



Published in final edited form as:

*Semin Oncol.* 2016 August ; 43(4): 453–463. doi:10.1053/j.seminoncol.2016.06.003.

## Translating Pharmacodynamic Biomarkers from Bench to Bedside: Analytical Validation and Fit-for-Purpose Studies to Qualify Multiplex Immunofluorescent Assays for use on Clinical Core Biopsy Specimens

Allison Marrero, PhD, Scott Lawrence, PhD, Deborah Wilsker, PhD, Andrea Regier Voth, Ph.D., and Robert Kinders, PhD

Pharmacodynamics Assay Development and Implementation Section, Laboratory of Human Toxicology and Pharmacology, Applied/Developmental Research Directorate, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, MD 21702

### Abstract

Multiplex pharmacodynamic (PD) assays have the potential to increase sensitivity of biomarker-based reporting for new targeted agents, as well as revealing significantly more information about target and pathway activation than single-biomarker PD assays. Stringent methodology is required to ensure reliable and reproducible results. Common to all PD assays is the importance of reagent validation, assay and instrument calibration, and the determination of suitable response calibrators; however, multiplex assays, particularly those performed on paraffin specimens from tissue blocks, bring format-specific challenges adding a layer of complexity to assay development. We discuss existing multiplex approaches and the development of a multiplex immunofluorescence assay measuring DNA damage and DNA repair enzymes in response to anti-cancer therapeutics and describe how our novel method addresses known issues.

### Introduction

The complexity of intracellular protein signaling, metabolic processes, and DNA replication and repair inherent in diseases such as cancer are well recognized; however, in measurements of clinical correlates from biopsies and patient specimens, analysis is still often limited to a single analyte, representing a single drug target within any one of these pathways. While this approach has the benefit of focusing preclinical development and pharmacodynamic (PD) marker selection, a critical limitation is that, in order to measure an effect, one must choose between upstream measurements of target activation and

**Corresponding Author:** Robert Kinders, Pharmacodynamics Assay Development and Implementation Section, Laboratory of Human Toxicology and Pharmacology, Applied/Developmental Research Directorate, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, MD 21702. Phone: (301) 846-6410; Fax: (301) 846-6536; kindersr@mail.nih.gov.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Conflicts of Interest:** None

**Research Performed At:** Frederick National Laboratory for Cancer Research

downstream measurements of pathway activation and/or intended treatment outcome at the cellular level. This, along with the additional difficulties associated with obtaining sufficient high-quality specimens for analysis, drives the current emphasis on multiplex analysis of clinical trial specimens.

There are numerous benefits to applying a multiplex format in support of a clinical trial. First, multiplex assays enable measurement of PD responses of multiple analytes on a single specimen, maximizing the amount of information obtained using a minimal amount of valuable patient tumor tissue. Second, multiplex assays can enable intracellular pathway activity reporting, measuring target engagement and the intended PD effectors and early sensors of the pathway as well as downstream markers of drug effect in the same tissue section; markers of commitment can potentially also be measured if they can be identified. A third critical aspect of a multiplex assay is that it reduces the possibility of missing a PD response due to factors such as specimen collection time, dose of the investigational agent(s), and genetic alterations in the tumor, as compared to a single marker being used as the assay readout. Finally, pathway reporting will be particularly useful in combination therapy approaches using two agents with different mechanisms of action.

One of the strengths of the multiplex assay is the ability to confirm a drug effect using a correlative marker in the event there is no modulation of the primary biomarker. A lack of modulation of the primary marker measured in a single analyte assay could be interpreted as either no drug effect or a genetic defect that prevents modulation of the target. For example, when profiling a DNA repair pathway, signal from the phosphorylated form of the DNA damage sensor Nbs1 (pS<sup>343</sup>-Nbs1) or histone H2AX phosphorylated at Ser139 ( $\gamma$ H2AX)<sup>1, 2</sup> could be absent in Ataxia telangiectasia mutated- (ATM) or DNA-dependent protein kinase- (DNA-PK) deficient models due to the genetic background. However, modulation of other markers included in a multiplex assay panel, such as Rad51 or ERCC1, could confirm drug effect on tumor. Importantly, the presence of additional markers provides information that allows a negative result in one marker to be distinguished from a lack of total response, and alternate interpretations to be generated. In addition, the use of combinations of markers for a particular PD pathway can also decrease false positive calls by clarifying a spurious positive signal from only one biomarker in a measured set. Using such approaches, molecular responses in clinical samples may come to light that could not have been predicted; however, there are a separate set of challenges associated with multiplexing assays, particularly those performed on solid tissues.<sup>3, 4</sup> Here we will discuss some popular technologies for multiplex assays and their utilization for PD studies, and then enumerate the challenges inherent in multiplex immunofluorescence assays, providing specific examples of how we dealt with them during the development of a multiplex analysis of the DNA repair activation pathway in patient biopsies.

## Multiplex Assays for Clinical Samples

From a technique standpoint, multiplex assays can be grouped into those requiring a homogenous sample (such as tissue lysates or blood samples) and those requiring an intact tissue section for analysis. Both types of multiplex assays present specific strengths and challenges.

## Assays for Tissue Lysates and Blood Samples

**The Luminex xMAP Platform**—One of the most popular multiplexing technologies is the bead-based flow cytometric xMAP<sup>®</sup> platform from Luminex. Assays developed for this platform use the two-site or sandwich immunoassay approach, employing a monoclonal antibody (mAb) conjugated to a fluorescently labeled bead to immobilize each analyte and a second, labeled mAb against the analyte to report its concentration. An assay calibrator is required for each analyte; usually a recombinant protein version of the analyte is used. In collaboration with the NCI, Myriad RBM has developed a number of multiplex assay panels using this technology,<sup>5, 6</sup> including Human OncologyMAP, which surveys 130 serum proteins that have been employed as cancer markers including established diagnostic markers such as CEA and CA125, and a number of important growth factors.<sup>7</sup> This assay can be run as a service by Myriad RBM, or validated assay kits can be purchased for use on Luminex<sup>®</sup> instruments. The specimen required for this assay is 500  $\mu$ L of patient serum, and is therefore readily applicable in most clinical situations. The Myriad RBM CytokineMAP A and B assays are also examples of widely used assays, in this case for immune response modeling.<sup>8</sup> Their use has even been extended back into preclinical development to assist in validating biomarkers for drugs under investigation. These assays have the additional advantage of requiring only a 50  $\mu$ L serum sample volume. With the current surge of interest in immunotherapy approaches to cancer treatment based on recent impressive clinical trials results, we anticipate a continuing increase in the use of these assay panels.

The NCI has recently contracted Myriad RBM to produce a Luminex platform assay for apoptosis signaling pathway proteins,<sup>9</sup> which is now commercially available from BioRad as a BioPlex kit and can be run on their xMAP multiplex magnetic bead-separation platform. Unlike the above-mentioned assays, this assay is intended to be run on tissue extracts (or cell extracts for preclinical work) and can be performed on a good quality 18 gauge core biopsy (average wet weight, 7 mg). In preclinical applications, our laboratories use a 20 mg tissue piece to provide enough material for repeat runs of all analytes. The analytes are divided into 3 panels, thus if a fit-for-purpose biomarker of therapy effect has been validated, the researcher may choose to use only the panel containing that biomarker, providing additional materials for repeat specimen analysis. Importantly, the kit includes a set of calibrators for each analyte in the assay.

**MesoScale Discovery MULTI-ARRAY Platform**—Another popular multiplex assay platform utilizing a two-site immunoassay approach is the electro-chemiluminescent MULTI-ARRAY technology from MesoScale Discovery. Here, each analyte is bound by a capture mAb that is precoated on a carbon electrode plate, and a second mAb conjugated to an electrochemiluminescent dye reports the concentration when voltage is applied to the carbon electrode plate. Commercially available multiplex assay kits for this platform are designed to assess biomarkers of cardiac, liver, kidney, or muscle injury, inflammation, cytokines and chemokines, and general toxicology, among others. For example, MesoScale kits measuring human growth factors and receptors have been used to correlate growth factor receptor inhibition and treatment with the receptor tyrosine kinase inhibitors foretinib and dovitinib in clinical trials.<sup>10, 11</sup> All kits are species-specific and include control materials for standard curve generation, and are examples of how multiplex assays consisting of a small

number of established disease biomarkers can be validated and implemented. The species restriction on kit utility is a reminder that antibody cross-reactivity between model system and human homologs will affect how the preclinical development of any biomarker is approached.

