For reprint orders, please contact:
reprints@expert-reviews.com

Expert Reviews

Diagnosis of uncertain primary tumors with the Pathwork® tissue-of-origin test

Expert Rev. Mol. Diagn. 10(1), 17–25 (2010)

Federico A Monzon† and Catherine I Dumur

†Author for correspondence
Department of Pathology, The Methodist Hospital and The Methodist Hospital Research Institute, 6565 Fannin Street, MS-205, Houston, TX 77030, USA
Tel.: +1 713 441 3291
Fax: +1 713 441 1565
famonzon@tmhs.org

Clinical workup of metastatic malignancies of unknown origin is an arduous and expensive process, which is reported to be unsuccessful in up to 30% of cases. Global gene expression-based molecular testing may offer accurate classification of metastatic tumors in which a primary site has not been identified. Recently, the US FDA cleared the Pathwork® tissue-of-origin test, which is a gene expression microarray-based test that quantifies the molecular similarity of tumor specimens to 15 known tissue types. A blinded, multicenter validation on poorly differentiated and undifferentiated tumors showed 87.8% sensitivity and 99.4% specificity in frozen tissue samples. The availability of ancillary gene expression-based molecular tests for tissue of origin determination represents a milestone in cancer patient management as part of the personalized medicine revolution.

Keywords: gene expression analysis • gene expression classifier • molecular testing • tissue of origin • uncertain primary cancer

Uncertain primary cancer (UPC) is a diagnostic problem that often occurs in the context of metastatic disease without an identifiable primary tumor. This problem is also encountered when the morphologic appearance of the tumor is unusual or unexpected for the anatomic site. This situation leads to a thorough investigation for a primary site, with diagnostic imaging strategies, serum tumor markers and immunohistochemical (IHC) stains of tumor tissue [1–4]. This workup is associated with considerable cost, time and expense [5]. Primary tumors eventually remain unidentified in up to 30% of patients, with an initial diagnosis of UPC [6]. Those tumors that remain without an identifiable tissue of origin (TOO) after efforts to determine the primary have failed, are considered carcinoma of unknown primary (CUP), which is estimated to account for 2–4% of all reported cancers [2,7].

Prognosis of patients with CUP is poor, with median survival ranging from 6–10 months in clinical studies of unselected CUP patients to 2–3 months in other studies [8]. Treatment strategies are based on categorizing CUP patients into favorable and unfavorable groups [9,10]. Better prognosis in the favorable group probably reflects the fact the treatments are based on the more likely primary for each patient [2,11]. In fact, UPC patients in whom the primary source of cancer is ultimately identified have been shown to have longer survival [12]. Thus, successful identification of the TOO significantly impacts patient prognosis and management. It is this type of clinical need and the emergence of specific and more effective therapy regimens designed to combat metastatic disease from unique identifiable sites that have resulted in the quest for the more accurate identification of these tumors. To address this need, new molecular tests that identify molecular signatures of a TOO have become available.

Identification of the tissue of origin with gene expression profiling

Development of DNA microarrays in the mid-1990s allowed researchers to profile gene expression patterns for hundreds to thousands of gene transcripts in human tissues [13]. It was soon evident that these gene expression profiles could differentiate neoplasms from known distinct biologic categories, such as myeloid and lymphoid leukemias [14]. In 2001, two independent groups showed the ability to classify multiple tumor types based on their gene expression profiles [15,16]. These first studies were mainly focused on demonstrating the feasibility of

10.1586/ERM.09.75
© 2010 Expert Reviews Ltd
ISSN 1473-7159

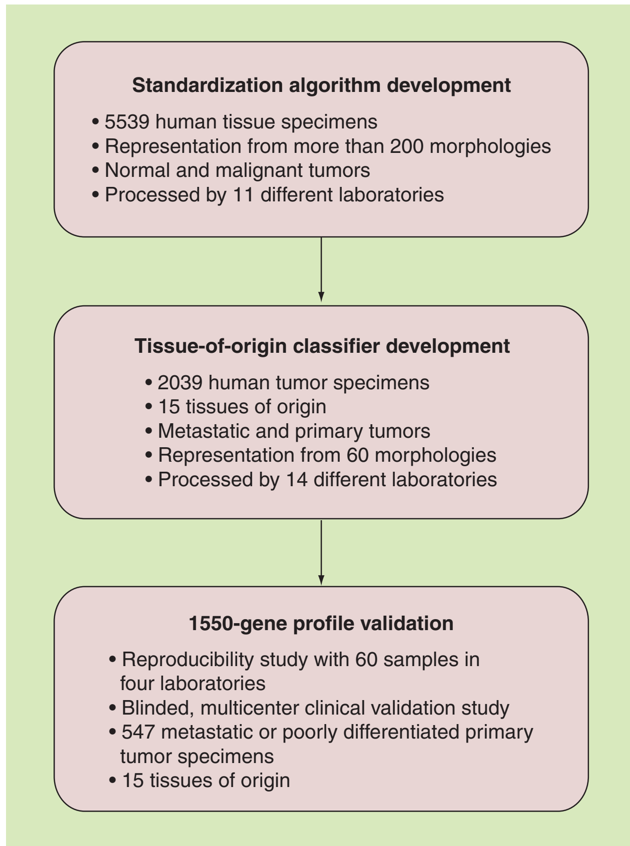

Figure 1. Development of the Pathwork® tissue-of-origin test.

classifying tumors based on their TOO and exploring the computing algorithms that were best suited for this purpose. Su and coworkers used supervised machine-learning algorithms to generate a tissue classifier based on a 110 mRNA transcript profile that showed 85% accuracy in classifying samples from an independent test set that included 11 tumor types (n = 75) [15]. Ramaswamy and coworkers developed a classifier for 14 tumor types based on support vector machines, utilizing 16,063 transcripts that achieved an overall prediction accuracy of 78% in the independent test set evaluated (n = 54) [16]. Interestingly, this algorithm failed to classify tumors of poorly differentiated morphology. Even at this early stage, Ramaswamy and collaborators recognized the potential clinical use of this technology for the diagnosis of clinically ambiguous tumors [16].

