The Journal of Molecular Diagnostics, Vol. 18, No. 3, May 2016

ELSEVIER

the Journal of Molecular Diagnostics
jmd.amjpathol.org

Technical Validation of a Next-Generation Sequencing Assay for Detecting Actionable Mutations in Patients with Gastrointestinal Cancer

Sophie R. Wang,* Simeen Malik,† Iain B. Tan,*†‡§ Yang Sun Chan,* Qiangze Hoi,* Jack L. Ow,* Cassandra Z. He,* Cindy E. Ching,* Dianne Y.S. Poh,* Hui Maan Seah,* Katie H.T. Cheung,* Dharuman Perumal,* Arun G. Devasia,* Lu Pan,* Shimin Ang,* Seow Eng Lee,* Rachel Ten,‡ Clarinda Chua,‡ Daniel S.W. Tan,*‡¶ James Z.Z. Qu,* Yasmin M. Bylstra,*‖ Lionel Lim,* Alexander Lezhava,* Pauline C. Ng,* Christopher W. Wong,* Tony Lim,*‖ and Patrick Tan*†**††

From the Genome Institute of Singapore, Singapore; the Cancer & Stem Cell Biology Program,† Duke-National University of Singapore Graduate Medical School, Singapore; the Department of Medical Oncology,‡ and the Divisions of Medical Sciences¶ and Cellular and Molecular Research,†† National Cancer Centre Singapore, Singapore; the Graduate School of Integrative Sciences and Engineering,§ and the Cancer Science Institute Singapore,** National University of Singapore, Singapore; and the Department of Pathology,‖ Singapore General Hospital, Singapore, Singapore

Accepted for publication January 8, 2016.

Address correspondence to Tony Lim, M.B.B.S., F.R.C.Path., F.R.C.P.A., Department of Pathology, Singapore General Hospital, Singapore, Singapore 169856; or Patrick Tan, M.D., Ph.D., Genome Institute of Singapore, Singapore, Singapore 138672. E-mail: lim.kiat.hon@sgh.com.sg or tanbop@gis.a-star.edu.sg.

Targeted next-generation sequencing is becoming increasingly common as a clinical diagnostic and prognostic test for patient- and tumor-specific genetic profiles as well as to optimally select targeted therapies. Here, we describe a custom-developed, next-generation sequencing test for detecting single-nucleotide variants (SNVs) and short insertions and deletions (indels) in 93 genes related to gastrointestinal cancer from routine formalin-fixed, paraffin-embedded clinical specimens. We implemented a validation strategy, based on the College of American Pathologists requirements, using reference DNA mixtures from cell lines with known genetic variants, which model a broad range of allele frequencies. Test sensitivity achieved >99% for both SNVs and indels, with allele frequencies >10%, with high specificity (97.4% for SNVs and 93.6% for indels). We further confirmed test accuracies using primary formalin-fixed, paraffin-embedded colorectal cancer specimens characterized by alternative and conventional clinical diagnostic technologies. Robust performance was observed on the formalin-fixed, paraffin-embedded specimens: sensitivity was 97.2% and specificity was 99.2%. We also observed high intrarun and inter-run reproducibility, as well as a low cross-contamination rate. Overall assessment using cell line samples and formalin-fixed, paraffin-embedded samples showed that our custom next-generation sequencing assay has consistent detection sensitivity down to 10% variant frequency. *(J Mol Diagn 2016, 18: 416–424; http://dx.doi.org/10.1016/j.jmoldx.2016.01.006)*

Precision medicine exploits individual genetic differences in patients to select the most appropriate therapies for each patient. In the context of cancer, precision medicine success stories include the targeted treatment of the following: i) epidermal growth factor receptor 2–positive breast cancer with the anti–human epidermal growth factor receptor 2 monoclonal antibody trastuzumab[1]; ii) epidermal growth factor receptor (*EGFR*)-mutated non–small-cell lung cancer with gefitinib and erlotinib[2]; iii) *BRAF* V600E-positive metastatic malignant melanoma with vemurafenib; and iv) anaplastic lymphoma kinase (*ALK*)-rearranged non–small-cell lung cancer with crizotinib.[3] In addition to predicting and improving clinical response rates,[4] precision medicine also holds the promise of reducing treatment toxicities[5–7] and optimizing clinical trials by focused recruitment of patients with specific genetic mutations.[8]

Because of the importance of selecting the right therapies for each patient, there has been increasing demand for

Supported by the Personalized OMIC Lattice for Advanced Research and Improving Stratification (POLARIS) program (Biomedical Research Council Strategic Positioning Fund, 2012/001) and Genome Institute of Singapore, Agency for Science Technology and Research.

S.R.W. and S.M. contributed equally to this work.