**Reverse Phase Protein Array Assays**—Reverse phase protein array (RPPA) assays provide functional proteomics analysis of complex signaling pathways by probing protein tissue or blood extracts spotted onto a slide with validated monoclonal antibodies under controlled conditions.<sup>12, 13</sup> A number of RPPA assays have been developed and validated at the NCI and at MD Anderson Cancer Center,<sup>14, 15</sup> and organizations such as TheraNostics Health and the MD Anderson Proteomics Core offer services for running patient specimen analysis.<sup>16, 17</sup> The results from this technique are best analyzed compared to a drug-treated control tissue analyzed on the same slide set as the clinical sample. The use of tissue controls allows scaling of assay values across multiple experiments, because calibrators are not available for the assays and only a single antibody is used to report each analyte. This approach has demonstrated utility and is especially well-suited for discovery work in complex systems.<sup>18–21</sup> In developing predictive molecular markers for dasatinib treatment, for example, RPPA identified ten potential markers that were differentially expressed in dasatinib-sensitive and -insensitive cell lines and researchers were able to build on this information to clarify the role of CAV-1-mediated interactions between EphA2 and BRaf on dasatinib sensitivity.<sup>14</sup>

**Advantages of Assays for Tissue Lysates and Blood Samples**—A major advantage of the two-site immunoassay format is the superior analyte specificity obtained by using two separate mAbs. Strategically chosen epitopes can report, for example, only full-length or only truncated proteins. The use of analyte calibrators in these assays also allows comparisons of assay values across laboratories and over time.

**Disadvantages of Assays for Tissue Lysates and Blood Samples**—Validated commercial assays tend to be very costly; unfortunately, the front-end costs of development and validation to the degree required by regulatory agencies are the primary drivers of these costs, making this limitation unavoidable. Additionally, multiplex formulations may reduce the dynamic range for some analytes in this type of assay, particularly in dealing with serum analytes where concentrations vary by orders of magnitude. Of particular concern for these assays, the origins of serum analytes are not known with certainty. The origin of analytes from tissue extracts is better described, but subject to the variability in quality of the starting material (for example, a biopsy) and it is not generally feasible to determine whether an analyte was extracted from tumor or healthy tissue due to the processing requirements for the assays, nor is it known what role serum contamination (an issue for certain analytes) played.

Analyte interactions are another possibility that should be taken into account. In cell extracts some proteins will automatically aggregate into complexes due to the law of mass action and this can alter readout compared to assay standards. As discussed above, not all multiplexes have standards, thus limiting their utility over time, unless the individual investigator has a useful calibrator or control available for longer-term studies.

Finally, while the methods described above allow for the concerted analysis of a large number of analytes, one must also consider the limitation of data analysis and interpretation. Even with excellent data analysis packages, it can be challenging to decipher treatment effects that are drug-mechanism or patient-response related. Our approach in these assay types has been to identify and qualify a particular biomarker or biomarker subset (usually in a specific signaling or effector pathway that the treatment is intended to target) and to establish the fitness-for-purpose of the biomarker/pathway in a preclinical model. This can also help to set rules for allowing or disqualifying assays if a particular assay calibrator or control fails. This issue of how failure of one analyte in a panel should affect reporting of the other analytes in a panel has not been resolved by the biomarker community, and currently the approach is to consider each analyte individually. This is acceptable when one or a few members of an assay panel fail, but there are no hard guidelines for what percent of a panel can be allowed to fail while allowing acceptance of all other assay values. The likelihood of a control failure is obviously much greater for a 130-analyte panel than for any single analyte. The controlling factor is still pre-analytical variables: specimen collection handling, shipping and storage. For these reasons, our laboratory prefers assays performed on tissue sections until more advanced systems of analysis can be developed and validated.

### Assays for Tissue Sections

There have been a number of multiplex technologies developed over the past 30 years attempting to impose a means of quantitation on the standard tissue slide (formalin fixed paraffin embedded [FFPE] or frozen sections). This general approach of imaging tissues that have been stained by some method to aid in morphological analysis, biomarker analysis, or morphometric analysis has a number of significant advantages, as well as limitations, compared to extraction-based assays.

**Multiplex IFA-AQUA technology**—The current state of the art methodology, particularly for diagnostic medicine is the AQUA (automated quantitative analysis) technology<sup>22-24</sup>. In this technology, subcellular compartments are defined using a fluorescence-labeled molecular tag, usually an antibody, that binds to a marker of a marker of that subcellular compartment, for example anti-cytokeratin. The processing is called “masking” and the mask defines the area in which a signal generated by another fluorescence-labeled antibody to a biomarker of interest is measured. In addition to restricting the area of signal measurement, the imaging method also collects intensity levels of the fluorescence signal. The measured biomarker signal is converted to an AQUA score, which can then be compared to the signal generated by the same fluorescent probe applied to a western blot. Developing assays for this methodology is time-consuming and expensive, and thus the leading application of the platform is with validated disease biomarkers such as the estrogen receptor (ER).

**Disadvantages of the AQUA Assay**—This elegant approach and its various derivatives have five main limitations. The first one is the issue of background fluorescence. Its variability in tissue sections is well known in pathology laboratories. Any method attempting to measure a biomarker signal in a tissue section must therefore compensate for this background in a consistent way that can also be flexible enough to allow for slide-to-slide

fluorescence background variability. This issue was clearly illustrated in the study by Welsh *et al.*<sup>24</sup> in which cells with known ER expression level were characterized by quantitative western blot and AQUA using the same anti-ER antibody. The correlation between absolute concentrations of ER (from western blot quantitation) and AQUA score in these cells with known ER level allowed for the identification of the cutoff point between highest negative value and lowest positive value.

The second limitation is due to denaturation of the specimen during processing. Antigen epitope retrieval is usually required for antibody recognition of tissue targets after the formalin fixation process, with combinations of heat and activating chemicals such as citrate or EDTA commonly used. For handling frozen sections embedded in OCT, antigen and RNA preservation are initially very good but can be impacted by downstream processing, for example, the use of an acetone fixation (preferred for mRNA recovery) step instead of formalin (preferred for morphology preservation), which results in significant differences in final epitope or RNA structure. These physical separation and retrieval processes are different enough to caution users regarding comparisons of immunochemical reaction strengths between the various methods.

A key limitation for quantitating phosphoproteins on AQUA and other IFA platforms is the use of a single binding epitope for analyte readout. Phosphoprotein quantitation requires two binding epitopes for each molecule measured: one pan-isoform epitope from the protein backbone to capture the various proteins, e.g., AKT1/2/3, and another epitope containing the desired phosphosite. This is particularly important in PD studies because a drug can impact either total protein level or degree of site-specific phosphorylation within the protein population. Technologies to address this are not currently available for IFA analysis.

The limited dynamic range of fluorescent cameras must also be considered. As we have reported previously, biomarker signals that increase with time, such as  $\gamma$ H2Ax can easily rise to intensities well above the dynamic range which CCD cameras can discriminate, resulting in significant underestimation of the amount of target present, even discounting effects of light spillover from one camera pixel to an adjacent pixel.<sup>25</sup>

Finally, a limitation for the AQUA technology and indeed all quantitative assays, is the control of preanalytical variables. As pointed out by Hewitt *et al.*, the requirements for assayed biomarkers of drug response, whether phosphorylated protein, DNA, or RNA, force much greater stringency onto specimen collection and assay validation than has traditionally been the case for diagnostic testing of slide specimens.<sup>26</sup> The greater stringency is dictated by the necessity of protecting labile biomarkers, such as phosphorylated protein epitopes and mRNA molecules to be used as in situ hybridization targets.

Despite these caveats, the IFA-AQUA method has been rigorously developed for a number of diagnostic biomarkers and is recognized as the standard for methods aimed at quantifying cellular targets in standard histochemical analysis. Commercial services for AQUA analysis are available from Genoptix ([www.genoptix.com](http://www.genoptix.com)).

## Applying a Multiplex Strategy to DNA Repair Enzymes

A large number of new molecularly targeted agents for cancer therapy, including some widely used natural-product derived chemotherapy agents such as camptothecins and their derivatives, are intended to disrupt DNA repair enzymes such as poly(ADP-ribose) polymerase (PARP) or DNA topoisomerase I. Development of such agents prompted us to create a multiplex assay capable of reporting their activity on DNA repair enzyme targets, particularly in the homologous recombination and base excision repair pathways. In this case, we chose to develop a quantitative IFA (qIFA) including markers from a variety of key proteins in the DNA damage repair chain, including upstream effectors and early sensors of the damage, such as the phosphorylated form of DNA damage sensor Nbs1 (pS<sup>343</sup>-Nbs1), combined with downstream markers of commitment to repair, such as Rad51. State of the art analysis software enables the imaged cell population to be categorized into distinct sub-populations, based on expression of one or more of the multiplex panel biomarkers, which may help identify possible tumor heterogeneity in PD response, identify different time courses of response for the biomarkers, or identify populations of altered response on a per cell basis using marker co-localization methods.

The validation work we discuss here for DNA repair pathway analysis includes model systems of varied genetic backgrounds selected to match the expected clinical trial patient population. For example, a *BRCA*-deficient cell line or xenograft and the appropriate matched wild type control were used in preclinical modeling of PD responses to drugs known to elicit a tumor response in *BRCA*-mutated ovarian or breast cancers. These model systems are critical in suggesting whether the presence of the specific genetic defect will enable the prediction of a PD response.

