

Precision Medicine Based in Epigenomics: The Paradigm of Carcinoma of Unknown Primary

Sebastián Moran^{1†}, Anna Martinez-Cardús^{1†}, Stergios Boussios², and

Manel Esteller^{1,3,4,*}

¹Cancer Epigenetics and Biology Program (PEBC), Bellvitge Biomedical Research Institute (IDIBELL), 08908 L'Hospitalet del Llobregat, Barcelona, Catalonia, Spain;

²Department of Medical Oncology, Ioannina University Hospital, 45110 Ioannina, Greece;

³Physiological Sciences Department, School of Medicine and Health Sciences, University of Barcelona (UB), Catalonia, Spain. ⁴Institució Catalana de Recerca i Estudis Avançats (ICREA) Barcelona, Catalonia, Spain.

† Equal contribution. *e-mail: mesteller@idibell.cat

ABSTRACT

Epigenomic disruption has emerged as a common hallmark of human cancer. Single epigenetic markers, such as the use of the promoter hypermethylation-associated inactivation of the DNA repair enzyme MGMT to predict response to temozolomide in gliomas, are starting to be incorporated in the clinical setting, however, at the “omics” level there are not validated examples of its translational use. We have recently provided one example of how epigenomics can enter the clinical arena by demonstrating its value in the identification of the site of origin in the cancer of unknown primary (CUP). Identifying the origin of metastatic tumors remains a challenge in the diagnostic and clinical management of CUP. In spite of the great diagnostic advances made in the last decade, traditional diagnostic procedures can only identify the origin of about 30% of CUP cases. Thus, development of diagnostic strategies in the field of molecular biology has emerged as a way of complementing traditional procedures, and thereby improving CUP patient management. DNA methylation signatures are suitable for measurement in easily obtained samples, such as paraffin-embedded tissue or liquid biopsies, which is a competitive advantage when trying to implement a molecular diagnostic tool. The aim of

the review is to systematically organize the most recent data about epigenetics, CUPs and the role of other molecular markers, and to highlight areas for further research to engage the medical community in this exciting subject.

Personalized oncology has been possible partly thanks to our improved understanding of cancer molecular biology and to the huge technical advances in methodology and analytical tools. These advances have allowed us to identify molecular targets that cause diseases and understand how to counteract them. Since the first human genome was determined in 1991, the amount of data generated using next-generation technologies has grown enormously.¹ Due to the need to manage all this valuable information, several consortia have been created, such as the Cancer Genome Atlas (TCGA), whose aim is to elucidate the most frequent molecular alterations in different kinds of cancer.^{2,3} The central hallmarks of cancer cells were defined on the basis of the data generated,⁴ providing a more comprehensive knowledge of cancer cell behavior and boosting the numbers of potential targets for therapy and of their associated biomarkers that may be of use in predicting treatment effectiveness. In the last decade, translational research in cancer has focused on finding molecular markers that may improve patient outcome and quality of life. However, besides identifying prognostic factors to characterize those tumors that more likely to relapse, it is crucial to identify biomarkers that allow the tumors to be correctly classified and their diagnosis and outcome thereby to be optimized. This will avoid the use of inefficient therapies and will reduce the health costs arising. These two aspects are of particular relevance to metastatic lesions of unknown origin, since these cannot be classified as a particular neoplasia and so have to be treated empirically, resulting in the notably poor survival rates in cancer of unknown primary (CUP).

In the era of targeted therapy, it is essential to obtain an accurate histopathological and molecular classification of tumors in order to administer the best tailored therapeutic strategy. Epigenetic alterations have served this purpose, since cancer cells are characterized by a massive overall loss of DNA methylation (20-60% decrease in global 5-methylcytosine),^{5,6} and simultaneously by the acquisition of specific patterns of hypermethylation at CpG islands of certain promoters, which can reversibly or irreversibly alter gene function and contribute to cancer progression.^{7,8} Given the importance of epigenetics in neoplasia development, it is worth highlighting its attractiveness in

personalized oncology, as well as its already demonstrated usefulness in the diagnosis of cancers. An association observed between epigenetic signatures and different tissues has enabled the discrimination among tumor types, not only according to their embryonic origin, but also on the basis of their tissue type.⁹ Recently, we have applied this approach to the generation of an epigenetic-based bioinformatic tool that allows the prediction of the primary tumor type from which a CUP metastatic lesion originates.¹⁰ An additional advantage of epigenetic translation to clinical practice comes from the nature of the material interrogated, DNA, a molecule that is stable over time, irrespective of the method of tissue fixation, and that it is not very reactive to change due to minimal external factors, unlike RNA expression levels. Moreover, interesting progress is being made in the applying epigenetic tests in newer and less intrusive diagnostic materials, such as liquid biopsies.

Taking into account the recent advances in the field of epigenomics, the purpose of this review is to put into context the state of our knowledge of CUP and the current diagnostic strategies, emphasizing the role of epigenetic signatures in clinical management, with the point of view focused on the present and the future. It is also important to inspire the interest of the health community in epigenetics, highlighting the great benefits it might provide to clinics.

