Virchows Arch (2014) 464:393–402
DOI 10.1007/s00428-014-1545-2

REVIEW AND PERSPECTIVES

Gene expression profiling in patients with carcinoma of unknown primary site: from translational research to standard of care

John D. Hainsworth · F. Anthony Greco

Received: 8 January 2014 / Accepted: 14 January 2014 / Published online: 1 February 2014

Abstract Carcinoma of unknown primary site (CUP) is diagnosed in approximately 3 % of patients with advanced cancer, and most patients have traditionally been treated with empiric chemotherapy. As treatments improve and become more specific for individual solid tumor types, therapy with a single empiric combination chemotherapy regimen becomes increasingly inadequate. Gene expression profiling (GEP) is a new diagnostic method that allows prediction of the site of tumor origin based on gene expression patterns retained from the normal tissues of origin. In blinded studies in tumors of known origin, GEP assays correctly identified the site of origin in 85 % of cases and compares favorably with immunohistochemical (IHC) staining. In patients with CUP, GEP is able to predict a site of origin in >95 % of patients versus 35–55 % for IHC staining. Although confirmation of the accuracy of these predictions is difficult, the diagnoses made by IHC staining and GEP are identical in 77 % of cases when IHC staining predicts a single primary site. GEP diagnoses appear to be most useful when IHC staining is inconclusive. Site-specific treatment of CUP patients based on GEP and/or IHC predictions appears to improve overall outcomes; patients predicted to have treatment-sensitive tumor types derived the most benefit. GEP adds to the diagnostic evaluation of patients with CUP and should be included when IHC staining is unable to predict a single site of origin. Site-specific treatment, based on tissue of origin diagnosis, should replace empiric chemotherapy in patients with CUP.

J. D. Hainsworth (✉)
Sarah Cannon Research Institute, 3322 West End Avenue, Suite 900, Nashville, TN, USA
e-mail: jhainsworth@tnonc.com

F. A. Greco
Sarah Cannon Research Institute, 250 25th Avenue South, Suite 110, Nashville, TN, USA
e-mail: fgreco@tnonc.com

Keywords Gene expression profiling · Cancer of unknown primary site · Immunohistochemical stains · Molecular diagnostics

Introduction and background

Patients with carcinoma of unknown primary site (CUP) account for approximately 3 % of cancer diagnoses and present management problems for both clinicians and pathologists. Unlike most cancers that are advanced at the time of diagnosis, standard clinical evaluation does not identify the anatomic primary site, leaving a much broader spectrum of diagnostic possibilities for the pathologist to consider. As evidenced by autopsy series, anatomic primary sites can be located in most patients with CUP, but usually remain small (less than 2 cm), even in the terminal stages of metastatic cancer (1, 2). The explanation for this unusual biologic behavior is unknown; although a molecular explanation seems likely, no molecular abnormalities or molecular profiles common to CUP have yet been identified.

The initial evaluation of a CUP patient involves a relatively brief clinical evaluation (CT scanning and specific evaluation of signs/symptoms); if this is unrevealing, the identification of a tissue of origin depends on the pathologic evaluation. In addition to an examination of histology, standard pathological evaluation in CUP includes immunohistochemical (IHC) staining, in order to narrow the diagnostic spectrum or, less commonly, to definitively identify the tissue of origin. Since most CUPs are adenocarcinomas, the distinction between specific adenocarcinomas is the most frequent diagnostic issue. Other than the IHC stain for PSA, which is quite sensitive and specific for prostate cancer, individual IHC stains are inadequate to differentiate between various adenocarcinomas. A number of IHC panels have been developed to increase the diagnostic capability. Selection of the initial group of IHC

stains is guided by histology and clinical presentation; in some cases, a second set of IHC stains is performed based on results of the initial group. In spite of improvements in the specificity of IHC stains, a single tissue of origin cannot be predicted in a substantial percentage of CUP patients.

Until recently, treatment recommendations for patients with CUP had changed little during the last 20 years. Approximately 20 % of CUP patients have clinical and/or pathologic features that fit into one of several defined "treatable subsets" (Table 1). In general, these patients have clinical features that strongly suggest a specific diagnosis, even though an anatomic primary site cannot be identified. Patients are treated following guidelines for the presumed cancer type, and treatment outcomes mirror results expected for these cancers. The remaining 80 % of patients with CUP do not fall into any of these favorable subsets and have been treated with empiric chemotherapy. Although some patients derive substantial benefit, the results of treatment for most patients in this group are poor, and the median survival is only 9 months (3, 4).

When empiric chemotherapy for CUP was developed, treatment for most kinds of metastatic carcinoma was poor. Some advanced cancers (e.g., breast and ovarian) derived substantial survival benefit from optimum systemic therapy, but most cancer types (e.g., pancreas, kidney, lung) derived minimal or no benefit from treatment. Effective chemotherapeutic agents were non-specific, and development of a "broad-spectrum" combination with reasonable activity against sensitive tumor types was feasible.

Current treatment for most types of advanced cancer has changed substantially. Not only has survival been improved for many cancer types, but treatments are also more individualized for each specific cancer type. Table 2 summarizes the changes that have occurred in the treatment of two common advanced cancer types, colorectal and non-small cell lung cancer, during the 20 years between 1993 and 2013. Survival has improved substantially in both cancer types. The combination regimens used for each cancer type are tumor specific. None of the agents listed in Table 2 is included in empiric treatment of CUP. Optimum treatment of patients with advanced cancer therefore becomes virtually impossible without the identification of a site of origin.

Instead of continuing to evaluate new empiric chemotherapy regimens, most recent clinical trials in patients with CUP have addressed the development of better diagnostics to enable accurate prediction of the tissue of origin. Gene expression profiling (GEP) has been the major new diagnostic method evaluated during recent years, and this review critically evaluates its current status. Before GEP is accepted as a standard part of CUP management, clinical experience must confirm that: (1) predictions rendered by GEP are accurate, (2) GEP adds to the diagnostic capability of the pathologic evaluation currently performed (i.e., histology and IHC), and (3) site-specific treatment directed by GEP results improves the outcome for patients with CUP. Important new data are available to address these issues, although further clinical experience is necessary to "fine-tune" management recommendations.