An example of the DNA repair enzyme multiplex assay is shown in Figure 1. We employed the human A375 melanoma xenograft model previously reported to demonstrate a targeted effect from the indenoisoquinoline topoisomerase 1 inhibitors. We used FFPE tumors collected from topotecan treated mice (1.5 mg/kg for 2 hours) as the positive biomarker control and vehicle (water) treated mice as the negative control both in the preclinical experiment and in generation of the response calibrators. Distinct expression patterns were seen for all 4 markers in the multiplex assay (Rad51, pS<sup>343</sup>-Nbs1, ERCC1, and  $\gamma$ H2AX) as well as considerable heterogeneity in expression between individual cells.

Factors such as sample collection time, dose of the investigational agent or agents, and genetic alterations in the sample can alter the PD response. For example, in Figure 2, a *BRCA*-deficient MX-1 xenograft was exposed to irinotecan for 24 hours at 7.5 mg/kg; however, no measureable  $\gamma$ H2AX induction was observed, raising the question of whether the drug reached the tumor tissue or the dose was sufficient to elicit a DNA damage response in the tumor. Because pS<sup>343</sup>-Nbs1, another early DNA damage response marker, was also measured, we were able to document the increased proportion of treated cells positive for this PD marker compared to untreated cells, indicating that the drug did induce DNA damage. Note that in both models (Figures 1 and 2), RAD51 and ERCC1 are detected by the multiplex assay, but neither was modulated by topotecan or irinotecan. We are currently

employing this DNA repair multiplex assay in a in support of drug development programs at both the pre- and clinical stage.<sup>27</sup>

## Methodological Considerations

Prior to use of a developed PD immunofluorescence assay on clinical specimens, we carry out full validation and fit-for-purpose testing of all assay reagents, including antibodies to selected PD markers, in appropriate *in vitro* and xenograft models using orthogonal methods of analysis (for example, Western blotting, ELISA of tissue extracts, and competition of antibody binding by cognate antigen). An example of competition testing is shown in Figure 3, using the peptide immunogen for the primary antibody.

## Assay Reagent Validation and Antibody Multiplexing Strategy

An analytical validation process is required to demonstrate the accuracy and reproducibility of an assay for the intended analyte, beginning with assay reagent validation. Most laboratories performing early stage drug development immunoassays are required to purchase antibodies to intended PD markers from commercial vendors. Very often, neither the antibody nor the target marker is analytically or clinically validated, and vendor release specifications are generally not as stringent as those required for diagnostic or USP grade materials. For initial validation of an antibody for use on clinical specimens, establishment of antibody specificity and sensitivity is required, minimally by titration of the antibody preparation on both positive and negative control cell lines and tissues to determine the optimal concentration for use, i.e., the best signal/noise ratio and dynamic range of the PD marker.

More than one lot of an antibody must be examined, since lot-to-lot variance in antibody production can alter the assay sensitivity and specificity. The quality of released product can degrade due to production and purification errors, resulting in decreased specificity or protein degradation, often seen as an increased background or zero signal in an assay. Research product vendors may also stop supplying a product, which is not an uncommon event in our experience. This last supply issue is the main reason we recommend against using polyclonal antisera, although there are notable exceptions, for example, Dako (Agilent Technologies, Santa Clara, CA) has proven to be a reliable source of clinically validated polyclonal antibodies.

## Criteria for Lot-to-Lot Quality Control Comparison in the DNA Damage Repair Multiplex qIFA

- The signal is nuclear
- The signal is visible as foci under confocal microscopy The concentration used has controllable autofluorescence in the cytosol
- The lots can be run at comparable concentrations (not antibody dilution)
- Signal observed in positive control tissue with constant exposure settings
- Signal is greater than isotype control run at the same concentration



- Antibody competition with cognate antigen successfully removes signal

If possible, antibody specificity should be tested with a syngeneic knockdown cell line. This approach will usually not present a perfect negative for DNA repair enzymes because they are critical for cell survival, but knockdowns of over 90% have been achieved for ERCC1 and PARP1, for example. This degree of rigor is required because neither the range or variability of biomarker expression in human tumors nor the drug effect will be known until after the first-on-human clinical trial is conducted.

Multiplexing antibodies for target staining on tissues poses a series of technical issues including cross-reactivity, channel bleed through, and increased background fluorescence. In the DNA repair enzyme multiplex assay, our strategy was to start with a retrieval optimized for one biomarker and then modified to obtain the best possible signal from the biomarker requiring the highest antibody sensitivity and image resolution. We then conjugated that antibody to a high energy dye, such as Alexafluor 488, or to a hapten that allows signal amplification with a conjugated anti-hapten antibody. The remaining antibodies to be used in the multiplex were then conjugated to a dye or hapten, with selection based on the signal intensity required to obtain a readout of suitable dynamic range and specificity in tissues. Practically speaking this means selecting antibody concentrations that are just at the signal saturation point for the most positive specimens but that are still positive for the least positive specimens. Appropriate selection of antibody conjugates and reporting dyes, with and without an amplification mechanism, is a useful way to extend the dynamic range of measurements for biomarkers that are low in expression levels (for example, ERCC1) or that are expressed as small foci (e.g., phosphorylated ATR). Use of anti-species antibodies as dye-conjugated reporters is only employed as a last resort, and we do not recommend use of these for multiplexing assays to eliminate the possibility of antibody cross-reactivity.

### **Assay Calibration and Controls and Their Impact on Data Reporting**

A response calibrator is a set of specimens, preferably xenografts from a drug-responsive model, treated at drug concentrations that elicit the maximum biomarker response, at least one intermediate level biomarker response, and a background biomarker level, usually a vehicle control. An example of our calibrator and control slide is shown in Figure 4. Ideally, calibrators define the range of signal that will be expected when the assay is run on clinical specimens, but in practice the highest signal obtainable on patient specimens is an unknown during preclinical assay development stages on an immunofluorescence platform due to inter-patient variation in unrelated factors such as background autofluorescence and specimen cellularity. This is particularly true for markers that have highly variable expression, such as cyclin-dependent kinases phosphorylated at Tyr15 (pY<sup>15</sup>-cdk). The low end of the assay range is defined by both the lowest limit of detection of the analyte above the background tissue autofluorescence that constitutes a specific binding of the reporter antibody to the cognate antigen, and the minimum biologically effective dose, which defines the low biological range of marker, reflecting the minimum drug dose required to produce a change in a PD marker that can be distinguished statistically from the no-treatment (or pre-treatment) control group.

While establishment of these criteria are not trivial for a single marker, additional factors must be considered when developing a multiplex assay. First, can all the PD markers in the multiplex be induced by a single drug? Second, what is the maximal expression time point of each individual biomarker in response to the drug? In our experience, it is unlikely that the maximal expression of multiple markers will coincide in time. In this case, the primary readout marker of the assay must be determined, and the calibrator tissue should be selected to reflect the drug effect on the primary biomarker. Correlative expression to the remaining biomarkers can be made and, importantly, positive and negative staining controls for the correlative markers should be included on the calibrator/control slides.

### **Immunofluorescence Imaging and Analysis Technical Considerations**

Imaging instrumentation (microscope components and camera) should be calibrated at least monthly (preferably weekly) or when a component changes are made, such as a new light bulb/light guide or a new objective being installed. If the microscope setup is used frequently or by multiple users, it is important to calibrate the instrumentation every time before running a specimen series, to ensure adjustments were not made that could significantly affect assay performance. We also image and analyze a stained calibrator slide according to the standard operating procedure (SOP) to determine if the microscopy setup is performing to specifications. Readings outside of calibrator specification suggest the light source or the light guide need to be changed. Regularly scheduled instrument calibration routines, including microscope, light source, and camera, should be performed using external calibrator materials.

Platform selection is critical for assay development in the circumstance where the objective is to implement and transfer the assay to multiple laboratories. Among the technologically advanced platforms currently available are whole slide scanning, high content wide field (HCWF) microscopy, and high-resolution confocal microscopy. Our approach has been a combination of all three (Table 1). Whole slide scanning has the advantage of rapidly digitizing glass slides and storing image data into a curated image database. To ensure slide to image accuracy, barcoding and automated metadata handling are employed. The resulting digital slides can be readily evaluated and annotated for prospective biomarker analysis on target areas (i.e., neoplastic tissue). We utilize whole slide scanning to pre-scan flanking H&E slides to assess the biopsy quality before moving on to biomarker immunofluorescence staining and analysis. Additionally, the digital slide scans are annotated by a pathologist to guide the microscopist.

HCWF microscopy allows for the acquisition of multiple biomarkers over several different slides. Critical to a HCWF system is the software to control the many facets of a typical microscopy setup including stage movement, camera exposures, and general microscope settings. Components of a HCWF system should include automation software, a multi-slide holder (4 or more slides) to increase the volume of automated image acquisition, high numerical aperture objectives, extended transmission filter cubes, and a digital monochrome camera. Camera selection should be determined based on the minimum need to acquire the biomarkers of interest. For instance, monochrome EMCCD cameras provide very high sensitivity and performance when it comes to visualizing weak expressing biomarkers and

are ideal for near infrared (NIR) wavelength. Finally, establishing an acquisition ruleset should include a low magnification pre-scan on a morphology channel (i.e., DAPI), fields to be imaged (including positioning information), an autofocus routine, multichannel exposure, and file saving criteria.