Introduction to cancer of unknown primary

The primary tissue of origin of 3-10% of all cancers diagnosed worldwide cannot be determined.^{11–13} This uncertainty has a direct impact on the outcome of patients. Under such circumstances, in which the metastases are clinically manifested, but the primary tumor originating them is unknown, the physician diagnoses patients as having cancer of unknown primary (CUP). Clinically, CUP is defined as a histologically confirmed malignant tumor, incompatible with a primary neoplasm in the biopsy area, whose origin is not clear, despite performing a thorough clinical history, a complete physical examination, and basic complementary studies.^{14–16} However, this definition is not an international standard as reflected in diagnostic codes, and varies between institutions and countries.^{17–19} As a result of this challenging situation, complete immunohistochemistry tests or post-mortem examination reveals the primary tumor in

only one in four (25%) CUP cases. The primary tumor of the majority of CUPs (75%) cannot be identified.²⁰

Epidemiological studies involving CUP cases are scarce and difficult to compare with each other, mainly due to the aforementioned differences in inclusion criteria. The available studies, mainly from northern European countries (Netherlands,¹³ Sweden^{21–23}, Scotland¹⁷ and Norway²⁴), Australia²⁵ and USA,²⁶ cover half a century of historical data (1960–2010), and highlight the distinctive trends of this disease. They reveal an increasing incidence of CUP over the period from 1960 to 1980, to a level of 16 per 100,000 inhabitants. Interestingly, although there are differences depending on the start year for each of the populations analyzed, there has nevertheless been a clear reduction in the incidence, whereby the figure now stands at 8 per 100,000 inhabitants. The reasons for such a reduction are thought to include advances in diagnostic methods at the biological, radiological and pathological levels, although it might also be a consequence of insufficient diagnostic inquiry, a phenomenon documented in the US population.

Histologically, well-differentiated or moderately differentiated adenocarcinomas account for 50-70% of CUP cases, while poorly differentiated carcinomas/adenocarcinomas make up 20-30% of them. The remaining cases are divided into squamous cell carcinomas (5-8%) and poorly differentiated or undifferentiated malignant neoplasms (2-3%).²⁷ CUP appears from the age 30 years, with a peak at around 70 years. In terms of survival, although current data are very discouraging, attaining overall survival (OS) of 2.75-11 months after diagnosis, with a 1-year survival rate of 15-20%, represents an improvement since the early 2000s where in a population based study the estimated median survival was just 11 weeks.¹³ The current improvement is thought to be a consequence of earlier detection and advances in the management of advanced-stage neoplasias.

Clinical management of patients with CUP entails taking different steps towards deciding the appropriate treatment. First, non-carcinoma neoplasms (sarcomas, melanomas and lymphomas) should be ruled out if possible. In 15-20% of cases there is a strong suspicion, either from immunohistochemical or molecular tests, of the presence of a primary cancer for which a specific and efficient treatment exists. In such circumstances,

patients should receive the corresponding site-specific treatment, since, with appropriate management, long-term disease control can be achieved in 30-60% of cases.¹⁶ Preliminary retrospective studies have shown an improvement in survival in patients treated with site-specific therapies for the equivalent primary tumor originating the metastasis.²⁸ The other 80-85% of CUP cases have a very poor response to treatment, and without any suspicion of its primary tumor, its treatment involves an empirical schedule consisting of either taxanes or platins or a combination of the two, in conjunction with gemcitabine.²⁹ The median survival of this group is a disheartening 6-9 months, response rates are around 20%, and the 1-year survival rate barely reaches 25%.³⁰ Nevertheless, within this non-specific subset of patients, two prognostic groups can be identified on the basis of their performance status and lactate dehydrogenase (LDH) serum levels. Patients in the first group, characterized by a good performance status and normal levels of LDH, are treated with empirical two-drug chemotherapy, resulting in a median life expectancy of 1 year. The favourable subset include women with adenocarcinoma involving axillary nodes, and papillary adenocarcinoma of peritoneal cavity, patients diagnosed with squamous cell carcinoma involving cervical nodes, poorly differentiated neuroendocrine carcinomas, and adenocarcinomas with a colon-profile (CK20+, CK7-, CDX2+). However, patients in the second group, who have elevated serum LDH levels and/or poor performance status, have a median OS of only 4 months, and depending on their performance status, best supportive care or chemotherapy is applied.³¹ The most common histological types are adenocarcinomas of moderate to poorly differentiated (64%), the rest been undifferentiated tumors. It involves mainly the liver in 40–50% of the cases, followed by lymph nodes (35%), lungs (31%), bones (28%) and the brain (15%).³² Additional independent prognostic factors include the age, number of metastatic sites, and neuroendocrine differentiation.

Diagnostic tools in CUPs

Until now, immunohistochemistry (IHC) based on the detection of particular antigens in tumor cells of a tissue section using specific antibodies, has been the main strategy for identifying the primary tumor of a given metastasis in routine clinical practice.³³ Of the most commonly used markers, cytokeratins 7 and 20, high- and low-molecular weight keratins (34βE12 and CAM.52, respectively) and tissue-specific markers, such as thyroid

transcription factor-1 (TTF1), mammoglobin, intestine-specific transcription factor (CDX2) and hepatocyte specific antigen 1 (HepPar-1), are of particular note.^{34,35} However, no hegemonic panel of IHC markers has been established, and instead multiple panels are used to address this challenge.³⁵⁻⁴⁰ Unfortunately, there is not a single specific marker which by itself can provide a conclusive diagnosis. Low reproducibility is another intrinsic limitation of this technique, in part due to the need to interpret IHC preparations, which, by their nature, involve subjective observation. The aforementioned reasons are those that lead to the poor 50-65% success rates of conventional histopathological methods in the diagnosis of metastases in known primary tumors, the rate dropping to a mere 25% of cases when considering only CUP cases.^{34,41}

Besides immunohistochemistry, image diagnosis is used to localize the origin and to reach a diagnosis of CUPs. Computer tomography (CT) and conventional magnetic resonance imaging have been extensively used to locate alterations compatible with the clinical manifestation of CUP. However, the small size of the primary tumor or the regression of the primary tumor could occur in CUP cases, hindering their diagnosis.³⁴ The use of 18F-fluoro-2-deoxyglucose positron emission tomography/computed tomography (FDG-PET/CT) successfully facilitates this task.⁴²⁻⁴⁸ However, although this technique is very specific (71%) and sensitive (87%) at detecting CUP primary lesions,⁴⁹ it has a poor detection rate that rarely exceeds 40% in identifying the primary tumors of patients with CUP.⁵⁰

Complementary to histological and imaging analysis, the dissemination patterns of tumors are taken into account in clinical protocols to identify the primary origin of CUP. For instance, metastases located in the liver most frequently originate in the gastrointestinal tract, lung, breast, genitourinary tract or uvea; the primary tumors of lung metastases are usually located in the breast, gastrointestinal tract, kidney or prostate.