Shortly after, Bhattacharjee *et al.* showed that gene expression profiling could identify previously unrecognized metastases of extrapulmonary origin and suggested the use of this approach to confirm the origin of metastatic tumors in the lung [17]. In addition, Weigelt and collaborators showed that distant metastases maintain the inherent genomic profile of the primary tumor [18]. In 2004, Bloom and others developed an artificial neural network-based gene expression classifier that was successful in identifying the TOO in 85% of tumors profiled in multiple different platforms and laboratories (n = 140) [19], thus indicating that a robust clinical assay could be developed using microarrays. In 2005, Tothill *et al.* developed a microarray-based gene expression classifier that achieved an internal accuracy of 89% in classifying 13 tissue classes [20]. In this study, they showed that having a diverse training sample set that included multiple histologic subtypes from each tissue class was essential for the development of a robust classifier.

These studies demonstrated the feasibility of using gene expression profiling with DNA microarrays to classify uncertain tumors according to their TOO. However, in 2003, Tan and colleagues published a microarray platform comparison study in which they indicated that gene expression profiles were not reproducible when using different commercial microarray platforms [21]. This and other work raised concerns regarding the reliability of microarray-based gene expression assays and highlighted the need for strict quality control in the development of microarray-based clinical tests [22,23]. These concerns were addressed by a consortium of academic institutions and industry, in coordination with the US FDA, who, in 2005, published seminal papers in which they showed that the concordance between different microarray platforms had improved substantially owing to advances in gene annotation and array design [24], and showed that high reproducibility in gene expression data from microarrays could be achieved among multiple laboratories with the use of standardized protocols and array platforms [25–27]. This confirmation that gene expression profiling with microarrays could be reproducible and reliable set the stage for the development of commercially available clinical tests for the purpose of identifying the TOO in patients with uncertain primary.

Pathwork® TOO test

As mentioned previously, one of the challenges of developing a microarray-based clinical test was the reported variability in gene expression measurements between multiple laboratories [21]. Thus, the ability to compare data from different laboratories was a critical step in the development of the Pathwork® TOO test (Pathwork Diagnostics, Redwood City, CA, USA). To achieve this, Moradela and coworkers used gene expression profiles from 5539 human tissue specimens to develop a 121-gene standardization algorithm that allowed comparison of gene expression data from different laboratories. They then developed a classification algorithm from gene expression profiles of 2039 tumors comprising 15 tissue types and 58 different morphologies [28]. The training set included both primary and metastatic tumors and well differentiated to undifferentiated tumors (Figure 1). This algorithm is the basis of the TOO test, which uses 1550 genes to classify tumors into 15 known tissue types, representing 58 morphologies (Table 1) [101]. The test generates gene expression profiles using standard Affymetrix 3′-based amplification strategy of RNA from frozen tumor specimens with hybridization to a proprietary microarray (PathChip™) manufactured by Affymetrix (CA, USA) which is then processed on Affymetrix's US FDA-approved clinical instrumentation (Figure 2) [29].

Gene expression data is analyzed by the test's proprietary algorithm, which determines molecular similarity of the tissue to the 15-tissue panel (Table 1) and calculates a numerical

similarity score (SS) that ranges from 0 to 100 for each of the tissues. For the frozen tissue version of the test, a SS of 30 or greater is considered evidence that the specific tissue is present in the sample. A SS of less than 5 allows the site of origin to be ruled out and a SS below 30 but above 5 is classified as indeterminate. The analytical performance and reproducibility of this test was evaluated in a study conducted at four laboratories using archival frozen tissue from 60 poor-to-undifferentiated primary and metastatic tumors [29]. In this study, the test showed very good reproducibility in the standardized expression values, SS and final tissue of origin calls between all four sites. Although the study was not powered to evaluate clinical performance, average percent agreement between the test result and the reference diagnosis was reported to be 86.7% (range: 84.9–89.3%). In a subsequent multicenter validation study with 547 samples, the overall accuracy (positive percent agreement with reference diagnosis) was found to be 87.8% (95% CI: 84.7–90.4%) and overall specificity (negative percent agreement) of 99.4% (95% CI: 98.3–99.9%) [30]. It is important to mention that the independent validation sample of 547 specimens contained a minimum representation of each tissue type of 25 specimens.

Table 1. Tissue types represented in the tissue-of-origin test.