HCWF microscopy allows us to automate the majority of our multiplex image capture removing much of the bias that comes with manual imaging. Field selection is guided by pathology annotations (from whole slide scanning) and done solely on the morphology assessment channel (typically DAPI). The autofocus routine determines the focus plane per channel, minimizing out of focus images. Determining the exposure time per channel is done by calibrating the microscopy to the average background intensity of the tissue for each slide imaged. Multiple regions of interest (ROI) are placed across the sample in non-biomarker stained regions where the exposure is increased incrementally until the background intensity reaches a previously determined threshold. Exposure values are applied to the capture settings for the current project but are recalculated for subsequent projects to account for instrument decay. Once the settings are applied to the rule set for image acquisition, the software runs the instrumentation. Run data, instrument settings, and the image captures are compiled and saved as a project set for rerunning experiments or auditing purposes.

When the biomarker signal(s) are hard to discern or the analysis requires intracellular localization, high-resolution confocal microscopy is employed. A key advantage to a confocal system is the exclusion of out of focus light (background emission) from the image capture, allowing for higher resolution images with greater spatial discrimination. This is pivotal when analyzing foci or other minute structures. Confocal imaging can be used in conjunction with HCWF microscopy where the same fields can be imaged in either modality depending on the necessity of the individual biomarker expression pattern within the multiplex of markers detected on the slide. Furthermore, confocal imaging can be configured for multispectral imaging to acquire greater than 5 independent channels for highly multiplexed image analysis.

We perform image analysis on Definiens® (Munich, Germany) software using scripts custom developed in our laboratory. This analysis technology drives the ability for quantitative measurements and enables analysis on a level not available before. Selection of a cell enumeration strategy allows switching between a cell count and percent nuclear area in a field, as appropriate, for the particular biomarker being analyzed. This may reduce the number of “un-evaluable” biopsies, which is critical due to the difficulty in securing the acquisition of paired (pre- and post-dose) biopsies.

### **Specimen Analysis and Interpretation of Results**

Quantitation by positive cell numeration has been reported in detail,<sup>25, 28</sup> and the same principles are used in the application of the Definiens software in our multiplexing approach. Prior to biomarker analysis, nuclei are segmented through a combination intensity thresholding in the DAPI channel and adjusting for nucleus size and roundness; not all nuclei will be segmented properly, most notably nuclei that are atypical in shape/size or are highly clustered. In our experience, the best accuracy for measurements of cell numbers

expressing biomarkers of interest is to employ a mask (such as DAPI for nuclear signals) and use this to restrict signal analysis to the cell compartment where the biomarker is proven to be localized and active. The signal for each DNA repair biomarker is read in a single channel, reading and scoring only the DAPI-positive region of the image (the DAPI mask). For each biomarker, pixel positivity is determined by measuring pixel intensity to the respective channel cutoff, which is set based on the autofluorescence signal for that channel on the tissue being imaged. For example, in our analyses we exclude measurements of ERRC1 presence in the cytosolic fraction of the cell. This may decrease overall assay sensitivity, but if the objective is to measure DNA repair processes then the increased specificity is worth the tradeoff.

Diffuse nuclear biomarker signal expression is analyzed by determining the percentage of positive pixels within a nucleus object. When the camera pixel density is high enough, this allows measurement of fine structure such as the chromosomal localization of ATR foci. Foci marker analysis is conducted in a similar manner except the positive pixels are additionally analyzed for clustering into small objects within a nucleus. Further morphology and size filters are utilized to count the number of foci and those nuclei that exceed a previously determined foci count are classified as positive.

Biomarker analysis is improved using a first step of tissue quality and morphology analysis. Chromogenic morphologic staining (i.e., H&E) is performed on all biopsies, using a sampling strategy that brackets the region to be analyzed with a first and a last slide. Whole tissue slide scanning is performed on an Aperio (Leica Biosystems, Wetzlar, Germany) system for morphological quality assessment and presence of tumor, under a pathologist's review. The objective of this exercise is to capture image data that represents the tissue component of the biopsy. Within the region bracketed by the H&E slides, a set of four non-overlapping sections can then be assessed for biomarker expression. All captured image data is then integrated into a single database including annotations from both the pathologist and pharmacologist.

Prior to acquisition of imaging data from patient specimens, a set of rules is drafted for each biomarker to be analyzed and this is encoded in the assay SOP as well as in the image acquisition software. We have varied these rules depending on the type of analysis being performed but always include: avoiding confounders (folds, bubbles, RBCs), avoiding necrosis, avoiding the edge (due to edge effects), automating acquisition on the selected fields using pre-established exposures, and archiving raw image data plus the image rule set. The exact set of rules to be used should be determined during preclinical fitness-for-purpose testing.

### **Fitness Modeling**

Because details such as biomarker localization in the cell and the type of matrix in which the tumor resides (e.g., the differences between a lymphoma and an osteosarcoma) will affect the rule set, validation and fitness-for-purpose testing in appropriate *in vitro* and xenograft models is necessary.<sup>29,30,31</sup> Figure 5 shows a block diagram of our workflow for our DNA damage repair image analysis project.

The objectives of this fitness modeling are to define the background variability of biomarker expression within and across tissues, the dynamic range of biomarker expression, and the degree to which drug treatment modulated that dynamic range. The lowest dose of the investigational drug that induced a biomarker change, and the time at which the change occurred after treatment must also be determined, along with the greatest measurable drug effect and whether or not that highest dose response is measurable within the range of the assay response curve. For immunofluorescence-based assays, this translates into the requirement that the highest value on the response curve should be less than the signal intensity that saturates the CCD chip pixels. The lowest point of the standard response curve should be one that is clearly more positive than no signal, and should be higher than the typical autofluorescence signal detected by the CCD. This process can be handled by setting a threshold signal acquisition time in the image acquisition software. This work to establish the dynamic range and total variability of the biomarker response is necessary to demonstrate that a particular method is “reliable for the intended application”.

Additional factors driving clinical suitability include the variability of the baseline signal within and across biopsies for each biomarker. If the baseline variability for the biomarker is high, the assay dynamic range must be correspondingly higher. Our starting assumption, based on detailed analysis of background signal variability within and across xenografts,<sup>30</sup> is that greater than a 50% change in biomarker signal is required to overcome background biomarker variability. In the case of PAR, for example, high baseline biomarker variability in the xenograft models was later confirmed in a clinical trial.<sup>32</sup> The dynamic range of the assay and the assay sensitivity (seen as the slope of the biomarker response curve) will therefore drive the utility of an assay in clinical specimens just as much as the specificity of the assay for the cognate biomarker. The DNA damage inducible markers  $\gamma$ H2AX and pS<sup>343</sup>-Nbs1 are examples of markers with low baseline variability, while pY<sup>15</sup>-cdk and ERCC1 have high variability; high variability limits the certainty with which the minimum drug effect level can be measured.

In addition to instrument control of image data acquisition, there is the further requirement to establish a suitable method to determine the limits of signal positive areas, i.e., which pixels define the outline of the positive structure. Two forms of this determination are used in the DNA repair multiplex assay discussed here. For  $\gamma$ H2AX signaling the positive nuclear area is large whereas for pS<sup>343</sup>-Nbs1 there are anywhere from 1 to  $\gg 20$  foci per cell, and the cutoff for detection of a drug-induced signal is based on counts of foci and therefore requires a robust and reliable counting script in the software. Morphological measurements of foci are made on the basis of object shape, size and punctate fluorescence spot counting.

Varied genetic backgrounds matching the expected clinical trial patient populations should be considered in selecting the xenograft models to be used in fitness testing. We prefer to test both a drug-responsive and a drug nonresponsive model at this stage and ideally wish to see tight correlation between biomarker modulation and response in both models.

### **Clinical Suitability**

Analysis of representative specimens from patients is critical to evaluate assay performance and determine whether the assay can both measure biological variability in the patient

population and detect drug action on the target. Xenografts are not always predictive of successful detection of a biomarker signal in clinical samples, and the mode of failure is unpredictable. For example, in preclinical modeling for the Wee1 inhibitor AZD1775, one xenograft model did not predict  $\gamma$ H2AX signal at clinically relevant doses, but there was a strong  $\gamma$ H2AX signal in patient biopsies.<sup>33</sup> Additionally, in clinical trials of veliparib plus irinotecan, clinical biopsy specimens demonstrated no  $\gamma$ H2AX signal despite preclinical xenograft models predicting a strong  $\gamma$ H2AX response. Whether the failure was due to biopsy timing or the genetics of the patient is unknown. Our data suggests timing is critical for  $\gamma$ H2AX detection, reinforcing the advantage of multiplex assays as other markers may be activated at earlier or later times, increasing the likelihood of observing a PD response somewhere in the pathway.

