



Best Practices for Technical Reproducibility Assessment of Multiplex Immunofluorescence

Caddie Laberiano-Fernández, Sharia Hernández-Ruiz, Frank Rojas and Edwin Roger Parra*

Department of Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, United States

OPEN ACCESS

Edited by:

Joe Yeong,
Institute of Molecular and Cell Biology
(A*STAR), Singapore

Reviewed by:

Dong Ren,
The First Affiliated Hospital of Sun
Yat-sen University, China
Brian Beliveau,
University of Washington,
United States

*Correspondence:

Edwin Roger Parra
erparra@mdanderson.org

Specialty section:

This article was submitted to
Molecular Diagnostics and
Therapeutics,
a section of the journal
Frontiers in Molecular Biosciences

Received: 29 January 2021

Accepted: 11 August 2021

Published: 31 August 2021

Citation:

Laberiano-Fernández C,
Hernández-Ruiz S, Rojas F and
Parra ER (2021) Best Practices for
Technical Reproducibility Assessment
of Multiplex Immunofluorescence.
Front. Mol. Biosci. 8:660202.
doi: 10.3389/fmolb.2021.660202

Multiplex immunofluorescence (mIF) tyramide signal amplification is a new and useful tool for the study of cancer that combines the staining of multiple markers in a single slide. Several technical requirements are important to performing high-quality staining and analysis and to obtaining high internal and external reproducibility of the results. This review manuscript aimed to describe the mIF panel workflow and discuss the challenges and solutions for ensuring that mIF panels have the highest reproducibility possible. Although this platform has shown high flexibility in cancer studies, it presents several challenges in pre-analytic, analytic, and post-analytic evaluation, as well as with external comparisons. Adequate antibody selection, antibody optimization and validation, panel design, staining optimization and validation, analysis strategies, and correct data generation are important for reproducibility and to minimize or identify possible issues during the mIF staining process that sometimes are not completely under our control, such as the tissue fixation process, storage, and cutting procedures.

Keywords: reproducibility, standardization, analytical evaluation, clinical application, multiplex immunofluorescence

INTRODUCTION

Multiplex immunofluorescence (mIF) tyramide signal amplification (TSA) is a new and useful tool for the study of cancer that combines the staining with multispectral imaging analysis technology, allows the design of mIF panels for up to six biomarkers, characterizes the co-expression of markers (cell phenotypes), and quantifies these markers overall with the use of a nuclear counterstain (DAPI) in a single slide (Parra et al., 2019; Francisco-Cruz et al., 2020b). Different mIF panels can be created using this technology to study the tissue microenvironment. The multispectral fluorescence microscope, along with the combined markers and individual fluorophores, is used to create a multispectral image that facilitates the analysis. By incorporating image analysis software, the images generated by the scanners can be easily analyzed and the cellular populations quantified (Parra et al., 2021a). mIF facilitates assessments at the cellular level of different proteins, as well as their spatial arrangement, and thus enables precision medicine in immuno-oncology, translational research, and clinical practice by elucidating the immune response of the human body to diverse tumors and showing differences in the pre- and post-treatment tissue.

Using mIF, it is possible to study the co-expression between markers to identify distinct cell populations and pathways and their relationships in different tissues and in turn to determine their roles in clinical outcomes (Parra, 2021). In that way, targetable biomarker pathways, such as PD-1/PD-L1, can be studied to verify the effect of immune therapies in the tumor microenvironment and

their clinical benefit (Velcheti et al., 2014; Schalper et al., 2015; Parra et al., 2018a; Barua et al., 2018). This technology therefore has an important role in translational oncology research (Stauber et al., 2010; Steiner et al., 2014; Sood et al., 2016; Rost et al., 2017; Gorris et al., 2018) and facilitating our understanding of the disease (Blom et al., 2017; Hofman et al., 2019). mIF also has applicability for diseases other than cancer, and it is well suited for prognostication at early stages of pathogenesis, when key signaling protein levels and activities are perturbed (Dejima et al., 2021). On the clinical side, there is high demand to incorporate mIF in a Clinical Laboratory Improvement Amendments (CLIA) certified as an innovative tool for diagnosis and prognosis.

The mIF-TSA workflow starts with antibody selection, optimization, and validation and ends with a digital image analysis (Parra et al., 2020). It is important to refine, standardize, optimize, and validate the end-to-end workflow in mIF to obtain reproducible results to support large-scale multi-site trials and individual principal investigator projects and to enable their possible clinical application.

The reproducibility of results remains the cornerstone of modern science (Hewitt, 2016). Given reproducible results, considering possible technical and human problems, with adequate protocols, each laboratory or institution can proceed in the same direction, using published experiences as a reference. Pre-analytic, analytic, and post-analytic variables that may influence reproducibility, quality, and staining procedure should be considered (Rojo et al., 2009; Okoye and Nnatuanya, 2015; Rudbeck, 2015; Meyerholz and Beck, 2018). Most of the descriptions related to these variables are focused on immunohistochemistry (IHC) on the basis of a study by Engel and others, who recognized more than 60 variables in the pre-analytic stage alone (Engel and Moore, 2011) and some variables which can be considered are pre-fixation, reagent conditions, and slide preparation, but those same variables can also be applied for IF and mIF.

It was recognized over a decade ago that standardization is vital for reproducible and reliable results in IHC (Goldstein et al., 2007). Agencies such as the Biological Stain Commission, Clinical and Laboratory Standards Institute, The U.S. Food and Drug Administration, and the manufacturing sector have established guidelines, standards, and recommendations for reagents and package inserts (Taylor, 1992; Taylor, 1998; Taylor, 1999; Goldstein et al., 2007). Although all of this effort has improved the quality of IHC, most of the causative responsibility rests with the individual laboratory performing the analysis, specifically the lack of standardization and attention to quality assurance programs (Rhodes, 2003; Varma et al., 2004; Goldstein et al., 2007).

CLIA requirements for determining test performance specifications apply to all laboratory tests. All the improvements related to reproducibility can positively affect the CLIA evaluation. For IHC assays, accuracy, analytic sensitivity, and specificity are determined by analytic assay validation, which is theoretically achieved by testing a validation tissue set against a gold standard (Fitzgibbons et al., 2014). In the last year, we saw an increase in the use of this

technique but the requirement aspects to be reproducible are not well established between the different centers and research groups. There are also few manuscripts about mIF reproducibility (Akturk et al., 2021; Taube et al., 2021) which have been published; thus, it is important to compare the results directly.

In the present article, we review and describe the difficulties in the reproducibility of the main workflow-related steps of the mIF technique and how to optimize the process.

PRE-ANALYTIC EVALUATION

To develop a reproducible mIF imaging platform, several technical requirements must be met: 1) rigorous tissue quality controls, 2) a balanced multiplex assay staining format, 3) the ability to quantitate multiple markers in a defined region of interest (considering a minimum number of areas selected), and 4) experimental reproducibility, both internally and across different laboratories (Shipitsin et al., 2014).

For all these considerations, the IHC and mIF staining and imaging protocols must be standardized, automated, and validated. Being able to adapt IHC workflows in mIF without extensive re-optimization saves time and avoids human error, making it useful for translational research and future clinical applications (Tumeh et al., 2014; Giraldo et al., 2018; Tan et al., 2020).

ANTIBODY SELECTION, OPTIMIZATION, AND CONTROLS GUIDING REPRODUCIBILITY

The staining protocol for mIF can begin with the selection of the antibodies and their optimization by IHC or IF according to the experience and confidence of the pathologist, especially when starting with IF instead of IHC (Carvajal-Hausdorf et al., 2015). In that way, the antibody selection for mIF panel design can be considered the first step for developing a panel and needs to be done by a multidisciplinary team, including pathologists, oncologists, and immunologists. Some antibodies can be selected because of their clinical implications, while other antibodies, such as those targeting immune checkpoint markers (Francisco-Cruz et al., 2020a) may be selected to answer specific scientific or research questions. Then, choosing the correct antibody's clones and their optimization by IHC or IF is crucial to detect specific epitopes. In parallel, the selection of correct controls, negative or positive, is essential to the valid interpretation of the staining (Engel and Moore, 2011), and it is one aspect by which methods can be systematically assessed in consecutive multiplexed assays to confirm reproducibility (Canadian Association of Pathologists-Association canadienne des pathologistes National Standards Committee et al., 2010; Stack et al., 2014). For antibody selection, each antibody's clonality must be considered regarding its advantages and disadvantages (**Table 1**). Monoclonal antibodies are often preferred for IHC and IF because of their higher specificity

TABLE 1 | Advantage and disadvantages of polyclonal and monoclonal antibodies.

| — | Antibody's type | | |
|--------------|--|--|--|
| | Polyclonal | Monoclonal | Recombinant |
| Advantage | <p>Low cost to produce</p> <p>Quick turnaround time from antigen preparation to antibody harvesting</p> <p>Ability to detect multiple epitopes on an antigen</p> <p>High affinity and sensitivity to detect low quantity proteins</p> <p>Preferred for detection of denatured proteins</p> | <p>Homogeneity is conserved between batches to ensure reproducible results</p> <p>High specificity for single epitope</p> <p>Less background</p> <p>Specificity of monoclonal antibodies make them efficient</p> <p>Cross-reactivity with other molecules is reduced</p> | <p>Improved reproducibility and control</p> <p>Antibodies can be produced rapidly</p> <p>No host animals are need</p> <p>Easier isotype conversion</p> |
| Disadvantage | <p>Higher tolerance for differences in antigen</p> <p>Variability in each batch</p> <p>Non-specific antibody</p> <p>Multiple epitopes cause high chance of cross-reactivity resulting in higher background</p> | <p>Significantly more expensive to produce</p> <p>Require more specialized training to create and have a much longer turnaround time</p> <p>Cover only one epitope</p> <p>More sensitive to buffer conditions</p> | <p>High cost to develop and produce</p> <p>High degree of technical skills of the professionals is required</p> |

and reproducibility and lower background and lot-to-lot variability. They are usually generated against unique peptides of the target antigen, located in regions that are less affected by formalin fixation. In contrast, polyclonal antibodies bind to different epitopes on the same protein and are obtained from experimental animals through repetitive stimulation of the antigen. Finally, recombinant antibodies, produced by recombinant DNA technology, should also be considered.