Accuracy of gene expression profiling in predicting the site of tumor origin

Validation studies in advanced cancers of known origin

Specific gene expression profiles are now recognized in cancers based on their site of origin, reflecting the different gene

Table 1 Carcinoma of unknown primary site—favorable subsets

Subset	Typical histology	Therapy
Women, isolated axillary LN	Adenocarcinoma	Treat as stage II breast cancer
Women, axillary LN + other metastases	Adenocarcinoma	Treat as metastatic breast cancer
Women, peritoneal carcinomatosis	Adenocarcinoma (often serous) PDC	Treat as stage III ovarian cancer
Men, blastic bone metastases or high serum PSA or PSA tumor staining	Adenocarcinoma	Treat as metastatic prostate cancer
Colon cancer profile (intra-abdominal metastases + typical histology/IHC)	Adenocarcinoma	Treat as metastatic colon cancer
Single metastatic site	Adenocarcinoma/PDC	Definitive local therapy
Isolated cervical LN	Squamous	Treat as locally advanced head/neck cancer
Isolated inguinal LN	Squamous	Definitive local therapy (inguinal node dissection and/or radiation therapy) ± chemotherapy
Extragonadal germ cell syndrome	PDC	Treat for poor prognosis germ cell tumor
Neuroendocrine carcinoma, low grade	Carcinoid/islet cell features	Treat as advanced carcinoid
Neuroendocrine carcinoma, aggressive	Small cell or PDC	Treat as small cell lung cancer

PDC poorly differentiated carcinoma, *PSA* prostate-specific antigen, *IHC* immunohistochemistry, *LN* lymph node

Table 2 Therapy for advanced cancer—improved and increasingly tumor specific

Cancer type	Median survival (months)		Effective drugs not included in empiric CUP regimens[a]
	1993	2013	
Colorectal, stage IV	8	23	Oxaliplatin, cetuximab, panitumumab, bevacizumab, aflibercept, regorafenib
Non-small cell lung cancer (NSCLC), stage IV	6	12	Pemetrexed, erlotinib, crizotinib, bevacizumab

[a] All drugs approved by US FDA

expression profiles present in their normal tissues of origin (5). The potential application of these findings to cancer diagnosis was first demonstrated when gene expression differences allowed the distinction of acute myeloid leukemia from acute lymphoblastic leukemia (6). By measuring the differential expression of different gene sets, this diagnostic method can potentially be applied to many cancer types, and is therefore a potentially important diagnostic tool in the evaluation of CUP.

During the last 15 years, a number of assays have been developed for the purpose of predicting the tissue of origin in patients with CUP. Early assays were limited by using relatively few gene expression markers, allowing the diagnosis of relatively few tumor types (7–10). Improved methodology, using either reverse transcriptase polymerase chain reaction (RT-PCR) or gene microarray techniques, coupled with improved bioinformatics have enabled the creation of assays to detect more than 40 distinct tumor types/subtypes (11–19). Most of the recent clinical data has been generated using one of three assays: CancerTYPE ID (bioTheranostics, Inc.), Cancer Origin Test (Rosetta Genomics), or the Tissue of Origin Test (Pathwork, Inc.). The first two of these assays are currently commercially available.

The CancerTYPE ID assay is a 92-gene RT-PCR assay that allows the identification of 30 main tumor types and 54 subtypes (17). The Cancer Origin Test uses 64 tissue-specific microRNAs to enable the identification of 42 tumor types, using microarray technology (19). Both of these assays were validated using biopsies of tumors from known primary sites. Biopsies were taken from either the primary site or a metastatic site; tumors with well-differentiated and poorly differentiated histology were included. In these validation studies, each of which included specimens from several hundred tumors, both assays correctly identified 85 % of the tumors included (17, 19). These two assays are compared and contrasted in Table 3.

These results with current GEP assays confirmed a high level of accuracy in identifying the origin of advanced cancers of known primary site and provided strong rationale for their continued evaluation in the diagnosis of CUP.

Accuracy of GEP in CUP—retrospective studies

The accuracy of GEP in patients with CUP is inherently difficult to assess, since the anatomic primary site is never identified in the large majority of patients. Although the accuracy of currently available GEP assays is impressive in patients with known cancer types, the unique biology of CUP has caused hesitation regarding the generalization of these findings. It has been speculated that gene expression may be altered in CUP so that the gene expression profiles diverge from the tissue of origin and are more difficult to recognize. Alternatively, the possibility of a profile "specific" for CUP has been postulated; the identification of which may lead to therapeutic opportunities, but may also hinder diagnosis by GEP.

The most direct evidence supporting the accuracy of GEP in CUP comes from a study of CUP patients who had an anatomic primary site ("latent" primary) identified later during the disease course (range 9–314 weeks, median 49 weeks after diagnosis) (23, 24). Twenty-four such patients were retrospectively identified and represented 3.7 % of a group of 652 patients with CUP seen between 2001 and 2010. Gene expression profiling was performed using the original biopsy material. In this group of 24 patients, the prediction by GEP matched the anatomic primary site in 18 patients (75 %). In two patients, the assay results were indeterminate, possibly due to limited available tumor tissue, while in four patients (17 %), the assay predictions were incorrect.

Further support for the accuracy of GEP predictions in CUP comes from patients who had additional pathologic or

Table 3 Currently available gene expression profiling assays—comparison and diagnostic accuracy in cancers of known primary site

	Cancer TYPE ID (17, 20–22)	Cancer Origin Test (19)
Assay platform	RT-PCR	Gene microarray
Genes assayed	92 (mRNA)	64 (microRNA)
Tumor types recognized	30 (54 subtypes)	42
Biopsy specimen required	FFPE	FFPE
Number of tumors in reference database	2,557	1,282
Accuracy in validation set	83 % (*N*=187)	85 % (*N*=509)
	87 % (*N*=790)	
Accuracy in high-grade cancers	78 % (*N*=132)	NR
Accuracy in neuroendocrine tumors	95 % (*N*=75)	NR

RT-PCR reverse transcriptase polymerase chain reaction, *mRNA* messenger RNA, *FFPE* formalin-fixed paraffin-embedded, *NR* not reported

clinical study following a GEP prediction that was unexpected. In one report, a group of 35 CUP patients had a specific tissue of origin predicted by GEP (24). Previous standard clinical and pathologic evaluations had been non-diagnostic. However, directed pathologic studies (specific IHC stains) or clinical evaluation provided additional support for the GEP predictions in 26 patients (74 %). Examples of these additional studies included RCC and PAX8 IHC stains after the GEP prediction of renal cell carcinoma or Hepar-1 stain and serum alpha-fetoprotein after the prediction of hepatocellular carcinoma. Although this evidence is circumstantial in most cases, since an anatomic primary site was not actually identified, the consistency among various diagnostic methods supports the accuracy of GEP.

Similar studies have been reported in a group of 22 CUP patients predicted to have renal cell carcinoma by GEP (papillary 8, clear cell 7, unknown subtype 7) (25). Histology in these patients included poorly differentiated carcinoma in 15 and adenocarcinoma in 7 (4 with papillary features). All 22 patients had normal kidneys on CT scans. Nine patients had sufficient tissue for additional IHC stains; in seven of nine patients, these additional IHC stains (RCC, PAX8, others) supported the diagnosis of renal cell carcinoma.

A group of 30 patients with poorly differentiated malignant neoplasms of unknown primary site has also been recently reported (26). In these patients, standard pathologic examination (including a median of 18 IHC stains) was unsuccessful in definitively identifying the tumor lineage. Gene expression profiling yielded a lineage prediction in 25 of 30 patients as follows: carcinoma 10 (germ cell 3, neuroendocrine carcinoma 2, others 5), melanoma 5, sarcoma 8, and hematopoietic neoplasm 2. Fifteen of these tumors were then studied with additional IHC stains, genetic testing for BRAF or i(12p) chromosomal abnormalities, or repeat biopsy. In 11, further studies supported the GEP prediction.

In several retrospective studies, GEP has been performed on biopsies from patients with CUP. Profiling results were then correlated with clinical presentation, standard pathologic evaluation, response to treatment, and subsequent disease course (19, 27–32). In these studies, a variety of profiling assays were used, including some capable of detecting only a few primary sites. Although the evidence presented in these studies is indirect and therefore of lesser importance than the studies already discussed, the following were consistent findings: (1) GEP resulted in the prediction of tissue of origin in the majority of patients with CUP; (2) in most of these patients, IHC results were atypical or non-diagnostic; and (3) clinical presentation and response to treatment were usually compatible with the GEP diagnosis.

At the same time that these results from multiple clinical studies supported the value of GEP, gene expression differences previously postulated to account for the unique biology of CUP have been difficult to demonstrate. In a recent study, biopsies from women with CUP involving either axillary nodes or the peritoneum (features defining two of the favorable treatment subsets, Table 1) were compared to reference patients with breast cancer or ovarian cancer, respectively (33). No differences in gene expression profiles could be detected using the Cancer Origin Test.

In summary, GEP accurately predicts the tissue of origin in the majority (approximately 75 %) of patients with CUP. In spite of the unique biology of CUP, these cancers apparently retain enough of the gene expression profile of their tissue of origin to allow identification with current GEP assays.

Comparisons of IHC staining and GEP

Accuracy of IHC staining in identifying advanced cancers of known primary site

Precise information regarding the accuracy of IHC staining in predicting the primary site is difficult to obtain, for a number of reasons. Panels of stains have been advocated for at least 20 years, and during that time, new and more specific IHC stains have been developed. A number of algorithms are currently proposed to guide the selection of stains (34–41). In general, these algorithms involve the use of a few initial stains (first round), followed if necessary by additional stains directed by initial IHC results and clinical features. However, specific stains recommended differ among algorithms, and the selection and number of stains performed can differ considerably among pathologists, particularly after the "first round" of stains. Furthermore, data for these algorithms was derived from tumors in which the number of IHC stains performed was not limited by availability of tissue, as is often the case in clinical practice.

Several studies have addressed the ability of panels of IHC stains to correctly identify advanced carcinomas of known primary site. In these studies, tumor biopsies were evaluated with panels of IHC stains, selected according to the algorithm being validated; pathologists were blinded as to the primary site. Although the IHC panels varied, correct primary site identification was achieved in 64–67 % of cases in four studies that included biopsies from metastatic lesions only (34, 35, 38, 40). In two studies that included biopsies from either metastatic or primary sites, the accuracy rose to >80 % (38, 39). In a meta-analysis, the mean expected accuracy was 65.6 % (95 % confidence interval 60.1–70.7 %) for analysis of metastatic lesions and 82.3 % (95 % confidence interval 77.4–86.3 %) for blended groups of metastatic and primary site biopsies (42). Several additional IHC panels have been proposed for the identification of the primary site, but have not been validated in prospective studies of similar design (43–46).

The consistent finding of an accuracy rate of approximately 65 % in the analysis of metastatic tumors highlights the limitations of IHC staining and indicates the need for additional diagnostic methods in a substantial proportion of tumors. Although similar studies with GEP reported higher accuracy rates (approximately 85 %), definitive conclusions are difficult without direct comparisons.

Comparisons of the accuracy of IHC staining versus GEP in advanced cancer of known primary site

Recently, two relatively large studies have assessed the accuracy of IHC staining versus GEP in the identification of a primary site in patients with metastatic cancer (21, 47). In both studies, pathologists were provided formalin-fixed paraffin-embedded biopsy specimens from patients with metastatic cancer. They were informed of the patients' gender and the site of the biopsy, but no other clinical information was provided. In both studies, pathologists were allowed to use as many IHC stains as they felt necessary for optimum evaluation (this approach to IHC staining is possible in only a minority of patients in the routine clinical practice, since many biopsies are small). After completion of the IHC evaluation, pathologists predicted a single primary site in all cases, even if the IHC staining was not completely specific. Different GEP assays were used in these two studies. The study reported by Weiss et al. (21) used the CancerTYPE ID assay (16, 17, 20) and the Handorf study (47) used the Tissue of Origin Test (Pathwork Diagnostics) (14, 15, 31).

The results of these two studies are summarized in Table 4. In both studies, IHC staining and GEP were able to correctly identify the primary site in the majority of patients. In the Weiss study, 122 cases had enough tissue available for complete IHC staining and gene expression profiling (21). Pathologists used a median of 7.9 IHC stains/case. IHC/morphology analysis correctly identified the site of origin in 84 of 122 cases (69 %), as compared to 96 of 122 (79 %) for GEP. Accuracy in several tumor types common in the CUP population is shown in Table 4; although the numbers are small in some categories, both assays correctly identified a majority of cases in each category.

In the Handorf study, 157 cases were examined with both techniques (47). The median number of IHC stains requested by pathologists was similar (8.3/case). In the 157 cases evaluated, IHC staining correctly identified the primary site in 83 versus 89 % for GEP. In a subset of 33 poorly differentiated carcinomas, GEP accuracy exceeded IHC staining (91 to 71 %). When the first round of IHC stains was successful in allowing a primary site prediction, both techniques exceeded 90 % in accuracy. The accuracy of IHC diagnosis fell when a second and/or third round of stains was requested by the pathologist (67 vs. 83 % for GEP). Accuracy was high in most of the individual cancer types common in the CUP population (Table 4).

Taken together, these studies show considerable accuracy for both IHC staining and GEP in identifying the primary sites in patients with metastatic cancer of known primary site. In each study, GEP performed slightly better than IHC staining; although the numbers are small, the differences in performance are probably greater in patients with poorly differentiated tumors. The accuracy of IHC staining may have been favorably influenced by the specific tumor types selected for these studies. Some tumor types (e.g., non-small cell lung, colorectal, breast, ovary, kidney) are more accurately detected by IHC staining and were well represented in these studies, while others for which IHC diagnosis is less specific (e.g., pancreas, gastric, biliary, urothelial) were present in small numbers.

Table 4 Accuracy of IHC versus gene expression profiling in identification of metastatic tumors of known primary site

Known primary site	Weiss et al. (21)			Handorf et al. (47)		
	Number of specimens	Accuracy (%)		Number of specimens	Accuracy (%)	
		IHC	GEP		IHC	GEP
All	122	69	79	157	83	89
Poorly differentiated histology	–	NR	NR	33	71	91
Lung	24	67	75	6	100	83
Colon	17	94	94	25	92	100
Breast	11	55	73	25	84	100
Ovary	5	100	100	8	75	88
Kidney	13	77	77	14	100	86
Bladder	11	45	82	10	43	60
Stomach/esophagus	5	60	60	7	29	29
Pancreaticobiliary	4	75	50	5	73	60
Prostate	4	50	100	3	56	100

IHC immunohistochemistry, *GEP* gene expression profiling, *NR* not reported

Studies using this design are impossible in patients with CUP, since the primary site usually remains unknown, but it is likely that differences in the accuracy of the two methods are magnified, since (1) poorly differentiated tumors, and (2) cancer types less easily diagnosed by IHC staining are common in this population.

IHC staining versus GEP in CUP

A few studies have addressed the results of GEP and IHC staining in patients with CUP. In the largest study, GEP was performed in 149 patients who had previously been evaluated with a standard IHC evaluation (median 8 stains) (24). A single site of origin was diagnosed by IHC staining in 52 patients (35 %), although all patients had retained the diagnosis of CUP, since an anatomical primary site was undetectable.

The major findings of this study are diagrammed in Fig. 1. When the IHC evaluation resulted in the diagnosis of single site of origin, GEP gave an identical prediction in 77 % of cases. However, IHC staining was unable to predict a single site of origin (i.e., two or more sites of origin were suggested, or results were non-specific) in 97 cases (65 %). In these patients, the correlation between IHC staining and GEP results was poor. In 43 cases (44 %), GEP predicted 1 of the primary sites suggested by IHC staining; however, in 54 cases (56 %), the GEP result did not match any of the suggestions made by IHC staining.

In 35 of the 54 patients with sufficient biopsy material remaining and in whom the GEP prediction did not match any of the IHC suggested diagnoses, attempts were made to determine whether the GEP prediction was accurate. Additional IHC stains and clinical evaluation supported the GEP prediction in 26 of 35 patients (additional details are in the "Accuracy of GEP in CUP—retrospective studies" section).

Four other studies have included a total of 65 CUP patients in whom IHC staining diagnosed a single tissue of origin; GEP performed in these patients gave an identical prediction in 51 patients (78 %) (28, 31, 48, 49). Therefore, in 78 % of 117 patients reported in 5 studies, the GEP diagnoses matched IHC diagnoses when IHC staining was able to predict a single tissue of origin. However, the reported ability of IHC staining to make a single diagnosis in CUP patients was less than 55 % in all studies.

Although it is difficult to make definitive conclusions based on the results of retrospective studies, several observations are clinically relevant and should be confirmed in future trials. First, the ability of IHC staining to predict a single site of origin in CUP patients (<55 % in all studies) is lower than previously documented in studies in metastatic cancers of known origin, where accuracy ranged from 64 to 83 % (21, 34, 35, 38, 40, 47). Second, in patients with a single prediction made by IHC staining, the correlation with GEP results is high. Oncologists have often been reluctant to make treatment decisions based on IHC predictions and in the past have frequently treated these patients with empiric chemotherapy. The finding that these IHC predictions match GEP predictions in 78 % of cases supports their accuracy and suggests the following: (1) site-specific treatment guided by IHC staining

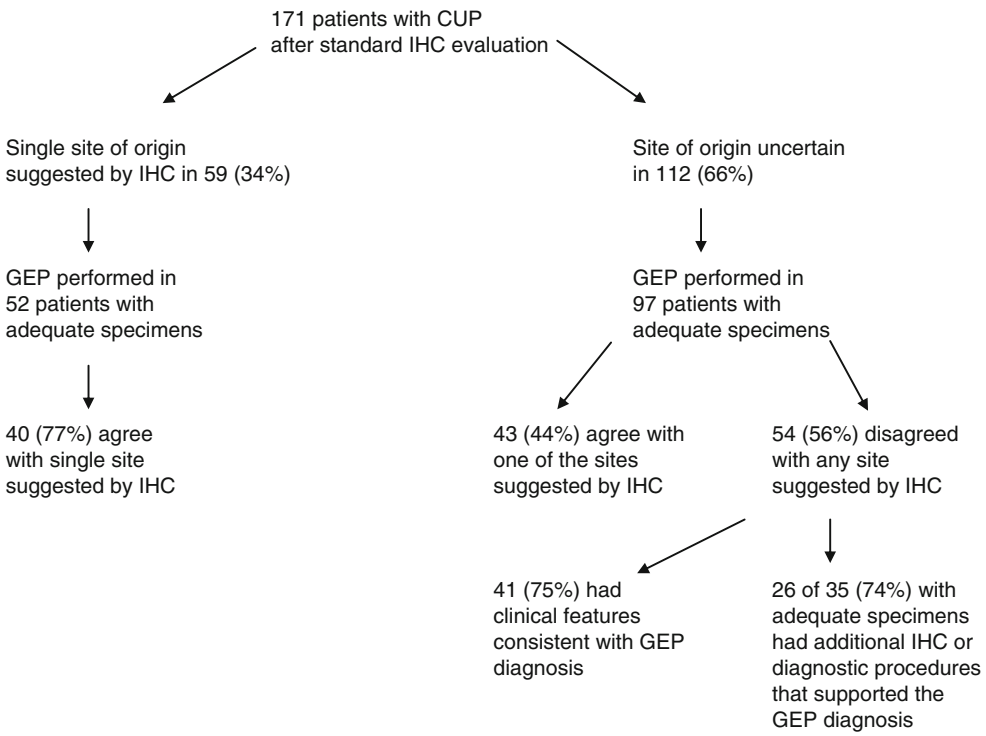

Fig. 1 Identifying the tissue of origin in patients with CUP: contribution of gene expression profiling

should be considered in these patients and (2) GEP may be unnecessary in patients with a single IHC diagnosis of the tissue of origin. Finally, GEP appears to add substantially to the diagnostic evaluation of patients with CUP who do not have a single primary site predicted by IHC staining. In this group, GEP rendered a prediction in a large majority; the prediction was supported in most by subsequent pathologic and clinical evaluation.

Site-specific therapy for patients with CUP—evidence for improved efficacy versus empiric chemotherapy

The final determinant of the clinical relevance of any diagnostic procedure is its ability to guide and improve the results of therapy. In the preceding sections, evidence has been presented that the tissue of origin can correctly be identified in most patients with CUP, using a combination of current IHC stains and GEP. One would assume that accurate identification of the site of tumor origin would effectively guide treatment and would improve therapy in many patients. However, clinical data confirming this hypothesis have developed only recently, and considerable skepticism remains. Although definitive randomized trials are not available, all clinical results to date support the use of site-specific therapy based on IHC staining and/or GEP predictions. These results are briefly summarized in this section.

Retrospective studies

For several of the tissues of origin identified in patients with CUP, including colon and kidney cancer, optimum therapy extends survival substantially but differs markedly from the empiric chemotherapy for CUP (which is ineffective in these tumor types). Three retrospective studies have focused on patients with CUP in whom a colorectal site of origin is predicted by either IHC staining or GEP (50–52). In one of these studies, CUP patients with adenocarcinoma and IHC staining typical of colorectal cancer were identified (48). Two groups were considered separately: group 1 contained 32 patients with CDX-2-positive and CK20-positive/CK7-negative stains and group 2 contained 36 patients with CDX-2-positive staining, regardless of CK20/CK7 results (either atypical or not done). Patients received treatment for metastatic colon cancer and had median survivals of 37 months (group 1) and 21 months (group 2). Both groups therefore had survival in the range expected for metastatic colon cancer and substantially better than the 8–10-month median survival achieved with empiric chemotherapy in CUP. To our knowledge, this is the only study in which IHC predictions have been used to select site-specific therapy for patients with CUP.

Two studies retrospectively identified CUP patients predicted to have a colorectal primary site by GEP (51, 52). Patients in both of these reports had also had IHC evaluations, and approximately 50 % had staining typical of colorectal cancer. Colonoscopies were uniformly negative. In patients who received treatment for metastatic colon cancer following the GEP prediction, median survivals were 21 and 27 months in the two series, again similar to the expected survival of patients receiving current therapy for metastatic colon cancer.

Prospective study

A single large, prospective study used GEP predictions to direct therapy in patients with CUP (53). In this trial, patients with a new diagnosis of CUP (after standard clinical and pathologic evaluation) had GEP with the CancerTYPE ID assay. When the assay predicted a site of origin, patients received standard treatment for the predicted cancer type.

In this study, 252 patients had successful GEP performed and 247 (98 %) had a tissue of origin predicted (Table 5). Twenty-six different primary sites were predicted; however, primary sites in the gastrointestinal tract or lung accounted for a majority, as expected. The frequency with which biliary tract and urothelial carcinomas were predicted was unexpected, based on previous autopsy series; these tissues of origin may be more common than previously appreciated, or this may represent a difficulty with the assay accuracy in predicting these primary sites.

Table 5 Gene expression profiling predictions in 252 consecutive patients with CUP

Predicted tissue of origin	Number of patients (%)
Biliary tract	52 (32)
Urothelium	31 (12)
Colorectum	28 (11)
Lung (non-small cell)	27 (11)
Pancreas	12 (5)
Breast	12 (5)
Ovary	11 (4)
Gastroesophageal	10 (4)
Kidney	9 (4)
Liver	8 (3)
Sarcoma	6 (2)
Cervix	6 (2)
Neuroendocrine	5 (2)
Prostate	4 (2)
Germ cell	4 (2)
Skin, squamous	4 (2)
Carcinoid, intestine	3 (1)
Mesothelioma	3 (1)
Others (8 sites represented)	12 (5)
Unclassifiable	5 (2)

One hundred ninety-four patients received site-specific assay-directed therapy; the median survival for these patients was 12.5 months. When patients were separated into groups of "more responsive" and "less responsive" tumor types based on the GEP diagnoses, the median survival for the more responsive patients was 13.4 versus 7.6 months for the less responsive subgroup ($p=0.04$). When individual tumor types were examined, median survivals in most instances mirrored those expected in patients with these tumor types (biliary tract 6.8 months, pancreas 8.2 months, colorectal 12.5 months, non-small cell lung 15.9 months, ovary >30 months, breast >30 months).

The results of this study support the premise that site-specific therapy, directed by GEP, improves the management of patients with CUP. Although the median survival of 12.5 months for the entire group may seem disappointingly close to the median survivals of 9 months previously achieved with empiric chemotherapy, it must be remembered that 41 % of patients in this group were predicted to have tumor types that respond poorly to standard therapies. For these tumor types, the impact of even the "best" therapy is modest. As predicted, patients with cancer types more responsive to standard therapies derived greater benefit from this approach, and the appropriate treatment of these patients is currently the strongest argument for site-specific therapy. However, an accurate diagnosis of an unresponsive tumor type is also of clinical usefulness and does not detract from the importance of GEP as a diagnostic procedure. Although a randomized trial in which patients are treated with either site-specific GEP-directed therapy versus empiric chemotherapy would provide the "definitive" answer to this question, such a study would require several hundred patients and seems unlikely to be completed.

Screening CUP for "actionable" molecular abnormalities

Most cancer therapies recently introduced or currently in development have been designed to exploit cancer-specific molecular abnormalities critical to cancer growth and metastasis. The identification of appropriate patient populations for an increasing number of these drugs depends not only on the identification of the primary site but also on the existence of specific genetic alterations within the cancer. Screening of patients with specific cancer types for these molecular abnormalities (e.g., BRAF v600e mutations in melanoma, HER2 overexpression in breast cancer, EGFR activating mutations in non-small cell lung cancer) is already a standard part of oncology clinical practice. Identification of these abnormalities allows treatment with specific agents and improves treatment outcome for these patient subsets. During the next few years, it is likely that broader screening panels for "actionable" molecular abnormalities will become commonplace in the evaluation of advanced cancer, and patients identified with specific molecular abnormalities will receive a trial of appropriate targeted therapy, regardless of the site of tumor origin.

Limited information currently exists regarding the prevalence of various "actionable" molecular abnormalities in patients with CUP. However, given the heterogeneity of cancers represented in the CUP population, it seems likely that some of these critical mutations and abnormalities exist. It is known that the incidence of mutations in CUP is relatively high, and previous studies and case reports have documented the existence of specific molecular abnormalities including HER2 overexpression and various mutations (EGFR, PI3K, MET, others) (54–58). Further characterization of these abnormalities and their frequency in the CUP population is important, and may lead to additional treatment options for some patients.

The incorporation of site-specific therapy directed by improved diagnostic techniques may also lead to assessment of specific molecular abnormalities based on the site of origin prediction. For example, further specific studies including assays for EGFR activating mutations and ALK and ROS1 rearrangements are indicated in CUP patients predicted to have non-small cell lung cancer. A small group of such patients has been reported where such studies identified ALK rearrangements, leading to effective treatment with crizotinib, an ALK inhibitor approved for the treatment of ALK-positive non-small cell lung cancer (59). The coordinated use of GEP for diagnosis and other molecular assays (e.g., next-generation sequencing techniques) to identify specific "actionable" molecular abnormalities is likely to become more common in the management of CUP in the future.

Summary and conclusions

During the last several years, improved diagnostic methods have improved the likelihood of accurately predicting the tissue of origin in patients with CUP. Although the roles of IHC evaluation and GEP in the evaluation of CUP patients remain incompletely defined, as do the optimal applications of these techniques in selecting treatment for individual patients, information accumulated during the last few years supports the following:

- Accurate prediction of the tissue of origin is possible in the majority of patients with advanced cancers of known primary site, using either panels of IHC stains or GEP. In two blinded comparisons, GEP was more often accurate than IHC staining and requires only a small amount of tissue to perform. Gene expression profiling appears to be more accurate in the diagnosis of poorly differentiated carcinomas.

- Most CUPs retain sufficient gene expression similarities to the tissue of origin as to allow identification by GEP. In a single study, GEP correctly identified 75 % of the primary tumors in a group of CUP patients who had latent primary sites discovered months to years later during their clinical course. Similar studies are not available using panels of IHC stains. However, the ability of IHC staining to predict a single site of origin is probably <55 % in patients with CUP.
- In evaluating CUPs, there is good correlation between IHC staining and GEP results (78 %) when IHC panels result in the prediction of a single primary site. These data reciprocally support the accuracy of both IHC staining and GEP in this setting.
- Gene expression profiling adds to the diagnostic capability in patients with CUP when IHC stains are unable to predict a single tissue of origin.
- Tissue management is important in CUP to ensure that the optimal diagnostic tests (IHC staining and GEP) are performed with the biopsy specimen available.
- Site-specific treatment directed by GEP predictions improves the results of CUP treatment in patients identified with responsive tumor types. Experience with site-specific therapy directed by IHC predictions is limited; however, this approach accurately identified a group of CUP patients with a "colorectal cancer profile" who responded well to treatment for advanced colon cancer. At present, it is unclear whether GEP should be included in the diagnostic evaluation of all patients with CUP; such an approach is probably unnecessary, but additional information is required. Perhaps CUP patients who have a single site of origin predicted by IHC staining can be treated on this basis, without also requiring GEP.

The era of empiric chemotherapy as a treatment for patients with CUP is ending. The increasing specificity of treatment for different types of advanced cancer requires an attempt at site-specific treatment for patients with CUP. Improved diagnostic methods, including IHC panels and GEP, enable accurate prediction of the site of origin in the majority of patients. GEP is a valuable addition to the diagnostic evaluation and should be included unless IHC evaluation allows a definite prediction of the tissue of origin. Ongoing clinical trials are required to refine these recommendations, to better delineate the advantages of site-specific therapy, to assess the cost-effectiveness of GEP, and to further evaluate the importance of profiling for additional "actionable" molecular abnormalities.

Conflict of interest Dr. Greco is a member of the bioTheranostics Speakers Bureau. Dr. Hainsworth has no conflicts of interest to report.

References

1. Nystrom JS, Werner JM, Heffelfinger-Juttner J et al (1977) Metastatic and histologic presentations in unknown primary cancer. Semin Oncol 4:53–58
2. Pentheroudakis G, Golfinopoulos V, Pavlidis N (2007) Switching benchmarks in cancer of unknown primary from autopsy to microarray. Eur J Cancer 43:2026–2036
3. Greco FA, Pavlidis N (2009) Treatment for patients with unknown primary carcinoma and unfavorable prognostic factors. Semin Oncol 100:65–74
4. Hainsworth JD, Greco FA: Cancer of unknown primary site. Brawley DW, et al. eds: ACS textbook of cancer. London: John Wiley and Sims LTD, pending publication
5. Su AI, Welsh JB, Sapinoso LM et al (2001) Molecular classification of human carcinoma by use of gene expression signatures. Cancer Res 61:7388–7393
6. Golub TR, Slonim DK, Tamayo P et al (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286:531–537
7. Talantov D, Baden J, Jatkoe T et al (2006) A quantitative reverse transcriptase-polymerase chain reaction assay to identify metastatic carcinoma tissue of origin. J Mol Diagn 8:320–329
8. Ramaswamy S, Tamayo P, Rifkin R et al (2001) Multiclass cancer diagnosis using tumor gene expression signatures. Proc Natl Acad Sci USA 98:15149–15154
9. Bloom GC, Eschrich S, Zhou JX et al (2007) Elucidation of a protein signature discriminating six common types of adenocarcinoma. Int J Cancer 120:769–775
10. Penland SK, Keku TO, Torrice C et al (2007) RNA expression analysis of formalin-fixed paraffin-embedded tumors. Lab Invest 87:383–391
11. Shedden KA, Taylor JMG, Giordano TJ et al (2003) Accurate molecular classification of human cancers based on gene expression using a simple classifier with a pathological tree-based framework. Am J Pathol 163:1985–1995
12. Yu J, Yu J, Almal AA et al (2007) Feature selection and molecular classification of cancer using genetic programming. Neoplasia 9: 292–303
13. Dumur CI, Lyons-Weiler M, Sciulli C et al (2008) Interlaboratory performance of a microarray-based gene expression test to determine tissue of origin in poorly differentiated and undifferentiated cancers. J Molecular Diagn 10:67–77
14. Monzon FA, Lyons-Weiler M, Buturovic LJ et al (2009) Multicenter validation of a 1550-gene expression profile for identification of tumor tissue of origin. J Clin Oncol 27:2503–2508
15. Pillai R, Deeter R, Rigl CT et al (2011) Validation and reproducibility of a microarray-based gene expression test for tumor identification in formalin-fixed paraffin-embedded specimens. J Molecular Diagn 13: 48–56
16. Ma XJ, Patel R, Wang X et al (2006) Molecular classification of human cancers using a 92-gene real-time quantitative polymerase chain reaction assay. Arch Pathol Lab Med 130:465–473
17. Erlander MG, Ma X-J, Kesty NC et al (2011) Performance and clinical evaluation of the 92-gene real-time PCR assay for tumor classification. J Molecular Diagn 13:493–503
18. Rosenwald S, Gilad S, Benjamin S et al (2010) Validation of a microRNA-based qRT-PCR test for accurate identification of tumor tissue origin. Mod Pathol 23:814–823
19. Meiri E, Mueller WC, Rosenwald S et al (2012) A second-generation microRNA-based assay for diagnosing tumor tissue origin. Oncologist 17:801–812
20. Kerr SE, Schnabel CA, Sullivan PS et al (2012) Multisite validation study to determine performance characteristics of a 92-gene molecular cancer classifier. Clin Cancer Res 18:3592–3960

21. Weiss LM, Chu PG, Schroeder BE et al (2013) Blinded comparator study of immunohistochemical analysis versus a 92-gene cancer classifier in the diagnosis of the primary site in metastatic tumors. J Molecular Diagn 15:263–269
22. Kerr SE, Schnabel CA, Sullivan PS et al (2012) Use of a 92-gene molecular classifier to predict the site of origin for primary and metastatic tumors with neuroendocrine differentiation. Lab Invest 92:147A
23. Greco FA, Spigel DR, Yardley DA et al (2010) Molecular profiling in unknown primary cancer: accuracy of tissue of origin prediction. Oncologist 15:500–506
24. Greco FA, Lennington WJ, Spigel DR, Hainsworth JD (2013) Molecular profiling diagnosis in unknown primary cancer: accuracy and ability to complement standard pathology. J Natl Cancer Inst. doi:10.1093/jnci/djt099
25. Hainsworth JD, Spigel DR, Greco FA: Renal cell carcinoma presenting as cancer of unknown primary: diagnosis by molecular tumor profiling. J Clin Oncol 2013; 31(Suppl): Abstract e15501
26. Greco FA, Spigel DR, Hainsworth JD: Molecular tumor profiling of poorly differentiated neoplasms of unknown primary site. J Clin Oncol 2013;31(Suppl):681 s (Abstract #11102)
27. Tothill RW, Kowalczyk A, Rischin D et al (2005) An expression-based site of origin diagnostic method designed for clinical application to cancer of unknown origin. Cancer Res 55:4031–4040
28. Horlings HM, Van Laar RK, Kerst J-M et al (2008) Gene expression profiling to identify histogenetic origin of metastatic adenocarcinoma of unknown primary. J Clin Oncol 26:4435–4441
29. Varadhachary GR, Talantov D, Raber MN et al (2008) Molecular profiling of carcinoma of unknown primary and correlation with clinical evaluation. J Clin Oncol 26:4442–4448
30. Bridgewater J, Van Laar R, Floore A, Van'tVeer L (2008) Gene expression profiling may improve diagnosis in patients with carcinoma of unknown primary. Br J Cancer 98:1425–1430
31. Monzon FA, Medeiras F, Lyons-Weiler M, Henner WD (2010) Identification of tissue of origin in carcinoma of unknown primary with a microarray-based gene expression test. Diagn Pathol 5:3
32. Pentheroudakis G, Pavlidis N, Fountzilas G et al (2013) Novel microRNA-based assay demonstrates 92 % agreement with diagnosis based on clinicopathologic and management data in a cohort of patients with carcinoma of unknown primary. Mol Cancer 12:57
33. Pentheroudakis G, Spector Y, Krikelis D et al (2013) Global microRNA profiling in favorable prognosis subgroups of cancer of unknown primary (CUP) demonstrates no significant expression differences with metastases of matched known primary tumors. Clin Exp Metastasis 30: 431–439
34. Gamble AR, Bell JA, Ronan JE et al (1993) Use of tumor marker immunoreactivity to identify primary site of metastatic cancer. BMJ 306:295–298
35. Brown RW, Campagna LB, Dunn JK et al (1997) Immunohistochemical identification of tumor markers in metastatic adenocarcinoma: a diagnostic adjunct in the determination of primary site. Am J Clin Pathol 107:12–19
36. Tot T (1999) Adenocarcinoma metastatic to the liver; the value of cytokeratins 20 and 7 in the search for unknown primary tumors. Cancer 85:171–177
37. DeYoung BR, Wick MR (2000) Immunohistologic evaluation of metastatic carcinomas of unknown origin: an algorithmic approach. Semin Diagn Pathol 17:184–193
38. Dennis JL, Hvidsten TR, Wit EC et al (2005) Markers of adenocarcinoma characteristic of the site of origin: development of a diagnostic algorithm. Clin Cancer Res 11:3766–3772
39. Lagendijk JH, Mullink H, VanDiest PJ et al (1999) Immunohistochemical differentiation between primary adenocarcinomas of the ovary and ovarian metastases of colonic and breast origin. Comparison between a statistical and an intuitive approach. J Clin Pathol 52: 283–290
40. Park SY, Kim BH, Kim JH et al (2007) Panels of immunohistochemical markers help determine primary sites of metastatic adenocarcinoma. Arch Pathol Lab Med 131:1561–1567
41. Oien K (2009) Pathologic evaluation of unknown primary cancer. Semin Oncol 36:8–37
42. Anderson GG, Weiss L (2010) Determining tissue of origin for metastatic cancers: meta-analysis and literature review of immunohistochemistry performance. Appl Immunohistochem Mol Morphol 18:3–8
43. Lester SC (2001) Special studies. In: Lester SC (ed) Manual of surgical pathology, 1st edn. Churchill Livingstone, Philadelphia, pp 73–74
44. Bugat R, Bataillard A, Lesimple T et al (2003) Summary of the standards, options and recommendations for the management of patients with carcinoma of unknown primary site (2002). Br J Cancer 89(suppl 1):559–566
45. Pavlidis N, Briasoulis E, Hainsworth J et al (2003) Diagnostic and therapeutic management of cancer of an unknown primary. Eur J Cancer 39:1990–2005
46. Varadhachary GR, Abbruzzese JL, Lenzi R (2004) Diagnostic strategies for unknown primary cancer. Cancer 100:1776–1785
47. Handorf CR, Kulkarni A, Grenert JD et al (2013) A multicenter study directly comparing the diagnostic accuracy of gene expression profiling and immunohistochemistry for primary site identification in metastatic tumors. Am J Surg Pathol 37:1067–1075
48. Varadhachary GR, Spector Y, Abbruzzese J et al (2011) Prospective gene signature study using microRNA to identify the tissue of origin in patients with carcinoma of unknown primary. Clin Cancer Res 17: 4063–4070
49. Morawietz L, Floure A, Stark-Sloots L et al (2010) Comparison of histopathological and gene expression-based typing of cancer of unknown primary. Virchows Arch 456:23–29
50. Varadhachary GR, Karanth GR, Qiao S et al (2013) Carcinoma of unknown primary with gastrointestinal profile, immunohistochemistry, and survival data for this favorable subset. Int J Clin Oncol 18: 226–231
51. Greco FA, Lennington WJ, Spigel DR et al (2012) Carcinoma of unknown primary site: outcomes in patients with a colorectal molecular profile treated with site-specific chemotherapy. J Cancer Ther 3:37–43
52. Hainsworth JD, Schnabel CA, Erlander MG et al (2012) A retrospective study of treatment outcomes in patients with carcinoma of unknown primary site and a colorectal cancer molecular profile. Clin Colorectal Cancer 11:112–118
53. Hainsworth JD, Rubin MS, Spigel DR et al (2012) Molecular gene expression profiling to predict the tissue of origin and direct site-specific therapy in patients with carcinoma of unknown primary site: a prospective trial of the Sarah Cannon Research Institute. J Clin Oncol 31:217–223
54. Hainsworth JD, Lennington WJ, Greco FA (2000) Overexpression of Her-2 in patients with poorly differentiated carcinoma or poorly differentiated adenocarcinoma of unknown primary site. J Clin Oncol 18:632–635
55. Tan DSW, Montoya J, Ng DS et al (2013) Molecular profiling for druggable genetic abnormalities in carcinoma of unknown primary. J Clin Oncol 31:e237–e239
56. Stella GM, Benvenuti S, Gramaglia D et al (2010) MET mutations in cancers of unknown primary origin (CUPs). Hum Mutat 32:44–50
57. Golfinopoulos V, Pentheroudakis G, Goussia A et al (2012) Intercellular signaling via the AKT axis and downstream effectors is active and prognostically significant in cancer of unknown primary (CUP): a study of 100 CUP cases. Ann Oncol 23(10):2725–2730
58. Gatalica Z, Millis S, Bender R, et al.: Molecular profiling of cancers of unknown primary site (CUP): paradigm shift in management of CUP. EJC 2011; 47(Suppl 2): Abstract LBA39
59. Penley WC, Spigel DR, Greco FA, Hainsworth JD: Confirmation of non-small cell lung cancer diagnosis using ALK testing and genetic profiling in patients presenting with carcinoma of unknown primary site. J Clin Oncol 2013; 31(suppl): Abstract e115004