## Future strategies

Multiplex PD assays hold great potential as a tool for both preclinical and clinical drug development and evaluation, and we anticipate several advances in technology to further boost the effectiveness of these methods and facilitate more widespread use in translational research. As noted above, measuring only total protein levels is rarely sufficient to observe a PD response, given that protein modifications such as phosphorylation are critical for modulating many of the key signaling pathways monitored by PD assays. Two-site immunoassays for tissue lysates and blood specimens can accurately measure these protein modifications, but these lack the positional information available from assays built on an IFA platform, such as the ability to segment tumor cells and even subcellular compartments in order to obtain biomarker information from only the relevant cell types and locations. Therefore, the development of double epitope assays for IFA that combine readouts for the protein backbone and the specific phosphosite of interest must be undertaken to fully characterize the PD response of many pathways. Alternatively, whole biopsy imaging could provide a different, if perhaps more labor intensive, approach to addressing this issue. The ability to use confocal microscopy to examine an 18-gauge core needle biopsy without first staining the tissue would allow the laser capture dissection of relevant areas of interest, specifically excluding necrotic and stromal areas. This advance would permit quantitative, two-site immunoassay analysis to be performed exclusively on confirmed tumor tissue, although it would not address the subcellular localization of certain biomarkers.

Finally, as the complexity of the cellular pathways being targeted for drug development grows and as we attempt to extend our understanding of the multiplicity of drug-driven molecular effects in tumor tissue, the number of biomarkers desired in a multiplex assay will have to keep pace. While certain tissue lysate-based assays are able to accommodate large numbers of markers, improvements in dyes and conjugation methods for intracellular biomarker labeling are urgently needed for IFA-based platforms. A recent paper by Grimm *et al.* described the synthesis of dyes with increased brightness, better photon yield, and good cell permeability, properties which may enable additional spectral channels and less cross-talk between channels.<sup>34</sup> Approaches such as this are far superior to spectral imaging, which although providing many more fluorescent channels, eliminates any reasonable specimen throughput. Additional work in this area will be important to continue increasing the capabilities of IFA platformbased assays.

## Conclusion

In summary, multiplex assays are an effective tool to measure pharmacodynamic response by monitoring simultaneously multiple markers of drug response or pathway activity at various levels (upstream and downstream), thus increasing the probability of capturing a PD response to a pharmacological agent. The application of multiplexing analysis of drug PD biomarker assays in patient tissues is already underway and will be made even more prevalent with the increase in commercially available multiplex assays. With careful application of assay development and analytical principals this approach has the potential to not only increase our understanding of drug effects on target in patient specimens, but even more importantly, can be used to reduce the number of false negatives and increase the number of reportable results within patient trials.

## Acknowledgments

This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. The authors thank Dr. Mariam Konate, Capital Consulting Corporation, for medical writing support in the preparation of this manuscript.

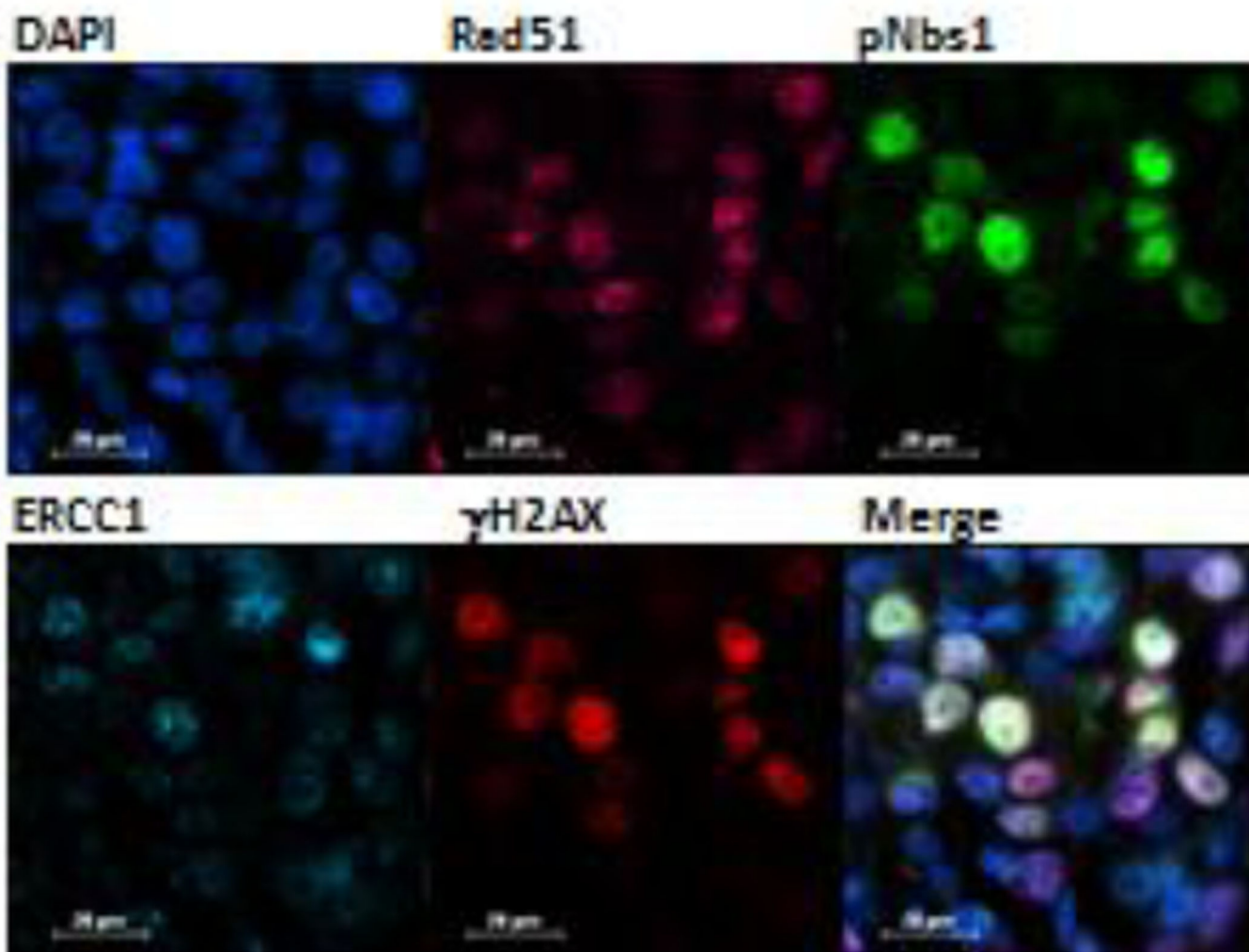
## References

1. Wang H, Wang M, Wang H, Bocker W, Iliakis G. Complex H2AX phosphorylation patterns by multiple kinases including ATM and DNA-PK in human cells exposed to ionizing radiation and treated with kinase inhibitors. *J Cell Physiol.* 2005; 202:492–502. [PubMed: 15389585]
2. Chanoux RA, Yin B, Urtishak KA, Asare A, Bassing CH, Brown EJ. ATR and H2AX cooperate in maintaining genome stability under replication stress. *J Biol Chem.* 2009; 284:5994–6003. [PubMed: 19049966]
3. Ellington AA, Kullo IJ, Bailey KR, Klee GG. Measurement and quality control issues in multiplex protein assays: a case study. *Clin Chem.* 2009; 55:1092–1099. [PubMed: 19372187]
4. Li D, Chiu H, Gupta V, Chan DW. Validation of a multiplex immunoassay for serum angiogenic factors as biomarkers for aggressive prostate cancer. *Clin Chim Acta.* 2012; 413:1506–1511. [PubMed: 22722017]
5. Sasaki K, Gotlib JR, Mesa RA, Newberry KJ, Ravandi F, Cortes JE, et al. Phase II evaluation of IPI-926, an oral Hedgehog inhibitor, in patients with myelofibrosis. *Leuk Lymphoma.* 2015:1–6.
6. Christiansson L, Soderlund S, Mangsbo S, Hjorth-Hansen H, Hoglund M, Markevorn B, et al. The tyrosine kinase inhibitors imatinib and dasatinib reduce myeloid suppressor cells and release effector lymphocyte responses. *Mol Cancer Ther.* 2015; 14:1181–1191. [PubMed: 25761894]
7. Schultheis B, Strumberg D, Santel A, Vank C, Gebhardt F, Keil O, et al. First-in-human phase I study of the liposomal RNA interference therapeutic Atu027 in patients with advanced solid tumors. *J Clin Oncol.* 2014; 32:4141–4148. [PubMed: 25403217]
8. Warnke E, Pietsch J, Wehland M, Bauer J, Infanger M, Gorog M, et al. Spheroid formation of human thyroid cancer cells under simulated microgravity: a possible role of CTGF and CAV1. *Cell Commun Signal.* 2014; 12:32. [PubMed: 24885050]
9. Srivastava AK, Jaganathan S, Stephen L, Hollingshead MG, Govindharajulu JP, Layhee A, et al. Effect of a Smac Mimetic (TL32711, Birinapant) on the Apoptotic Program and Apoptosis Biomarkers Examined with Validated Multiplex Immunoassays Fit for Clinical Use. *Clin Cancer Res.* 2015
10. Shah MA, Wainberg ZA, Catenacci DV, Hochster HS, Ford J, Kunz P, et al. Phase II study evaluating 2 dosing schedules of oral foretinib (GSK1363089), cMET/VEGFR2 inhibitor, in patients with metastatic gastric cancer. *PLoS One.* 2013; 8:e54014. [PubMed: 23516391]