Tissue	Morphologies
Bladder	Adenocarcinoma and transitional cell carcinoma
Breast	Ductal, lobular, medullary, mucinous and papillary carcinomas
Colorectal	Adenocarcinoma (including mucinous)
Gastric	Adenocarcinoma and signet ring carcinoma
Germ cell	embryonal carcinoma, mixed germ cell tumor, seminoma and teratoma
Kidney	Clear cell, papillary and chromophobe carcinoma
Hepatocellular	Hepatocellular carcinoma and hepatoblastoma
Non-small-cell lung	Adenocarcinoma, adenosquamous, papillary, squamous and large cell carcinoma
Non-Hodgkin's lymphoma	Diffuse large B-cell, extranodal marginal zone, follicular, mantle cell and peripheral T-cell lymphoma
Melanoma	Malignant melanoma
Ovarian	Adenocarcinoma, carcinosarcoma (malignant mixed mullerian), cystadenocarcinoma, clear cell, endometriod and papillary serous carcinoma
Pancreas	Adenocarcinoma, acinar cell and intraductal papillary carcinoma
Prostate	Adenocarcinoma
Soft-tissue sarcoma	Angiosarcoma, carcinosarcoma, chondrosarcoma, desmoplastic small round cell tumor, fibromyxosarcoma, fibrosarcoma, gastrointestinal stromal tumor, leiomyosarcoma, liposarcoma, malignant schwannoma and osteosarcoma
Thyroid	Follicular, Hurthle cell, medullary and papillary carcinoma

The validation data for the frozen version of the test was reviewed by the FDA [12] and approved in July 2008 to be marketed as an *in vitro* diagnostic (IVD) device [102]. The FDA approval allows the company to commercialize a reagent kit that would allow molecular diagnostics laboratories to perform the test, provided they have the adequate equipment for array hybridization, washing and scanning. In this mode, a clinical laboratory would perform the wet-bench aspects of the test (including array processing) and send an electronic file with the gene expression data to the company to be processed with the diagnostic algorithm and issuing a report to be interpreted by the local pathologist. However, at the time of writing of this manuscript, no IVD kit has been released yet. Currently, Pathwork Diagnostics offers a version of the test for formalin-fixed, paraffin-embedded (FFPE) tissues in their Clinical Laboratory Improvement Amendments-accredited laboratory, and the frozen version is not currently available.

Other molecular tests for tissue-of-origin determination

Two main strategies have been followed by the commercial developers of TOO tests: the development of assays using microarray platforms and the use of information generated from gene expression microarray studies to develop quantitative reverse transcription real-time PCR assays (qRT-PCR) (Table 2). Microarray platforms have the advantage of being able to measure hundreds to thousands of transcripts in a single assay but they usually require more expensive instrumentation and long, complex protocols. qRT-PCR assays are, in practice, limited to assay tens of transcripts, so profiles that require hundreds to thousands of transcripts cannot be translated to this platform. Theoretically, hundreds of transcripts could be evaluated with real-time PCR in instruments that support microfluidic cards. Advantages of qRT-PCR assays are that they are easily translatable to FFPE specimens and qRT-PCR has shorter, simpler protocols and can be performed in equipment already present in most molecular diagnostics laboratories.

Apart from the Pathwork TOO test the other microarray-based test is the 1900-gene CupPrint assay (Agendia BV, Amsterdam, The Netherlands), which is offered clinically in Europe but not in the USA [31–33,103]. Other tests available in the USA as qRT-PCR laboratory-developed tests are Theros CancerTYPE ID® (Biotheranostics, CA, USA), a 92-gene qRT-PCR assay [31,104], and the miRview™ mets test (Rosetta Genomics, PA, USA), a 48-micro RNA qRT-PCR assay [34,105]. These currently available molecular tests for TOO determination have been reviewed in detail elsewhere [31]. Although some of these tests are currently being offered (and used) for clinical purposes,

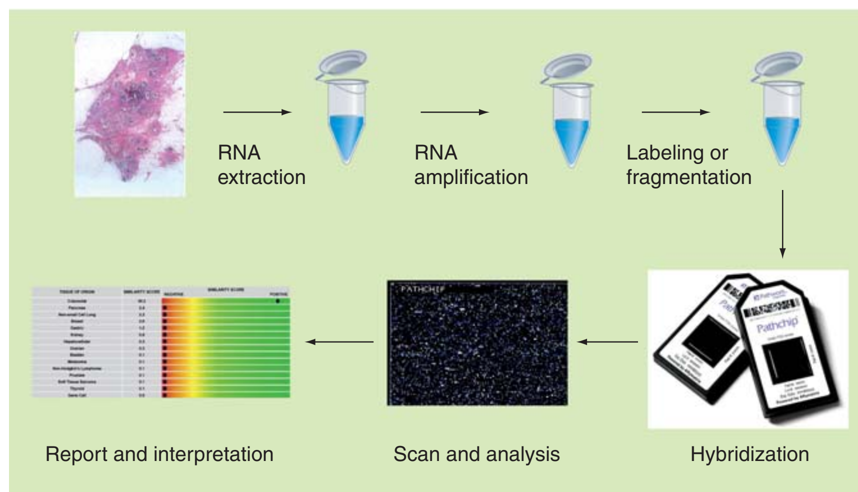

Figure 2. Laboratory workflow for the Pathwork® tissue-of-origin test.

the available peer-reviewed published data on their performance characteristics suggests that some do not meet published criteria for adequate translation of genomic classifiers or validation of clinical molecular tests (Table 2) [35–37]. Unfortunately, a conclusion regarding which test is better for establishing TOO in patients with uncertain primary cancer cannot be reached at this moment. A head-to-head comparison of test performance among the various TOO tests with the same tissue samples would need to be conducted in order to truly compare performance among these tests.

