproducibility (ie, concordance within and across runs, instruments, and technical personnel). The samples used for validation will contain previously confirmed variants, or the identified variants may be confirmed post bioinformatics analysis. The validation may confirm that the bioinformatics tools and parameters are performing satisfactorily (eg, high specificity and sensitivity if the assay is a stand-alone assay for variant detection and reporting versus high sensitivity if it is a screening assay followed by a second assay that is used for confirmation) per laboratory requirements and clinical criteria for reporting, or adjustments or alternative tools may need to be further evaluated.

Once a satisfactory bioinformatics validation has been achieved, translation of the NGS assay into the clinical laboratory requires that laboratories document all aspects of the bioinformatics processes used for clinical diagnostics and implement a quality management program for these

steps. Further highlights of the bioinformatics requirements for NGS are discussed below.

### NGS Bioinformatics Pipeline Documentation

**The Laboratory Uses an SOP to Document the Bioinformatics Pipeline Used to Analyze, Interpret, and Report NGS Results.**—Laboratories must document all algorithms, software, and databases (referred to as *components*) used in the analysis, interpretation, and reporting of NGS results.<sup>45</sup> The versions of each of these components in the overall bioinformatics pipeline must be recorded and traceable for each patient result (Version Control). For each component, the laboratory may use a baseline, default installation, or may customize the pipeline by using alternate configuration parameters in deploying individual bioinformatics tools or in running specific algorithms. In either case, laboratories must document any customizations that vary from default configuration or should indicate which parameters, cutoffs, and values are used. Most NGS bioinformatics analyses are conducted by aligning sequence reads to a reference sequence. The reference sequence version number and assembly details need to be identified. When describing the bioinformatics pipeline, laboratories should document the overall workflow of data analysis and include the input and output files for each process step. For each step, laboratories should also develop and document quality control parameters for optimal performance. For example, in the primary step, a laboratory would determine acceptable criteria such as the number of reads passing instrument-specified quality filters. Criteria for variant calling are essential and parameters that are invoked include thresholds for read coverage depth, variant quality scores, and allelic read percentages. Each of these requirements applies to multigene panel applications as well as to exome and genome sequencing. Laboratories should also document the bioinformatics processes that are used for reducing a large variant data set to a list of causal and/or candidate genes and/or variants. For example, in inherited disease assays, laboratories should document approaches used to identify recessive, dominant, and de novo variants. Evidence of compliance for this requirement would be demonstration of appropriate documentation and that the laboratory follows its outlined procedures.

### NGS Bioinformatics Pipeline Validation

**The Laboratory Validates the Bioinformatics Pipeline and Revalidates the Entire Pipeline and/or Confirms the Performance of the Components of the Pipeline as Satisfactory When Modifications Are Made. The Extent of Revalidation and/or Confirmation Is Modification Dependent.**—As with wet bench processes, laboratories use an iterative process during the establishment of a bioinformatics pipeline that involves analyzing sequence read files containing known variants and demonstration that the pipeline can identify the variants.<sup>17</sup> For laboratories offering the entire process from wet bench through bioinformatics analysis, the validation of the bioinformatics pipeline should be included in the overall test validation. Once the laboratory has developed and empirically determined optimal performance, and performed adequate testing of its pipeline, the next step is to perform and document a comprehensive validation, again using sequence reads generated from samples with variants that cover the spectrum of the diagnostic testing that the laboratory intends to perform. These steps are essential for



both in-house–developed tools and in those cases where a vendor–provided tool or pipeline is used in a manner where it is locked down, for example, the laboratory does not modify or alter any components or parameters of the underlying tools. As with wet bench processes, a sufficient number of samples need to be analyzed to assess the pipeline’s analytic and diagnostic sensitivity and specificity as well as the assay’s reproducibility. The number of samples assessed should be determined from the assay. Parameters such as the number of genes assessed, which regions of a gene are assessed, and types of variants that need to be detected should ultimately be used to determine the number of control, well-characterized samples (eg, HapMap samples or cell lines with known inherent or engineered variants) and previously analyzed diagnostic samples. The presence of pseudogene sequences and other sequences highly homologous to the target are known to interfere with accurate sequence mapping, alignment, and, by extension, variant calling. The degree of interference, if applicable in a given diagnostic assay, needs to be determined. While it may be possible to address the challenge of coalignment of highly homologous sequences bioinformatically, laboratories may need to set up independent alternative method assays for these problematic regions. The NGS Work Group acknowledged that it was not feasible to comprehensively and exhaustively define the error rates for false positives and false negatives in variant calls. However, the laboratory should assess the error rates for several representative examples by variant type. These rates may be assessed analytically by using well-characterized, control samples. False-positive error rates can be ascertained by sequencing using an alternative method. False-negative error rates are more difficult to ascertain because they may originate from several sources, including insufficient read coverage in a given target region, a lack of variant calling due to parameters such as lower variant quality scores, or a distribution of sequence read directions (ie, forward and reverse reads) that does not meet quality control criteria. In the process of validating variant calls by an alternative method, it may be possible to analyze flanking regions to determine if the variant calling pipeline is identifying all possible variants. For example, when using Sanger sequencing to confirm a variant, primer pairs can be designed to sequence the region of the variant as well as generous portions of flanking regions, which can then be inspected for the presence of variants and correlated with those identified by the bioinformatics pipeline.

A now common practice in NGS is the use of molecular barcodes or indexes during the preparation of libraries. Indexed sequences need to be validated with respect to their uniqueness in a pool and the pipeline must be able to accurately bin (segregate) such indexed sequences. In the analysis of indexed and pooled samples, it is essential to establish criteria for retention or exclusion of sequence reads. For example, some laboratories will only accept sequence reads with indexes in which the index sequence is identical to the index that was used during library preparation. Other indexed reads that do not align in a completely identical fashion are not assigned to the respective sample. The monitoring of the percentage of indexed reads that maintained full identity can be a measure of the presence of contamination from other index sequences. For those assays in which limit of detection is relevant, such as identification of somatic mutations in tumor samples, the bioinformatics pipeline needs to be

assessed for that parameter. One approach that can be used to validate limit of detection is to sequence samples with decreasing concentrations of target variants that have been created from a cell line or DNA dilution series.