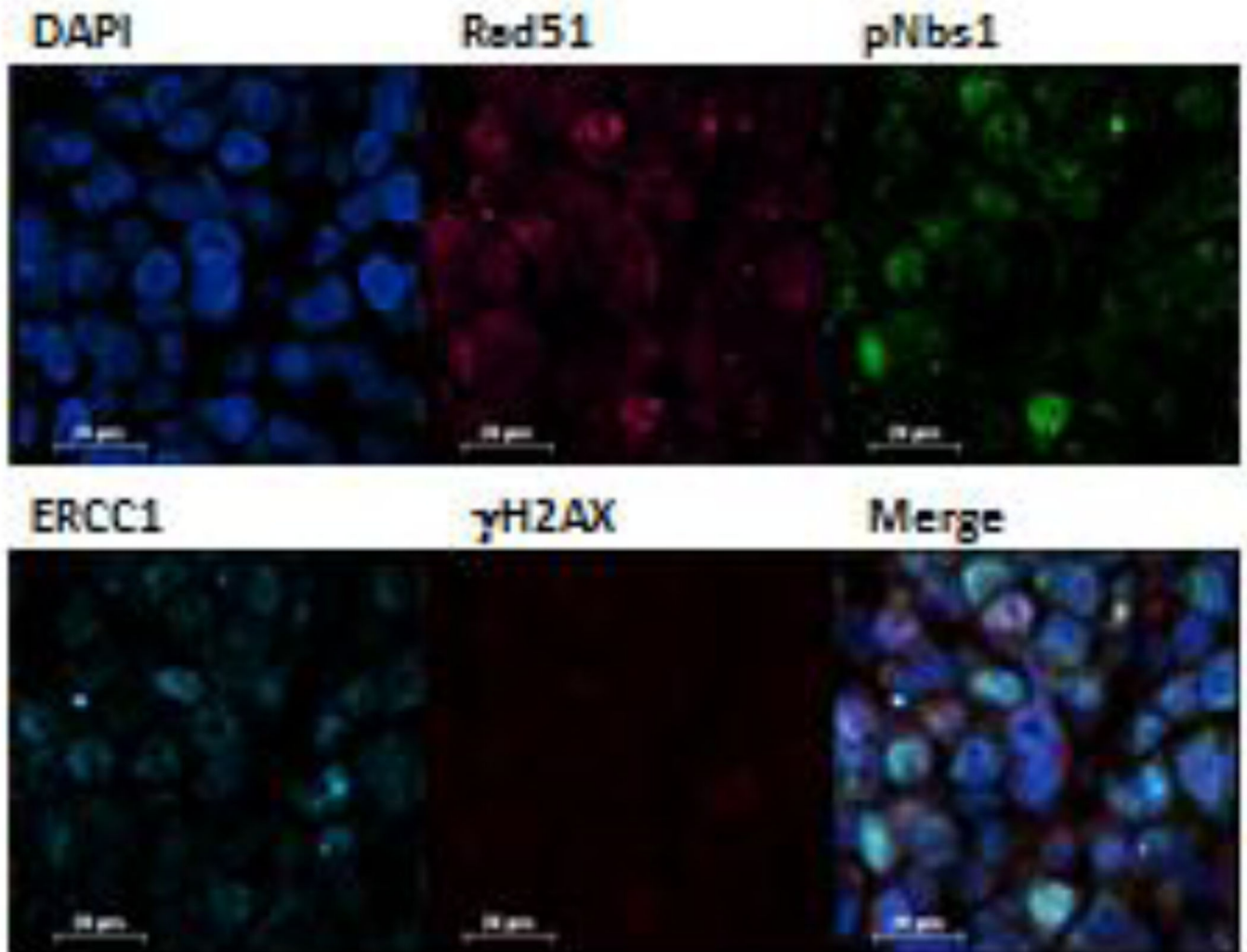
11. Angevin E, Lopez-Martin JA, Lin CC, Gschwend JE, Harzstark A, Castellano D, et al. Phase I study of dovitinib (TKI258), an oral FGFR, VEGFR, and PDGFR inhibitor, in advanced or metastatic renal cell carcinoma. *Clin Cancer Res.* 2013; 19:1257–1268. [PubMed: 23339124]
12. Akbani R, Becker KF, Carragher N, Goldstein T, de Koning L, Korf U, et al. Realizing the promise of reverse phase protein arrays for clinical, translational, and basic research: a workshop report: the RPPA (Reverse Phase Protein Array) society. *Mol Cell Proteomics.* 2014; 13:1625–1643. [PubMed: 24777629]
13. Lu Y, Ling S, Byers LA, Coombes K, Mills GB, Akbani R. Using Reverse Phase Protein Array (RPPA) as a Pharmacodynamic Assay for Functional Proteomics, Biomarker Discovery, and Drug Development in Cancer. *Semin Oncol.* 2015
14. Winters M, Dabir B, Yu M, Kohn EC. Constitution and quantity of lysis buffer alters outcome of reverse phase protein microarrays. *Proteomics.* 2007; 7:4066–4068. [PubMed: 17952872]
15. Mendes KN, Nicorici D, Cogdell D, Tabus I, Yli-Harja O, Guerra R, et al. Analysis of signaling pathways in 90 cancer cell lines by protein lysate array. *J Proteome Res.* 2007; 6:2753–2767. [PubMed: 17564428]
16. The University of Texas MD Anderson Cancer Center RPPA Core Facility. Advantages of function proteomics and reverse phase protein array. [accessed May 2015] available from <http://www.mdanderson.org/education-and-research/resources-for-professionals/scientific-resources/core-facilities-and-services/functional-proteomics-rppa-core/advantages/index.html>.
17. Theranostics Health. Information on reverse-phase protein array proteomic profiling. [accessed May 1, 2015] available from <http://www.theranosticshealth.com/proteomicservices/proteomic-profiling/>.
18. Huang YJ, Frazier ML, Zhang N, Liu Q, Wei C. Reverse-phase protein array analysis to identify biomarker proteins in human pancreatic cancer. *Dig Dis Sci.* 2014; 59:968–975. [PubMed: 24248418]
19. Konopleva MY, Walter RB, Faderl SH, Jabbour EJ, Zeng Z, Borthakur G, et al. Preclinical and early clinical evaluation of the oral AKT inhibitor, MK-2206, for the treatment of acute myelogenous leukemia. *Clin Cancer Res.* 2014; 20:2226–2235. [PubMed: 24583795]
20. Malinowsky K, Nitsche U, Janssen KP, Bader FG, Spath C, Drecoll E, et al. Activation of the PI3K/AKT pathway correlates with prognosis in stage II colon cancer. *Br J Cancer.* 2014; 110:2081–2089. [PubMed: 24619078]
21. Ummanni R, Mannsperger HA, Sonntag J, Oswald M, Sharma AK, Konig R, et al. Evaluation of reverse phase protein array (RPPA)-based pathway-activation profiling in 84 non-small cell lung cancer (NSCLC) cell lines as platform for cancer proteomics and biomarker discovery. *Biochim Biophys Acta.* 2014; 1844:950–959. [PubMed: 24361481]
22. Camp RL, Chung GG, Rimm DL. Automated subcellular localization and quantification of protein expression in tissue microarrays. *Nat Med.* 2002; 8:1323–1327. [PubMed: 12389040]
23. McCabe A, Dolled-Filhart M, Camp RL, Rimm DL. Automated quantitative analysis (AQUA) of in situ protein expression, antibody concentration, and prognosis. *J Natl Cancer Inst.* 2005; 97:1808–1815. [PubMed: 16368942]
24. Welsh AW, Moeder CB, Kumar S, Gershkovich P, Alarid ET, Harigopal M, et al. Standardization of estrogen receptor measurement in breast cancer suggests false-negative results are a function of threshold intensity rather than percentage of positive cells. *J Clin Oncol.* 2011; 29:2978–2984. [PubMed: 21709197]
25. Kinders RJ, Hollingshead M, Lawrence S, Ji J, Tabb B, Bonner WM, et al. Development of a validated immunofluorescence assay for gammaH2AX as a pharmacodynamic marker of topoisomerase I inhibitor activity. *Clin Cancer Res.* 2010; 16:5447–5457. [PubMed: 20924131]
26. Hewitt SM, Badve SS, True LD. Impact of preanalytic factors on the design and application of integral biomarkers for directing patient therapy. *Clin Cancer Res.* 2012; 18:1524–1530. [PubMed: 22422404]
27. LoRusso P, Ji JJ, Li J, Heilbrun LK, Shapiro G, Sausville EA, et al. Phase I study of the safety, pharmacokinetics (PK), and pharmacodynamics (PD) of the poly(ADP-ribose) polymerase (PARP) inhibitor veliparib (ABT-888; V) in combination with irinotecan (CPT-11; Ir) in patients (pts) with advanced solid tumors [abstract]. *J Clin Oncol (Meeting Abstracts).* 2011; 29:3000.



