Molecular Pathology Informatics

Somak Roy, MD

KEYWORDS

• Molecular informatics • Next-generation sequencing • Clinical informatics
• Bioinformatics

OVERVIEW

Molecular pathology has emerged as a major arsenal for promoting and sustaining personalized health care in modern medicine. In contrast to other ancillary diagnostic tools, molecular pathology delivers a wide spectrum of theranostic information starting from as little as a minute fragment of tissue or a simple blood draw. The results of molecular testing are significantly more complex than a simple numerical value or a binary outcome (yes/no). The post–Human Genome Project era has witnessed the most significant developments in genomics that has impacted the way clinical medicine is practiced today. Personalized medicine, which enables patient-specific therapeutic and disease management protocols, is being increasingly boosted and supported by the results of molecular testing performed on tissue biopsies and resections. In addition, it also serves as a valuable diagnostic tool in morphologically challenging cases. Genomic profiling of certain cancers, such as lung, colon, and melanoma, are now considered to be the standard of care for guiding appropriate treatment protocols. Very recently, a global molecular testing guideline was published for molecular testing of lung cancers by the collaborative efforts of College of American Pathologists, International Association for the Study of Lung Cancer, and Association for Molecular Pathology.[1] The Cancer Genome Atlas project (http://cancergenome.nih.gov/) over the past few years has released a wealth of genomic profiling information about major tumor types that had significantly fueled the discovery of potential actionable biomarkers for routine clinical use.[2–8]

A wide array of assay methodologies and platforms are used in laboratories for performing molecular testing, such as Sanger sequencing, real-time polymerase chain reaction (PCR), allele-specific PCR, DNA and RNA in situ hybridization (ISH), multiplex ligation-dependent probe amplification, microarray technology, and, more recently, next-generation sequencing (NGS) technology. The choice of platform depends on the question(s) answered by the clinical test and its appropriate applicability to patient care.

This article originally appeared in Surgical Pathology Clinics, Volume 8, Issue 2, June 2015.
Department of Pathology, Molecular and Genomic Pathology, University of Pittsburgh Medical Center, 3477 Euler way, Pittsburgh, PA 15213, USA
E-mail address: roys@upmc.edu

Clin Lab Med 36 (2016) 57–66
http://dx.doi.org/10.1016/j.cll.2015.09.007 **labmed.theclinics.com**
0272-2712/16/$ – see front matter

MOLECULAR INFORMATICS

Given the increasing popularity and relevance of high-complexity molecular testing in surgical pathology, where does informatics fit into the equation? To answer this question, it is important to understand the major operational domains of informatics, namely bioinformatics (BI) and clinical informatics (CI). Bioinformatics is a dynamic and rapidly evolving multidisciplinary field of informatics that has emerged out of the application of computational, mathematical, and statistical methods to investigate biological phenomena. BI dates back to as early as the 1960s, when fundamental discoveries in molecular biology were made. With more recent data deluge and development of massive repositories of biological and genomic data, the symbiotic relationship between molecular biology and BI has fueled mutual exponential growth.[9–11] Large and complex data (so-called big data), data-driven analytics, statistical modeling, development of new algorithms, and discovering meaningful and relevant information from raw biological data is inherent to BI.[10] In contrast, CI is a more established discipline that has developed and matured over a longer period of time with rigorous testing and validation in a clinical environment. CI implementations in health care systems are typically on an enterprise scale, using industry standard software and hardware with principal focus on secure and reliable data transmission, storage, retrieval, and interoperability between information systems and wide variety of medical devices rather than data-driven analytics and discovery. One of the key differences between BI and CI is application development environment and life cycle, which gives rise to several other differences between the two disciplines. Application development in BI is centered on a fundamental genomic or biological phenomenon that is under investigation in a given research project. The pace of design, testing, implementation, and troubleshooting (software patches) is typically coupled with the progress of the research project. If a BI application is successful and is usable across a wider user community, further improvement and maintenance is supported by open-source software development. The overall life cycle is fast paced but with variable support, documentation, and version control. CI applications (such as electronic medical record [EMR], laboratory information system [LIS]), in contrast, have a much tighter and longer development and implementation life cycle. Vendors usually provide formal application support, tight version control, and extensive documentation. Updates for feature enhancements and patches for troubleshooting are relatively infrequent and often expensive.

This conceptual distinction has existed for several years in clinical laboratories with minimal interactions between the two disciplines. With conventional molecular testing instrumentation, vendors were able to mask the complexity of underlying BI with a point-and-click user interface. However, with the speedy introduction of NGS testing in clinical laboratories, molecular pathology has finally broken this barrier, forming a unique collaborative environment for development of informatics solutions that bridge the gap between BI and CI. The term molecular informatics (MI) is often loosely used to refer to various informatics principles and operations in a molecular laboratory. The rapid advance in this domain during a short period of time has manifested its presence in the practice of molecular pathology.

Conceptually, MI forms the core conduit for generating, processing, porting, and management of the molecular data points from each step of every assay performed in the molecular laboratory. With increasing sophistication of the assay platform and workflow elements, the magnitude and complexity of the generated information increases exponentially. Surprisingly, MI has largely been an unrealized and neglected part of laboratory workflow, until the introduction of NGS technology in clinical laboratories.