Validation of the bioinformatics pipeline for identification of variants is application specific and the above discussion is broadly pertinent, with the exception of the limit of detection analyses being specific to samples with heterogeneous genotypes. When using exome and genome sequencing for causal and candidate gene identification, the laboratory must additionally validate its bioinformatics pipeline for this purpose. For example, in the case of inherited diseases, laboratories may approach this by analyzing sequence read sets with known pathogenic variants that are present in several deleterious variant configurations, such as recessive, dominant, and de novo.

Once a bioinformatics pipeline has been validated to meet laboratory requirements and has been implemented, revalidation is required when any changes are made in the pipeline. A practical approach that can be used to revalidate a sequencing pipeline is to use sequence read files from the original validation and simply reanalyze them with the new parameters. This approach may result in identical, smaller, or larger numbers of identified variants and these findings would need to be confirmed. For exome and genome sequencing, changes in bioinformatics pipelines can also result in a new list of presumptive causal or candidate genes. Evidence of compliance for a bioinformatics pipeline validation/revalidation would include the records of validation and any subsequent revalidation and their documented approval for clinical use.

#### **NGS Bioinformatics Pipeline—Quality Management Program**

**The Laboratory Has a Documented Quality Management Program for the NGS Bioinformatics Pipeline.**—The routine application of a validated bioinformatics pipeline must be accompanied by monitoring of laboratory-determined quality control metrics.<sup>46</sup> Divergence from expected quality metrics during the analysis of clinical samples requires investigation and resolution. Some examples include the following situations: the bioinformatics output of NGS data analysis may demonstrate that an insufficient number of sequence reads passed the expected or required base quality score threshold. Alternatively, the number of variants identified in a data set may deviate substantively from an expected value, based on prior information regarding known frequencies of variation in the human genome. Another example may be an inappropriately high number of indexed sequencing reads that cannot be specifically segregated. Such deviations may indicate a technical aberration or process failure occurring during technical wet bench procedures or during a step in the bioinformatics pipeline. An appropriate quality management program provides the structure and process for investigating these divergences to pinpoint possible causes, and institute appropriate corrective measures. Laboratories must maintain a record of deviations from expected results and document the investigative measures that were used to determine the cause as well as the corrective measures that were implemented. Evidence of compliance would include documentation of monitoring quality control metrics as well as records describing any divergences, including appropriate investigative measures and subsequent corrective actions.

## Bioinformatics Pipeline—Updates

**The Laboratory Has a Policy for Monitoring, Documenting, and Implementing Patch-Releases, Upgrades, and Other Updates to the Bioinformatics Pipeline.**—This checklist item addresses the requirement for laboratories to establish and follow a procedure for identifying and implementing updates to components of the bioinformatics pipeline. Next-generation sequencing bioinformatics pipelines often use multiple packages of open-source software with additional scripts and databases for managing content and aspects of analysis and reporting. Owing to the ongoing evolution of the field, laboratories must have a policy for monitoring updates, patch-releases, and other upgrades to the bioinformatics pipeline. This policy should also address when such updates will be implemented. For example, the laboratory may decide to do this at the time of the update release or at specified intervals (such as quarterly, biannually, or annually), depending on the nature and relevant urgency of the update. Since such updates require revalidation of some or all of the bioinformatics pipeline (see “Bioinformatics Process” section on validation/revalidation), the latter approach of incorporating updates at set intervals may be more efficient, although again this depends on the update. Finally, the laboratory should maintain records that clearly document regular monitoring and implementation of updates. It should be emphasized that this requirement mandates a policy, but it is up to the discretion of the laboratory director if and when a particular update should be incorporated into the laboratory’s bioinformatics pipeline.

## Data Storage

**The Laboratory Has a Policy Regarding the Storage of Input, Intermediate, and Final Data Files Generated by the Bioinformatics Pipeline.**—Laboratories must establish and follow a procedure for the storage of data files generated by the bioinformatics pipeline. Large data files are generated by NGS and the associated data analysis, including flow cell imaging files, sequence read files containing base calls and associated quality scores, other intermediate files generated after subsequent analysis steps, and variant text files. It is generally not practical to retain all such files for an extended period, so this checklist requirement mandates that the laboratory establish a policy for data storage that specifies data file retention times and which files will be retained after a final report has been generated. The NGS Work Group recommends that, if feasible, laboratories retain sequence files with corresponding quality scores (eg, FASTQ files<sup>47</sup>) or retain an archival format from which these files can be regenerated (eg, BAM files<sup>48</sup>). These formats will likely allow reanalysis at a later date, if indicated. For genome or other large-scale sequencing data, retention of FASTQ files or standard archival formats for long periods of time may be cost prohibitive with current storage technologies; however, newer compression formats provide one near-term solution.<sup>49,50</sup> How long to store such files is a more complicated decision that depends on numerous issues, including the size of the data set, laboratory storage capacity, medicolegal considerations, as well as other institutional, local, state, or national requirements for data storage. Finally, it should be emphasized that the laboratory’s policy for data storage and file retention times must be in accordance with local, state, and national requirements for storage of data.

## Version Traceability

**The Specific Version(s) of the Bioinformatics Pipeline Used to Generate NGS Data Files Are Traceable for Each Patient Report.**—The specific versions of each component and, where available, associated configurations (eg, command line parameters or other configuration items) of the bioinformatics pipeline used to generate NGS data should be traceable for each patient report. As noted before, the bioinformatics pipeline for analyzing NGS data, especially when based primarily on open-source software, is often composed of a combination of different software packages, scripts, and databases. The performance of a single software package or script and the composition of an internal or external database can significantly impact the overall performance of the bioinformatics pipeline. Consequently, it is important for the laboratory to be able to connect each patient report to the particular bioinformatics pipeline used to generate the report. For in-house-generated scripts and software packages, changes in the script or software should also be documented, but documentation of each component of the pipeline does not need to appear in the patient report. Rather, it is acceptable to refer to the pipeline as a whole, using a laboratory-specific designation (eg, NGS Pipeline v1.0.1). Laboratory-specific designations should be unique to a single combination of pipeline components and configurations. Therefore, any change to a different version of a software package, script, or internal or external database, or change to the configuration of any software, would require a new unique laboratory-specific designation and would require assay revalidation.