Overall, it is not surprising that daily clinical practice for the treatment of CUP suffers from a high degree of diagnostic uncertainty, mainly due to the strong discrepancies between the cytological, histological and immunohistochemical diagnoses performed in a given CUP case.⁵¹ Moreover, the idiosyncrasy of the disease itself predisposes towards diagnostic confusion, since the tumors are very often poorly differentiated, or completely

undifferentiated,⁵² hindering their correct diagnosis. This challenging situation translates directly into a low success rate of diagnostic work-up in CUP, with only 20-30% of the cases being successfully diagnosed.

Unlike immunohistochemical panels, in which the expression of only a small number of genes is tested, and whose interpretation is subjective more than quantitative, gene expression profiling was introduced for cancers of unknown diagnosis in 2005 as a new approach to study in a quantitative manner the expression profiles of larger panels of genes characteristically and differentially expressed in different tumor types.⁵³⁻⁵⁹ This approach enabled the prediction of the tissue of origin in a CUP by comparing its expression profile with that of a set of primary/metastatic tumors of known origin. Applying this method to primary and metastatic tumors of known origin enabled the correct identification of 83-89% of cases. Moreover, 91% accuracy in the prediction of tissue of origin in poorly differentiated and undifferentiated carcinoma metastases is achieved, although the figure drops to 71% of cases when IHC is the chosen diagnostic technique.⁵⁷ These results show what could happen in the case of applying the prediction of tissue of origin in CUPs, in which results compatible with clinicopathological features and response to treatment are obtained in 94% of cases.^{54,57} Besides gene expression prediction, another interesting approach is the use of microRNA assays to identify the tissue of origin of CUPs; this gives similar results to those based on gene expression.^{60,61}

Table 1 summarizes the features of the most relevant studies that use expression profiles in CUP diagnosis. It is also important to note that all molecular assays assigning a primary tissue of origin to CUP are constructed in order to look for similarities (but not differences) between CUP and metastatic solid tumours. However, the peculiar behavior of CUP hints for the existence of molecular differences between CUP and Known Primary metastases that have been overlooked so far in the developed diagnostic pipelines. In addition, for the described assays, there are limitations in the cohort patient size and the tumor types, the requested amount and state of preservation of the analyzed clinical sample, and the overall cost of the assessment, that warrant the development of additional molecular profiling tools for cancer of unknown primary. To address the unmet medical need in this area with the dismal survival of the CUP patients, the use of epigenetic biomarkers^{62,63} is worth to be explored.

Epigenetics in cancer diagnosis

Epigenetic mechanisms are generally understood to be those that regulate gene expression. The epigenetic code is characterized as being inherited throughout cell division and even transgenerationally.⁶⁴ Covalent addition of methyl groups to the carbon-5 of cytosines in DNA is the predominant epigenetic mechanism in mammal cells, although chromatin remodeling also has an important role in controlling genes and ncRNAs expression, DNA-protein interactions, mobility of transposable element suppression, cellular differentiation, embryogenesis, X-chromosome inactivation and genomic imprinting.⁶⁵ The ability of epigenetic marks to persist during development and potentially be transmitted to offspring may be necessary to generate the wide range of phenotypes that arise from the same genotype.⁶⁴

The importance of epigenetics in normal cell maintenance is reflected by the observation that the introduction of the wrong type of epigenetic marks at an inappropriate time or place is associated with development of many diseases, such as cancer.⁶⁶ In this context, the global genomic loss of DNA methylation leads to alterations in chromosomal integrity. In addition, hypomethylation at a specific promoter activates oncogenes and induces loss of imprinting (LOI) at some loci. DNA methylation is also important in that it causes a loss of expression of the tumor suppressor genes involved in cell cycle, apoptosis, and proliferation, which may in turn contribute to cancer progression.⁷

In spite of huge advances in cancer diagnosis, there are still many cancer cases that have unexpected progress or that are unclassified as known tumor subtypes. Indeed, tumors with the same histopathological and genetic characteristics show completely dissimilar behaviors. This phenomenon could be the consequence of different epigenetic backgrounds. A recent example is the effect of the methylation status of TBC1D16 gene, which leads to differential prognosis in cohorts of melanoma BRAF V600E-mutated tumors.⁶⁷ DNA methylation profiling has revealed cancer-specific signatures of hypermethylated CpG islands, by which we can distinguish tumor types and predict antineoplastic treatment responsiveness and patient outcome.^{9,68} Related to therapies, Pharmacoeugenetics started with the original observation of the association between hypermethylation of MGMT and response to alkylating agents in gliomas^{69,70}, a test that it

is used nowadays in clinical care. Other epigenetic markers have emerged in a similar manner such as BRCA1 hypermethylation to predict response to PARP inhibitors,^{71,72} DERL3 hypermethylation as predictor of sensitiveness to glycolysis inhibitors,⁷³ as well TBC1D16 hypomethylation as biomarker of higher sensitivity to BRAF and MEK inhibitors.⁶⁷ Furthermore, epigenetic markers are also proficient to foresee response to treatment, such as SLFN11 or SRBC hypermethylation that is associated with resistance to platinum drugs.^{74,75}