Conclusion

Carcinoma of unknown primary is an important clinical problem that generates frustration in surgeons, oncologists and pathologists, in addition to the uncertainty and stress it imposes on the patient. Incidence of CUP is 2–4% of all malignancies in two European countries [7,38] and, in the USA, it has been estimated that there will be 31,490 cases of cancer with unspecified primary site in 2009 [39]. The ability to molecularly classify tumors, and the advances in clinical microarray and PCR technologies, have resulted in the development of assays for TOO identification intended for clinical application. Oien has estimated that cases that undergo a UPC workup might be approximately double the number of CUP cases reported [1]. Given the frequency of this problem, TOO identification is a dilemma that oncologists and pathologists face frequently. Although some cases that require a UPC workup can be adequately resolved by the use of IHC, as well as consultations with expert pathologists, radiologists and oncologists, in many cases there is still uncertainty even after a full evaluation (a diagnosis might be favored but not with 100% certainty).

Based on published evidence, the classifier used in the Pathwork TOO test meets criteria for an adequate clinical validation – reproducibility of the test in a clinical environment was shown and the test was validated on an independent sample set of sufficient size (547 samples), class representation (≥25 samples per tissue type) and inclusion of indeterminate results, to meet the criteria for successful translation [35,36] and clinical validation [37].

We note that the performance characteristics of the Pathwork TOO test discussed here pertain to the test performed on frozen tissues, which is the version approved by the FDA. An FFPE version is now offered but limited performance information on this version is available. An abstract from the American Society of Clinical Oncology 2009 meeting describes an 89% agreement with reference diagnosis in 352 specimens [40]. However, a peer-reviewed publication of these data is not available yet. One of the effects of formalin fixation is the formation of cross links between proteins or between proteins and nucleic acids [41], which causes lower purity RNA when compared with frozen tissues. It has been shown that total RNA extracted from FFPE samples has been modified by the addition of monomethylol groups, especially to adenine, and it is fragmented as a result [42]. It is, thus, expected that RNA from FFPE tissues will be of lower quality than that obtained from frozen tissue. In the reproducibility study for the test's frozen version, we observed a decrease in sensitivity/specificity when specimens with low-quality RNA were processed [29]. Interestingly, the report for the FFPE version does not have an indeterminate category of results. Thus, all tissue types with a SS of 5 or greater can be considered possible sites of origin (with the highest of them being the more likely one). This has the potential to make interpretation of results difficult, since frozen specimens with low-quality RNA (from the reproducibility study) would often show more than one tissue type with SS results in the indeterminate range (between 5 and 30) and, in many cases, with two or more tissue types showing SSs close to each other. Importantly, these low-quality RNA samples were the ones with least reproducible results [29]. The high specificity of the frozen version of the test is, in part, due to the high stringency of the SS 30 or greater cutoff, which effectively reduces the appearance of false-positives. It is possible that modification of the bioinformatics algorithm utilized with the FFPE samples is able to compensate for this loss of nucleic acid quality. Studies showing direct comparison between FFPE specimens and matched frozen samples, reproducibility and clinical validation of the FFPE version have to be conducted in order to thoroughly assess the performance of the TOO test on formalin-fixed specimens.

Expert commentary

Molecular testing for TOO identification is now a reality and has the potential to become an important tool in the management of patients with UPC. However, there are still concerns regarding the reliability and value of these molecular tests. Since these

assays are intended to assist in samples with uncertain origin, some have voiced concerns about how well gene expression profiles from known tumors, used in the development and validation of these tests, reflect the biology of CUP samples [43]. By necessity, all studies reporting the development and evaluation of tests for TOO need to establish performance on samples from tissues of known type. Recently, Greco and colleagues used a retrospective analysis of UPC patients in which a primary tumor was eventually identified to evaluate the accuracy of one of the molecular tests for TOO identification (Biotheranostics' CancerType ID). Even though they evaluated 501 UPC patients, they could only find 16 patients with a subsequent primary tumor identification who had adequate tissue and that yielded a positive result with the test (CancerType ID) [44]. Thus, a large validation study of tissue identification accuracy with UPC samples of known origin is, in practice, not feasible. Furthermore, CUP samples cannot be used to establish the accuracy of tissue calls for these tests since, for the majority of these samples, the TOO is not known and, thus, a gold standard for tissue type determination cannot be established. This means that the current gold-standard procedures for TOO determination remain histopathology and imaging methods. Importantly, validation of antibodies used in IHC methods has historically been performed on tissues of known type [3,45]. Thus, the sample type used to validate molecular tests is the same one used to judge current methods in their specificity for a tumor or tissue type. Performance on known tissue types is a necessary first step in the clinical validation of molecular approaches for TOO determination and should reflect the accuracy of a tissue call when one tests a sample of uncertain origin.

A few recent studies have evaluated gene expression assays on CUP specimens. Horlings *et al.* used the CupPrint assay to analyze gene expression profiles from tumor samples from patients with CUP subdivided into three groups:

- Patients presenting with CUP and TOO identified by IHC (n = 16), in whom the test showed concordance with IHC diagnosis in 93.8%;
- Carcinoma of unknown primary patients with differential diagnosis of two or three sites after IHC (n = 12), in whom the test predicted a single-origin concordant with clinicopathologic information in eight of 12 cases;
- Uncertain primary cancer cases with no suspected primary site, in whom the test predicted a single-origin concordant with the clinical suspicion in six out of ten cases [46].