Another aspect to evaluate is the potential impact of antibody sensitivity and specificity during the optimization process considering the antibodies must be verified by the user (Taylor, 1999). Besides, in the optimization process, staining intensity can be modified according to the results of a pre-analytic study, which may be affected by methodological variables such as tissue fixation, antibody specificity and dilution, antigen retrieval duration and type, and detection systems (Ng et al., 2018). For this reason, it is crucial to compare samples using external or internal control. While cell lines are useful for testing individual markers and defining their expression level, they are not completely appropriate to use as positive controls; the most rigorous are tissue controls (Hewitt et al., 2014), which can contain multiple proteins, unlike pure cell line preparations. In addition, negative controls are used to demonstrate that the reaction visualized is a result of the interaction of the epitope of the target molecule and the paratope of the antibody or affinity reagent, demonstrating the specificity of the antibody (Hewitt et al., 2014) during the run staining. Although antibodies must be prepared according to the vendors' instructions, the experience of laboratory members, under pathologist supervision, is important to determine optimal staining conditions and correct marker expression as part of quality control **Figure 1**). In this regard, the primary antibody should be titrated to an appropriate concentration that retains the specificity of the stain while removing any background signal or non-specific staining of the tissue. Antibodies that are prepared at a too high concentration can result in off-target staining (Anagnostou et al., 2010; Toki et al., 2017); an optimal

concentration results in better accuracy and reproducibility (Toki et al., 2017; Taube et al., 2020). The adequate expression must be tested because some markers are able to stain more than one compartment of cells or other types of cells (e.g., PD-L1 could have cytoplasmic expression, but only membrane expression is considered positive staining, and it could be expressed in inflammatory cells besides the malignant cells) (Parra and Hernández Ruiz, 2021a) (**Figure 2**).

STRATEGIES FOR ANTIBODY VALIDATION

One of the key factors for mIF panel reproducibility is to use antibodies that have been thoroughly optimized and validated for their application in research studies or for clinical applications. After antibody optimization by IHC or IF in control tissues, a good practice is applying those antibodies in a set of different tissues and organs including different common cancer types contained in tissue microarrays (TMAs) for quantitative measurement and antibody testing and validation. Although the construction of TMAs is often expensive for some laboratories (Taube et al., 2020), it is highly recommended to test the antibodies that will be integrated with an mIF panel in at least a set of cases for validation purposes, as a minimum requirement (Parra et al., 2017). The International Working Group for Antibody Validation proposed in total five different “pillars” to use for antibody validation with 1) genetic, 2) orthogonal, 3) independent antibody strategies, 4) expression of tagged proteins, and 5) immunocapture followed by mass spectrometry. It is recommended to consider at least one of these pillars as a minimum criterion for claiming that a selected protein has been adequately valid for a particular application (Uhlen et al., 2016). The most common and mainstay strategies are the orthogonal and the independent antibody strategy (Sivertsson et al., 2020). In the case of orthogonal validation (the most common), for an mIF panel validation, we use a non-antibody-based method to identify any effects or artifacts that

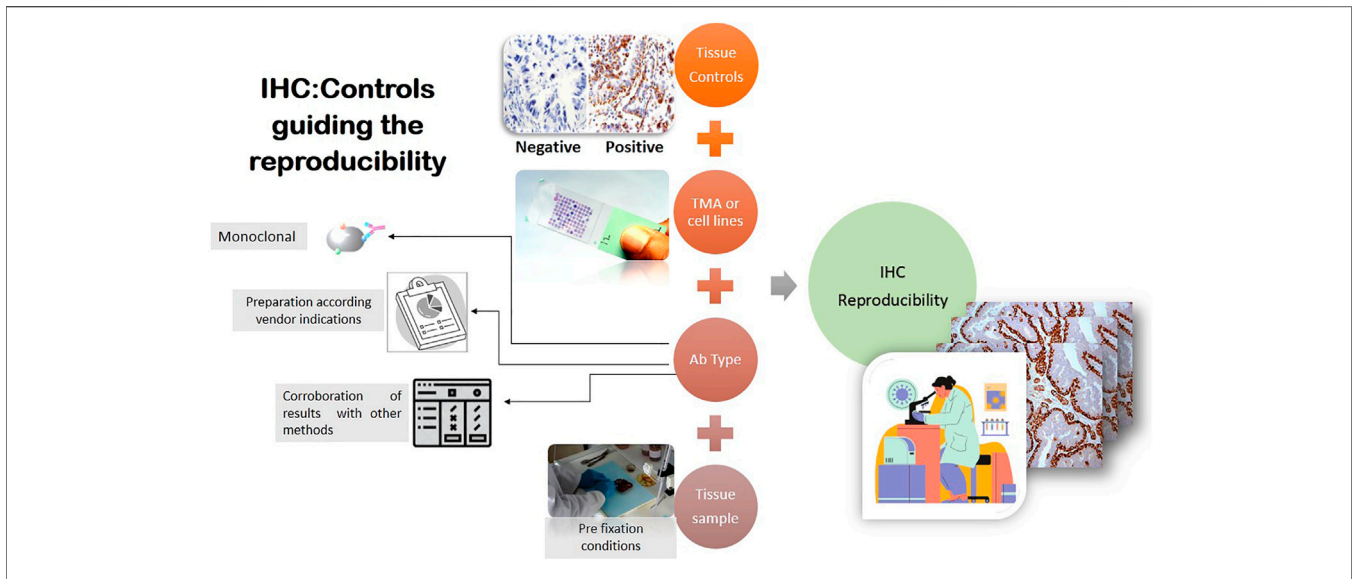


FIGURE 1 | One of the most important steps in obtaining reproducibility in mIF is evaluating antibodies in IHC. Different factors must be considered to provide the best results; studies of tissue controls, TMAs, or cell lines are needed to analyze the staining of each marker using vendors’ instructions or via corroboration by other methods, such as Western blot analysis. The pre-analytical process can also affect the marker expression results; all of these factors together are part of IHC reproducibility.

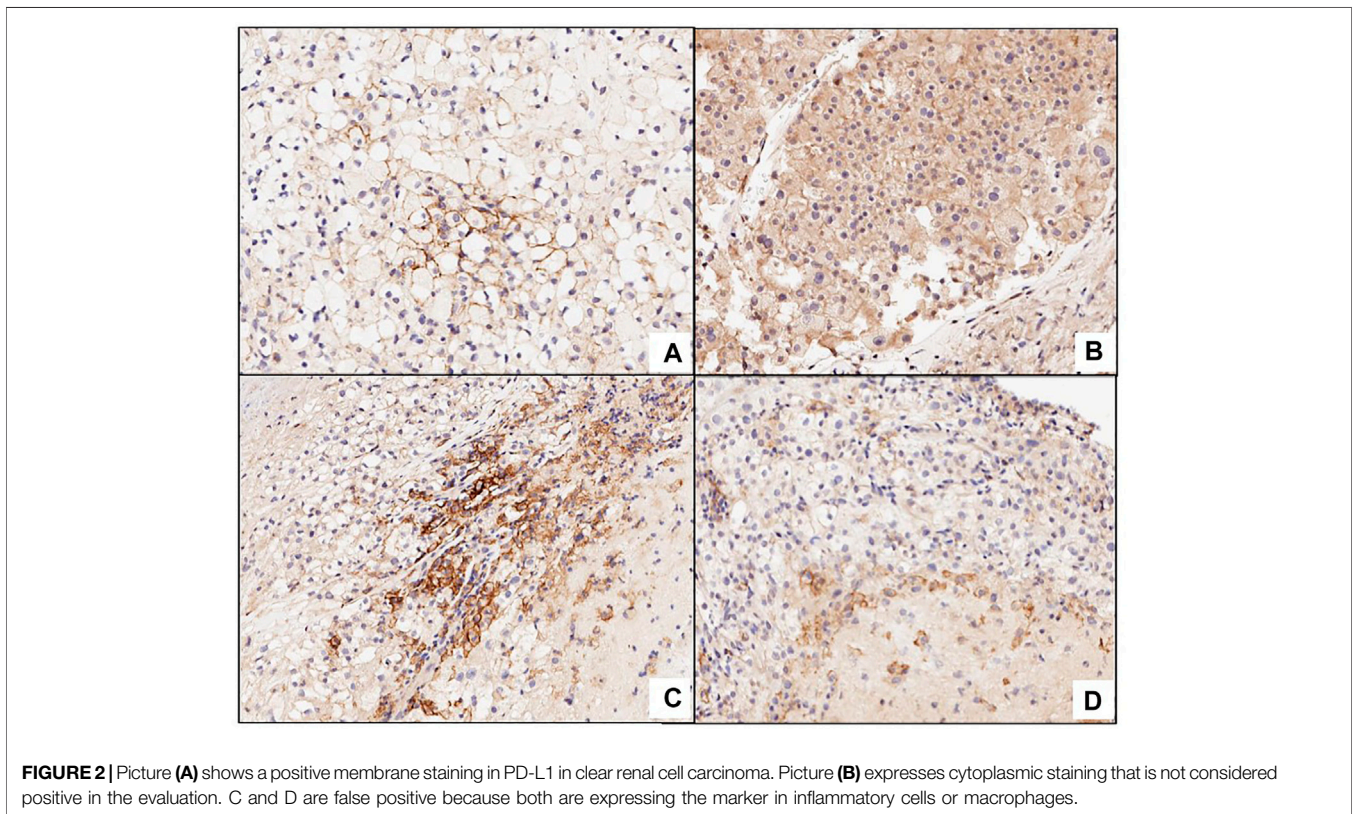


FIGURE 2 | Picture (A) shows a positive membrane staining in PD-L1 in clear renal cell carcinoma. Picture (B) expresses cytoplasmic staining that is not considered positive in the evaluation. C and D are false positive because both are expressing the marker in inflammatory cells or macrophages.

are directly related to the antibody or panel in question (Sivertsson et al., 2020). Depending on the antibodies targeted in a panel, non-antibody-based methods can include mining

previously published results. Overall, it is possible studying expression analysis *via* genomics, transcriptomics, and proteomics techniques; or employing other established

TABLE 2 | Strategies and methods for antibody/multiplex immunofluorescence panel validation.

| Strategy | Method | Advantage | Disadvantage |
|---|--|---|--|
| Genetic | <i>In situ</i> hybridization (ISH), CRISPR/CAS9 or siRNA/shRNA, Western blot | <ul style="list-style-type: none"> - Novel genes in spatial context - The use of genome editing techniques is preferred - Provide a direct link between the gene, the target protein, and its detection by the antibody - Useful for examining antibody specificity for proteins that come from related genes | <ul style="list-style-type: none"> - Limited co-expression - Need functional knockdown reagents - Cannot be used for human tissue samples and body fluids (plasma and serum) - Time-consuming |
| Orthogonal | Fluorescent <i>in situ</i> hybridization (FISH), quantitative PCR, RNA-seq, Western blot | <ul style="list-style-type: none"> - Expression of the target protein is compared with an antibody-independent method - Co-expression in spatial context | <ul style="list-style-type: none"> - Limited probes and parameters - Need differential expression of target protein |
| Independent antibody | Immunofluorescence imaging, Immunohistochemistry, Western blot | <ul style="list-style-type: none"> - Co-expression can be in spatial context - The data generated using several antibodies (different epitopes) in the same protein is compared | <ul style="list-style-type: none"> - Limited parameters - Need antibodies with different epitopes |
| Tagged protein expression | Immunohistochemistry, Western blot | <ul style="list-style-type: none"> - Novel target in spatial context - Tagged proteins should be expressed at endogenous levels | <ul style="list-style-type: none"> - Limited co-expression - Overexpression of the target protein might mask the detection of off-target binding events - Limitations of this method are similar to those of the genetic approaches - Avoid potential artifacts introduced by the tag itself |
| Immunocapture followed by mass spectrometry | Immunoprecipitation, chromatin immunoprecipitation | <ul style="list-style-type: none"> - Fast, easily co-expression - This is one of the best methods for identifying off-target protein binding | <ul style="list-style-type: none"> - Many proteins have similar size - Difficulty in distinguishing direct interactors with the antibody versus proteins that form relevant complexes with the target protein - Some of the antibodies validated still do not perform in immunofluorescence assays |

antibody-independent methods such as *in situ* hybridization or RNA sequencing. This strategy can also be used to ensure that any antibody validation performed in-house uses the most relevant biological models for the targets of interest. Although immunostaining techniques that are established in a lab, such as Western blot (Parra et al., 2018b), in positive and negative cell lines (Bordeaux et al., 2010) for research antibodies, can help provide a quick visual indication of antibody specificity (Parra et al., 2020; Parra and Hernández Ruiz, 2021b), it is always important and recommended that the antibody's data generated be supported by orthogonal testing. One way of achieving this is to mine publicly available databases (e.g., CCLE, BioGPS, Human Protein Atlas, DepMap Portal, COSMIC) for genomic and transcriptomic profiling information to clarify whether observed immunostaining results are relevant or are instead due to antibody-related artifacts (Cell Signaling, 2019; Ghandi et al., 2019; Broad Institute, 2021; COSMIC, 2021; The Human Protein Atlas, 2021).

About the independent antibody strategy, this is characterized by the use of independent antibodies, defined as a similar expression pattern determined by an independent antibody targeting a non-overlapping region of the similar protein (Sivertsson et al., 2020). Two or more independent antibodies

that acknowledge a similar target may be used to assess antibody specificity in a range of assays. This approach requires that the expression patterns generated by the two antibodies correlate within a given application environment, which means that the two antibodies are able to bind to totally different regions of the protein and thus have different epitopes, minimizing the likelihood of off-target binding to a similar unrelated protein (Uhlen et al., 2016). Although diverse techniques can be used for antibody validation according to the necessities of the studies as described above, it is important to consider, when choosing one, its advantages and disadvantages, which are described in **Table 2**

MIF OPTIMIZATION AND CONTROL SELECTION

mIF panel development is essentially the consolidation of a single IF protocol in a multiplex protocol (Taube et al., 2020); it should ideally be performed using tissues with a full range of known expression patterns for the targets of interest, using the same positive and negative controls as described above for antibody optimization and validation. Careful project design is mandatory, as well as choosing correct, reliable, and very well optimized

antibodies to create a panel; other important variables for optimizing results include fresh tissue sections and regular or thin tissue slices (maximum, 4 μm) and adequately charged slides to avoid tissue detachment. It is important to use very well-known control tissues during each run of staining to detect possible errors in the mIF panel; for example, human reactive tonsil is frequently used during mIF optimization because we know the exact distribution of its different cell populations (Parra et al., 2017). Although it has been demonstrated that we can design panels containing up to eight antibody targets (Parra et al., 2021b), the complexity of handling will increase with the number of markers introduced in a panel. For the pre-analytical step, it is also necessary to consider individual marker signals; the subcellular location of the targets' expression (nuclear, membrane, and cytoplasmic); optimization of antigen retrieval conditions (pH and temperature); reagent titration (e.g., primary antibody, secondary antibody, and fluorophores); incubation conditions (time and temperature); and blocking of non-specific binding, following similar rigor to that described in the antibody validation.

Besides the factors mentioned before, two important aspects remain. First, because TSA reagents covalently bind to sites surrounding the antigen, they can potentially inhibit the binding of a subsequent primary antibody through steric hindrance. This phenomenon is considered an umbrella effect and tends to occur in situations where multiple markers reside in a single cell compartment, such as a CD3⁺ CD8⁺ PD-1⁺ T cell, where all three markers are expressed on the cell membrane. It is conceivable that if CD3 and/or CD8 comes before PD-1 in the panel, sufficient tyramide could be deposited to block the PD-1 antigen. If present, this phenomenon might also be diagnosed when the evaluation to singleplex IHC/IF is performed. A useful strategy to determine antibody/fluorophore interference or blocking is the drop controls method to find which one is causing the interference (Surace et al., 2019). To correct this situation, we can increase the primary antibody concentration(s), reduce TSA fluorophore concentration(s), and/or change the order of targets in the panel (Taube et al., 2020).

The second aspect to consider is crosstalk, which is an additional signal from the non-target fluorophore captured by the microscopic system (60, (Arppe et al., 2017)). There are commonly recommended practices to cut back this effect; for example, crosstalk is often considerably reduced by choosing fluorophores whose excitation and emission spectra match those of the corresponding channels but minimally overlap those of non-corresponding ones. Alternatively, optimizing the filters of imaging channels, such as the adoption of excitation and emission filters with narrower bandwidths, can very effectively alleviate the crosstalk, although the signal strength might be sacrificed (Tie and Lu, 2020).

MULTISPECTRAL LIBRARY AND OPTICAL DETECTION

Multispectral libraries and their optical detection play an important role in determining the correct extraction of the photophore's signal according to their fluorescence

wavelength. Exposure times need to be set up carefully to maintain a balance of the signal intensity across markers in the panel (Parra et al., 2020). Because we are working with multispectral imaging, additional considerations required for capturing the images include the generation of a spectral library, which will facilitate the discrimination and capture of the individual fluorescence signal using the correct spectra from each fluorophore (Francisco-Cruz et al., 2020b; Parra et al., 2021b; Viratham Pulsawatdi et al., 2020). The creation of the spectral library with a single stained sample for each individual fluorophore corresponding primary antibody will be important for the signal extraction (Figure 3). Also recommended for signal extraction is a marker with a highly prevalent antigen such as CD20, anti-sodium potassium ATPase, or vimentin, as well as rechecking this spectral library regularly depending on whether the scanner system uses a fluorescence bulb or LED light sources for the excitation. Finally, it is important that the signal extraction is from exogenous and endogenous autofluorescence in this methodology (Francisco-Cruz et al., 2020b). Other components in the scanner systems used for acquiring the images that must be considered when choosing the scanner system are multispectral range, fluorescence throughput, automation, and multiplexing capability, among others, to obtain high-quality images (Table 3).

PANEL VALIDATION

The final validation of the mIF panel requires the performance of intra-site and inter-site reproducibility studies prior to clinical use (Taube et al., 2020). At this point, the same TMA as used in the antibody validation is an optimal material for validation purposes. The experience with IHC is diverse, according to the marker, without a universal consensus, because each marker is different; in mIF, this knowledge is still being developed. Although automated staining can give us high reproducibility and is recommended for mIF staining, manual staining can be considered to process small quantities of slides at the same time to avoid errors and antibody variability caused by manual manipulation (Parra et al., 2017). Finally, similar strategies as mentioned for antibody validation can be applied for panel validation.

ANALYTIC EVALUATION

Diverse factors could affect the pre-analytical step, as mentioned previously; reagents, autostainer performance, section thickness variation, scanner performance, and change in quality and quantity of the cells between serial sections can influence the mIF analysis (Lee et al., 2020). For an IHC or mIF assay to be considered validated, at a minimum, it must be demonstrated to be accurate and precise, as well as reproducible from an analytic perspective and on pathologist interpretation (Taube et al., 2020).

Marker evaluation is a key aspect of reproducibility. Markers with abundant and specific cell expression, such as

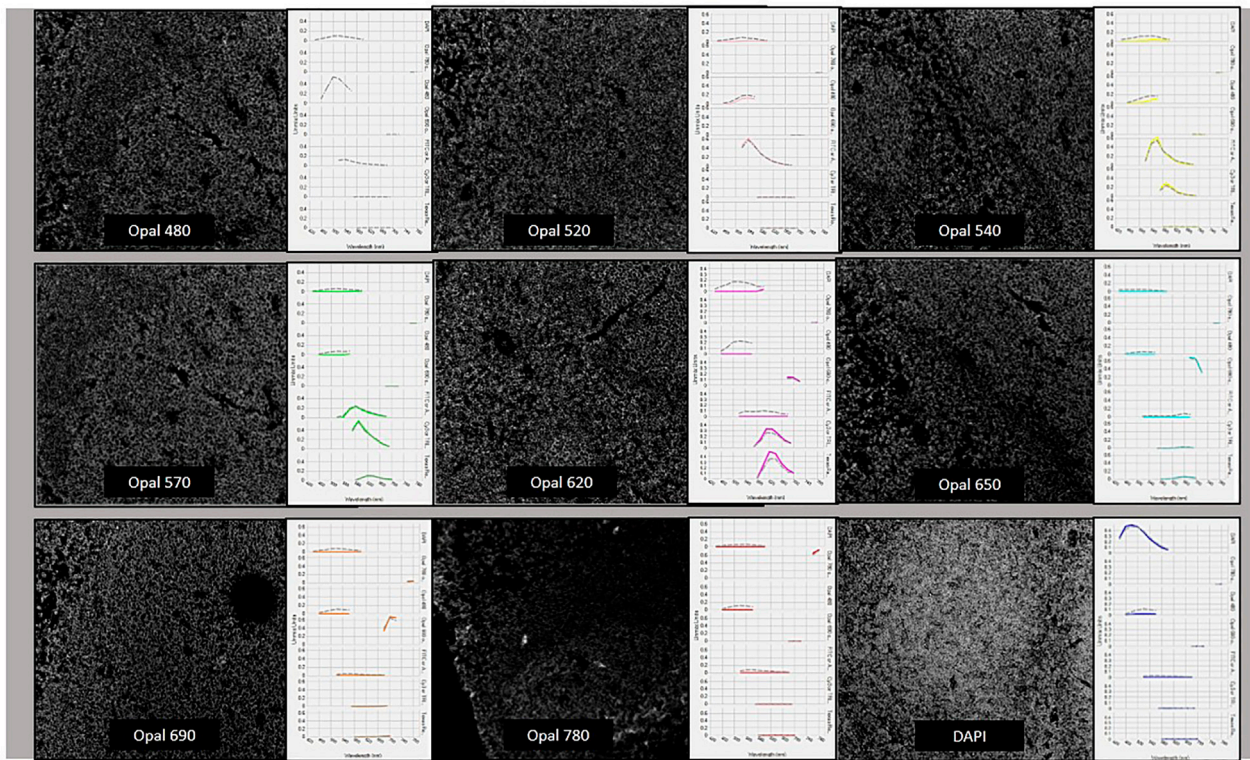
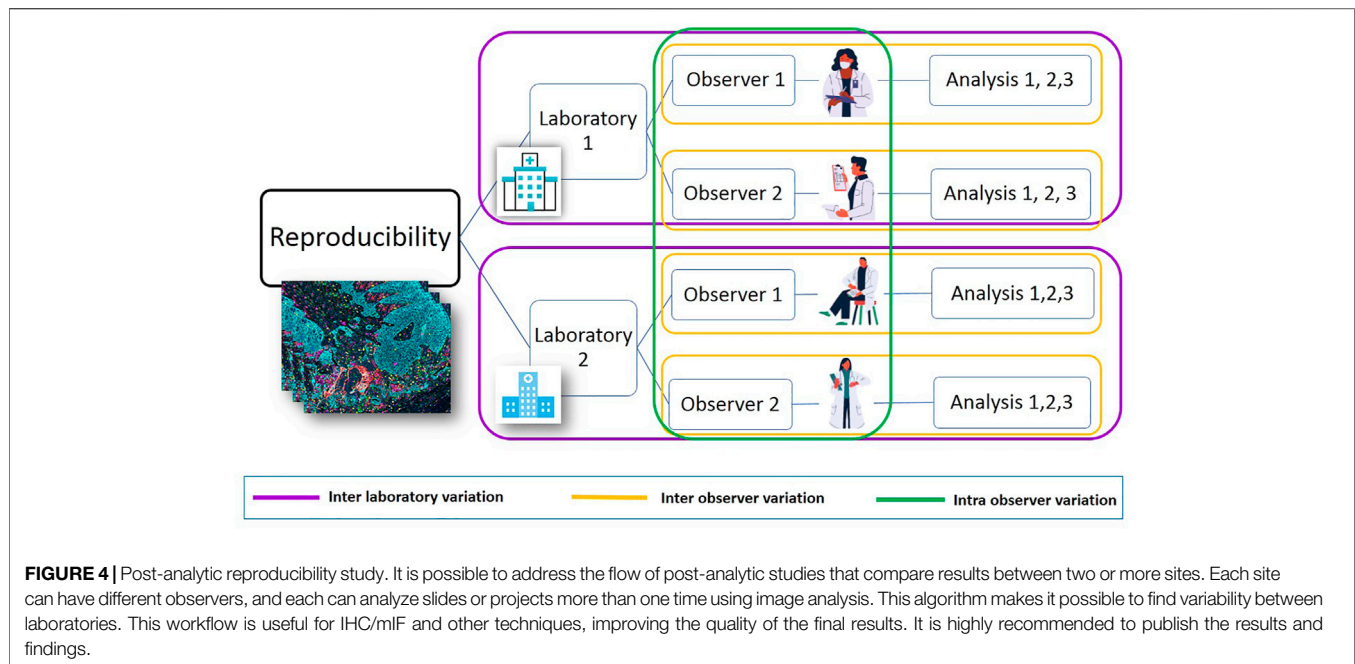


FIGURE 3 | Spectral library creation with the different fluorophores from the Opal-9 kit. Opal 480, Opal 520, Opal 540, Opal 570, Opal 620, Opal 650, and Opal 690, Opal 780, and DAPI were stained, optimized, and validated until we obtained similar dynamic ranges and specific wave peaks as is possible to see in this picture.

TABLE 3 | Differences scanners used for multiplex analysis.

| Company | Scanner type | Image acquisition and scanning instrument | Corporate location/notes | Resolution | Image extraction | File type | Automatization |
|---------------------------|------------------------------|---|---|--|------------------|--|--|
| Leica Biosystems | BF and FL | Aperio Versa | Illinois, United States | 0.468 μm per pixel with $\times 20$ objective. 0.2 μm at 40x | WS | TIFF, JPEG | Semi-auto and auto |
| 3DHistech | BF and FL | Pannoramic 250 FLASH III | Budapest, Hungary | 0.172 and 0.087/0.325 and 0.162 pixels | WS | --- | Auto |
| Ventana/Roche PerkinElmer | BF and FL MSI (BF and FL) | iScan Vectra/Vectra Polaris | United States, International Boston, United States | 24-bit true color 10 \times (1.0 $\mu\text{m}/\text{pixel}$), 20 \times (0.5 $\mu\text{m}/\text{pixel}$) and 40 \times (0.25 $\mu\text{m}/\text{pixel}$) | WS ROI | TIFF/BIF QTIFF, IM3, JPEG, single-layer TIFF, BMP, PNG | Auto and manual Touchless automation with walk-away image acquisition |
| Olympus America | BF and FL | VS110, Nanozoomer (United States) | Japan, International | 0.32 $\mu\text{m}/\text{pixel}$ (20 \times /NA 0.75) - 0.16 $\mu\text{m}/\text{pixel}$ (40 \times /NA 0.95) | WS | Compress images and save images in different file formats | Auto |
| Zeiss | BF and FL | AxioVision MosaiX | United States, Germany | 5 \times (2.11 $\mu\text{m}/\text{pixel}$), 10 \times (1.05 $\mu\text{m}/\text{pixel}$), and 20 \times (0.53 $\mu\text{m}/\text{pixel}$) | WS | AVI, BMP, J2K, JP2, JPG, LSM, MOV. PCT, PCX, PNG, PSD, TGA, TIF, WMF | Auto/manual |

BF: bright field; FL: fluorescence; MALDI: matrix-assisted laser desorption/ionization; FOV: field of view; WS: whole section; ROI: region of interest.



CD3, are easy to evaluate and will probably be consistent across serial sections when the expression is evaluated. For markers with variable geographic distribution across tissues and variable tumoral expression, such as PD-L1, reproducibility will be more challenging across serial sections (Parra et al., 2018b). To determine the reproducibility of markers in the mIF panel, we must consider that a group of markers is being evaluated and that those markers have specific cell phenotypes (marker co-expression) across different sections, according to the abundance of specific cell phenotypes (Lee et al., 2020). Marker reproducibility studies are easier in IHC compared with mIF because the evaluation is performed one by one; in mIF, it is harder to evaluate an entire panel using only one method, so the variability is related to the number of markers and their expression is extensive when specific phenotypes are evaluated.

Another drawback for mIF is high inter-observer variability for the same marker (Gerdes et al., 1984; Vincent-Salomon et al., 2007; Mohammed et al., 2012; Munzone et al., 2012; Cheng et al., 2015; Matsumoto et al., 2015). For instance, Ki-67 is a widely endorsed marker for a range of cancers Tumeh et al. (2014), but an issue has been raised concerning the reproducibility of IHC for Ki-67 and the implications of variability in clinical decision-making (Curigliano et al., 2017). Multiple research groups have demonstrated that inter-observer variability can be negated using digital analysis (Tan et al., 2020). There are different ideas as to the causes of between-pathologist variation; it may be the result of differences in each pathologist's clinical experience and technological competence (Barnes et al., 2017). In this case, the best approach may be to create a protocol of interpretation, with a consensus across all the groups. It will be useful to perform an objective analysis of each marker, or at

least most markers. Having clear examples of false positives or false negatives can also be fundamental.

The selection of representative regions (hot spots) to score, cellular expression or intensity thresholding, binning, overall positive and negative slide rating, and cut-offs are additional challenges to consider in the post-analytic study.

While training and various quality systems have increased pathologists' scoring repeatability, reproducibility, and accuracy, there is still significant room for improvement (Terrenato et al., 2013; Lin and Chen, 2014; Nielsen, 2015), and the same challenges can arise even in image analysis (Barnes et al., 2017), especially when different laboratories use different image analysis systems. Although computational quantitation using digital image analysis algorithms may improve reader precision performance (Rexhepaj et al., 2008; Ghaznavi et al., 2013; Barnes et al., 2017), it is important to harmonize those systems between laboratories and create protocols to make the data more reproducible.

In digital analyses, the pathologist evaluates a digital image of the glass slide on a computer monitor and uses a computational algorithm to provide a result. The reader selects representative fields of view or regions of interest (ROIs) of the tumor that the algorithm analyzes to yield a score that is intended to represent the whole tumor (Barnes et al., 2017).

As tumors often harbor substantial cellular and spatial heterogeneity, it is essential to perform high-resolution multiplexed analysis across entire tumor sections. Other factors must also be considered when determining whether to select representative ROIs or the entire tissue. Analyzing the entire tumor can be time- and resource-consuming, so it is best to select areas that are representative of the tumor's heterogeneity. The analysis of small ROIs or small tissue areas generates important variations and errors in the assessment of tumor and immune markers in cancer (Hofman et al., 2019). Other

TABLE 4 | Stages, challenge, and possible solutions for the best reproducibility of multiplex immunofluorescence panel.

| Stage | Problem | Solution | Advantages of solutions | Disadvantages of solutions |
|-----------------|---|---|--|---|
| Pre-analytical | Antibody specificity and staining | Use of positive and negative controls. Review of publications and experiences related to the Ab | Comparison with standardized process and other experiences | If the Ab does not have a previous protocol, it could result not reproducible |
| | Type of antibody and preparation | Preference by monoclonal antibody. Use specification of the vendor to prepare it | Better results | Not always is possible to monoclonal antibodies |
| Analytical | Optimization of panels in mIF | Test and work all the markers previously with IHC. | Comparison between IHC and mIF results | Some markers could not stain as the IHC |
| | Interpretation of markers | Standardized the interpretation of the most common markers | Interpretation well established | Some markers do not have protocols |
| | Consideration of areas of analysis and hotspots | Decide the number of representative areas of analysis and avoid select hotspots | Better representativeness | To have the right representative areas not always is possible. Number of ROIs could change depending on the type of tumor |
| | Type of image analysis | Do not expect to have the same result in all the different types of analysis technique. Consider the differences between software. Each one has its advantage and disadvantages | | Experience-dependent |
| Post-analytical | Variability of intra- and inter-observer | Create protocols. If still persisting some variability, identify the problem | Standardization | Time-consuming and requires additional effort of the collaborators |
| | External and internal variability | Publish the results of each project. Take the experience of other laboratories to improve | Share knowledge | The new technologies do not have other experiences in other laboratories because they can be expensive |

tumor types may have a higher degree of molecular heterogeneity, which may contribute to outcome (Rizzardi et al., 2016); analyzing a minimum area according to the complexity of each case is the most reasonable solution, but it is also important to have consensus between groups.

Given all of these challenges, laboratories that use mIF should standardize a minimum ROI or tissue area for analysis to generate accurate and reproducible results, considering the bibliographic data already available. The criteria to select areas of analysis should be compared in select representative areas, using the same method for evaluating each marker. In addition, although the algorithm can be locked, it will not always fit all the tumors; thus, it is possible to use an algorithm model as a base and make small changes according to the heterogeneity or type of tumor.

As each sample is complex, it is necessary to determine what factors should be excluded from the analysis (such as necrotic areas, hemorrhagic areas, non-preserved areas, unsatisfactory samples) and to standardize the reasons for exclusion to identify those that are unsatisfactory for mIF; in this way, only the cases without these considerations will be analyzed. The fewer the confusing factors are involved in the results, the easier it will be to standardize the workflow; each analysis could have less interobserver variability.

POST-ANALYTIC EVALUATION

After the pre-analytic and analytic evaluations, it is important to consider inside and outside evaluations. On the basis of our experience with IHC, although internal quality control procedures address daily reproducibility and are fundamental for monitoring performance in individual laboratories, external quality assessment is necessary to compare results from many laboratories by means of an external agency. This step allows the

identification of insufficient stains and inappropriate protocols, as well as the identification of possible issues with interpretation (Vyberg et al., 2005; Copete et al., 2011). An external evaluation can provide an objective evaluation of staining results from many laboratories for a given epitope or biomarker, identify the best practice protocols to obtain optimal results, and systematically identify causes of insufficient results (Nielsen, 2015). A similar evaluation is expected to be performed for mIF panels.

Some of the challenges in the pre-analytical and analytical steps have included standardizing the post-analytic component of mIF quantitation, including the interpretation approach, representative region (hot spot) selection, cellular expression, intensity thresholding, and cut-offs. While training and quality systems have increased pathologists' scoring repeatability, reproducibility, and accuracy, there is still significant room for improvement (Barnes et al., 2017). The experiences of different institutions should be combined in a common effort to standardize tissue scoring.

The final device design and configuration should be verified, including accuracy, technical sensitivity, and specificity and precision (i.e., intra-assay run, inter-assay run, inter-lot variability, inter-reader variability, and inter-instrumentation variability). External analytical validation studies should then be performed to document reproducibility (Figure 4).

Several published reports have described mIF optimization panel methods for solid tumors, but few are fully automated or reproducible for large numbers of samples (Lee et al., 2020) or between multiple institutions. One study described a collaboration between six institutions to develop an automated six-plex assay that is focused on the PD-1/PD-L1 axis and assesses inter- and intra-site reproducibility, on the basis of the percentage of expression by immune cells, in serial sections of tonsils and a lung cancer TMA. This approach improved the reproducibility of PD-L1 and immune cells (Hoyt et al., 2019).

It is necessary to create groups or committees that include experts in mIF from different institutions to generate guidelines and recommendations for staining, optimization, and validation procedures for mIF technology that can help to harmonize this assay across different research laboratories and standardize its clinical application (Table 4). Finally, the goal is to establish only one protocol for all of the institutions that use this technology, making it possible to identify issues even when each lab has its own differences in the items related to pre-analytical and analytical evaluation; however, these differences must not be an excuse to not improve internal protocols or to justify incorrect results.