Regarding this utility for more accurately classifying the tumors into already existing categories, it has been reported that aberrant DNA hypermethylation events in cancer-related genes, such as BRCA1, are common features in the breast cancer triple-negative subtype.^{76,77} Key epigenetic alterations involved in the initiation and progression of cancer have also been described in other neoplasias, such as acute myeloid leukemia and renal cell carcinoma, highlighting their utility in early detection and differential diagnosis.^{78–80} The detection of hypermethylated genes in stool and blood samples is a sensitive and specific way of assessing the risk of colon cancer, and for its early detection.^{81–83} Moreover, hypermethylation events at APC, MGMT, RASSF2 and WIF1 genes can be detected in the plasma of patients with colon cancer with high sensitivity and specificity (87%).⁸⁴ Another example is that prostate cancer can be diagnosed through the detection of hypermethylation at the GSTP1 gene in tissue biopsies, urine, plasma or ejaculate samples,^{85,86} with a high sensitivity and specificity (82% and 95%, respectively), and also in precursor lesions for the development of this tumor. The use of DNA methylation as a potential biomarker for diagnosing other neoplasias, such as esophageal squamous cell carcinoma,⁸⁷ non-small cell lung cancer⁸⁸ and rectal cancer has also been described,⁸⁹ highlighting the value of studying methylation as a diagnostic tool in cancer. For colorectal cancer^{90–92} and neuroblastoma,^{93,94} the described existence of several CpG islands that undergo a shared DNA methylation gain (“CpG island methylator phenotype”) also represents a useful biomarker tool for these malignancies.

Epigenetic approach to CUP diagnosis

The huge advances in understanding how the epigenetic machinery works, acts and interacts, and the implications of these for the proper functioning of cells, and how

epigenetic deregulation may lead to the development of cancer, has encouraged the search for DNA methylation features that, being present in CUP, might maintain the signature of the presumed primary origin.

The observation that DNA methylation fingerprints provide important clues concerning the tissue-type identity was first made in a study of more than 1,000 human samples using a high-throughput approach that interrogated 1,505 CpG sites and was able to discriminate 24 different tumor types.⁹ Until then, most of the studies characterizing DNA methylation patterns had been restricted to examining particular genomic loci in a limited number of human samples. The study analyzed 42 CUPs and it was able to assign a given tumor type for these CUPs in 69% of cases using L1-regularized logistic regression with misclassification to create a prediction heatmap.⁹ A proposed foster primary in these 29 cases was also achieved by conventional clustering analysis. The tumor type prediction of the CUPs based on the DNA methylation analyses was fully confirmed in 78% of cases for which detailed pathological analysis developed at a later stage in a blind fashion was able to provide a diagnosis (**Table 1**).⁹ Certainly, the DNA methylation patterns identified in that study across the largest spectrum of samples constituted a baseline for developing higher-resolution DNA methylation maps and providing important clues about the contribution of DNA methylation to tissue uniqueness. Based on these results and with the aim of developing a sufficiently robust diagnostic tool for use in clinical practice, a CUP classifier based on DNA methylation profiling (EPICUP) was generated that enabled the tissue of origin to be predicted in 87% of cases,¹⁰ (**Table 1**) representing a huge improvement on the 25% of cases identified using light microscopy and IHC testing.²⁰ This could not have been achieved without the dramatic improvements made in the field of new-generation technologies in DNA methylation analysis,^{95,96} which have generated large amounts of information highlighting the role of epigenetics in personalized oncology.⁹⁷ In an initial discovery phase, 2,790 primary tumor and metastasis samples from 38 tumor types of known origin, including the most common human cancers, were analyzed using Infinium HumanMethylation450K BeadChip (Illumina), a high-resolution method that interrogates more than 485,577 CpG sites and covers 99% of RefSeq genes.⁹⁵ A classifier algorithm was generated and evaluated with a secondary validation cohort that included 7,691 samples, resulting in a highly accurate predictive tool (99.6% specificity and 97.7% sensitivity), with a positive predictive value of 88.6%. These are

higher values than those obtained with other molecular platforms previously generated for the diagnosis of CUP.^{55,59} For clinical application, primary and metastatic tissues were classified with equivalent precision, as were both frozen and formalin-fixed paraffin-embedded (FFPE) tissue samples. Analysis of these did not reveal any different results, as had been previously demonstrated.⁹⁸ Further, the effectiveness of EPICUP as predictor of tumor origin was demonstrated using a cohort of 216 CUP cases collected from various countries.¹⁰ The median age of the cohort was 63 years, with most of its members presenting multiple metastasis sites at diagnosis (60%); adenocarcinomas/carcinomas accounted for 66% of cases and the median survival was 9 months, these values being similar to those for a representative CUP population.²⁷ Lymph nodes were the main site of biopsy at the moment of diagnosis (29% of cases), and the tumor type most frequently predicted by EPICUP was non-small cell lung cancer (20.6% of cases), in accordance with data previously reported about the origin of CUP in necropsies (**Figure 1**).²⁷

EPICUP has proved to be reliable according to the best available knowledge of the clinical and pathological description of each CUP case. In one case, the tissue of origin was determined at necropsy and was corroborated by immunohistochemistry. Interestingly, the tissue of origin was found to be the same as that previously predicted by the DNA methylation-based classifier (sarcoma) (**Figure 2A**). In other cases, EPICUP guided IHC analysis with proper specific markers that confirmed the prediction. An unusual case was that of a male with lesions in the axillary lymph nodes, who was diagnosed as CUP after clinical and histopathological evaluation, and was then treated with empirical chemotherapy. At progression, the CUP was subjected to the classifier, which yielded a prediction of breast cancer that was subsequently corroborated by mammoglobin-positive staining. The condition was then treated appropriately (**Figure 2B**). Similar observations were made of other tumor types, such as the determination of CDX2-positive staining in a colon cancer case after being correctly predicted by EPICUP (**Figure 2C**). Another clear advantage of this epigenetic approach is that is based on DNA, a molecule that is stable over time, regardless of the method of tissue fixation, and so not one predisposed to change due to minimal external factors, unlike RNA expression levels. Moreover, tumor type predictions were reproducible using DNA methylation profiling with two methylation microarrays, 450K⁹⁵ and 850K.⁹⁶ Both of these are useful

for analytical purposes. In summary, the use of methylation to predict the tissue of origin is a potent strategy that may be applied in CUP diagnosis in conjunction with clinical and histopathological tests.²⁰