## Exception Log

**The Laboratory Maintains an Exception Log for Patient Cases Where Steps Used in the Bioinformatics Pipeline Deviate From Standard Operating Procedures.**—Deviations from the laboratory SOP during any step used in the bioinformatics pipeline are documented in an exception log file, including any alterations in software packages, script, version number, database, command line, or parameters.<sup>51</sup> Any failures arising during the bioinformatics process should also be recorded in the exception log and include documentation of the issues, the results of any investigations of these issues, any corrective actions taken, and pertinent communications, with sign-off by the laboratory director or designee. The exception log is also required to retain links to the patient reports, and the laboratory director may choose to communicate any clinically relevant SOP deviations to the ordering physician. Exception log documentation may also be incorporated into the monthly quality assurance report.

Deviation such as needing to rerun the analytic pipeline due to network, computer, or storage failure or memory issues, to run a particular step with different parameters or cutoffs than that used to validate the assay, must be documented along with the outcome and explanation. For example, a laboratory may need to alter settings on specific tools or components of its bioinformatics pipeline to adequately analyze particular regions or variants in a given patient case. The reason for the deviation should be described in the exception log, as well as the specific components of the deviation. Each deviation should be linked to the associated patient case and be reviewed by the appropriate laboratory director or designee(s). As warranted, the deviation, or aspects of it, may be included in the final report or in specific communication with the ordering physician.

Deviations related to bugs or failures in the bioinformatics pipelines also need to be recorded in the exception log. The bug, affected cases, and proposed corrective action must be approved, signed, and dated by the laboratory director or designee. Outright failures of the bioinformatics pipelines, which could result from hardware as well as software or operator error, should also be recorded in the exception log to document errors that may have occurred in analyzing individual patient cases.

Evidence of compliance for the exception log requires the ability to demonstrate appropriate documentation of review of the exception log by the laboratory director, demonstration that the laboratory records any issue arising during the bioinformatics procedure, and adequate documentation of subsequent corrective actions taken as a result of these reviews.

### NGS Data Transfer Confidentiality Policy

**The Laboratory Has a Policy and Procedures Describing Processes to Ensure That Internal and External Storage and Transfer of Sequencing Data Maintains Patient Confidentiality and Security.**—Next-generation sequencing generates significant amounts of data, particularly of gene sequences that, with other information such as name, date of birth, medical record number, and other components of protected health information, can potentially be used to identify individual patients. Laboratories must establish rigorous processes to ensure the protection and privacy of this information. Laboratories need robust policies regarding the transfer of genomic information to other health care entities and third-party vendors such as those providing cloud-based computing resources or reference laboratory services. Procedures to ensure confidentiality should include data encryption, secure data transfer, user authentication with controlled access to protected health information, and audit trails that track the transmission of data as well as the receiving entities and/or users. Laboratories should also follow standard requirements in the Health Insurance Portability and Accountability Act,<sup>52</sup> such as establishing business agreements with external vendors that include sufficient due diligence to verify that appropriate methods are used to ensure confidentiality in the sending and receipt of patient clinical and genomic data.

### Sequence Variants—Interpretation/Reporting

**Interpretation and Reporting of Sequence Variants Follows Professional Organizations' Recommendations and Guidelines.**—With the adoption of NGS technology, clinical laboratories are expanding their test menus from single gene testing to gene panels, and more recently, to exome and whole genome sequencing. It is evident that laboratories using NGS-based tests will come across a multitude of novel variants that have not been previously reported or classified as being causative of disease.

Currently, most laboratories report gene variants by using the Human Genome Variation Society nomenclature guidelines ([www.hgvs.org](http://www.hgvs.org); accessed January 6, 2014) and follow variant classification guidelines from the American College of Medical Genetics (ACMG)<sup>53</sup> for inherited diseases. The nomenclature guidelines were originally published by Antonarakis and den Dunnen in 2001,<sup>54</sup> but significant changes and additions have been made since then. Therefore, it is recommended that laboratories use the Web-based versions of these guidelines as they represent

the latest revisions. The ACMG guidelines for variant classification are currently under revision and the new version will include interpretation guidelines. It is recommended that the ACMG classification system for inherited diseases be used as reference to increase consistency in variant classification. Disease and gene-specific modification may be necessary but should be documented. For other clinical genomic testing (eg, tumor or pathogen diagnosis), the laboratory should use its best judgment to categorize variants and adopt guidelines as they emerge.

Laboratories must also be aware of the lack of consensus in how transcript versions are used for variant numbering, an area that creates confusion in the literature, and can do the same in clinical reports. An example of this is provided for multiple transcripts produced from the *MUTYH* gene, which has complicated the nomenclature used to describe mutations identified in the gene. The 2 major transcripts are *hMYHα1* (NM\_012222.2) and *hMYHα3* (NM\_001048171.1), encoding polypeptides of 546 and 535 amino acids, respectively. The *hMYHα3* transcript is 33 nucleotides shorter than the *hMYHα1* transcript and results from alternative splicing of exon 3, which eliminates 11 amino acids from the 5' end of exon 3 (GMIAECPGAPA). All other codons, and therefore amino acids, are identical between the 2 isoforms. Most literature uses the *hMYHα3* variant when naming mutations; however, some reports do use the full-length transcript (*hMYHα1*). When reporting results for *MUTYH* testing, or comparing reports from different laboratories, it is imperative to note which transcript has been used to name the alteration(s) found in the gene. The laboratories have called the 2 most common alterations in this gene p.Tyr165Cys and p.Gly382Asp using the reference transcript *hMYHα3* (NM\_001048171.1), and p.Tyr176Cys and p.Gly393Asp using the reference transcript *hMYHα1* (NM\_012222.2). Laboratories should have a mechanism to monitor for such changes and use great caution when any variant designation changes are made in clinical reports. It is therefore useful to provide the transcript accession number and version along with the protein syntax in clinical reports to help avoid confusion.