Disclosures: None declared.

http://dx.doi.org/10.1016/j.jmoldx.2016.01.006

robust companion diagnostic assays to deliver accurate genetic information of a patient's tumor sample. Traditionally, such companion diagnostic assays largely have relied on single-gene technologies such as Sanger sequencing, mass spectrometric genotyping, and fluorescence *in situ* hybridization.[9–11] One companion diagnostic routinely used in clinical practice is the therascreen *KRAS* RGQ PCR Kit (Qiagen, Manchester, UK), which detects seven mutations in the *KRAS* gene in colorectal cancer. The presence of these *KRAS* mutations renders certain drugs, such as panitumumab (Amgen, Thousand Oaks, CA) and cetuximab (Merck Serono, Rockland, MA), ineffective for colorectal cancer treatment.[12,13] At present, however, most companion diagnostic assays can assess only a limited number of oncogenic markers, and because of technical limitations cannot be scaled to include the growing number and variety of clinically informative genomic alterations, shown through efforts such as The Cancer Genome Atlas and the International Cancer Genome Consortium.[14–18]

Recently, the availability of next-generation sequencing (NGS) has enabled comprehensive mutational profiling of large numbers of genes in a single assay that promises to radically decrease the associated time, cost, and amount of precious biopsy tumor DNA needed for multigene profiling.[19,20] For example, we recently described an NGS-based cancer genome profiling assay composed of 750 cancer-associated genes for exploring patterns of colorectal cancer progression.[21] However, implementation of a research-grade assay as a routine clinical-grade test still faces many challenges. These include the assay's ability to robustly handle the following: i) fragmented and chemically degraded DNA extracted from formalin fixed, paraffin-embedded (FFPE) samples, which comprise the majority of clinical cancer specimens; ii) limited DNA quantities extracted from fine-needle aspirations, small-core needle biopsies, or effusions; and iii) low tumor content. To this effect, the following are needed: i) standardized fixation, DNA extraction, and sequencing library construction protocols; ii) reproducible protocols using limited amounts of DNA; and iii) uniform high coverage across all clinically relevant test regions. In addition, rigorous bioinformatics analytic approaches are needed to show the accuracy and reproducibility of the assay to conform to the College of American Pathologists (CAP) standards.

To this end, we have developed a 93-gene panel for gastrointestinal cancers and have validated this for clinical use in a CAP-accredited facility using 48 colorectal cancer FFPE samples. In accordance with guidelines established by CAP, we established quality metrics and reportable ranges as well as assessed the analytic accuracy, sensitivity, specificity, and precision of the assay. Both single-nucleotide variants (SNVs) and insertions and deletions (indels) were tested.

Materials and Methods

NGS Assay Design

We performed a comprehensive literature and database review to identify genes biologically and clinically relevant to gastrointestinal cancers, including genes involved in key oncogenic signaling pathways, oncogenes, tumor-suppressor genes, and genes from kinase and chromatin remodeler families. Genes with biomarkers that may predict a response to drugs in clinical practice or drugs in development were curated from this literature search. A total of 93 genes were identified (Supplemental Table S1), corresponding to a total capture region of approximately 500 kb, within the limits of the SureSelect Tier 1 assay (Agilent, Santa Clara, CA). The panel comprised all exons of the 93 genes. However, this report focuses on specific clinically actionable alterations, found in exons of the five relevant genes (*KRAS*, *BRAF*, *NRAS*, *KIT*, and *PDGFRA*), which can be validated in the clinical setting.

DNA Sample Preparation and Sequencing

The test was performed on FFPE DNA extracted from colorectal cancer samples, using the QIAamp DNA FFPE Tissue Kit (Qiagen). Multiplexed library preparation was performed based on the Agilent SureSelectXT2 protocol, starting from 1 μg DNA, according to the manufacturer's instructions. Each indexed library was quantified accurately using the KAPA Quantification Kit (catalog no. KK4824; KAPA Biosystems, Wilmington, MA) before pooling. This library was hybridized to RNA-based capture probes in solution to genetic regions of interest, and captured using a magnetic bead-based method. After library clean up, sequencing was performed on the MiSeq (Illumina, San Diego, CA) by performing two runs of 300 cycles paired-end sequencing to achieve a minimum average coverage of 100× for each sample.

Bioinformatics Pipeline and Data Analysis

Paired-end alignments were performed using a Burrows-Wheeler Aligner version 0.7.5.a (*http://bio-bwa.sourceforge.net*, last accessed December 16, 2014) against the human genome 19. Alignments were processed using the Genome Analysis Toolkit version 1.0.5974 (*https://www.broadinstitute.org/gatk*, last accessed November 2, 2012)-based workflow to perform local realignments and to recalibrate base quality scores. Other analyses included variant calling using VarScan version 2.3.5 (options used: min-coverage, 30; min-reads2, 2; min-avg-qual, 30; and min-var-freq, 0.02; Genome Institute at Washington University, St. Louis, MO) and annotation using Variant Effect Predictor version 75 (Ensembl, Hinxton, Cambridgeshire, UK).

Validation Specimens

For validation of the NGS assay, we tested 48 tumor specimens derived from FFPE samples and 23 additional cell line DNA samples. Extracted DNAs were prepared and sequenced multiple times for some specimens to assess precision (seven cell line samples were sequenced in two runs, replicated within each run; and four FFPE samples were sequenced in three runs, replicated within each run).