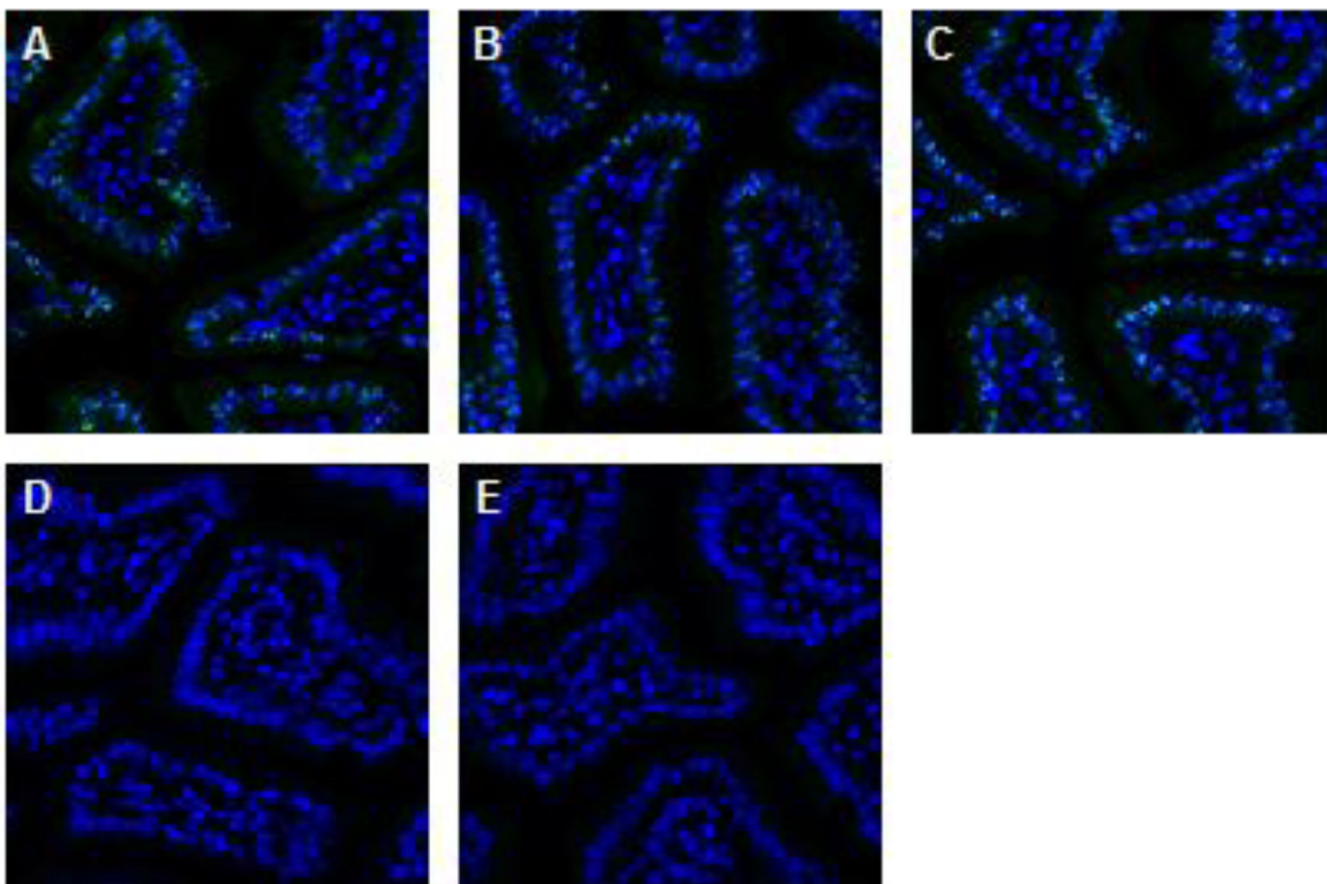
28. Redon CE, Nakamura AJ, Sordet O, Dickey JS, Gouliava K, Tabb B, et al. gamma-H2AX detection in peripheral blood lymphocytes, splenocytes, bone marrow, xenografts, and skin. *Methods Mol Biol.* 2011; 682:249–270. [PubMed: 21057933]
29. Kinders RJ, Hollingshead M, Khin S, Rubinstein L, Tomaszewski JE, Doroshow JH, et al. Preclinical modeling of a phase 0 clinical trial: qualification of a pharmacodynamic assay of poly (ADP-ribose) polymerase in tumor biopsies of mouse xenografts. *Clin Cancer Res.* 2008; 14:6877–6885. [PubMed: 18980982]
30. Kinders R, Ferry-Galow K, Wang L, Srivastava AK, Ji JJ, Parchment RE. Implementation of validated pharmacodynamic assays in multiple laboratories: challenges, successes, and limitations. *Clin Cancer Res.* 2014; 20:2578–2586. [PubMed: 24831280]
31. Lee JW. Method validation and application of protein biomarkers: basic similarities and differences from biotherapeutics. *Bioanalysis.* 2009; 1:1461–1474. [PubMed: 21083095]
32. Kummar S, Kinders R, Gutierrez ME, Rubinstein L, Parchment RE, Phillips LR, et al. Phase 0 clinical trial of the poly (ADP-ribose) polymerase inhibitor ABT-888 in patients with advanced malignancies. *J Clin Oncol.* 2009; 27:2705–2711. [PubMed: 19364967]
33. Do KT, Wilsker D, Balasubramanian P, Zlott J, Jeong W, Lawrence SM, et al. Phase I trial of AZD1775 (MK1775), a wee1 kinase inhibitor, in patients with refractory solid tumors [abstract]. *J Clin Oncol (Meeting Abstracts).* 2014; 32:2503.
34. Grimm JB, English BP, Chen J, Slaughter JP, Zhang Z, Revyakin A, et al. A general method to improve fluorophores for live-cell and single-molecule microscopy. *Nat Methods.* 2015; 12:244–250. [PubMed: 25599551]



**Figure 1. DNA damage multiplex assay utilized on an A375 melanoma xenograft sample 2 hours after beginning treatment with topotecan at 1.5 mg/kg**  
False color assignments are as follows: Rad51 (pink), pS<sup>343</sup>-Nbs1 (green), ERCC1 (cyan), and γH2AX (red). A representative 60× image is shown. Cells expressing both pS<sup>343</sup>-Nbs1 (green) and γH2AX (red) are yellow on the merged image. Scale bar = 20 μm.