Accurate interpretation of the combination of sequence variants observed in a specimen is a critical component of clinical testing, as it integrates variants that are potentially disruptive for gene function with the patient's clinical phenotype in order to determine whether identified variants may be causative for the disease for which the patient is undergoing testing. During the past few years, various terminologies have been used in clinical reports to denote the consequence of sequence changes, including pathogenic, deleterious, and disease associated, with qualifiers such as possible, probably, and likely, or VUS and VOUS (variant of unknown clinical significance and variant of uncertain clinical significance, respectively). Standardized sequence variant guidelines have been recommended for inherited diseases, while those for tumor or infectious pathogen diagnosis are still under flux. For inherited diseases, the most commonly applied classification is divided into 5 categories: (1) pathogenic, (2) likely pathogenic, (3) uncertain clinical significance, (4) likely benign, and (5) benign.

Laboratories should be cautious when interpreting the potential clinical consequences of sequence changes and carefully consider evidence for disease causation, frequency in the general population (including race/ethnicity considerations), and functional studies. With the freely available exome and genome sequencing data from many large-scale

Useful Resources for Variant Assessment		
Utility	Software and Databases	URL <sup>a</sup>
Annotation of disease causative variants	GeneReviews <sup>65</sup>	<a href="http://www.ncbi.nlm.nih.gov/books/NBK1116">http://www.ncbi.nlm.nih.gov/books/NBK1116</a>
	Human Gene Mutation Database (HGMD) <sup>66</sup>	<a href="http://www.hgmd.cf.ac.uk/ac/index.php">http://www.hgmd.cf.ac.uk/ac/index.php</a>
	Locus Specific Mutation Databases <sup>67</sup>	<a href="http://grenada.lumc.nl/LSDB_list/lsdbs">http://grenada.lumc.nl/LSDB_list/lsdbs</a>
	Online Mendelian Inheritance in Man (OMIM) <sup>43</sup>	<a href="http://omim.org">http://omim.org</a>
Browsing tools	Catalogue of Somatic Mutations in Cancer (COSMIC) <sup>68</sup>	<a href="http://cancer.sanger.ac.uk/cancergenome/projects/cosmic">http://cancer.sanger.ac.uk/cancergenome/projects/cosmic</a>
	Ensembl <sup>69</sup>	<a href="http://www.ensembl.org/index.html">http://www.ensembl.org/index.html</a>
Evolutionary conservation	UCSC Genome Browser <sup>70</sup>	<a href="http://genome.ucsc.edu">http://genome.ucsc.edu</a>
	Genomic Evolutionary Rate Profiling (GERP) <sup>71</sup>	<a href="http://mendel.stanford.edu/SidowLab/downloads/gerp">http://mendel.stanford.edu/SidowLab/downloads/gerp</a>
Prediction of missense variants	NCBI <sup>72,73</sup>	<a href="http://www.ncbi.nlm.nih.gov">http://www.ncbi.nlm.nih.gov</a>
	PhastCons <sup>74</sup>	<a href="http://compugen.bscb.cornell.edu/phast">http://compugen.bscb.cornell.edu/phast</a>
	PhyloP <sup>75</sup>	<a href="http://compugen.bscb.cornell.edu/phast">http://compugen.bscb.cornell.edu/phast</a>
	Align GVGD <sup>76</sup>	<a href="http://agvgd.iarc.fr/agvgd_input.php">http://agvgd.iarc.fr/agvgd_input.php</a>
	CONsensus DEleteriousness score of missense SNVs (CONDEL) <sup>77</sup>	<a href="http://bg.upf.edu/condel/analysis">http://bg.upf.edu/condel/analysis</a>
	MutationAssessor <sup>78</sup>	<a href="http://mutationassessor.org">http://mutationassessor.org</a>
Prediction of splice-site variants	MutationTaster <sup>79</sup>	<a href="http://www.mutationtaster.org">http://www.mutationtaster.org</a>
	Polyphen2 <sup>80</sup>	<a href="http://genetics.bwh.harvard.edu/pph2">http://genetics.bwh.harvard.edu/pph2</a>
	Sorting Intolerant From Tolerant (SIFT) <sup>81,82</sup>	<a href="http://sift.jcvi.org">http://sift.jcvi.org</a>
	GeneSplicer <sup>83</sup>	<a href="http://ccb.jhu.edu/software/genesplicer">http://ccb.jhu.edu/software/genesplicer</a>
	Human Splicing Finder <sup>84</sup>	<a href="http://www.umd.be/HSF/">http://www.umd.be/HSF/</a>
Variant databases	MaxEntScan <sup>85</sup>	<a href="http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html">http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html</a>
	NNSplice <sup>86</sup>	<a href="http://www.fruitfly.org/seq_tools/splice.html">http://www.fruitfly.org/seq_tools/splice.html</a>
	ClinVar <sup>87</sup>	<a href="http://www.ncbi.nlm.nih.gov/clinvar">http://www.ncbi.nlm.nih.gov/clinvar</a>
	dbSNP <sup>88</sup>	<a href="http://www.ncbi.nlm.nih.gov/projects/SNP">http://www.ncbi.nlm.nih.gov/projects/SNP</a>
	Human Genome Variation Society (HGVS) <sup>89</sup>	<a href="http://www.hgvs.org">http://www.hgvs.org</a>
	NHLBI Exome Variant Server <sup>90</sup>	<a href="http://evs.gs.washington.edu/EVS">http://evs.gs.washington.edu/EVS</a>
	1000 Genomes Project <sup>91</sup>	<a href="http://www.1000genomes.org">http://www.1000genomes.org</a>

Abbreviations: NCBI, National Center for Biotechnology Information; NHLBI, National Heart, Lung, and Blood Institute; SNP, single nucleotide polymorphism; SNVs, single nucleotide variants; UCSC, University of California Santa Cruz.

<sup>a</sup> All accessed on June 19, 2014.

projects (eg, Exome Variant Server, 1000 Genomes Project), laboratories should make use of relevant databases and computational tools<sup>39,55,56</sup> (Table) when interpreting the effects of sequence changes. For large-scale tests such as exome and genome sequencing, it is also critical to assess the evidence implicating a gene in disease, understand the types of variants that have been implicated in disease as well as the mode of inheritance if known. For example, a novel loss-of-function variant in a gene with no established role in disease or a well-characterized variant spectrum should not be assumed to cause disease, despite the severity of the predicted impact to the protein.

### Reporting of Incidental Genetic Findings

**The Laboratory Has a Policy for Reporting Incidental Genetic Findings Unrelated to the Clinical Purpose for Testing.**—Clinically significant genetic findings that are unrelated to the phenotype for testing can occur when performing single gene, gene panel, exome, and whole genome sequencing. Limiting sequence analysis to a panel of genes that are relevant to the diagnosis of a particular disease state (either with targeted sequencing or targeted bioinformatics analysis) may limit, but not eliminate, the potential for incidental findings.<sup>57–64</sup> This may include identification of variants relevant to autosomal dominant disease, carrier status for recessive diseases, predisposition to adult-onset dominant conditions (including cancer and neurodegenerative conditions), and drug response alleles commonly known as *pharmacogenetic markers*. Laboratories

embarking on use of NGS for clinical testing should be aware of the potential for finding incidental, clinically significant results and should have a policy in place for whether and how these results will be reported for those assays where such incidental findings are expected (eg, exome). The recently published ACMG recommendations for reporting medically actionable incidental findings include a minimum gene list for which, if a known mutation is found, it should be reported.<sup>57</sup> Laboratories may choose to follow the ACMG recommendations but are not expected to necessarily only report findings in these genes. Laboratories may also develop their own policies regarding return of incidental results. If the laboratory's policy is not to report incidental findings or to limit reporting to a subset of variants related to a particular disease state, this should be clearly stated in the laboratory report for assays where incidental findings are expected.

Ethical considerations must also be taken into account when deciding whether to reveal certain genetic information to patients. The level of risk associated with disclosing incidental findings depends on the severity of the disease, clinical actionability, and other risk-benefit indicators. For example, common disease risk alleles, such as for type 2 diabetes or cardiovascular disease, which have a small effect size (low relative risks), or pharmacogenetic risk information, may have different severity of consequence, compared to genetic information indicating a predisposition to cancer or a Mendelian disorder that may or may not be medically treatable. All of these facets must be considered before

returning results to patients. Laboratories performing large-scale genomic sequencing analysis for clinical testing should be aware of efforts to study the medical and ethical implications of returning incidental results of NGS and consider these when developing their reporting policies.

### NGS Test Referral Policy

**The Laboratory Has a Policy for Selection of Reference Laboratories and Other Service Providers for NGS Test Referral. Referral May Include the Entire NGS Test Process or Only the Wet Bench or Bioinformatics Processes.**—The complexity of NGS clinical testing with challenging data analysis has led to some interesting transformation of the traditional laboratory testing model. Although most clinical laboratories are conducting the full NGS testing process in-house, there are a growing number of laboratories that have started outsourcing parts of the testing workflow to external facilities. For example, CLIA laboratories that do not have extensive data analysis capabilities are sending out the bioinformatics data analysis to external facilities. On the other hand, bioinformatic vendor companies, or institutions that are entering the clinical testing arena and have sophisticated data analysis capabilities but do not have a CLIA laboratory facility to process DNA for sequence generation, are outsourcing the wet bench workflow to external laboratories.

With this emerging trend in fragmentation of the clinical workflow, the NGS Work Group decided to include a new requirement for the 2014 version of the NGS Checklist concerning test referral policy. The CAP Laboratory General Checklist requirement 41350 states that the laboratory director or designee is responsible for the selection of an external reference laboratory or other service provider. It is expected that the laboratory director or designee ensure the quality of performance of the external NGS wet bench and/or bioinformatics service provider. Some of the specific aspects that the laboratory director needs to consider in selecting external reference laboratories or service providers are to ensure that (1) the turnaround times are acceptable for the clinical needs for which testing is being done; (2) the external laboratory providing the analytic wet bench information, (ie, sequence generation) is a CLIA-certified laboratory or a laboratory meeting the selection criteria as per CAP requirements; and (3) the quality of the results from the external bioinformatics service provider is verified to be accurate and of high standards. For evidence of compliance, copies of valid CLIA certificates from CLIA-certified external reference laboratories are required for those who outsource the wet bench sequencing workflow. Copies of in-house validation of non-CLIA-certified entities providing bioinformatics analysis are required for those entities that outsource the bioinformatics workflow.

### COMMENT

The translation of NGS from basic to clinical research and adoption for clinical diagnostics has occurred over a relatively short period of time. A growing number of clinical laboratories are implementing NGS-based diagnostic assays, mostly in the form of multigene panels, although an increasing number of laboratories are performing exome and genome sequencing. CAP identified that the adoption of NGS by clinical laboratories required the development of accreditation requirements specific to NGS. This report highlights the content of the accreditation requirements that were developed by the CAP NGS Work Group in an effort