Present and future challenges

The use of molecular platforms in the clinical management of human cancer is now a reality. Our better knowledge of cancer biology and behavior has allowed a more accurate classification of neoplasias, resulting in optimized cancer diagnosis, prognosis and tailored therapy. A good example is the case of patients with CUP, for whom this DNA methylation-based classifier, as well other types of molecular platforms, is and will be a key component of diagnostic workflow and clinical management. It is expected to have a significant effect on the quality of life and longevity of these patients. Nevertheless, essential and crucial steps are still necessary in order to facilitate its use in daily clinical practice and to guarantee its inclusion in CUP management guidelines.

Impact on survival

A retrospective study performed in a cohort of CUP patients with suspected tissue of origin in the colon provided support to the hypothesis that the administration of site-specific chemotherapy consistent with the molecularly predicted tissue of origin produces a benefit in terms of survival rates in patients with CUP.⁹⁹ Such an observation has also been supported by our retrospective work,¹⁰ in which we observed that those patients whose treatment was specific to the methylation-based tissue of origin prediction showed longer OS than those CUP patients who did not receive the best treatment available (empirical treatment) for their epigenetically predicted origin (13.6 versus 6 months, respectively). Nevertheless, a prospective non-randomized study has been reported,¹⁰⁰ in which RT-PCR molecular tumor profiling was performed on biopsy specimens from patients diagnosed with CUP. Site-specific treatment was administered on the basis of the tissue of origin predicted by the assay, while those patients for whom the assay was unable to predict a tissue of origin were treated with standard empirical chemotherapy. The results indicated that CUPs treated with site-specific therapeutic schedules improved their median survival to 12.5 months, which a 38% improvement is compared with the

value of 9 months noted in the empirically treated CUP patients. Moreover, patients whose predicted origin was a responsive tumor type and whose molecular profile was confidently predicted (probability of tissue of origin $\geq 80\%$) achieved longer survival (median survival of 15.4 months; 49% 2-year survival rate) when treated with a site-specific therapy compared with patients with less responsive tumor types and less certain predictions (median survival of 7.0 months). Indeed, with an increasing number of effective molecularly targeted therapies available, an assay prediction may lead to the identification of additional therapeutic options, such as targeted therapies. Although this prospective clinical trial has not fully answered the question of whether site-specific treatment in CUP patients is unequivocally favorable in terms of survival, this could be resolved by a fully randomized clinical trial in which patients are randomly divided into two arms, one being treated with site-specific schedules, the other arm receiving empirical CUP chemotherapy. However, although it might be technically feasible, this clinical trial approach has some difficulties. For example, a percentage of CUP patients have cancer types that could be not significantly affected by either standard therapy or empirical chemotherapy. Moreover, even among tumor types such as those of ovarian, breast, non-small cell lung and bladder cancers, which are considered to be more responsive to chemotherapy, their first-line empirical CUP therapy is similar to standard site-specific therapy. Nonetheless, and recently summarized in an editorial in the journal,¹⁰¹ there are enough data to support the concept that molecular profiling in CUP diagnosis leading to the administration of a specific treatment according to its predicted origin improves survival rates of these patients.^{58,60,100,102} Future perspectives include the creation of a clinically useful diagnostic algorithm, which would incorporate pathological findings and molecular profiling tests along with crucial clinical judgment to maximise clinical benefit and limit costs. However, it is a must the development of high quality, prospective and randomized clinical trials to confirm that primary-tissue tailored therapy does indeed result in improvement of patient survival in CUP.

Screening of drug-actionable alterations

Only 25% of cases of CUP receive a single putative primary tumour diagnosis using light microscopy and immunohistochemical testing. An algorithm that integrates immunohistochemistry, tissue-of-origin profiling, and comprehensive genomic profiling in

some combination may maximize clinically meaningful benefit and minimize costs. A sizable percentage of patients with CUP will not currently benefit from assay-directed therapy, because effective therapy for these tumor types does not currently exist. The correct determination of the primary site of origin by EPICUP could guide the screening of drug-actionable mutations. The observation that patients with CUP who received a tumour type-oriented treatment did better than those receiving non-specific therapy might also be associated with an inherently different prognosis, regardless of the received treatment.

In fact, molecular-based prediction of tissue of origin in orphan metastases could have significant implications for the management of CUPs, given that revealing the tissue of origin might be useful for determining whether the sample carries any drug-actionable alteration that would make patients candidates for achieving better survival rates.¹⁰³ In this regard, next generation sequencing of CUP cases have identified that many of these lesions harbor targetable genomic alterations,^{104,105} such as those shown in **Figure 3**. Strikingly, a higher incidence of mutations in the C-MET gene (around 30%) has been found in the CUP population compared with the frequency in solid tumors (around 4%).¹⁰⁶ Other clinically relevant genetic alterations that have been detected in metastases of unknown primary that are targetable, include EGFR mutations, HER2 amplification and ALK translocation, among others,^{10,107} Importantly, the correct determination of the primary site of origin by EPICUP and other methods could guide targeted therapies, because if we search for actionable mutations without knowledge of the precise cellular context some surprises may arise.¹⁰⁸ For example, a K-RAS mutation in a patient that EPICUP predicts to be a non-small cell lung cancer with hilar nodes plus brain metastasis may have different clinical implications compared with an EPICUP-diagnosed colorectal cancer patient sharing that K-Ras mutation. Another illustrative case would be the discovery of a BRAFV600 mutation: if the DNA methylation profile predicts melanoma or thyroid carcinoma, the targeted therapy would be more adequate than if the EPICUP system indicated that the primary site was a colorectal tumor. The same scenario maybe true for epigenetic markers of drug sensitivity: if MGMT promoter hypermethylation has proved to be the best predictor of good clinical response to alkylating agents in gliomas,^{69,70} it only shows a more modest capacity to predict sensitivity to the same

agents in colorectal tumors.^{109,110} **Figure 3** illustrate how the determination of the primary site can be helpful for the right selection of the tailored-treatment.