In another study using the same test, Bridgewater *et al.* reported clinically compatible results in 18 out of 21 tumors, which, in the authors' judgment, would have changed patient management if they had been available at the time of clinical decision-making [33]. We have evaluated the Pathwork TOO test in 21 CUP cases [Monzon et al.; Unpublished Data], in which the test identified a probable single primary site in 76% of cases, with all identified sites compatible with the available clinical information. In a 120-patient study that evaluated the ten-gene CUP assay, Varadhachary *et al.* identified

Table 2. Molecular tests for tissue-of-origin determination.

Test Name	Laboratory/ manufacturer	Availability	Technology	Accepted sample type	Reported accuracy* (%)	Independent test set	Ref.
Pathwork® tissue-of-origin test	Pathwork Diagnostics, Redwood City, CA, USA	Clinical US FDA cleared (frozen)	1550-gene microarray	Fresh frozen tissue	87.8	547 frozen samples	[29,30,101]
		LDT (FFPE)		FFPE tissue	89	352 FFPE samples	[40]
CupPrint	Agendia, BV, Amsterdam, The Netherlands	Clinical (in Europe)	1900-gene microarray	FFPE tissue	83	84 FFPE samples	[31,33,45,103]
Theros CancerTYPE ID®	Biotheranostics, San Diego, CA, USA	Clinical as LDT	92-gene qRT-PCR	FFPE tissue	86	119 FFPE samples	[31,104]
miRview™ mets	Rosetta Genomics, Philadelphia, PA, USA	Clinical as LDT	48-miRNA qRT-PCR	FFPE tissue	Not completely reported for qRT-PCR validation‡	65 FFPE samples	[34,105]
CUP assay	Veridex, La Jolla, CA, USA	Research	10-gene qRT-PCR	FFPE tissue	75.6	37 FFPE samples	[46,47]

*In peer-reviewed literature.
‡Reported only for specific binary decisions (i.e., liver vs nonliver or gastrointestinal vs nongastrointestinal).
CUP: Carcinoma of unknown primary; FFPE: Formalin-fixed paraffin-embedded; LDT: Laboratory-developed test; qRT-PCR: Quantitative reverse transcription PCR.

a putative TOO in 61% of patients [47]. Although the number of CUP samples analyzed is still small, these studies suggest that gene-expression based tests can obtain a molecular signature of a possible primary in 60–85% of CUP cases.

It is important to take into account that the accuracy of the call is essential for these test results to translate into patient benefit, both in terms of clinical outcomes and cost. As mentioned earlier, patients in whom a TOO is identified using current diagnostic approaches fare better than patients that remain with an unknown primary [12]. Thus, it is reasonable to expect that if TOO identification can be achieved by molecular profiling in up to 80% of patients currently classified as CUP, it would lead to better therapeutic selection, which could decrease the use of costly, ineffective therapies and improve patient outcomes. A recent study conducted by Varadhachary and coworkers showed that patients who were treated on the basis of a colorectal origin profile (based on Veridex's CUP assay) had better outcomes than those treated with conventional CUP management [48]. Furthermore, if the use of molecular approaches for TOO determination substitutes some of the currently performed diagnostics, it would have additional potential to reduce costs. In 1995, the cost for a UPC workup was estimated to be US$18,000, which is considerably higher than the current cost of molecular profiling (which ranges from $3350 to $3750). Thus, it is possible that patient management guided by results from molecular TOO tests could be reflected in better patient outcomes and reduced costs. Importantly, these hypotheses on potential patient benefit of molecular profiling of unknown primary tumors need to be explored with more prospective studies. It is desirable that these future studies compare the performance of the different molecular tests of TOO determination for accuracy of tissue calls and for the ability to improve patient outcomes. Only with studies that compare all the available options, will one be able to determine which one of them has the performance characteristics that will result in highest patient benefit.

When using the Pathwork TOO test, or any other molecular assay for that matter, on CUP samples, careful interpretation of the reported result within the clinical context of the case in hand is mandatory. One important variable to keep in mind is that the real TOO for a given CUP case might not be part of the tissue types included in the test panel. Currently, no available molecular test covers all the possible organs in the human body. Thus, one needs to take into account the clinically possible tissue types involved in the tumor sample in relation to the 15 included tissue types, as well as other tissue types and morphologies considered in the differential diagnosis, that are not covered by the assay. Ideally, one could expect to obtain indeterminate results when the TOO on the CUP sample belongs to those not included in the test. In our experience, this is not always true and a number of cross reactivities have been seen, especially on squamous cell carcinomas, such as those from the head and neck area. These tumors may be derived from multiple local organs (head and neck squamous cell carcinomas) or may represent metastases from other organs, including non-small-cell lung cancers. Since head and neck squamous cell carcinomas are not covered by the assay, results of molecular similarity to in-panel tissues with common squamous differentiation, such as non-small-cell lung or bladder cancers, are typically obtained with the Pathwork TOO test. This illustrates the importance of thoroughly understanding the test's advantages and limitations, along with the pathologic and clinical context of the patient, for the adequate interpretation of this and other molecular tests.

On the other hand, when suspecting two or more tissue sites as putative primaries on poorly differentiated metastases, the Pathwork TOO test offers a dual advantage: confirming one of the possibilities and excluding the other choices. In our experience, the ability of ruling out specific tissue types as the primary site for CUP patients has proven to be very useful in clinical scenarios involving patients with a prior history of cancer presenting with new tumor masses.

Other limitations for the routine implementation of this test are: the availability of frozen tissue, the requirement for more than 60% tumor content and the requirement to perform this test at a single laboratory. Collection of frozen tumor tissue is a relatively common practice in most academic settings, due to the need to bank tissues for research; however, these hospitals may only cover a minority of cancer cases and, thus, most cases of uncertain primary cancers will not have frozen tissue available. This limitation should be addressed by the availability and validation of an FFPE-based assay, as described earlier. Often, poorly differentiated tumors show extensive necrosis and/or an infiltrative pattern into normal tissue, thus making it sometimes difficult to obtain tissue that fulfills the requirements of greater than 60% tumor and less than 20% necrosis. As reported in the reproducibility study for this assay, accuracy of calls decreased when samples failing quality-control criteria were analyzed [29]. Thus, although these criteria will limit the number of suitable specimens for testing, they appear to be critical to ensure the accuracy of tissue calls.

Five-year view

We are now clearly in a new era in the diagnosis of tumors of uncertain origin that is part of the personalized medicine revolution. Gene expression-based molecular tests for TOO determination are available and have the potential to significantly impact patient management by potentially decreasing the number of unidentified primaries by approximately 70%. Since these tests address an important clinical need, one can forecast increased adoption in the near future. It is also reasonable to expect that further development of existing tests, and maybe appearance of new ones, will lead to better performance. For example, an expansion in the number of tissue types covered by the Pathwork TOO test could happen in the coming years. Such an increase in tissue coverage, if adequately validated, would overcome some of the current limitations of this and other available molecular assays. In addition, new data are emerging that are rapidly changing the homogeneity of tumor classification by morphologic appearance alone. Genomic tests that identify breast cancer molecular subgroups with different prognosis are already available. Thus, it is quite possible that, in the future, we will use molecular tools not only for identification of the site of origin, but also to identify molecular subsets of poorly differentiated and undifferentiated tumors that

will better respond to specific therapeutic approaches. In this regard, tests that are performed on microarray platforms would have an advantage, since the majority of the microarray content is currently not used and would enable these platforms to easily upgrade their test to include therapeutic selection.

Financial & competing interests disclosure

Federico A Monzon and Catherine I Dumur have received funding from Pathwork Diagnostics for the performance of the reproducibility and clinical validations studies of the tissue of origin test. Federico A Monzon has received honoraria for speaking engagements discussing the results of these studies and for consultation regarding development of tests unrelated to the tissue of origin test. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Key issues

- Uncertain primary cancer occurs when metastatic disease arises without an identifiable primary tumor, resulting in a poor prognosis scenario and a diagnostic dilemma. The emergence of specific and more effective therapy regimens that are designed to combat metastatic disease from unique identifiable sites has resulted in the quest for better and more accurate identification of these tumors.
- Genome-wide gene expression profiling using microarray technology can differentiate neoplasms from known distinct biologic categories, including distant metastases.
- Gene expression classifiers with a diverse training sample set that includes multiple histologic subtypes from each tissue class are essential in the development of robust clinical assays for tissue-of-origin determination, while high reproducibility of such assays are achieved with the use of standardized protocols and microarray platforms.
- The Pathwork® tissue-of-origin test was developed based on a classification algorithm from gene expression profiles of 2039 tumors, comprising 15 tissue types and 58 different morphologies, and was independently validated on 547 specimens containing at least 25 specimens for each tissue type.
- The test showed an overall accuracy (positive percent agreement with reference diagnosis) of 87.8% (95% CI: 84.7–90.4%) and overall specificity (negative percent agreement) of 99.4% (95% CI: 98.3–99.9%).
- Preliminary studies have shown that this and other tissue-of-origin tests can identify a putative tissue of origin in up to 85% of cases currently classified as carcinoma of unknown primary.
- Interpretation of results from such molecular testing on uncertain primary cancer cases require understanding of the test characteristics and careful evaluation by the pathologists and oncologists within the clinical context of the case.
- Further studies on the ability of molecular tests for tissue-of-origin determination to improve patient outcomes are necessary, as well as studies that compare performance of different assays with the same tissue samples and patients.

References

Papers of special note have been highlighted as:
• of interest
•• of considerable interest

1. Oien KA. Pathologic evaluation of unknown primary cancer. *Semin. Oncol.* 36(1), 8–37 (2009).
2. Pavlidis N, Fizazi K. Carcinoma of unknown primary (CUP). *Crit. Rev. Oncol. Hematol.* 69(3), 271–278 (2009).
   - **Excellent review on the clinical problem of carcinoma of unknown primary.**
3. Bahrami A, Truong LD, Ro JY. Undifferentiated tumor: true identity by immunohistochemistry. *Arch. Pathol. Lab. Med.* 132(3), 326–348 (2008).
4. Varadhachary GR, Abbruzzese JL, Lenzi R. Diagnostic strategies for unknown primary cancer. *Cancer* 100(9), 1776–1785 (2004).
5. Schapira DV, Jarrett AR. The need to consider survival, outcome, and expense when evaluating and treating patients with unknown primary carcinoma. *Arch. Intern. Med.* 155(19), 2050–2054 (1995).
6. Pavlidis N, Fizazi K. Cancer of unknown primary (CUP). *Crit. Rev. Oncol. Hematol.* 54(3), 243–250 (2005).
7. van de Wouw AJ, Janssen-Heijnen ML, Coebergh JW, Hillen HF. Epidemiology of unknown primary tumours; incidence and population-based survival of 1285 patients in Southeast Netherlands, 1984–1992. *Eur. J. Cancer* 38(3), 409–413 (2002).
8. Abbruzzese JL, Abbruzzese MC, Lenzi R, Hess KR, Raber MN. Analysis of a diagnostic strategy for patients with suspected tumors of unknown origin. *J. Clin. Oncol.* 13(8), 2094–2103 (1995).
9. Greco FA, Pavlidis N. Treatment for patients with unknown primary carcinoma and unfavorable prognostic factors. *Semin. Oncol.* 36(1), 65–74 (2009).
10. Hainsworth JD, Fizazi K. Treatment for patients with unknown primary cancer and favorable prognostic factors. *Semin. Oncol.* 36(1), 44–51 (2009).
11. Winn RJ, McClure JS. Occult Primary. V.I.2009. In: *NCCN Clinical Practice Guidelines in Oncology.* National Comprehensive Cancer Network, GA, USA (2009).
   - **Current guidelines for management of patients with occult primary.**
12. Bishop JF, Tracey E, Glass P, Jelfs P, Roder D. Prognosis of sub-types of cancer of unknown primary (CUP) compared to metastatic cancer. *J. Clin. Oncol.* 25(Suppl. 18), 21010 (2007).
13. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270(5235), 467–470 (1995).
14. Golub TR, Slonim DK, Tamayo P *et al.* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286(5439), 531–537 (1999).

15 Su AI, Welsh JB, Sapinoso LM *et al.* Molecular classification of human carcinomas by use of gene expression signatures. *Cancer Res.* 61(20), 7388–7393 (2001).

16 Ramaswamy S, Tamayo P, Rifkin R *et al.* Multiclass cancer diagnosis using tumor gene expression signatures. *Proc. Natl Acad. Sci. USA* 98(26), 15149–15154 (2001).

17 Bhattacharjee A, Richards WG, Staunton J *et al.* Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc. Natl Acad. Sci. USA* 98(24), 13790–13795 (2001).

18 Weigelt B, Glas AM, Wessels LF, Witteveen AT, Peterse JL, van't Veer LJ. Gene expression profiles of primary breast tumors maintained in distant metastases. *Proc. Natl Acad. Sci. USA* 100(26), 15901–15905 (2003).

19 Bloom G, Yang IV, Boulware D *et al.* Multi-platform, multi-site, microarray-based human tumor classification. *Am. J. Pathol.* 164(1), 9–16 (2004).

20 Tothill RW, Kowalczyk A, Rischin D *et al.* An expression-based site of origin diagnostic method designed for clinical application to cancer of unknown origin. *Cancer Res.* 65(10), 4031–4040 (2005).

21 Tan PK, Downey TJ, Spitznagel EL Jr *et al.* Evaluation of gene expression measurements from commercial microarray platforms. *Nucleic Acids Res.* 31(19), 5676–5684 (2003).

22 Johnson K, Lin S. QA/QC as a pressing need for microarray analysis: meeting report from CAMDA'02. *BioTechniques* (Suppl.), 62–63 (2003).

23 Ma C, Lyons-Weiler M, Liang W *et al.* *In vitro* transcription amplification and labeling methods contribute to the variability of gene expression profiling with DNA microarrays. *J. Mol. Diagn.* 8(2), 183–192 (2006).

24 Larkin JE, Frank BC, Gavras H, Sultana R, Quackenbush J. Independence and reproducibility across microarray platforms. *Nat. Methods* 2(5), 337 (2005).

25 Bammler T, Beyer RP, Bhattacharya S *et al.* Members of the toxicogenomics research consortium. standardizing global gene expression analysis between laboratories and across platforms. *Nat. Methods* 2(5), 351 (2005).

26 Dobbin KK, Beer DG, Meyerson M *et al.* Interlaboratory comparability study of cancer gene expression analysis using oligonucleotide microarrays. *Clin. Cancer Res.* 11(2 Pt 1), 565–572 (2005).

27 Irizarry RA, Warren D, Spencer F *et al.* Multiple-laboratory comparison of microarray platforms. *Nat. Methods* 2(5), 345 (2005).

28 Moraleda J, Grove N, Tran Q *et al.* Gene expression data analytics with interlaboratory validation for identifying anatomical sites of origin of metastatic carcinomas. *J. Clin. Oncol.* 22(Suppl. 14), 9625 (2004).

29 Dumur CI, Lyons-Weiler M, Sciulli C *et al.* Interlaboratory performance of a microarray-based gene expression test to determine tissue of origin in poorly differentiated and undifferentiated cancers. *J. Mol. Diagn.* 10(1), 67–77 (2008).

30 Monzon FA, Lyons-Weiler M, Buturovic LJ *et al.* Multicenter validation of a 1,550-gene expression profile for identification of tumor tissue of origin. *J. Clin. Oncol.* 27(15), 2503–2508 (2009).