to respond to the clinical adoption of NGS and to articulate the perspectives of the NGS Work Group in developing the requirements. The NGS field continues to evolve at a rapid pace and this evolution reflects continuing improvements in NGS instrumentation, chemistries, and analytic tools and software. The work group took the perspective that the NGS accreditation requirements should strike a balance that would ensure patient safety and also foster the responsible adoption of NGS in clinical laboratories, using fundamental diagnostic laboratory principles. These principles include documentation of process steps, demonstration that laboratory procedures have been well characterized through proper validation, and institution of quality management programs. Given the reality of this rapidly changing technology, and because NGS-based tests can be varied depending on type of tests and scale of analysis, the NGS Work Group developed general requirements for a variety of NGS clinical testing scenarios to allow laboratories flexibility and latitude in individual approaches to meeting the requirements, while at the same time providing much needed regulatory standards. Some examples used in this article regarding nomenclature, interpretation, and incidental findings highlight and are most applicable to NGS for inherited disease testing. Practically, it was more straightforward to develop laboratory standards for these topics as they have been previously addressed in accreditation requirements for single gene tests in inherited disease. Notably, guidance documents on topics of nomenclature, interpretation of variants, and incidental findings have been previously published. The work group elected not to prematurely introduce requirements on NGS topics that would benefit from further technology evolution and/or deliberation and consensus building at the professional society level for specific clinical disciplines. An example is that guidelines for interpretation of somatic variants need to be developed before introducing a correlate accreditation requirement. Whereas this topic is not unique to NGS-based molecular oncology testing, NGS multigene panel and exome testing have heightened the need for somatic variant interpretation guidelines. Additional, future topics for accreditation requirements for NGS-based molecular oncology could include the role of preanalytic sample assessment and processing of oncology samples and their influence on NGS testing results. Preanalytic metrics, such as formalin-fixed, paraffin-embedded variables influencing NGS results, have not been established in the field and therefore this would be another example where the committee felt it would be premature to include specific requirements. Further accreditation requirements could address the quantitative aspects of both molecular oncology and infectious diseases NGS testing. In essence, there are an emerging number of discipline-specific topics for new accreditation requirements that will be introduced as the field of NGS testing matures and consensus on practice is built through professional experience.

The NGS Work Group acknowledged that accreditation requirements for NGS would need to be revisited and revised as part of an ongoing process as the field of NGS-based diagnostics evolved and matured. With the publication of the accreditation requirements, CAP has subsequently fielded many questions from laboratories as they implement NGS. Specifically, laboratories are periodically seeking clarification with respect to the requirements, and the questions being posed are providing the work group with feedback that is guiding our discussions focused on

improving and clarifying these requirements. In addition, the NGS work group has recognized that the initial requirements could not cover all of the NGS applications that clinical laboratories might pursue. Operationally, the NGS work group meets throughout the year via teleconferences and face-to-face meetings. As a part of these activities, discussions focused on revising and expanding accreditation requirements occur. Within the CAP accreditation requirements revision and publication cycle, the NGS work group has the ability to formally submit revised and expanded NGS requirements. The current requirements reflect those that have undergone 1 formal cycle of revisions, and as of this writing, additional revisions and expansions are underway. Again, feedback and experience from the field will be incorporated into the work group's current and future deliberations, resulting in further explanation and expansion of these requirements in future versions of the NGS-related accreditation requirements.