However, the field is young and the available quality data establishing that therapy targeting molecular aberrations of a CUP does indeed result in improvement of patient survival are very scarce. For example, in an EPICUP studied sample,¹⁰ the epigenetically predicted origin was a non-small cell lung cancer and it carried an EGFR-mutation (G719X). The patient had received a specific treatment with erlotinib and achieved long survival (**Figure 4**).¹⁰ In another study, the treatment of a CUP case harboring C-MET gene amplification led to a complete clinical response greater than 3 years,¹¹¹ whereas a CUP patient with an ALK fusion responded dramatically to crizotinib therapy.¹¹² Furthermore, CUP cases can also be great candidate to participate in aberration-specific clinical studies, the so called “basket trials” to improve the survival of these patients. Importantly, it is now possible to apply technologies based on next-generation sequencing that allow the interrogation at the same time of different phenomena at the DNA level, such as methylation fingerprinting and drug-actionable alterations that will provide a more complete picture of the molecular profile of any given CUP case.

The establishment of an international CUP consortium with a virtual tissue bank storing clinically annotated CUP biological material in various centres would enhance our ability to perform correlative translational research projects within or outwith trials. There is retrospective evidence that the biologically genuine CUP is the visceral CUP subgroup. An inherent problem is that these patients harbor tumours not markedly impacted by either standard therapy or empiric chemotherapy. Ideally, a prospective trial would classify visceral CUP by EPICUP to a tissue of origin, followed by randomization to tissue-independent empiric chemotherapy versus tissue-guided optimal chemotherapy combined with targeted therapy. Intention-to-treat analysis of overall survival, quite dismal in this subgroup of visceral CUP patients, would be an endpoint not confounded by salvage therapies and mature enough in relatively short follow-up times. Alternatively, EPICUP could be used in retrospective or prospective cohort studies looking at outcome differences between typical metastatic solid tumours of known primary versus CUPs with EPICUP-identified matched primaries. Such a cohort study could provide hints at biological peculiarities of CUP.

Optimization of DNA methylation in single-cell and liquid biopsies

We are currently in an era of technological transition leading to a reduction in the invasiveness and a greater personalization of diagnostic tests. The times of using a sledgehammer to crack a nut are over, and more refined approaches are now gaining momentum. As an example, obtaining the DNA methylome of the whole tumor, although providing very useful information, might not be the optimal approach since each tumor might have a different cell composition.¹¹³ Obtaining single cell profiles, which requires a minimal sample, is now feasible,^{114,115} and with sequencing costs falling, its use in diagnostic testing is becoming more attractive. Single-cell approaches have not been applied in CUP, although they might lead us to a better understanding and means of diagnosing CUP. An intermediate approach would be to analyze circulating tumor cells, since this might enable us to understand the absence of primary tumors originating CUP. The value of this approach in providing predictive biomarkers of patient outcome has been reported,¹¹⁶ although we currently lack an approach with which to characterize their genomic material with respect to distinctive alterations. In this regard, other types of epigenetic patterns, such as histone modifications, can be studied in this context. It is known that histone modification patterns undergo significant aberrations in human tumors.^{117,118} Importantly, we have now powerful techniques to study not only particular histone modifications starting with few cells,¹¹⁹⁻¹²¹ but also to measure tissue-type specific patterns of chromatin accessibility.^{122,123} Another interesting advance is the incipient use of liquid biopsies, in which circulating cell-free DNA is interrogated for biomarker alterations. This approach will obviate or minimize the need for intrusive diagnostic biopsies, a less-intrusive blood extraction method being available instead to provide enough material to perform the diagnostic procedures. We have been able to detect the presence of aberrant hypermethylation of tumor suppressor genes in serum DNA since the late 90s.^{124,125} However, it has been the introduction of more sensitive and quantified techniques to analyze DNA methylation, such as it has been as shown for diffuse large B cell lymphoma,¹²⁶ and hepatocellular carcinoma,¹²⁷ that can facilitate its translation to the clinical setting. In this regard, although this approach has not been directly reported in relation to defined CUP samples, very promising results have been recently obtained to determine the tissue-specific origin of circulating DNAs.^{128,129}

Conclusions

Cancer of unknown primary is a heterogeneous disease defined by the presence of metastatic lesions with no identified primary tumor at presentation, resulting in an extremely poor outcome for patients, in part due to the lack of information with which to choose the most suitable treatment. The huge advances in molecular platforms have improved our ability to recognize the tumor of origin in CUP, which is correctly identified by conventional diagnostic tools in only 30% of cases. This progress has had a demonstrable influence on extending survival. In particular, the DNA methylation-based classifier EPICUP has correctly identified the tumor of origin in a higher proportion of patients than have older established approaches, and with more confidence than with other gene expression-based platforms. Prediction of tissue of origin might serve as a clue to interrogate particular genetic alterations for which specific treatments exist (drug-actionable targets). Prospective clinical trials are underway to assess the real impact of these platforms on survival.