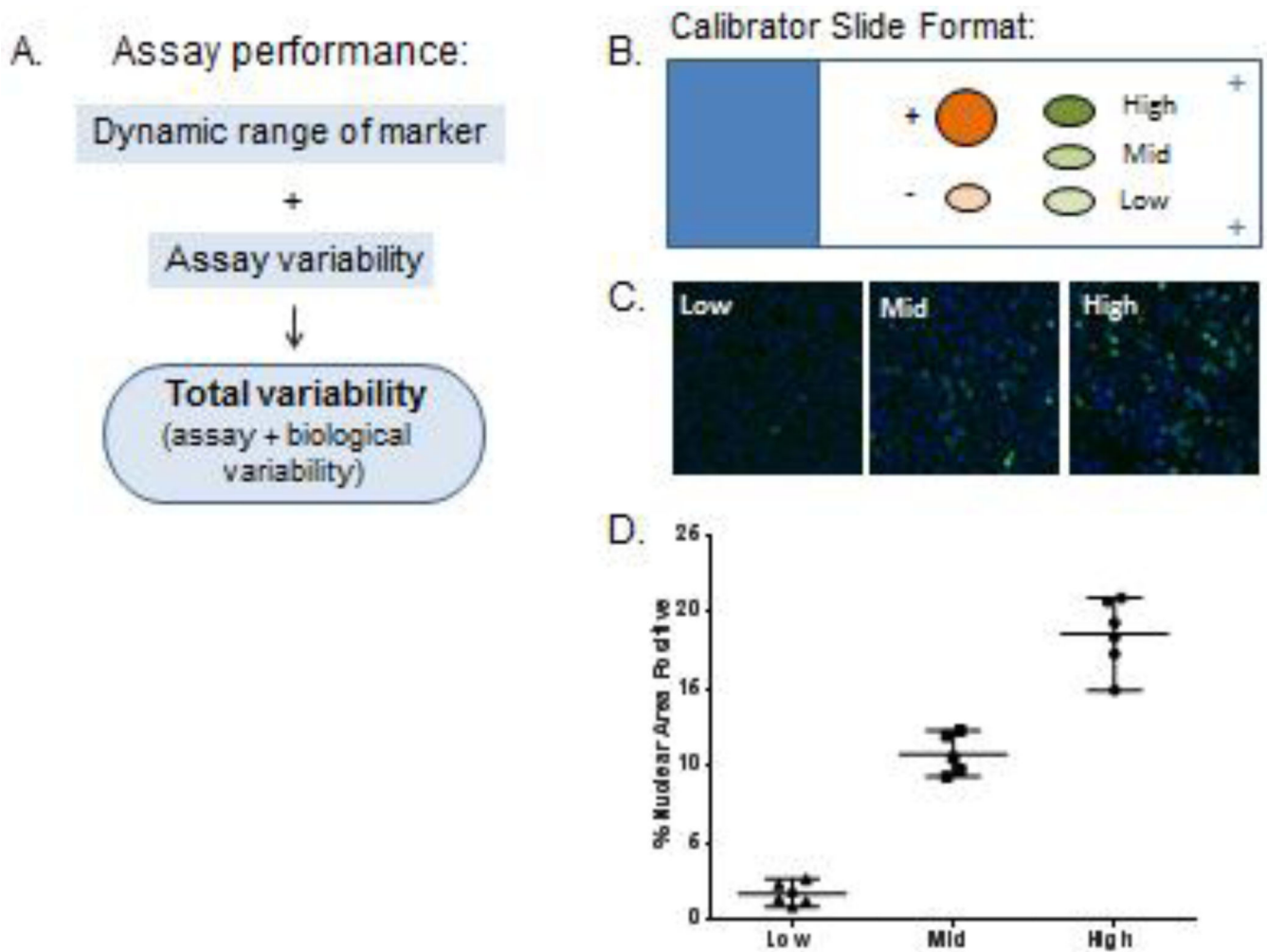


**Figure 2. DNA damage multiplex assay utilized on an MX-1 *BRCA*-deficient breast cancer xenograft 24 hours after beginning treatment with irinotecan at 7.5 mg/kg** False color assignments are as follows: Rad51 (pink), pS<sup>343</sup>-Nbs1 (green), ERCC1 (cyan), and  $\gamma$ H2AX (red). A representative 40 $\times$  image is shown. Cells expressing both pS<sup>343</sup>-Nbs1 (green) and  $\gamma$ H2AX (red) are yellow on the merged image. Scale bar = 20  $\mu$ m.



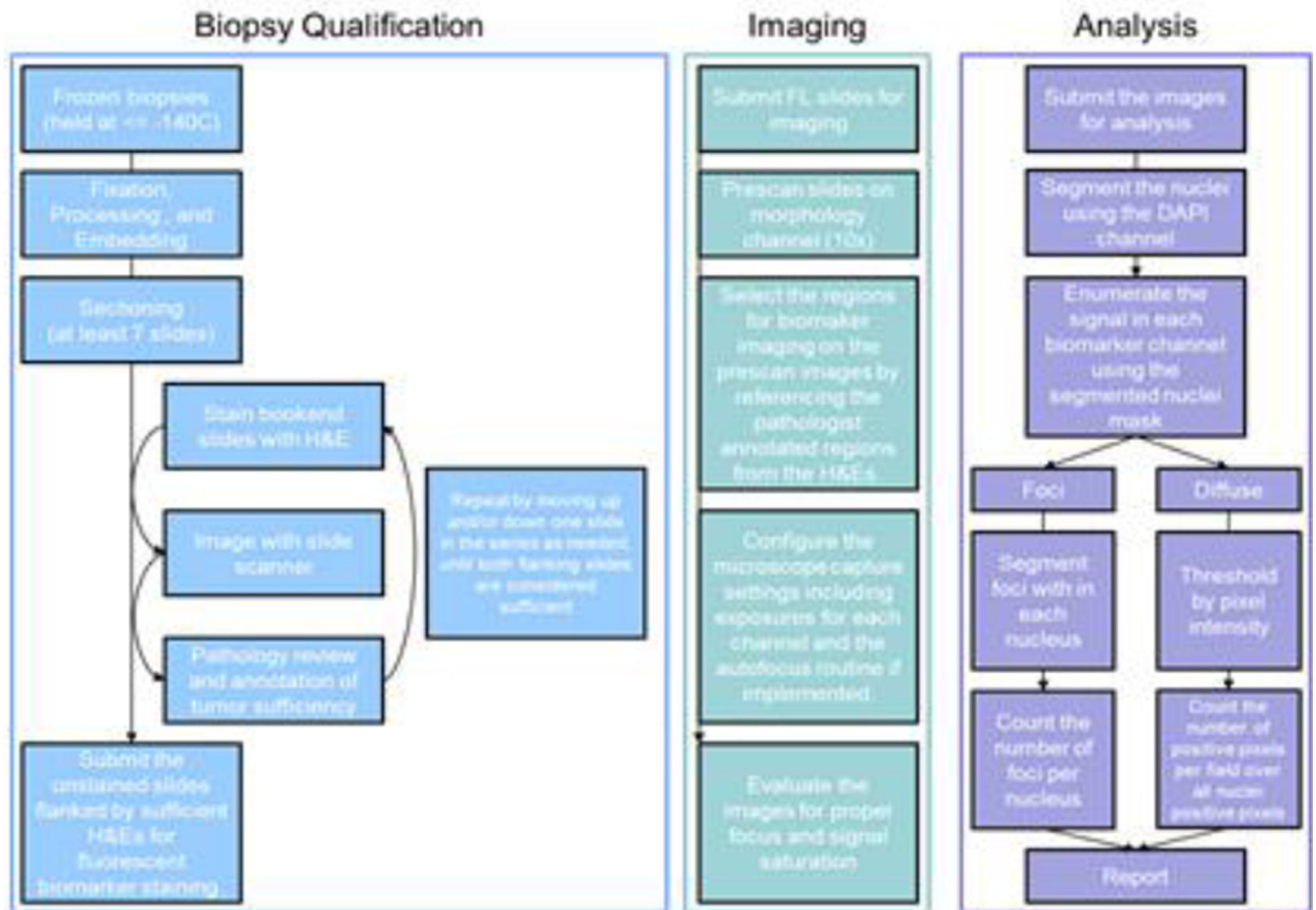
**Figure 3. Peptide competition assay demonstrates pS<sup>343</sup>-Nbs1 antibody specificity on the positive control tissue**

Representative images of formalin fixed, paraffin embedded mouse jejunum from serial cut slides. Tissue was stained with (A) 3 µg/ml of pS<sup>343</sup>-Nbs1 primary antibody, or pS<sup>343</sup>-Nbs1 protein incubated overnight with (B) peptide solution buffer, (C) 10 fold molar excess of the unphosphorylated peptide sequence or (D) 10 fold molar excess of the peptide sequence phosphorylated at S343. (E) A slide was also stained with a monoclonal rabbit isotype control. Images were extracted from a 20× Aperio scan with DAPI (blue) and pS<sup>343</sup>-Nbs1 (green) shown.



**Figure 4. Establishment of assay performance and variability**

(A) Factors that influence assay performance. (B) Diagram of typical calibrator slide layout for an assay measuring a single marker. Positive and negative control tissue are designated “+” and “-”, respectively. Low, mid and high range calibrator tissue depicted with ovals. (C) Representative images of calibrator tissue for a nuclear marker. (D) Representative plot of assay performance and variability of the calibrator material used to define the specifications. Mean and one standard deviation plotted for each calibrator along with individual values.



**Figure 5. Biopsy specimen workflow**  
Major steps performed from biopsy receipt through data reporting are highlighted. The processes are designed for 18-gauge core biopsies.

**Table 1**

List of equipment and software employed in image acquisition

<b>Equipment</b>	<b>Manufacturer</b>	<b>Model Number</b>
Wide Field Microscope	Nikon	Ti-E
Confocal Microscope	Nikon	A1
Objective – 10×	Nikon	CFI Plan Fluor 10×
Objective – 20×	Nikon	CFI Plan Apo VC 20×
Objective – 40×	Nikon	CFI Plan Apo Lambda 40×
Camera	Andor	DU-888
Light Source	Sutter	Lambda LS
Filter Cube (ET – DAPI)	Chroma	49000ET
Filter Cube (ET – eGFP)	Chroma	49002ET
Filter Cube (ET – R&B Phycoerythrin)	Chroma	49010ET
Filter Cube (ET – Cy5)	Chroma	49006ET
Filter Cube (ET – Cy7)	Chroma	49007ET
Automation Software	Nikon	NIS-Elements + JOBS
Whole Slide Scanner Brightfield	Leica	ScanScope AT Turbo
Whole Slide Scanner Fluorescence	Leica	ScanScope FL

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript