1 Introduction to Predictive Biomarkers: Definitions and Characteristics

Clive R. Taylor

Biomarkers

The concept of "biomarkers" as indicators of health or disease is not new. Under the broadest interpretation, the use of biomarkers extends back to the "ancients," who elicited medical signs, measured the pulse, observed, and even tasted the urine and the like [1]. However, the use of the term biomarker is relatively recent in the field of medicine, where the definition continues to shift with context.

Certainly many clinical laboratory tests fall under a broad definition. Examples include hormone levels for endocrine disease, a succession of enzymes and proteins, up to present day troponin for myocardial infarction, and prostatic acid phosphatase, then PSA (prostate-specific antigen), for prostate cancer. Extending the definition to its limits, the structural changes observed in anatomic pathology, or in radiology, also meet the definitional criteria; a tissue diagnosis of prostate cancer, plus or minus grading (e.g., Gleason), is a biomarker in a very real sense. Other "biomarkers" of diverse variety also have long been applied in unrelated fields, such as archeology, geology, and the petrochemical industry.

This introductory chapter has a more restricted focus, namely, the utilization of "biomarkers" as identified by laboratory tests in relation to cancer; still more specifically, the focus is upon biomarkers detected directly in tissues from cancer patients (Table 1.1). Within this context of tissue and cancer, biomarkers include proteins and nucleic acids and derivatives and parts thereof. While the focus is narrow, the levels of complexity are manifold and growing day by day.

Biomarkers in Cancer

Tests for biological markers in malignant disease, for diagnosis, prognosis, and monitoring of progression, can be traced back at least a century and a half to the example of Bence-Jones protein in urine (Henry Bence-Jones 1813–1873) [1] for Kahler's disease (Otto Kahler 1849–1893), a surrogate for the detection and measurement of monoclonal (malignant-M) proteins that identify the condition that we now know as multiple myeloma. The modern era of biomarkers with respect to cancer in general may, on the one hand, be traced back to the discovery and use of CEA (carcinoembryonic antigen), a protein biomarker, and, on the other, to the Philadelphia chromosome, a genetic marker of chronic myeloid leukemia [1]. While CEA did not meet initial hopes of diagnostic utility in terms of sensitivity or specificity, measurement of CEA in the serum did find

C. R. Taylor
Department of Pathology, University of Southern California, Los Angeles, CA, USA

© Springer Nature Switzerland AG 2019
S. Badve, G. L. Kumar (eds.), *Predictive Biomarkers in Oncology*,
https://doi.org/10.1007/978-3-319-95228-4_1

Table 1.1 Biomarkers in the context of cancer

Biomarker: general definition	A characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacological responses to therapeutic intervention
Diagnostic	Design and usage; primarily to assist diagnosis; commonly in IHC on tissue sections, but also sometimes indicative in serum
Prognostic	Design and usage; primarily as a guide to prognosis; the course and progress of disease –therapy unspecified
Predictive	Design and usage; specifically for classification of responders vs. nonresponders for a defined (usually targeted) therapy; assay and threshold developed in conjoint clinical trial with the specified drug
Companion	Predictive; co-developed with a specified therapy and "required" prior to use of said therapy
Complementary	Predictive; co-developed with a specified therapy; accepted as providing guidance for therapy but not required
Pharmacodynamic	Definitional within the pharmaceutical field, such as providing a surrogate marker for disease status, as in remission or progression
Monitoring	Design and usage; for evaluation of status, progression, and/or recurrence of established disease process

a place in monitoring of established disease and as a "biomarker" of recurrence, likewise for CA-125 and arguably PSA. Notably, in a different context that still is within the field of cancer, all three of these biomarkers maintain a (variable) role as diagnostic biomarkers when demonstrated in situ within tissue or cell by immunohistochemistry (IHC). Thus context matters.

The decade of the 1990s saw major developments in the measurement of estrogen (and progesterone) receptors (ER and PR) in breast cancer, with applications that were prognostic and, to a degree, predictive in terms of choice of therapy. Cytosol-based competitive assays, relying upon extracts of purported tumor tissue, gradually gave way to a different methodology based on the detection of ER (and or PR) in situ within tissue sections by labeled antibody methods, with IHC (immunohistochemistry) using FFPE (formalin-fixed paraffin-embedded) sections emerging as the standard.

This transition occurred in spite of the arguments levied against FFPE tissue, because of the unknown effects of protein "masking," and against IHC, because of subjectivity in interpretation and hence variability in scoring, and also because of the nonlinear relationship between signal intensity and target antigen (in this instance the estrogen receptor protein) [2]. The efforts of Craig Allred and others in the development of defined (but semi-quantitative) scoring methods were critical to acceptance of the IHC method for this purpose.

In the presence of proper controls of assay performance [2, 3], IHC brings exquisite specificity, by scoring only recognizable cancer cells, and extraordinary technical sensitivity, with the ability to detect one ER-positive cell among a 100 identifiable cancer cells (1%; the current threshold of a positive ER IHC test) or in fact 1 positive cell among 1000 or 10,000 or more cells. Expressed in these terms, namely, detection of positive cells, this level of sensitivity is far beyond anything that can be achieved by any method using an extract of tissue, which is necessarily an imperfectly known extract of an imperfectly known mixture of normal and cancer cells, themselves imperfectly identified.

In this mode of performance, the IHC ER "test" may be considered to represent the beginning of the current era of employment of biomarkers in cancer, for prognostic and predictive purposes.

The "First" Predictive Biomarker

However, the moment of critical impetus for the current explosion in interest and variety of cancer biomarkers was the day (September 25, 1998) upon which the FDA approved the HercepTest

(Dako, now Agilent, CA, USA) and simultaneously gave approval for the use of the companion drug Herceptin (Genentech, now Roche) for the treatment of patients with Her2-positive breast cancer (as measured by the HercepTest). A vitally important corollary message from the FDA was that drug and test should be developed in concert, during a combined clinical study, hence "companion diagnostic" (Table 1.1) (Fig. 1.1) [4–10].

From the beginning of the millennium to the present time, US and European regulatory and working groups [4–8] offered various definitions of a biomarker, including the following: "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacological responses to therapeutic intervention." Subsequently the FDA went further with the definition of a "valid biomarker" – including that it should:

- Be measured in a test system with well-established performance characteristics
- Have a scientific background of evidence including clinical significance
- Be "fit to purpose"

A final consideration extended to a "clinically useful biomarker," which should in addition be reliable and clinically actionable in the specified setting.

The subsequent two decades have seen ongoing evolution of the term, with sub-definitions according to the design and use (Tables 1.1 and 1.2), accompanied by growing emphasis upon objectivity, reproducibility, and elements of true quantification, which reflect back upon methodology and ultimately performance of the "total test" from inception to interpretation, whichever the test modality employed (Table 1.3) [2, 3, 10, 11].

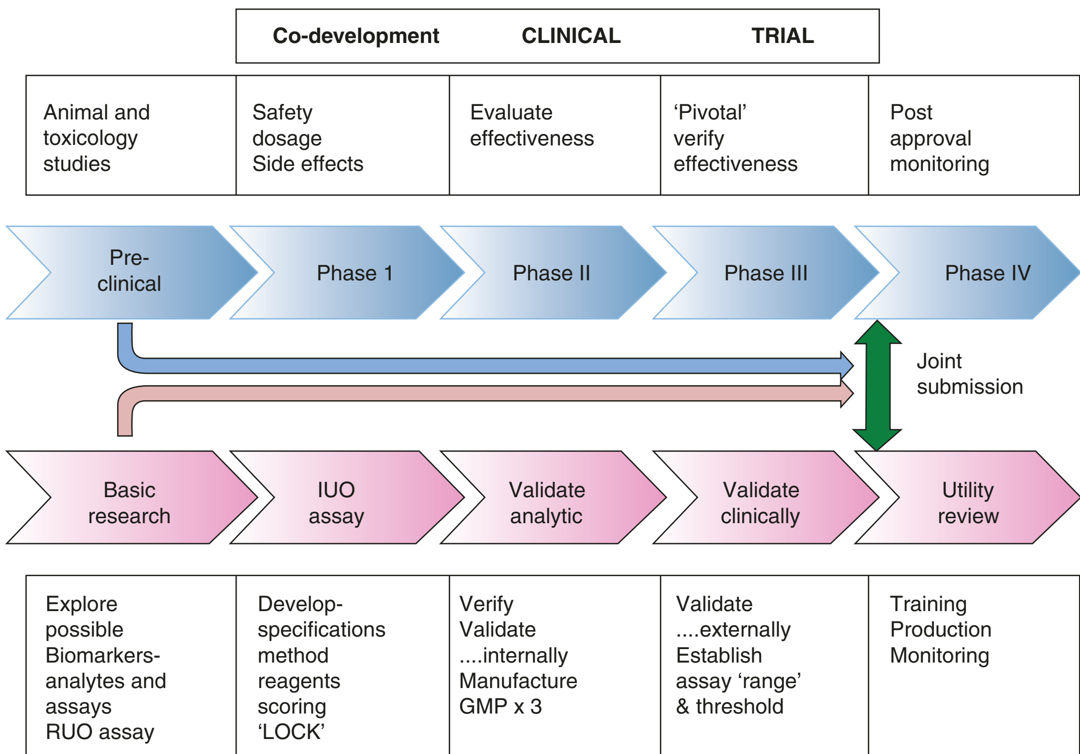

Fig. 1.1 Co-development process for "drug" and companion diagnostic. Time frame, up to 10 years; cost, up to 100 million dollars

Table 1.2 Laboratory reagents and tests; FDA categories

ASR	RUO	IUO	IVD	LDT
Analyte-specific reagent	Research use only	Investigational Use only	In vitro device	Lab developed test
No diagnostic claims	No diagnostic claims	No diagnostic claims	Specified claims FDA approved	Lab responsible for any claims[a]
FDA regulations	FDA regulations	FDA regulations	FDA regulations	CLIA[b] regulations FDA discretion
May be used as reagents for RUO, IUO, IVD, and LDT tests	Not for clinical use	Use restricted to specified study	Intended use define by trial Specified in labeling	For use only in the lab that developed the test

https://www.cms.gov/Clia/

[a]LDT may require FDA approval if used as a predictive marker; clinical utility must be validated

[b]*CLIA* Clinical Laboratory Improvement Amendments

Table 1.3 The "total test" approach

Pre-analytical (Sample preparation)	Test selection: indication for the test
	Specimen handling, from operating room to histology laboratory
	Fixation: total fixation time and type of fixative
	Paraffin embedding, storage, and sectioning
	Deparaffinization
Analytical (Reagents and protocol)	Antigen retrieval (exact method)
	Assay (staining) method and protocol
	Reagent validation
	Controls (reference standards)
	Technologist and laboratory certification
	Proficiency testing and quality assurance
Post-analytical (Interpretations and reporting)	Reading of result(s)/scoring/quantification
	Diagnostic, prognostic, or predictive significance
	Report
	Turnaround time
	Outcomes analysis/economics/reimbursement Pre-analytical

Based on data from Taylor [16]

Predictive Biomarkers: Companion Versus Complementary

The distinction of companion versus complementary biomarkers (Table 1.1) emerged from conjoint clinical studies, determined by the level of prediction of clinical response that the test rendered.

With a companion diagnostic, a positive result indicates treatment with the companion drug; a negative result indicates no treatment; and the test is required before the use of the corresponding drug.

With a complementary diagnostic, a positive result usually indicates treatment, but a patient having a negative result may or may not be treated according to an informed clinical decision.

For example, with PD-L1 tests, some "tests" emerged as companion diagnostics, and others as

complementary, varying according to which anti-PD-L1 antibody was employed [8, 12, 13], by which method, and in which specified tumor type.

Intrinsic to the FDA definition of an approved IVD (in vitro diagnostic) companion diagnostic is that it “provides information that is essential for the safe and effective use of a corresponding therapeutic product” and that its use is “stipulated in the instructions for use in the labeling of both the diagnostic device and the corresponding therapeutic agent” (Table 1.2) [6–8]. The current EU definition is less rigorous, but similar in intent, and interestingly admits both “quantitative and qualitative determination of specific markers identifying subjects” [5, 8]. It specifically excludes monitoring.

The FDA definition carries with it an assignment of the IHC IVD to Class III (the highest level) requiring PMA (pre-market approval) in a co-development mode with the drug [4, 6–8, 12], whereas the EU regulations appear to leave companion diagnostics in the current general IVD category [5]; new regulations are afoot that likely will raise the level and may preclude the current self-certification route (for discussion of the subtleties of these definitions, see references 4 and 12 and later chapters in this book). The above statements apply specifically to companion diagnostics; there are as yet no corresponding written rules for complementary diagnostics; the definition of which is at present by precedent and usage, although proposals have been aired.

Method Development

These types of predictive biomarker tests have come to be of critical import in the context of targeted drug therapies, such that the majority of such agents now in clinical studies are following a co-development plan for “test” and “corresponding therapy.” Detailed discussion of this co-development process is outside the scope of this chapter but is summarized in Fig. 1.1, examined in detail elsewhere in this book, and well-reviewed in a recent National Policy Workshop [4]. For drug development generally the process includes preclinical (animal) studies: phase 1, toxicity, in which potential biomarkers may also be assessed; phase 2, preliminary efficacy of drug, plus biomarker evaluation; phase 3, definitive efficacy and validation of biomarker; and phase 4, post market surveillance. Total patient accrual will be in the hundreds.

For the biomarker there is a preceding period of basic research and discovery that provides initial evidence of the potential utility of a molecule (biomarker) in the context of diagnosis or prognosis of cancer or a relationship to a potential therapeutic modality (drug – predictive) (Fig. 1.1). This discovery process is followed by evolution of a prototypic test using analyte-specific reagents (ASRs), through an investigational use only (IUO) test, on to an FDA-approved IVD (Table 1.2), which category includes all companion diagnostics. In some instances clinical laboratories may separately develop assays for clinical use, with internal validation under CLIA regulations (LDT, laboratory-developed test) (Table 1.2). The FDA has provided notice that it holds discretionary authority to regulate LDTs and has published guidelines, but not yet enforced them.

The total time span from bench discovery to approval and general clinical application is measured in years, and the total cost is counted in tens of millions of dollars, to be weighed by clinicians, and eventually by society at large, against the undoubted good sense of administering a targeted therapy only to those patients likely to benefit, and the avoidance of side effects and costs of inappropriate treatment of the remainder. This route to approval developed with reference to IHC tests, the most common method adopted for companion diagnostics to date; but other methods as they appear are constrained by similar rules.

As targeted therapies have proliferated, so of course have the corresponding biomarkers, and the methods applied for their detection

Table 1.4 FDA-approved biomarkers and LDTs

Test	Commonly applied tumor types
HER2	Breast, gastric
PD-L1	Melanoma, lung, kidney, head and neck, uterus
CTLA-4	Melanoma
CD 20	B lymphoma, CLL
CD 30	ALCL, Hodgkin L
ALK	Lung
TOPO1	Bladder, breast, colon, uterus, ovary
MMR (MLH1,MSH2, MSH6,PMS2)	Colon
EGFR	Colon, lung, pancreas, thyroid
VEGF	Lung, kidney, glioblastoma, colon,
TUBB3	Lung, bladder, uterus, kidney, prostate
PTEN	Breast, uterus, head and neck, lung, prostate
ER, PR	Breast, uterus, ovary
K-ras	Lung, colon
myc	Lymphoma
BCR-ABL 1	CML, (Ph chromosome)
BRCA 1	Breast, others
c-KIT protein	GIST
ERCC1	Bladder lung
BRAF	Melanoma, lung, colon, others
Immune cell profile[a]	Melanoma, lung, colon, breast, others
PSA CEA, p53, p21, Ki67	Various tumors, prognostic mainly
Multiple tissue biomarkers	Several hundred molecules demonstrated by IHC are use in diagnostic surgical pathology[b]

Multiple methods are applied [9–13]; to date the majority of FDA-approved biomarkers are demonstrated directly in tissues by IHC for diagnostic and or predictive use
[a]Immune cell profile, including CD3, CD4, CD8, CD20, CD68, FoxP3, and others (e.g., see Fig. 1.3)
[b]IHC tests (stains) used in surgical pathology as aids to diagnosis are considered Class 1 by the FDA. They require in lab validation

(Table 1.4). The practice of surgical pathology is being forced to change to meet these new demands (Fig. 1.2) [9–11]. Commensurately with these new assays, there has been a growing recognition of the need for higher standards of testing, in particular higher levels of control and reproducibility of test results from lab to lab (Tables 1.3 and 1.5). At long last the anatomic or surgical pathology laboratory that performs these tests, or at a minimum is involved in providing and preparing the tissues for these tests, is being held to the standards of the clinical laboratory.

Method Validation

For blood-based assays in the clinical laboratory, including serum biomarkers, a reference range usually is established that includes 95% of the "normal" population, with the "reference range" becoming the de facto definition of normalcy. Establishing a reference range is part of "routine" practice in the clinical laboratory and usually involves the testing of a defined population of "normal" subjects (may be a 100 or more), but not so in tissue-based anatomic pathology and not so with many of the newly developed companion diagnostics, where often only subcomponents of the "total test" (Tables 1.3 and 1.5) are validated, in spite of quite large case numbers incorporated into clinical trials.

In the validation of any new assay, and companion diagnostics are no exception, sample size is a matter of the clinical sensitivity and specificity of the test, variation in the population, confidence levels, and statisticians; it usually is accomplished during the transition from discovery (investigational use only (IUO)) to a validated assay (approved IVD) (Table 1.2) [4, 12, 13]. The matter is complex, beyond the compass of this introductory chapter, but is discussed in greater depth in succeeding chapters.

Suffice to say that for all assays that rely upon the use of tissue from cancer patients, the challenges in meeting these demands have been great, but not quite insurmountable. Effective sample (tissue) preparation has emerged as a neglected but key consideration for all assays, both IHC and those dependent upon extracts of FFPE tissues

(Table 1.3). In accommodating these demands, the practice of pathology has changed forever [9].

The Range of Methods

Viewed retrospectively, the first companion diagnostic of this present era was, as noted previously, an IHC–FFPE-based test for Her2 that incorporates cell line-based technical controls, a defined protocol and scoring guidelines derived from conjoint clinical studies. Subsequently, this prototypic IHC Her2 test has served as the model for a multitude of newly developed predictive biomarker tests, developed to match the burgeoning repertoire of targeted therapies [4, 6, 8, 9]. In addition, other technologies have been introduced to the companion diagnostic arena (Table 1.6), including ISH (in situ hybridization), PCR (polymerase chain reaction), and sequencing (Sanger or NGS – next-generation sequencing), with clear and imminent extension into RNA expression methods and proteomics (usually mass spectrometry or reverse-phase protein array) [9–12]. To date these methods have mostly been designed to detect molecular biomarkers, DNA (mutations), RNA (expression), or proteins (receptors, ligands, enzymes), either singly or in exploratory panels,

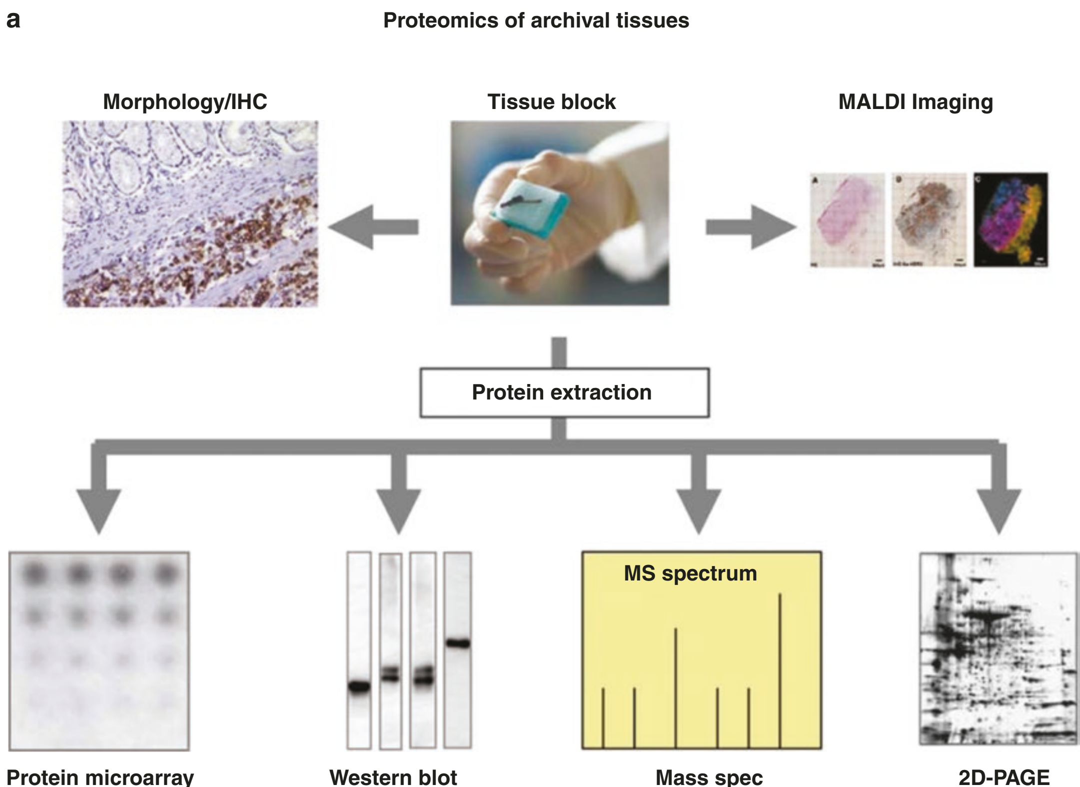

Fig. 1.2 Proteomics of archival tissue, and correlation with morphology, to capture cell origin of proteins of interest. (**a**) Many protein assay methods that are routinely used for frozen tissues can also be applied for FFPE tissues including immunohistochemistry (IHC), matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (MS), Western blot, protein microarray, and two-dimensional (2D) gel electrophoresis. (**b**) Extraction-based protein analysis with parallel IHC studies to capture exact cell(s) of origin of protein(s) of interest. (Reprinted from Taylor and Becker [11]. With permission from Wolters Kluwer Health)

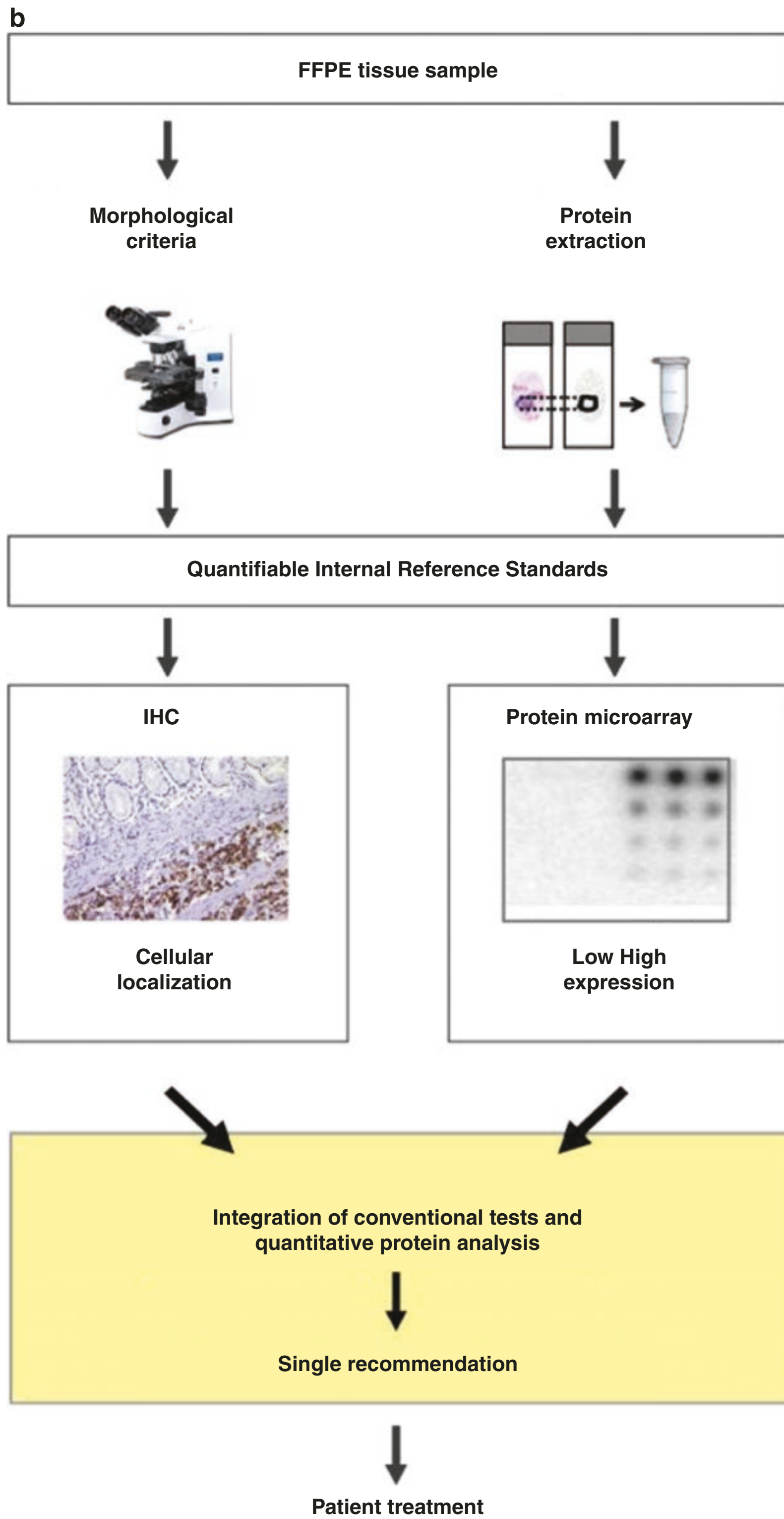

Fig. 1.2 (continued)

Table 1.5 Requirements for laboratory assays of cancer biomarkers

Total test approach – all aspects of test system should be encompassed, including sample preparation (Table 1.3)
Test method and analyte should have well-established performance characteristics
Test should be objective for read out/interpretation
Test ideally should produce a quantitative result (objective)
Threshold and reference range should be established
Test should be "fit to purpose", that is, designed and validated for the defined application
There should be well-developed control systems that are universally available
Test should be reproducible; run to run, day to day, lab to lab
Test should be readily performed and inexpensive

Based on data from Refs. [2, 3, 10, 11]

Table 1.6 Biomarker tests: commonly applied and developing methods

Sequencing: Sanger and NGS (next-generation sequencing)
Epigenetic differentiation
Laser capture microscopy
T and B cell receptor deep sequencing
Mass spectroscopy
Reverse phase protein arrays
RNA expression arrays
In situ hybridization (ISH)
Multiplex immunohistochemistry (IHC)

Based on data from Refs. [9–12]

exemplified by 40 plus mutation screens included in some NGS "tests" [4, 9, 12].

DNA and RNA sequencing methods can be traced back to the work of Frederick Sanger at the MRC Unit in Cambridge, England, in the 1970s [1]. Direct derivatives of his method provided the basis for the first sequencing of the human genome at the turn of the millennium. The achievement, time, and cost were extraordinary, but this success contributed to the development of multiple new approaches including the commercial availability of high-throughput sequencers, all of which together are known as next-generation sequencing (NGS). As a result the cost of sequencing a "cancer genome" has fallen dramatically and continues to fall, while availability, utility, and range of applications have enlarged so as to bring NGS from a discovery research mode into the realm of companion diagnostics. While the word genomics had been used half a century earlier, in practical terms this was the birth of the burgeoning field of "genomics" in medicine and in the public lexicon. Details of these various NGS approaches, instrumentation, reagents, methods, relative advantages, and disadvantages form the major topics of later chapters of this book.

The discovery of PCR, the polymerase chain reaction, is generally attributed to Kary Mullis in the 1980s [1]. It provided a means of almost infinite replication of defined DNA sequences that rapidly found an interface with Sanger DNA sequencing. Again numerous variants and derivative approaches have been described, and many have found major roles in the biomarker field, for highly sensitive detection of specific oncogenes, mutations, translocations, and the like in cancer, contributing to diagnosis, as well as much broader application in genetics as a whole.

DNA methods remain open to criticism in terms of clinical application, because not every change in DNA sequence is reflected in a change of cell function, a deficit that the biomarker field has attempted to repair through the use of RNA expression analysis, and studies of intermediate and end protein expression dubbed "proteomics." In the arena of cancer biomarkers, both transcriptional and posttranscriptional regulation have been studied extensively as described in later chapters. Proteomics as a concept, signifying both extensive and detailed analysis of tissue and cellular proteins, evolved also around the turn of the millennium as a companion of "genomics."

Detailed analysis of proteins has in many ways lagged behind related DNA and RNA analysis, for cogent reasons. Just as not every DNA sequence is translated to RNA, so not every RNA molecule is translated to protein, and RNA expression does not always correlate with protein expression. The whole process is increasingly recognized as being dynamic beyond earlier beliefs; in short, while the genome is relatively

fixed and constant across time and across all of the cells of the organism (excepting a "few" mutations in cancer and in aging), the proteome dramatically is not, varying from tissue to tissue, cell to cell, and time to time.

Paradoxically, analysis of proteins by immunologic techniques has a long history, including, as noted at the beginning of this chapter, early biomarkers [1]. For example the ELISA (enzyme-linked immuno-sorbent assay) method devised by Stratis Avrameas has served as a gold standard for measuring individual proteins in fluids for well over half a century [1]. Detection of protein in a frozen section tissue environment by immunofluorescence was described by Albert Coons 80 years ago [1] and was adapted to FFPE sections for general routine use in the author's laboratory 40 years later and 40 years ago [1, 2]. However, these methods dependent as they are on the use of a specific antibody were directed to the protein of interest, typically detected only one protein at a time, until more recent developments as described subsequently.

Thus the advent of proteomics, in the context of "massive" analysis, awaited the use of techniques such as mass spectrometry, protein "chips," and reversed-phase protein arrays described in later chapters [11] (Fig. 1.2a, b). These methods initially proved difficult to standardize, for reasons of cell diversity and physiology as noted above and for technical reasons relating to extraction from FFPE tissue, principally unknown levels of degradation and loss, and in mass spectrometry, variable peptide recovery and detection (Fig. 1.3). Last but not least, interpretation of the huge data sets that were

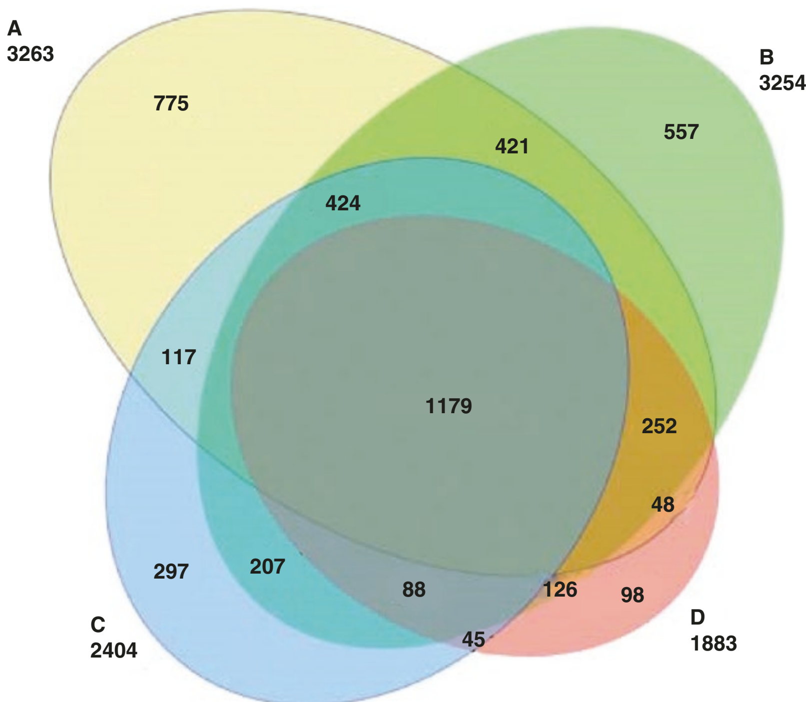

Fig. 1.3 Importance of validated sample preparation for mass spectrometry extraction-based proteomics. Four differently prepared extracts of the same renal carcinoma showing the number of distinct protein entries mapped by mass spectrometry using capillary isoelectric focusing (CIEF) with capillary reversed-phase liquid chromatography (RLPC). Samples A and B were extracted from FFPE tissue sections by using protocol of heat-induced retrieval with Tris-HCl buffer containing 2% SDS under different pH (pH 9 for A; pH 7 for B). Sample C was extracted from fresh tissue of the same case. Sample D was extracted from FFPE tissue by a protocol without heating treatment. (Reprinted from Shi et al. [17]. With permission from Sage Publications)

generated was a challenge. Much as with NGS, advancement of these methods was contingent upon the manifold increases in computer data analysis that occurred concurrently.

Each of these very different methods has inherent advantages and disadvantages. Most have been applied to extracts of FFPE tissues, or directly to FFPE tissue sections (IHC, ISH); all methods employed FFPE tissues – "because that is what we have" when the need for the test is recognized. Pathologists have long known that the process of formalin fixation and paraffin embedment compromises the integrity of all of the analytes tested by each of these methods, to differing degrees that are not yet completely understood. It is a significant problem that must be recognized and controlled whatever the method employed.

Extraction methods also require that the tissue that is subject to extraction contains a sufficient proportion of tumor cells versus normal cells (usually >20–30% for NGS), and mutated versus germ line DNA among the tumor cells (usually >10% depending upon method), in order to avoid a false-negative result [12]. Also for certain "biomarkers," such as "immune cell profiles," there are data that the use of tissue extracts necessarily sacrifices morphologic cellular and spatial information that may be critical to therapy choice and outcome. Selective extraction of tissue sections by microdissection or laser capture microscopy may also discard the very cell populations that subsequent tests seek to measure (e.g., immune cells). IHC has exquisite sensitivity on a cell to cell basis as already referenced but in the past has suffered from choice and quality of reagents, inefficient labeling methods, and subjective reading of the result. These shortfalls may be addressed by proper use of the method, coupled with computer-based analysis [2, 3, 10, 13].

With the current realization that the patient's immune response to their tumor, or lack thereof, affects the therapeutic efficiency of many drugs, it has become critically important to assess the patient's "immune cell profile." Determination of the immune profile is currently believed to be important for a broad range of new therapies, for which patient selection is critical to outcome (e.g., PD-1, PD-L1) (Fig. 1.4) (Table 1.4) [8, 9, 12]. While information on the nature and extent of any immune response to tumor may be derived from sequencing and proteomics studies, such information is inferential and may be compromised by extraction methods. The immune response and its constituent cells and molecular signals may be directly visualized in situ within the tissue by multiplex IHC, which accordingly has been added to the repertoire of methods now available (Figs. 1.4 and 1.5) [13].

Also notable are recent ventures into an area that has been by some termed "liquid biopsy," usually implying examination of blood components and or blood cells, although others have used liquid biopsy for various methods of examining tissue extracts [11]. Analysis of circulating DNA fragments and circulating tumor cells falls under the former definition. These methods hold great promise. Initial work is reviewed in later chapters but is yet to enter the mainstream of clinical care in a major way.

Multiple Biomarker Analysis

Until recently most of the approved companion diagnostics, as well as those in current ongoing trials, have been based upon detection of a single biomarker, although NGS and proteomics increasingly provide the potential for multiple parallel analysis. Now new demands have emerged, with an even higher order of complexity. The notion that clinical decisions may be based upon identification of the presence, or absence, of a single molecular target (exemplified by HER2, or PD-1) has extended to attempts at stratifying patients with respect to more than one biomarker. For example, with some targeted therapies the "drug labeling" states that it is necessary in arriving at a clinical decision to evaluate not only PD-L1 but also ALK and EGFR. The ultimate expression of this multi-marker trend has found immediate application in methods to assess the immune cell environment in and around the tumor. In real terms, this approach seeks to evaluate not simply the tumor itself but

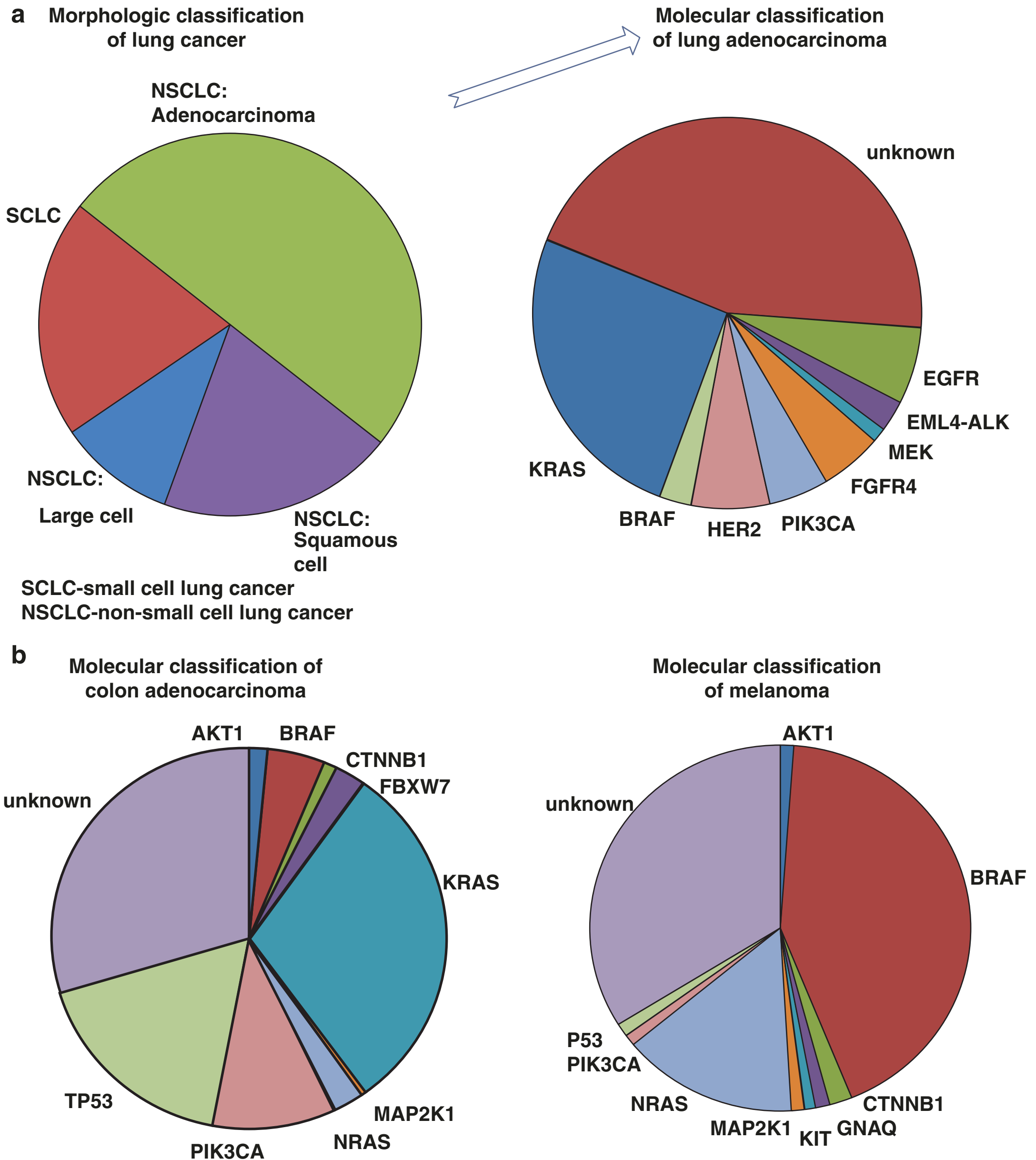

Fig. 1.4 Multiple "predictive biomarkers," exemplified by lung cancer, colon cancer, and melanoma. The "molecular" classification of these tumor types is superseding traditional morphologic classification as shown for lung cancer in (**a**); molecular profiles are shown for colon cancer in (**b**) and melanoma in (**c**). (Reprinted from Gu and Taylor [9]. With permission from Wolters Kluwer Health)

also the patient's immunologic response to the tumor, or lack thereof.

These studies have emerged primarily from evidence and resurgent enthusiasm for the "immunotherapy" of cancer, including the use of checkpoint inhibitors, exemplified by antibodies to CTLA-4 and PD-1, or its ligand PD-L1. Clinical trials, beginning with melanoma and extending rapidly to other solid tumors, indicated that patient responsiveness (or not) is dependent not only upon whether or not the tumor expresses the target (for the drug) but also whether there is an underlying immune response and whether such response is active or ineffective (suppressed).

Given the great complexity of the immune system in terms of both cellular and molecular

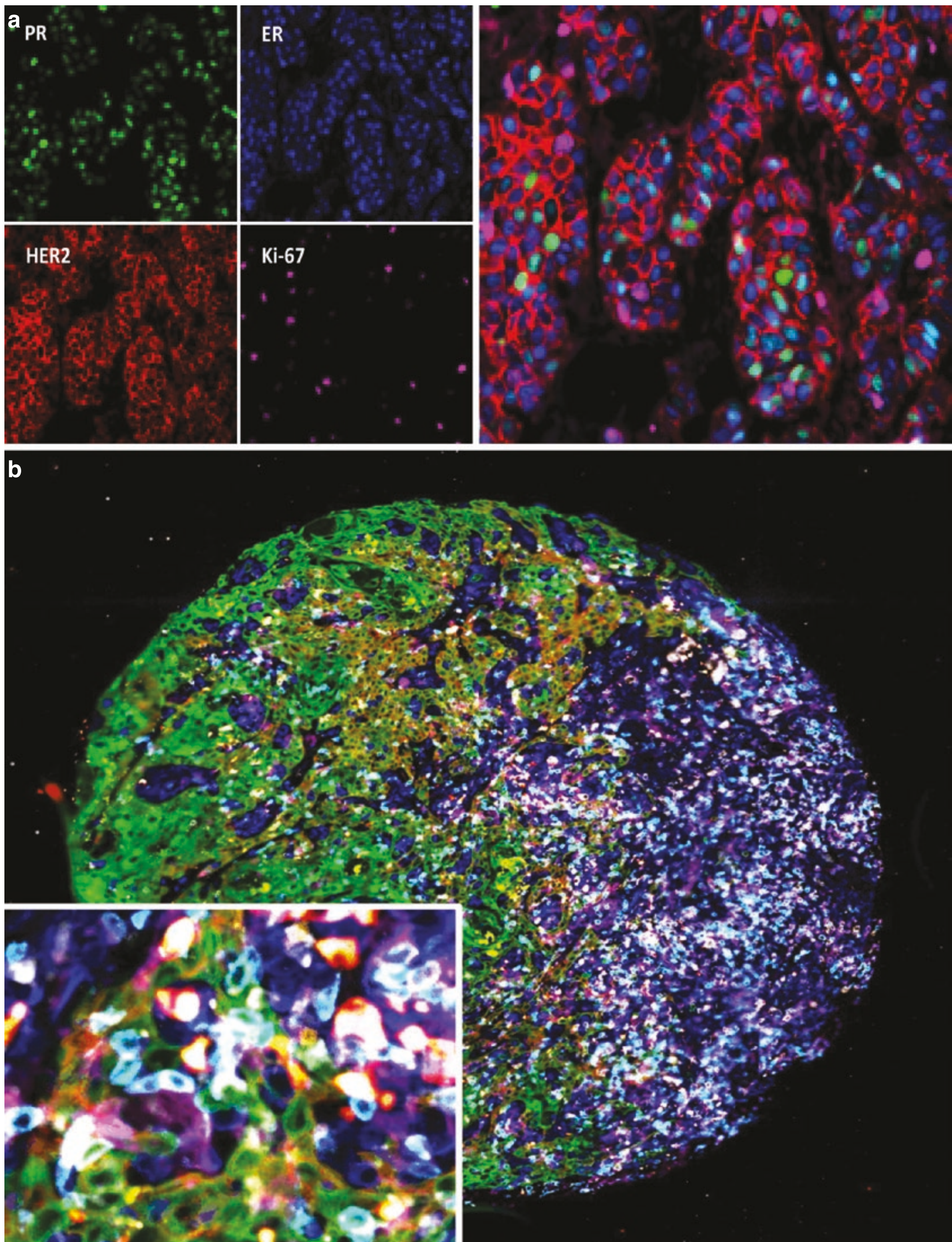

Fig. 1.5 Multiplex IHC ("Ultraplex"). (**a**) Quadraplex (four biomarkers) method. Triple positive breast cancer. On the left the four targets (colors) are displayed individually by the computer, allowing separate analysis. The composite image is on the right. PR, green; ER , blue; HER2, red; Ki67, magenta. (**b**) Decaplex (ten marker) method demonstrating cell identification, companion diagnostic and immune profile markers; squamous carcinoma, head and neck. Markers – cell identification: CK5, green; vimentin, blue. Companion diagnostic: EGFR, red. Immune cell profile: CD3, cyan; CD4, magenta; CD8, yellow; CD20, sepia; CD68, hot red; PD-1, gray; FoxP3, hot yellow. (Courtesy of David Schwartz, CEO, CSO (Cell IDx) with TMA samples provided by Mark Lingen, University of Chicago)

interactions, any "test" that evaluates only a single "biomarker" is unlikely to suffice. In addition, a means of evaluating the direct interface between the multiple types of immune cells and the tumor cells to which they are responding appear to be critically important. Lastly heterogeneity of biomarker expression in tumors has been recognized as a critical issue in terms of predictive value of testing, a concern that certainly includes evaluation of the immune cell infiltrate, not only variations in its intensity but also its character, focal or diffuse, and its location, intratumoral or at the invasive margin.

As noted, the presence of various immune cells and their state of activation may be inferred from proteomics or sequencing studies, including T cell receptor analysis, and information may be derived to class tumors as inflamed (hot) or non-inflamed [12]. However, numerical immune cell assessment, heterogeneity, and spatial relationships of multiple types of immune cells to each other and to tumor are necessarily compromised in any extraction-based assay and can only be fully assessed when considered in an undisturbed tissue-based context.

A Role for Multiplexed IHC Methods

"Multiplex" tissue-based IHC tests when performed in situ on FFPE sections of tumor tissue have the capability of displaying the "immune cell profile" (e.g., CD4, CD8, T regulatory lymphocytes, macrophages, myeloid-derived suppressor cells, etc.) and at the same time demonstrating the expression and distribution of regulatory molecules of interest, such as PD-1 and PD-L1, on tumor cells and associated immune cells (Fig. 1.5). On this basis tumors have been grouped into two broad categories, immunologically active (inflamed, hot) or immunologically silent (non-inflamed, ignorant, cold) (Table 1.7), which in turn have major implications for selection of classes of therapy, whether checkpoint inhibitors on the one hand or immune vaccines on the other.

Table 1.7 Two major classes of cancer as identified by immune profiling

	Immune silent/'ignorant'	Immunogenic/response suppressed
Class	"Non-inflamed"	'Inflamed'
Mechanisms	Lack of or tolerance to (self) tumor antigens (HLA)	"Tumor-induced" intrinsic suppression:
		Check point; PD-1; CTLA-4, Tim3, LAG3
		"Extrinsic" suppression: Tregs
		(CD25, FOXP3, Ki67), MDSC, blocking Abs
Tests		
Prognostic/predictive		
NGS/PCR	Low mutation load	High mutation load
NGS/RNA, protein, ISH/IHC	Targetable mutations – few	Targetable mutations – likely present
RNA, protein, IHC	Low check point expression	High check point expression; PD-1, PD-L1, CTLA-4, Tim3, LAG3
RNA, protein, IHC	Lack chemokines; immunomodulators	High immune modulators; suppressors dominate
Multiplex IHC	Lack – critical immune cells	High number critical immune cells; Tregs (CD25, FoxP3), MDSCs, macrophages (CD68)
Possible therapies	"Vaccines," immune activation modulators, BCG	Specific targeted therapy
		Checkpoint inhibitor blockade (PDL-1; PD-L1 block/deplete suppressor cells
	Recruit activated immune cells	Recruit and/or activate immune cells
	CAR T, CAR NK	CAR T CAR NK
Monitoring	Monitor immune profile change	Monitor immune profile change
	Monitor biodistribution	Monitor biodistribution
	CAR T, CAR NK, etc.	CAR T, CAR NK, etc.

These types of "immune profile" analyses clearly represent an entirely new class of assays for consideration, but equally clearly they are powerful "biomarkers" with both predictive and prognostic import.

Multiplex IHC is an extension of the basic IHC method, whereby several separate IHC protocols (four to eight or more) that are designed to detect different antigens (and cell types) are run on a tissue section in such a way that the results of all can be displayed and analyzed simultaneously. Several different approaches exist, either applying each separate antibody reaction sequentially, as in "Opal" (PerkinElmer), or "MultiOmyx" (Neogenomics) methods, the process taking 2 or more days to complete, or "UltraPlex" (Cell IDx) and "SigErMabs" (Calico Labs) that runs all reagents synchronously to complete a four- or ten-plex analysis in just 3 h (Fig. 1.5). Details of these methods are beyond the scope of this introductory chapter and are discussed elsewhere.

In brief, typically four or more differently colored fluorescent (or chromogenic) labels, each representing a different targeted molecule (protein, or nucleotide when combined with FISH), are developed on a single section. However, the human eye cannot distinguish the resultant kaleidoscope of colors (four to eight or more). Thus, this method has achieved practical utility only with the advent of high-resolution, high-speed tissue "scanners" that permit whole slide imaging and computer-based analysis of the complex multiple labels (Fig. 1.3a, b), coupled with sensitive, properly controlled, automated immune staining methods. Multiplex methods are evolving rapidly but are of course subject to similar standardization and total test requirements (Tables 1.3 and 1.5) as exist for other biomarker assays, including not only enhanced imaging and analysis methods but also high level controls for standardization [3].

The End of the Beginning

The challenges that this constellation of new test modalities presents to pathologists and clinicians should not be underestimated [2, 3, 10, 12, 13]. Neither should aspects of test availability and cost be neglected, for they may become the primary determining factors [14, 15]. There is ongoing debate with respect to choice of test, between "discovery-type tests" that assess multiple possible markers and generate huge data sets, but are very expensive, and tests that are specifically designed to answer a single question, to give the drug, or not, and are much less expensive. Some authors have explored the approach of using inexpensive, easy to perform tests, such as IHC, as screening tests, then following up with a more complex and expensive assay, only where clinically indicated [14].

Nonetheless, "precision" or "personalized medicine" appears to be an irresistible force, in turn requiring "precision pathology," which may be expected to result from further refinement and development of the methods, described briefly here, and discussed at greater length in the body of this book. Already the practice of pathology has been radically changed in the management of many malignant tumors (Fig. 1.4). Today we stand only at the end of the beginning of these changes; the ultimate end none of us as yet can foresee [9].

References

1. Van den Tweel J, Jiang J, Taylor CR. From magic to molecules: an illustrated history of disease: Beijing University Press; 2016.
2. Taylor CR. Quantitative in situ proteomics; a proposed pathway for quantification of immunohistochemistry at the light-microscopic level. Cell Tissue Res. 2015;360:109–20.
3. Cheung CC, D'Arrigo C, Dietel M, et al.; From the International Society for Immunohistochemistry and Molecular Morphology (ISIMM) and International Quality Network for Pathology (IQN Path). Evolution of quality assurance for clinical immunohistochemistry in the era of precision medicine: part 4: tissue tools for quality assurance in immunohistochemistry. Appl Immunohistochem Mol Morphol. 2017;25: 227–230.
4. Nass SJ, Phillips J, Patlak. Policy issues in the development and adoption of biomarkers for molecularly targeted cancer therapies. National Cancer Policy Forum. Workshop Summary. The National Academies Press. 2015. NAP.edu/10766.
5. European Parliament. Directive 98/79/EC of the European Parliament and of the Council of 27 October

1998 on in vitro diagnostic medical devices. 1998. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:1998:331:0001:0037:EN:PDF.
6. U.S. Food and Drug Administration. List of cleared or approved companion diagnostic devices (in vitro and imaging tools). 2016. Other nucleic acid based tests are listed separately under an included link. Updated 6/09/16. http://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/InVitroDiagnostics/ucm301431.htm.
7. U.S. Food and Drug Administration. Guidance for industry and FDA staff. In vitro diagnostic (IVD) device studies – frequently asked questions. 2010. http://www.fda.gov/downloads/MedicalDevices/.../ucm071230.pdf.
8. Mahoney K, Atkins MB. Prognostic and predictive markers for the new immunotherapies. Oncology. 2014;28(suppl 3):39–48.
9. Gu J, Taylor CR. Practicing pathology in the era of big data and personalized medicine. Appl Immunohistochem Mol Morphol. 2014;22:1–9.
10. Taylor CR. Predictive biomarkers and companion diagnostics. The future of immunohistochemistry: "in situ proteomics," or just a "stain"? Appl Immunohistochem Mol Morphol. 2014;22:555–61.
11. Taylor CR, Becker KF. "Liquid morphology": immunochemical analysis of proteins extracted from formalin fixed paraffin embedded tissues: combining proteomics with immunohistochemistry. Appl Immunohistochem Mol Morphol. 2011;19:1–9.
12. Yuan J, Hegde PS, Clynes R, et al. Novel technologies and emerging biomarkers for personalized cancer immunotherapy. J Immunol Ther Cancer. 2016;4:3.
13. Gaule P, Smithy JW, Toki M, et al. A quantitative comparison of antibodies to programmed cell death 1 ligand 1. JAMA Oncol. 2016; https://doi.org/10.1001/jamaoncol.2016.3015. Published online August 18, 2016.
14. Murphy DA, Ely HA, Shoemaker R, et al. Detecting gene rearrangements in patient populations through a 2-step diagnostic test comprised of rapid IHC enrichment followed by sensitive next generation sequencing. Appl Immunohistochem Mol Morphol. 2017;25: 513–523.
15. Yaziji H, Taylor CR. PD-L1 assessment for targeted therapy testing in Cancer: urgent need for realistic economic and practice expectations. Appl Immunohistochem Mol Morphol. 2017;25(1):1–3.
16. Taylor CR. Quality assurance and standardization in immunohistochemistry. A proposal for the annual meeting of the Biological Stain Commission. Biotech Histochem. 1992;67:110–7.
17. Shi S-R, Liu C, Balgley BM, Lee C, Taylor CR. Protein extraction from Formalin-fixed, paraffin-embedded tissue sections: quality evaluation by mass spectrometry. J Histochem Cytochem. 2006;54:739–43.