EPICUP has begun to be implemented, in conjunction with current diagnostic strategies. Their application in clinical care supposes an advance in classification of these tumors, homogenizing their management in the different health centers and increasing the current poor survival rates associated with this disease. The suitability of the analysis of methylation as diagnostic tool in the clinics had already been demonstrated, in part because the stability of DNA means that it does not easily react to changes caused by external factors, and because it is relatively straightforward to detect in samples that are easily obtained in the course of daily clinical practice, such as FFPE or liquid biopsies. Likewise, the great capacity for expansion of this epigenetic-based tool could give rise to a wide range of additional applications in the future in the arena of cancer diagnosis.

KEY POINTS

- Knowing the tissue of origin leads to improved management and prognosis of CUP patients.
- Advances in molecular platforms have enabled CUP diagnosis to be optimized.
- DNA methylation is a valid biomarker for clinical diagnosis and has the capacity to discriminate different tissues.
- Epigenetic prediction of tissue of origin improves CUP diagnosis, being more robust than other molecular-based platforms.
- Prediction of tissue of origin might serve as a first clue to determine particular genetic alterations for which specific treatments exist (drug-actionable targets).

REVIEW CRITERIA

The bibliographic search was performed using the PubMed search engine, using “carcinoma unknown primary” or “cancer unknown primary” as the query terms. Results were limited to those written in English and published between December 1st 2001 and December 1st 2016. The search produced 8,260 hits, which, after filtering out those that were not related to CUP, or were duplicated entries, or for which no abstract was available, the list was reduced to 837 scientific articles.

COMPETING INTERESTS

SM has no competing interests to declare. AM-C reports personal fees from Ferrer outside of this work. ME reports grants and a patent from Ferrer International SA related to the herein summarized work.

REFERENCES

1. Forbes, S. A. *et al.* COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* **39**, D945-950 (2011).
2. Bell, D. *et al.* Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609–615 (2011).
3. Boehm, J. S. & Hahn, W. C. Towards systematic functional characterization of cancer genomes. *Nature Rev.Genet.* **12**, 487–498 (2011).
4. Hanahan, D. & Weinberg, R. A. A. Hallmarks of cancer: The next generation. *Cell* **144**, 646–674 (2011).
5. Carmona, F. J. & Esteller, M. Epigenomics of human colon cancer. *Mutation Res.* **693**, 53–60 (2010).
6. Ushijima, T. Detection and interpretation of altered methylation patterns in cancer cells. *Nat. Rev.Cancer* **5**, 223–231 (2005).
7. Esteller, M. Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat. Rev.Genetics* **8**, 286–298 (2007).
8. Esteller, M., Corn, P. G., Baylin, S. B. & Herman, J. G. A gene hypermethylation profile of human cancer. *Cancer Res.* **61**, 3225–3229 (2001).
9. Fernandez, A. F. *et al.* A DNA methylation fingerprint of 1628 human samples. *Genome Res.* **22**, 407–419 (2012).
10. Moran, S. *et al.* Epigenetic profiling to classify cancer of unknown primary: a multicentre, retrospective analysis. *Lancet Oncol.* **17**, 1386-1395 (2016).
11. Löffler, H., Neben, K. & Krämer, A. CUP-Syndrom. *Der Radiologe* **54**, 107–111 (2014).
12. Stewart, J., Tattersall, M., Woods, R. & Fox, R. Unknown primary adenocarcinoma: incidence of over investigation and natural history. *Br.. Med. J.* **1**, 1530–1533 (1979).
13. van de Wouw, A. J., Janssen-Heijnen, M. L. G., Coebergh, J. W. W. & Hillen, H. F. P.

Epidemiology of unknown primary tumours; incidence and population-based survival of 1285 patients in Southeast Netherlands, 1984-1992. *Eur. J. Cancer* **38**, 409–413 (2002).

14. Fizazi, K. *et al.* Cancers of unknown primary site: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **26**, 3479–3482 (2015).
15. Fizazi, K., Greco, F. A., Pavlidis, N. & Pentheroudakis, G. Cancers of unknown primary site: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **22**, 64–68 (2011).
16. Collado Martin, R., Garcia Palomo, A., de la Cruz Merino, L., Borrega Garcia, P. & Baron Duarte, F. J. Clinical guideline SEOM: cancer of unknown primary site. *Clin. Transl. Oncol.* **16**, 1091–1097 (2014).
17. Brewster, D. H., Lang, J., Bhatti, L. A., Thomson, C. S. & Oien, K. A. Descriptive epidemiology of cancer of unknown primary site in Scotland, 1961-2010. *Cancer Epidemiol.* **38**, 227–234 (2014).
18. Ferlay, J. *et al.* GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11. *Int. Agency Res. Cancer* (2013). Available at: <http://globocan.iarc.fr>. (Accessed: 5th August 2016)
19. Parkin, D. M., Whelan, S. L., Ferlay, J., Thomas, T. & Thomas, D. B. *Cancer Incidence in Five Continents. IARC Sci. Publication* **143**, (IARC Sci. Publications, 1997).
20. Varadhachary, G. R., & Raber, M. N. Cancer of Unknown Primary Site. *New Engl. J. Med.* **371**, 757–765 (2014).
21. Riihimäki, M., Hemminki, A., Sundquist, K. & Hemminki, K. Time trends in survival from cancer of unknown primary: small steps forward. *Eur. J. Cancer* **49**, 2403–2410 (2013).
22. Bevier, M., Sundquist, J. & Hemminki, K. Incidence of cancer of unknown primary in Sweden: analysis by location of metastasis. *Eur. J. Cancer Prev.* **21**, 596–601 (2012).
23. Shu, X., Sundquist, K., Sundquist, J. & Hemminki, K. Time trends in incidence, causes of death, and survival of cancer of unknown primary in Sweden. *Eur. J. Cancer Prev.* **21**, 281–288 (2012).