CONCLUSIONS

Reproducibility must be evaluated at each step of the process. Small mistakes could have a large impact on the final results and on reproducibility within and between laboratories. The use of standardized protocols is a good approach to avoid wrong results, poor workflow, or whatever issue could affect the quality and results.

AUTHOR CONTRIBUTIONS

CL wrote most of the manuscript. FR and SH contributed to the writing with their expertise on digital image analysis and immune profiling. EP conceived the idea of the manuscript, developed the

technology in our laboratory, and edited the manuscript according to his experience.

FUNDING

This study was supported in part by the scientific and financial support for the CIMAC-CIDC Network provided through the National Cancer Institute (NCI) Cooperative Agreement U24CA224285 of The University of Texas MD Anderson Cancer Center CIMAC and for the Translational Molecular Pathology Immunoprofiling Platform, as well as by National Institutes of Health/NCI through Cancer Center Support Grant P30CA016672 (Institutional Tissue Bank) and SP0RE grant 5P50CA070907-18 from the Cancer Prevention and Research Institute of Texas through MIRA RP160668.

ACKNOWLEDGMENTS

The authors would like to acknowledge the Translational Molecular Pathology Immunoprofiling Laboratory members, who contribute daily to quality multiplex IF and IHC processing. The manuscript was edited by Sarah Bronson of the Research Medical Library at The University of Texas MD Anderson Cancer Center.

REFERENCES

- Akturk, G., Parra, E. R., Gjini, E., Lako, A., Lee, J. J., Neuberger, D., et al. (2021). Multiplex Tissue Imaging Harmonization: A Multicenter Experience from CIMAC-CIDC Immuno-Oncology Biomarkers Network. *Clin. Cancer Res.* doi:10.1158/1078-0432.ccr-21-2051
- Anagnostou, V. K., Welsh, A. W., Giltane, J. M., Siddiqui, S., Liceaga, C., Gustavson, M., et al. (2010). Analytic Variability in Immunohistochemistry Biomarker Studies. *Cancer Epidemiol. Biomarkers Prev.* 19 (4), 982–991. doi:10.1158/1055-9965.epi-10-0097
- Arpe, R., Carro-Temboury, M. R., Hempel, C., Vosch, T., and Just Sørensen, T. (2017). Investigating Dye Performance and Crosstalk in Fluorescence Enabled Bioimaging Using a Model System. *PLoS One* 12 (11), e0188359. doi:10.1371/journal.pone.0188359
- Barnes, M., Srinivas, C., Bai, I., Frederick, J., Liu, W., Sarkar, A., et al. (2017). Whole Tumor Section Quantitative Image Analysis Maximizes Between-Pathologists' Reproducibility for Clinical Immunohistochemistry-Based Biomarkers. *Lab. Invest.* 97 (12), 1508–1515. doi:10.1038/labinvest.2017.82
- Barua, S., Solis, L., Parra, E. R., Uraoka, N., Jiang, M., Wang, H., et al. (2018). A Functional Spatial Analysis Platform for Discovery of Immunological Interactions Predictive of Low-Grade to High-Grade Transition of Pancreatic Intraductal Papillary Mucinous Neoplasms. *Cancer Inform.* 17, 1176935118782880. doi:10.1177/1176935118782880
- Blom, S., Paavola, L., Bychkov, D., Turkkii, R., Mäki-Teeri, P., Hemmes, A., et al. (2017). Systems Pathology by Multiplexed Immunohistochemistry and Whole-Slide Digital Image Analysis. *Sci. Rep.* 7 (1), 15580. doi:10.1038/s41598-017-15798-4
- Bordeaux, J., Welsh, A. W., Agarwal, S., Killiam, E., Baquero, M. T., Hanna, J. A., et al. (2010). Antibody Validation. *Biotechniques* 48 (3), 197–209. doi:10.2144/000113382
- Broad Institute (2021) Cancer Cell Line Encyclopedia. Available at: <https://sites.broadinstitute.org/ccl/>.
- Canadian Association of Pathologists-Association canadienne des pathologistes National Standards Committee/Torlakovic, E. E., Riddell, R., Banerjee, D., El-Zimaity, H., Pilavdzic, D., et al. (2010). Canadian Association of Pathologists-Association canadienne des pathologistes National Standards Committee/Immunohistochemistry: Best practice recommendations for standardization of immunohistochemistry tests. *Am. J. Clin. Pathol.* 133 (3), 354–365. doi:10.1309/AJCPDYZ1XMF4HJWK
- Carvajal-Hausdorf, D. E., Schalper, K. A., Neumeister, V. M., and Rimm, D. L. (2015). Quantitative Measurement of Cancer Tissue Biomarkers in the Lab and in the Clinic. *Lab. Invest.* 95 (4), 385–396. doi:10.1038/labinvest.2014.157
- Cell Signaling (2019). Hallmarks of Antibody Validation. Available at: <https://www.korambiotech.com/wp-content/uploads/2021/06/Hallmarks-of-Antibody-Validation.pdf>.
- Cheng, C. L., Thihe, A. A., Tan, S. Y. J., Chua, P. J., Bay, B. H., and Tan, P. H. (2015). Expression of FGFR1 Is an Independent Prognostic Factor in Triple-Negative Breast Cancer. *Breast Cancer Res. Treat.* 151 (1), 99–111. doi:10.1007/s10549-015-3371-x
- Copete, M., Garratt, J., Gilks, B., Pilavdzic, D., Berendt, R., Bigras, G., et al. (2011). Inappropriate Calibration and Optimisation of Pan-Keratin (Pan-CK) and Low Molecular Weight Keratin (LMWCK) Immunohistochemistry Tests: Canadian Immunohistochemistry Quality Control (CIQC) Experience. *J. Clin. Pathol.* 64 (3), 220–225. doi:10.1136/jcp.2010.085258
- COSMIC (2021). Catalogue of Somatic Mutations in Cancer (COSMIC). Available at: <https://cancer.sanger.ac.uk/cosmic>.
- Curigliano, G., Burstein, H. J., Winer, E. P., Gnant, M., Dubsy, P., Loibl, S., et al. (2017). De-escalating and Escalating Treatments for Early-Stage Breast Cancer: the St. Gallen International Expert Consensus Conference on the Primary Therapy of Early Breast Cancer 2017. *Ann. Oncol.* 28 (8), 1700–1712. doi:10.1093/annonc/mdx308
- Dejima, H., Hu, X., Chen, R., Zhang, J., Fujimoto, J., Parra, E. R., et al. (2021). Immune Evolution from Preneoplasia to Invasive Lung Adenocarcinomas and Underlying Molecular Features. *Nat. Commun.* 12 (1), 2722. doi:10.1038/s41467-021-22890-x