## References

- Maxam AM, Gilbert W. A new method for sequencing DNA. *Proc Natl Acad Sci U S A*. 1977;74(2):560–564.
- Sanger F, Coulson AR. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol*. 1975;94(3):441–448.
- Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*. 1977;74(12):5463–5467.
- Bentley DR, Balasubramanian S, Swerdlow HP, et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*. 2008;456(7218):53–59.
- Mardis ER. The impact of next-generation sequencing technology on genetics. *Trends Genet*. 2008;24(3):133–141.
- Mardis ER. Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet*. 2008;9:387–402.
- Margulies M, Egholm M, Altman WE, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*. 2005;437(7057):376–380.
- Metzker ML. Sequencing technologies: the next generation. *Nat Rev Genet*. 2010;11(1):31–46.
- Rothberg JM, Hinz W, Rearick TM, et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature*. 2011;475(7356):348–352.
- Schuster SC. Next-generation sequencing transforms today's biology. *Nat Methods*. 2008;5(1):16–18.
- Shendure J, Ji H. Next-generation DNA sequencing. *Nat Biotechnol*. 2008;26(10):1135–1145.
- Valouev A, Ichikawa J, Tonthat T, et al. A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Res*. 2008;18(7):1051–1063.
- Hutchison CA III. DNA sequencing: bench to bedside and beyond. *Nucleic Acids Res*. 2007;35(18):6227–6237.
- Metzker ML. Emerging technologies in DNA sequencing. *Genome Res*. 2005;15(12):1767–1776.
- Mardis ER. A decade's perspective on DNA sequencing technology. *Nature*. 2011;470(7333):198–203.
- Schrijver I, Aziz N, Jennings LJ, Richards CS, Voelkerding KV, Weck KE. Methods-based proficiency testing in molecular genetic pathology. *J Mol Diagn*. 2014;16(3):283–287.
- Gargis AS, Kalman L, Berry MW, et al. Assuring the quality of next-generation sequencing in clinical laboratory practice. *Nat Biotechnol*. 2012;30(11):1033–1036.
- Jennings L, Van Deerlin VM, Gulley ML; College of American Pathologists Molecular Pathology Resource Committee. Recommended principles and practices for validating clinical molecular pathology tests. *Arch Pathol Lab Med*. 2009;133(5):743–755.
- Mattocks CJ, Morris MA, Matthijs G, et al. A standardized framework for the validation and verification of clinical molecular genetic tests. *Eur J Hum Genet*. 2010;18(12):1276–1288.
- Rehm HL, Bale SJ, Bayrak-Toydemir P, et al. ACMG clinical laboratory standards for next-generation sequencing. *Genet Med*. 2013;15(9):733–747.
- Schrijver I, Aziz N, Farkas DH, et al. Opportunities and challenges associated with clinical diagnostic genome sequencing: a report of the Association for Molecular Pathology. *J Mol Diagn*. 2012;14(6):525–540.
- Chin EL, da Silva C, Hegde M. Assessment of clinical analytical sensitivity and specificity of next-generation sequencing for detection of simple and complex mutations. *BMC Genet*. 2013;14:6.
- Valencia CA, Rhodenizer D, Bhide S, et al. Assessment of target enrichment platforms using massively parallel sequencing for the mutation detection for congenital muscular dystrophy. *J Mol Diagn*. 2012;14(3):233–246.
- Valencia CA, Ankala A, Rhodenizer D, et al. Comprehensive mutation analysis for congenital muscular dystrophy: a clinical PCR-based enrichment and next-generation sequencing panel. *PLoS One*. 2013;8(1):e53083.
- Cui H, Li F, Chen D, et al. Comprehensive next-generation sequence analyses of the entire mitochondrial genome reveal new insights into the molecular diagnosis of mitochondrial DNA disorders. *Genet Med*. 2013;15(5):388–394.
- Hagemann IS, Cottrell CE, Lockwood CM. Design of targeted, capture-based, next generation sequencing tests for precision cancer therapy. *Cancer Genet*. 2013;206(12):420–431.
- Ankala A, Hegde M. Genomic technologies and the new era of genomic medicine. *J Mol Diagn*. 2014;16(1):7–10.
- Williams ES, Hegde M. Implementing genomic medicine in pathology. *Adv Anat Pathol*. 2013;20(4):238–244.
- Teekakirikul P, Kelly MA, Rehm HL, Lakdawala NK, Funke BH. Inherited cardiomyopathies: molecular genetics and clinical genetic testing in the post-genomic era. *J Mol Diagn*. 2013;15(2):158–170.
- Orlando MA, Rhodenizer D, da Silva C, et al. Molecular diagnostic testing for congenital disorders of glycosylation (CDG): detection rate for single gene testing and next generation sequencing panel testing. *Mol Genet Metab*. 2013;110(1–2):78–85.
- Wong LJ. Next generation molecular diagnosis of mitochondrial disorders. *Mitochondrion*. 2013;13(4):379–387.
- Cottrell CE, Al-Kateb H, Bredemeyer AJ, et al. Validation of a next-generation sequencing assay for clinical molecular oncology. *J Mol Diagn*. 2014;16(1):89–105.
- Zook JM, Chapman B, Wang J, et al. Integrating human sequence data sets provides a resource of benchmark SNP and indel genotype calls. *Nat Biotechnol*. 2014;32(3):246–251.
- National Center for Biotechnology Information. GeT-RM Homo sapiens: GRCh37.p13(CGF\_000001405.25). US National Library of Medicine. <http://www.ncbi.nlm.nih.gov/variation/tools/get-rm/>. Accessed June 3, 2014.
- Code of Federal Regulations. Laboratory requirements, 42 CFR 493.1249, 1289, and 1299. 2014.
- College of American Pathologists. *Standards for Laboratory Accreditation*. Northfield, Illinois: College of American Pathologists; revised August 2010.
- Sikkema-Raddatz B, Johansson LF, de Boer EN, et al. Targeted next-generation sequencing can replace Sanger sequencing in clinical diagnostics. *Hum Mutat*. 2013;34(7):1035–1042.
- Strom SP, Lee H, Das K, et al. Assessing the necessity of confirmatory testing for exome-sequencing results in a clinical molecular diagnostic laboratory [published online ahead of print January 9, 2014]. *Genet Med*. doi:10.1038/gim.2013.1183.
- Moorthie S, Hall A, Wright CF. Informatics and clinical genome sequencing: opening the black box. *Genet Med*. 2013;15(3):165–171.
- Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat Rev Genet*. 2010;11(6):415–425.
- Gibson G. Rare and common variants: twenty arguments. *Nat Rev Genet*. 2011;12(2):135–145.
- Stenson PD, Mort M, Ball EV, Shaw K, Phillips A, Cooper DN. The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. *Hum Genet*. 2014;133(1):1–9.
- Amladi S. Online Mendelian Inheritance in Man 'OMIM'. *Indian J Dermatol Venereol Leprol*. 2003;69(6):423–424.
- Coonrod EM, Durtschi JD, Margraf RL, Voelkerding KV. Developing genome and exome sequencing for candidate gene identification in inherited disorders: an integrated technical and bioinformatics approach. *Arch Pathol Lab Med*. 2013;137(3):415–433.
- Gullapalli RR, Desai KV, Santana-Santos L, Kant JA, Becich MJ. Next generation sequencing in clinical medicine: challenges and lessons for pathology and biomedical informatics. *J Pathol Inform*. 2012;3:40.
- Davis MP, van Dongen S, Abreu-Goodger C, Bartonicek N, Enright AJ. Kraken: a set of tools for quality control and analysis of high-throughput sequence data. *Methods*. 2013;63(1):41–49.
- Cock PJ, Fields CJ, Goto N, Heuer ML, Rice PM. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Res*. 2010;38(6):1767–1771.
- Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–2079.
- Cochrane G, Alako B, Amid C, et al. Facing growth in the European Nucleotide Archive. *Nucleic Acids Res*. 2013;41(database issue):D30–D35.
- Hsi-Yang Fritz M, Leinonen R, Cochrane G, Birney E. Efficient storage of high throughput DNA sequencing data using reference-based compression. *Genome Res*. 2011;21(5):734–740.
- Code of Federal Regulations. Laboratory requirements, 42 CFR 493.1254 and 1281. 2014.
- Code of Federal Regulations. Public welfare, 45 CFR 160 and 164. 2014.
- Richards CS, Bale S, Bellissimo DB, et al. ACMG recommendations for standards for interpretation and reporting of sequence variations: revisions 2007. *Genet Med*. 2008;10(4):294–300.
- den Dunnen JT, Antonarakis SE. Nomenclature for the description of human sequence variations. *Hum Genet*. 2001;109(1):121–124.
- Johnston JJ, Biesecker LG. Databases of genomic variation and phenotypes: existing resources and future needs. *Hum Mol Genet*. 2013;22(R1):R27–R31.
- Duzkale H, Shen J, McLaughlin H, et al. A systematic approach to assessing the clinical significance of genetic variants. *Clin Genet*. 2013;84(5):453–463.

57. Green RC, Berg JS, Grody WW, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med*. 2013;15(7):565–574.
58. Christenhusz GM, Devriendt K, Dierickx K. Disclosing incidental findings in genetics contexts: a review of the empirical ethical research. *Eur J Med Genet*. 2013;56(10):529–540.
59. Bennette CS, Trinidad SB, Fullerton SM, et al. Return of incidental findings in genomic medicine: measuring what patients value—development of an instrument to measure preferences for information from next-generation testing (IMPRINT). *Genet Med*. 2013;15(11):873–881.
60. Huang JT, Heckenlively JR, Jayasundera KT, Branham KE. The Ophthalmic Experience: Unanticipated Primary Findings in the Era of Next Generation Sequencing [published online ahead of print January 8, 2014]. *J Genet Couns*. doi:http://dx.doi.org/10.1007/s10897-013-9679-y.
61. Krier JB, Green RC. Management of incidental findings in clinical genomic sequencing. *Curr Protoc Hum Genet*. 2013;Chapter 9:Unit 9.23. doi:10.1002/0471142905.hg0923s77.
62. Lupski JR, Gonzaga-Jauregui C, Yang Y, et al. Exome sequencing resolves apparent incidental findings and reveals further complexity of SH3TC2 variant alleles causing Charcot-Marie-Tooth neuropathy. *Genome Med*. 2013;5(6):57.
63. Rigter T, Henneman L, Kristoffersson U, et al. Reflecting on earlier experiences with unsolicited findings: points to consider for next-generation sequencing and informed consent in diagnostics. *Hum Mutat*. 2013;34(10):1322–1328.
64. Rigter T, van Aart CJ, Elting MW, Waisfisz Q, Cornel MC, Henneman L. Informed consent for exome sequencing in diagnostics: exploring first experiences and views of professionals and patients. *Clin Genet*. 2014;85(5):417–422.
65. National Center for Biotechnology Information. GeneReviews. <http://www.ncbi.nlm.nih.gov/books/NBK1116/>. Seattle, WA: University of Washington; 1993–2014. Accessed June 19, 2014.
66. Stenson PD, Ball EV, Mort M, Phillips AD, Shaw K, Cooper DN. The Human Gene Mutation Database (HGMD) and its exploitation in the fields of personalized genomics and molecular evolution. *Curr Protoc Bioinformatics*. 2012;Chapter 1:Unit 1.13. doi:10.1002/0471250953.bi0113s39.
67. Claustres M, Horaitis O, Vanevski M, Cotton RG. Time for a unified system of mutation description and reporting: a review of locus-specific mutation databases. *Genome Res*. 2002;12(5):680–688.
68. Forbes SA, Bindal N, Bamford S, et al. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res*. 2011;39(database issue):D945–D950.
69. Flicek P, Ahmed I, Amode MR, et al. Ensembl 2013. *Nucleic Acids Res*. 2013;41(database issue):D48–D55.
70. Kuhn RM, Haussler D, Kent WJ. The UCSC genome browser and associated tools. *Brief Bioinform*. 2013;14(2):144–161.
71. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput Biol*. 2010;6(12):e1001025.
72. Benson D, Boguski M, Lipman D, Ostell J. The National Center for Biotechnology Information. *Genomics*. 1990;6(2):389–391.
73. Dooley EE. National Center for Biotechnology Information. *Environ Health Perspect*. 2004;112(12):A674.
74. Siepel A, Bejerano G, Pedersen JS, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res*. 2005;15(8):1034–1050.
75. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res*. 2010;20(1):110–121.
76. Tavtigian SV, Deffenbaugh AM, Yin L, et al. Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *J Med Genet*. 2006;43(4):295–305.
77. Gonzalez-Perez A, Lopez-Bigas N. Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel. *Am J Hum Genet*. 2011;88(4):440–449.
78. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res*. 2011;39(17):e118.
79. Schwarz JM, Rodelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods*. 2010;7(8):575–576.
80. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7(4):248–249.
81. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*. 2009;4(7):1073–1081.
82. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res*. 2003;31(13):3812–3814.
83. Pertea M, Lin X, Salzberg SL. GeneSplicer: a new computational method for splice site prediction. *Nucleic Acids Res*. 2001;29(5):1185–1190.
84. Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res*. 2009;37(9):e67.
85. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol*. 2004;11(2–3):377–394.
86. Reese MG, Eeckman FH, Kulp D, Haussler D. Improved splice site detection in Genie. *J Comput Biol*. 1997;4(3):311–323.
87. National Center for Biotechnology Information. ClinVar. US National Library of Medicine. <http://www.ncbi.nlm.nih.gov/clinvar/>. Accessed June 9, 2014.
88. Liu X, Jian X, Boerwinkle E. dbSNFP v2.0: a database of human non-synonymous SNVs and their functional predictions and annotations. *Hum Mutat*. 2013;34(9):E2393–E2402.
89. Taschner PE, den Dunnen JT. Describing structural changes by extending HGVS sequence variation nomenclature. *Hum Mutat*. 2011;32(5):507–511.
90. NHLBI Exome Sequencing Project (ESP). Exome Variant Server. <http://evs.gs.washington.edu/EVS/>. Accessed June 9, 2014.
91. 1000 Genomes Project Consortium, Abecasis GR, Auton A, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491(7422):56–65.

## Submissions Now Accepted for CAP '15 Abstract Program

Abstract and case study submissions are now being accepted for the College of American Pathologists (CAP) 2015 meeting, which will be held October 4th through the 7th in Nashville, Tennessee. Submissions for the CAP '15 Abstract Program will be accepted from:

**Monday, February 9, 2015 through 6 p.m. CT Friday, April 10, 2015**

Accepted submissions will be published as a Web-only supplement to the October 2015 issue of the *Archives of Pathology & Laboratory Medicine* and will be posted on the *Archives* Web site. Visit the CAP '15 Web site at [www.cap.org/cap15](http://www.cap.org/cap15) to access the abstract submission site and additional abstract program information as it becomes available.