•• **Largest validation study to date for a gene expression-based test for tissue-of-origin determination.**

31 Ma XJ, Patel R, Wang X *et al.* Molecular classification of human cancers using a 92-gene real-time quantitative polymerase chain reaction assay. *Arch. Pathol. Lab. Med.* 130(4), 465–473 (2006).

32 Glas A, Floore A, Delahaye L *et al.* Converting a breast cancer microarray signature into a high-throughput diagnostic test. *BMC Genomics* 7(1), 278 (2006).

33 Bridgewater J, van Laar R, Floore A, Van TVL. Gene expression profiling may improve diagnosis in patients with carcinoma of unknown primary. *Br. J. Cancer* 98(8), 1425–1430 (2008).

34 Rosenfeld N, Aharonov R, Meiri E *et al.* MicroRNAs accurately identify cancer tissue origin. *Nat. Biotechnol.* 26(4), 462–469 (2008).

35 Simon R. Roadmap for developing and validating therapeutically relevant genomic classifiers. *J. Clin. Oncol.* 23(29), 7332–7341 (2005).

• **Outlines and justifies requirements for adequate validation of genomic classifiers.**

36 Simon R, Radmacher MD, Dobbin K, McShane LM. Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification. *J. Natl Cancer Inst.* 95(1), 14–18 (2003).

37 Jennings L, Van Deerlin VM, Gulley ML. Recommended principles and practices for validating clinical molecular pathology tests. *Arch. Pathol. Lab. Med.* 133(5), 743–755 (2009).

38 Levi F, Te VC, Erler G, Randimbison L, La Vecchia C. Epidemiology of unknown primary tumours. *Eur. J. Cancer* 38(13), 1810–1812 (2002).

39 Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer Statistics, 2009. *CA Cancer J. Clin.* 59(4), 225–249 (2009).

40 Pillai R, Deeter R, Rigl CT, Halks-Miller M, Henner WD, Buturovic L. Validation of a microarray-based gene expression test for tumors with uncertain origins using formalin-fixed paraffin-embedded (FFPE) specimens. *J. Clin. Oncol.* 27(Suppl. 15), e22015 (2009).

41 Werner M, Chott A, Fabiano A, Battifora H. Effect of formalin tissue fixation and processing on immunohistochemistry. *Am. J. Surg. Pathol.* 24(7), 1016–1019 (2000).

42 Masuda N, Ohnishi T, Kawamoto S, Monden M, Okubo K. Analysis of chemical modification of RNA from formalin-fixed samples and optimization of molecular biology applications for such samples. *Nucleic Acids Res.* 27(22), 4436–4443 (1999).

43 Pentheroudakis G, Greco FA, Pavlidis N. Molecular assignment of tissue of origin in cancer of unknown primary may not predict response to therapy or outcome: a systematic literature review. *Cancer Treat Rev.* 35(3), 221–227 (2009).

44 Greco FA, Spigel DR, Yardley DA *et al.* Unknown primary cancer (UPC): accuracy of tissue of origin prediction by molecular profiling. *J. Clin. Oncol.* 27(Suppl. 15), 11070 (2009).

45 Dennis JL, Hvidsten TR, Wit EC *et al.* Markers of adenocarcinoma characteristic of the site of origin: development of a diagnostic algorithm. *Clin. Cancer Res.* 11(10), 3766–3772 (2005).

46 Horlings HM, van Laar RK, Kerst JM *et al.* Gene expression profiling to identify the histogenetic origin of metastatic adenocarcinomas of unknown primary. *J. Clin. Oncol.* 26(27), 4435–4441 (2008).

47 Varadhachary GR, Talantov D, Raber MN *et al.* Molecular profiling of carcinoma of unknown primary and correlation with clinical evaluation. *J. Clin. Oncol.* 26(27), 4442–4448 (2008).

48 Varadhachary GR, Raber MN, Matamoros A, Abbruzzese JL. Carcinoma of unknown primary with a colon-cancer profile-changing paradigm and emerging definitions. *Lancet Oncol.* 9(6), 596–599 (2008).

•• **First study to demonstrate outcome benefit for patients treated on the basis of a molecular profile for site of origin.**

Websites

101 Pathwork Diagnostics, Tissue of Origin Ltd www.pathworkdx.com/TissueOfOriginTest (Accessed 6/29/2009)

102 US FDA News Release. FDA clears test that helps identify type of cancer in tumor sample www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/2008/ucm116931.htm (Accessed 29 June 2009)

103 Cupprint. Decoding primary tumors http://row.agendia.com/en/cupprint.html (Accessed 29 June 2009)

104 THEROS CancerTYPE ID®. Changing the way you identify cancer www.biotheranostics.com/products-services/hcp/ctid (Accessed 29 June 2009)

105 Rosetta Genomics. MiRview www.mirviewdx.com (Accessed 29 June 2009)

Affiliations

- Federico A Monzon
  Department of Pathology, The Methodist Hospital and The Methodist Hospital Research Institute, 6565 Fannin St, MS-205, Houston, TX 77030, USA
  and
  Department of Pathology, Weill Cornell Medical College, 1300 York Avenue – C302, New York, NY 10065, USA
  Tel.: +1 713 441 3291
  Fax: +1 713 441 1565
  famonzon@tmhs.org
- Catherine I Dumur
  Department of Pathology, Virginia Commonwealth University, 403 North 13th St, CSC 247, Richmond, VA 23298, USA