- Engel, K. B., and Moore, H. M. (2011). Effects of Preanalytical Variables on the Detection of Proteins by Immunohistochemistry in Formalin-Fixed, Paraffin-Embedded Tissue. *Arch. Pathol. Lab. Med.* 135 (5), 537–543. doi:10.5858/2010-0702-rair.1
- Fitzgibbons, P. L., Bradley, L. A., Fatheree, L. A., Alsabeh, R., Fulton, R. S., Goldsmith, J. D., et al. (2014). Principles of Analytic Validation of Immunohistochemical Assays: Guideline from the College of American Pathologists Pathology and Laboratory Quality Center. *Arch. Pathol. Lab. Med.* 138 (11), 1432–1443. doi:10.5858/arpa.2013-0610-cp
- Francisco-Cruz, A., Parra, E., Tetziaff, M., and Wistuba, I. (2020a). “Multiplex Immunofluorescence Assays,” *Biomarkers for Immunotherapy of Cancer* (New York: Springer), 467–495.
- Francisco-Cruz, A., Parra, E. R., Tetzlaff, M. T., and Wistuba, II (2020b). Multiplex Immunofluorescence Assays. *Methods Mol. Biol.* 2055, 467–495. doi:10.1007/978-1-4939-9773-2_22
- Gerdes, J., Dallenbach, F., Lennert, K., Lemke, H., and Stein, H. (1984). Growth Fractions in Malignant Non-hodgkin's Lymphomas (NHL) as Determined in Situ with the Monoclonal Antibody Ki-67. *Hematol. Oncol.* 2 (4), 365–371. doi:10.1002/hon.2900020406
- Ghandi, M., Huang, F. W., Jané-Valbuena, J., Kryukov, G. V., Lo, C. C., McDonald, E. R., et al. (2019). Next-generation Characterization of the Cancer Cell Line Encyclopedia. *Nature* 569 (7757), 503–508. doi:10.1038/s41586-019-1186-3
- Ghaznavi, F., Evans, A., Madabhushi, A., and Feldman, M. (2013). Digital Imaging in Pathology: Whole-Slide Imaging and beyond. *Annu. Rev. Pathol. Mech. Dis.* 8, 331–359. doi:10.1146/annurev-pathol-011811-120902
- Giraldo, N. A., Nguyen, P., Engle, E. L., Kaunitz, G. J., Cottrell, T. R., Berry, S., et al. (2018). Multidimensional, Quantitative Assessment of PD-1/PD-L1 Expression in Patients with Merkel Cell Carcinoma and Association with Response to Pembrolizumab. *J. Immunother. Cancer* 6 (1), 99. doi:10.1186/s40425-018-0404-0
- Goldstein, N. S., Hewitt, S. M., Taylor, C. R., Yaziji, H., and Hicks, D. G. Members of Ad-Hoc Committee On Immunohistochemistry Standardization (2007). Recommendations for Improved Standardization of Immunohistochemistry. *Appl. Immunohistochem. Mol. Morphol.* 15 (2), 124–133. doi:10.1097/pai.0b013e31804c7283
- Gorris, M. A. J., Halilovic, A., Rabold, K., van Duffelen, A., Wickramasinghe, I. N., Verweij, D., et al. (2018). Eight-Color Multiplex Immunohistochemistry for Simultaneous Detection of Multiple Immune Checkpoint Molecules within the Tumor Microenvironment. *J. Immunol.* 200 (1), 347–354. doi:10.4049/jimmunol.1701262
- Hewitt, S. M., Baskin, D. G., Frevert, C. W., Stahl, W. L., and Rosa-Molinar, E. (2014). Controls for Immunohistochemistry. *J. Histochem. Cytochem.* 62 (10), 693–697. doi:10.1369/0022155414545224
- Hewitt, S. M. (2016). Reproducibility. *J. Histochem. Cytochem.* 64 (4), 223. doi:10.1369/0022155416636547
- Hofman, P., Badoual, C., Henderson, F., Berland, L., Hamila, M., Long-Mira, E., et al. (2019). Multiplexed Immunohistochemistry for Molecular and Immune Profiling in Lung Cancer—Just about Ready for Prime-Time? *Cancers (Basel)* 11 (3), 283. doi:10.3390/cancers11030283
- Hoyt, C., Roman, K., Engle, L., Wang, C., Ballesteros-Merino, C., Jensen, S. M., et al. (2019). “Abstract LB-318: Multi-Institutional TSA-Amplified Multiplexed Immunofluorescence Reproducibility Evaluation (MITRE Study): Reproducibility Assessment of an Automated Multiplexed Immunofluorescence Slide Staining, Imaging, and Analysis Workflow. Cancer Research.” Proceedings of AACR Annual Meeting 2019, Atlanta, GA, March 29–April 3, 2019. doi:10.1158/1538-7445.sabcs18-lb-318
- Lee, C.-W., Ren, Y. J., Marella, M., Wang, M., Hartke, J., and Couto, S. S. (2020). Multiplex Immunofluorescence Staining and Image Analysis Assay for Diffuse Large B Cell Lymphoma. *J. Immunol. Methods* 478, 112714. doi:10.1016/j.jim.2019.112714
- Lin, F., and Chen, Z. (2014). Standardization of Diagnostic Immunohistochemistry: Literature Review and Geisinger Experience. *Arch. Pathol. Lab. Med.* 138 (12), 1564–1577. doi:10.5858/arpa.2014-0074-ra
- Matsumoto, H., Koo, S.-I., Dent, R., Tan, P. H., and Iqbal, J. (2015). Role of Inflammatory Infiltrates in Triple Negative Breast Cancer: Table 1. *J. Clin. Pathol.* 68 (7), 506–510. doi:10.1136/jclinpath-2015-202944
- Meyerholz, D. K., and Beck, A. P. (2018). Principles and Approaches for Reproducible Scoring of Tissue Stains in Research. *Lab. Invest.* 98 (7), 844–855. doi:10.1038/s41374-018-0057-0
- Mohammed, Z. M. A., McMillan, D. C., Elsberger, B., Going, J. J., Orange, C., Mallon, E., et al. (2012). Comparison of Visual and Automated Assessment of Ki-67 Proliferative Activity and Their Impact on Outcome in Primary Operable Invasive Ductal Breast Cancer. *Br. J. Cancer* 106 (2), 383–388. doi:10.1038/bjc.2011.569
- Munzone, E., Botteri, E., Sciandivasci, A., Curigliano, G., Nolè, F., Mastropasqua, M., et al. (2012). Prognostic Value of Ki-67 Labeling index in Patients with Node-Negative, Triple-Negative Breast Cancer. *Breast Cancer Res. Treat.* 134 (1), 277–282. doi:10.1007/s10549-012-2040-6
- Ng, S.-B., Fan, S., Choo, S.-N., Hoppe, M., Mai Phuong, H., De Mel, S., et al. (2018). Quantitative Analysis of a Multiplexed Immunofluorescence Panel in T-Cell Lymphoma. *SLAS TECHN. Transl. Life Sci. Innov.* 23 (3), 252–258. doi:10.1177/2472630317747197
- Nielsen, S. (2015). External Quality Assessment for Immunohistochemistry: Experiences from NordiQC. *Biotech. Histochem.* 90 (5), 331–340. doi:10.3109/10520295.2015.1033462
- Okoye, J. O., and Nnatuanya, N. I. (2015). Immunohistochemistry: a Revolutionary Technique in Laboratory Medicine. *Clin. Med. Diagn.* 5, 60–69. doi:10.5923/j.cmd.20150504.02
- Parra, E. R., Francisco-Cruz, A., and Wistuba, II (2019). State-of-the-Art of Profiling Immune Contexture in the Era of Multiplexed Staining and Digital Analysis to Study Paraffin Tumor Tissues. *Cancers (Basel)* 11 (2). doi:10.3390/cancers11020247
- Parra, E. R., Jiang, M., Solis, L., Mino, B., Laberiano, C., Hernandez, S., et al. (2020). Procedural Requirements and Recommendations for Multiplex Immunofluorescence Tyramide Signal Amplification Assays to Support Translational Oncology Studies. *Cancers (Basel)* 12 (2), 255. doi:10.3390/cancers12020255
- Parra, E. R., Ferrufino-Schmidt, M. C., Tamegnon, A., Zhang, J., Solis, L., Jiang, M., et al. (2021a). Immuno-profiling and Cellular Spatial Analysis Using Five Immune Oncology Multiplex Immunofluorescence Panels for Paraffin Tumor Tissue. *Sci. Rep.* 11 (1), 8511. doi:10.1038/s41598-021-88156-0
- Parra, E. R., and Hernández Ruiz, S. (2021a). Detection of Programmed Cell Death Ligand 1 Expression in Lung Cancer Clinical Samples by an Automated Immunohistochemistry System. *Methods Mol. Biol.* 2279, 35–47. doi:10.1007/978-1-0716-1278-1_4
- Parra, E. R., and Hernández Ruiz, S. (2021b). Western Blot as a Support Technique for Immunohistochemistry to Detect Programmed Cell Death Ligand 1 Expression. *Methods Mol. Biol.* 2279, 49–57. doi:10.1007/978-1-0716-1278-1_5
- Parra, E. R. (2021). Methods to Determine and Analyze the Cellular Spatial Distribution Extracted from Multiplex Immunofluorescence Data to Understand the Tumor Microenvironment. *Front. Mol. Biosci.* 8, 668340. doi:10.3389/fmolb.2021.668340
- Parra, E. R., Uraoka, N., Jiang, M., Cook, P., Gibbons, D., Forget, M.-A., et al. (2017). Validation of Multiplex Immunofluorescence Panels Using Multispectral Microscopy for Immune-Profiling of Formalin-Fixed and Paraffin-Embedded Human Tumor Tissues. *Sci. Rep.* 7 (1), 13380. doi:10.1038/s41598-017-13942-8
- Parra, E. R., Villalobos, P., Behrens, C., Jiang, M., Pataer, A., Swisher, S. G., et al. (2018a). Effect of Neoadjuvant Chemotherapy on the Immune Microenvironment in Non-small Cell Lung Carcinomas as Determined by Multiplex Immunofluorescence and Image Analysis Approaches. *J. Immunother. Cancer* 6 (1), 48. doi:10.1186/s40425-018-0368-0
- Parra, E. R., Villalobos, P., Mino, B., and Rodriguez-Canales, J. (2018b). Comparison of Different Antibody Clones for Immunohistochemistry Detection of Programmed Cell Death Ligand 1 (PD-L1) on Non-Small Cell Lung Carcinoma. *Appl. Immunohistochem. Mol. Morphol.* 26 (2), 83–93. doi:10.1097/pai.0000000000000531
- Parra, E. R., Zhai, J., Tamegnon, A., Zhou, N., Pandurengan, R. K., Barreto, C., et al. (2021b). Identification of Distinct Immune Landscapes Using an Automated Nine-Color Multiplex Immunofluorescence Staining Panel and Image Analysis in Paraffin Tumor Tissues. *Sci. Rep.* 11 (1), 4530. doi:10.1038/s41598-021-83858-x
- Rexhepaj, E., Brennan, D. J., Holloway, P., Kay, E. W., McCann, A. H., Landberg, G., et al. (2008). Novel Image Analysis Approach for Quantifying Expression of Nuclear Proteins Assessed by Immunohistochemistry: Application to Measurement of Oestrogen and Progesterone Receptor Levels in Breast Cancer. *Breast Cancer Res.* 10 (5), R89. doi:10.1186/bcr2187

- Rhodes, A. (2003). Quality Assurance in Immunohistochemistry. *Am. J. Surg. Pathol.* 27 (9), 1284–1285. doi:10.1097/0000478-200309000-00015
- Rizzardi, A. E., Zhang, X., Vogel, R. I., Kolb, S., Geybels, M. S., Leung, Y.-K., et al. (2016). Quantitative Comparison and Reproducibility of Pathologist Scoring and Digital Image Analysis of Estrogen Receptor β Immunohistochemistry in Prostate Cancer. *Diagn. Pathol.* 11 (1), 63. doi:10.1186/s13000-016-0511-5
- Rojo, M. G., Bueno, G., and Slodkowska, J. (2009). Review of Imaging Solutions for Integrated Quantitative Immunohistochemistry in the Pathology Daily Practice. *Folia Histochem. Cytobiol.* 47 (3), 349–354. doi:10.2478/v10042-008-0114-4
- Rost, S., Giltman, J., Bordeaux, J. M., Hitzman, C., Koeppen, H., and Liu, S. D. (2017). Erratum: Multiplexed Ion Beam Imaging Analysis for Quantitation of Protein Expression in Cancer Tissue Sections. *Lab. Invest.* 97 (10), 1263. doi:10.1038/labinvest.2017.94
- Rudbeck, L. (2015). Adding Quality to Your Qualitative IHC. *MLO Med. Lab. Obs* 47 (12), 18–21.
- Schalper, K. A., Brown, J., Carvajal-Hausdorf, D., McLaughlin, J., Velcheti, V., Syrigos, K. N., et al. (2015). Objective Measurement and Clinical Significance of TILs in Non-small Cell Lung Cancer. *J. Natl. Cancer Inst.* 107 (3), dju435. doi:10.1093/jnci/dju435
- Shiptsin, M., Small, C., Giladi, E., Siddiqui, S., Choudhury, S., Hussain, S., et al. (2014). Automated Quantitative Multiplex Immunofluorescence *In Situ* Imaging Identifies Phospho-S6 and Phospho-PRAS40 as Predictive Protein Biomarkers for Prostate Cancer Lethality. *Proteome Sci.* 12, 40. doi:10.1186/1477-5956-12-40
- Sivertsson, Å., Lindström, E., Oksvold, P., Katona, B., Hikmet, F., Vuu, J., et al. (2020). Enhanced Validation of Antibodies Enables the Discovery of Missing Proteins. *J. Proteome Res.* 19 (12), 4766–4781. doi:10.1021/acs.jproteome.0c00486
- Sood, A., Miller, A. M., Brogi, E., Sui, Y., Armenia, J., McDonough, E., et al. (2016). Multiplexed Immunofluorescence Delineates Proteomic Cancer Cell States Associated with Metabolism. *JCI Insight* 1 (6), e87030. doi:10.1172/jci.insight.87030
- Stack, E. C., Wang, C., Roman, K. A., and Hoyt, C. C. (2014). Multiplexed Immunohistochemistry, Imaging, and Quantitation: a Review, with an Assessment of Tyramide Signal Amplification, Multispectral Imaging and Multiplex Analysis. *Methods* 70 (1), 46–58. doi:10.1016/j.ymeth.2014.08.016
- Stauber, J., MacAleese, L., Franck, J., Claude, E., Snel, M., Kaletas, B. K., et al. (2010). On-tissue Protein Identification and Imaging by MALDI-Ion Mobility Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* 21 (3), 338–347. doi:10.1016/j.jmass.2009.09.016
- Steiner, C., Ducret, A., Tille, J. C., Thomas, M., McKee, T. A., Rubbia-Brandt, L., et al. (2014). Applications of Mass Spectrometry for Quantitative Protein Analysis in Formalin-Fixed Paraffin-Embedded Tissues. *Proteomics* 14 (4-5), 441–451. doi:10.1002/pmic.201300311
- Surace, M., DaCosta, K., Huntley, A., Zhao, W., Bagnall, C., Brown, C., et al. (2019). Automated Multiplex Immunofluorescence Panel for Immuno-Oncology Studies on Formalin-Fixed Carcinoma Tissue Specimens. *J. Vis. Exp.* 143. doi:10.3791/58390
- Tan, W. C. C., Nerurkar, S. N., Cai, H. Y., Ng, H. H. M., Wu, D., Wee, Y. T. F., et al. (2020). Overview of Multiplex Immunohistochemistry/immunofluorescence Techniques in the Era of Cancer Immunotherapy. *Cancer Commun.* 40 (4), 135–153. doi:10.1002/cac2.12023
- Taube, J. M., Akturk, G., Angelo, M., Engle, E. L., Gnjatich, S., Greenbaum, S., et al. (2020). The Society for Immunotherapy of Cancer Statement on Best Practices for Multiplex Immunohistochemistry (IHC) and Immunofluorescence (IF) Staining and Validation. *J. Immunother. Cancer* 8 (1), e000155. doi:10.1136/jitc-2019-000155
- Taube, J. M., Roman, K., Engle, E. L., Wang, C., Ballesteros-Merino, C., Jensen, S. M., et al. (2021). Multi-institutional TSA-Amplified Multiplexed Immunofluorescence Reproducibility Evaluation (MITRE) Study. *J. Immunother. Cancer* 9 (7), e002197. doi:10.1136/jitc-2020-002197
- Taylor, C. R. (1999). FDA Issues Final Rule for Classification and Reclassification of Immunohistochemistry Reagents and Kits. *Am. J. Clin. Pathol.* 111 (4), 443–444. doi:10.1093/ajcp/111.4.443
- Taylor, C. R. (1998). Report from the Biological Stain Commission: FDA Issues Final Rule for Classification/reclassification of Immunohistochemistry (IHC) Reagents and Kits. *Biotech. Histochem.* 73 (4), 175–177. doi:10.3109/10520299809141107
- Taylor, C. R. (1992). Report of the Immunohistochemistry Steering Committee of the Biological Stain Commission. "Proposed Format: Package Insert for Immunohistochemistry Products". *Biotech. Histochem.* 67 (6), 323–338. doi:10.3109/10520299209110045
- Terrenato, I., Arena, V., Pizzamiglio, S., Pennacchia, I., Perracchio, L., Buglioni, S., et al. (2013). External Quality Assessment (EQA) Program for the Preanalytical and Analytical Immunohistochemical Determination of HER2 in Breast Cancer: an Experience on a Regional Scale. *J. Exp. Clin. Cancer Res.* 32, 58. doi:10.1186/1756-9966-32-58
- The Human Protein Atlas (2021). The Human Protein Atlas Project 2021. Available at: <https://www.proteinatlas.org/>.
- Tie, H. C., and Lu, L. (2020). How to Quantitatively Measure, Assess and Correct the Fluorescence Crosstalk in the Wide-Field Microscopy. *bioRxiv* [Epub ahead of print]. doi:10.1101/2020.05.20.105627
- Toki, M. I., Cecchi, F., Hembrough, T., Syrigos, K. N., and Rimm, D. L. (2017). Proof of the Quantitative Potential of Immunofluorescence by Mass Spectrometry. *Lab. Invest.* 97 (3), 329–334. doi:10.1038/labinvest.2016.148
- Tumeh, P. C., Harview, C. L., Yearley, J. H., Shintaku, I. P., Taylor, E. J. M., Robert, L., et al. (2014). PD-1 Blockade Induces Responses by Inhibiting Adaptive Immune Resistance. *Nature* 515 (7528), 568–571. doi:10.1038/nature13954
- Uhlen, M., Bandrowski, A., Carr, S., Edwards, A., Ellenberg, J., Lundberg, E., et al. (2016). A Proposal for Validation of Antibodies. *Nat. Methods* 13 (10), 823–827. doi:10.1038/nmeth.3995
- Varma, M., Berney, D. M., Jasani, B., and Rhodes, A. (2004). Technical Variations in Prostatic Immunohistochemistry: Need for Standardisation and Stringent Quality Assurance in PSA and PSAP Immunostaining. *J. Clin. Pathol.* 57 (7), 687–690. doi:10.1136/jcp.2003.014894
- Velcheti, V., Schalper, K. A., Carvajal, D. E., Anagnostou, V. K., Syrigos, K. N., Sznol, M., et al. (2014). Programmed Death Ligand-1 Expression in Non-small Cell Lung Cancer. *Lab. Invest.* 94 (1), 107–116. doi:10.1038/labinvest.2013.130
- Vincent-Salomon, A., Gruel, N., Lucchesi, C., MacGrogan, G., Dendale, R., Sigal-Zafrani, B., et al. (2007). Identification of Typical Medullary Breast Carcinoma as a Genomic Sub-group of Basal-like Carcinomas, a Heterogeneous New Molecular Entity. *Breast Cancer Res.* 9 (2), R24. doi:10.1186/bcr1666
- Viratham Pulsawatdi, A., Craig, S. G., Bingham, V., McCombe, K., Humphries, M. P., Senevirathne, S., et al. (2020). A Robust Multiplex Immunofluorescence and Digital Pathology Workflow for the Characterisation of the Tumour Immune Microenvironment. *Mol. Oncol.* 14 (10), 2384–2402. doi:10.1002/1878-0261.12764
- Vyberg, M., Torlakovic, E., Seidal, T., Risberg, B., Helin, H., and Nielsen, S. (2005). Nordic Immunohistochemical Quality Control. *Croat. Med. J.* 46 (3), 368–371.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Laberiano-Fernández, Hernández-Ruiz, Rojas and Parra. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.