24. Brustugun, O. T. & Helland, Å. Rapid reduction in the incidence of cancer of unknown primary. A population-based study. *Acta Oncol.* **53**, 134–137 (2014).
25. Luke, C. *et al.* Exploring the epidemiological characteristics of cancers of unknown primary site in an Australian population: implications for research and clinical care. *Australian and New Zealand J Public Health* **32**, 383–389 (2008).
26. Mnatsakanyan, E., Tung, W.-C., Caine, B. & Smith-Gagen, J. Cancer of unknown primary: time trends in incidence, United States. *Cancer causes & control: CCC* **25**, 747–757 (2014).
27. Pentheroudakis, G., Golfopoulos, V. & Pavlidis, N. Switching benchmarks in cancer of unknown primary: From autopsy to microarray. *Eur. J. Cancer* **43**, 2026–2036 (2007).
28. Hainsworth, J. D., Schnabel, C. A., Erlander, M. G., Haines, D. W. & Greco, F. A. A retrospective study of treatment outcomes in patients with carcinoma of unknown primary site and a colorectal cancer molecular profile. *Clin. Col. Cancer* **11**, 112–118 (2012).
29. Hainsworth, J. D. *et al.* Paclitaxel/carboplatin with or without belinostat as empiric first-line treatment for patients with carcinoma of unknown primary site: A randomized, phase 2 trial. *Cancer* **121**, 1654–1661 (2015).
30. Hainsworth, J. D. *et al.* Treatment for patients with unknown primary cancer and favorable prognostic factors. *Seminars in Oncol.* **36**, 44–51 (2009).
31. Culine, S. *et al.* Development and validation of a prognostic model to predict the length of survival in patients with carcinomas of an unknown primary site. *J. Clin. Oncol.* **20**, 4679–4683 (2002).
32. Pavlidis N., Pentheroudakis G. Cancer of unknown primary site. *Lancet* **379**, 1428–1435 (2012).
33. Green, A. C. Cancer of unknown primary: Does the key lie in molecular diagnostics? *Cytopathology* **26**, 61–63 (2015).
34. Ariza, A. *et al.* Update on the diagnosis of cancer of unknown primary (CUP) origin. *Clin. Transl. Oncol.* **13**, 434–441 (2011).

35. Kandalaft, P. L. & Gown, A. M. Practical applications in immunohistochemistry: Carcinomas of unknown primary site. *Arch. Pathol. and Lab. Med.* **140**, 508–523 (2016).
36. Onofre, A. S. C., Pomjanski, N., Buckstegge, B. & Böcking, A. Immunocytochemical diagnosis of hepatocellular carcinoma and identification of carcinomas of unknown primary metastatic to the liver on fine-needle aspiration cytologies. *Cancer* **111**, 259–268 (2007).
37. Hashimoto, K. *et al.* Immunohistochemical profile for unknown primary adenocarcinoma. *PLoS ONE* **7**, e31181 (2012).
38. Conner, J. R. & Hornick, J. L. Metastatic carcinoma of unknown primary. *Advances In Anatom. Pathol.* **22**, 149–167 (2015).
39. Rubin, B. P., Skarin, A. T., Pisick, E., Rizk, M. & Salgia, R. Use of cytokeratins 7 and 20 in determining the origin of metastatic carcinoma of unknown primary, with special emphasis on lung cancer. *European J. Cancer Prev.* **10**, 77–82 (2001).
40. Gunia, S., Koch, S. & May, M. Is CDX2 immunostaining useful for delineating anorectal from penile/vulvar squamous cancer in the setting of squamous cell carcinoma with clinically unknown primary site presenting with histologically confirmed inguinal lymph node metastasis? *J. Clin. Pathol.* **66**, 109–112 (2012).
41. Anderson, G. G. & Weiss, L. Determining tissue of origin for metastatic cancers. *Appl. Immunohistochem. Mol. Morphol.* **18**, 3-8 (2010).
42. Ruiz-Ruiz, F. J., Saenz-Abad, D., Hualde-Enguita, A. M. & Morales-Rull, J. L. Positron emission tomography: useful in detecting metastatic cancer of unknown primary site. *Singapore Med. J.* **46**, 302–303 (2005).
43. Karapolat, I. & Kumanlioğlu, K. Impact of FDG-PET/CT for the detection of unknown primary tumours in patients with cervical lymph node metastases. *Mol. Imag. and Radionuclide Ther.* **21**, 63–68 (2012).
44. Rades, D. *et al.* Localised disease in cancer of unknown primary (CUP): the value of positron emission tomography (PET) for individual therapeutic management. *Ann. Oncol.* **12**, 1605–1609 (2001).
45. Lee, J. R. *et al.* Detection of occult primary tumors in patients with cervical metastases of

unknown primary tumors: comparison of (18)F FDG PET/CT with contrast-enhanced CT or CT/MR imaging-prospective study. *Radiology* **274**, 764–771 (2015).

46. Gutzeit, A. *et al.* Unknown primary tumors: detection with dual-modality PET/CT--initial experience. *Radiology* **234**, 227–234 (2005).
47. Fencel, P. *et al.* Prognostic and diagnostic accuracy of [18F]FDG-PET/CT in 190 patients with carcinoma of unknown primary. *Eur. J. of Nucl. Med. Molec. Imag.* **34**, 1783–1792 (2007).
48. Nanni, C. *et al.* Role of 18F-FDG PET-CT imaging for the detection of an unknown primary tumour: preliminary results in 21 patients. *ur. J. of Nucl. Med. Molec. Imag.* **32**, 589–592 (2005).
49. Delgado-Bolton, R. C., Fernández-Pérez, C., González-Maté, A. & Carreras, J. L. Meta-analysis of the performance of 18F-FDG PET in primary tumor detection in unknown primary tumors. *J. Nucl. Med.* **44**, 1301–1314 (2003).
50. Kwee, T. C. & Kwee, R. M. Combined FDG-PET/CT for the detection of unknown primary tumors: Systematic review and meta-analysis. *Eur. Radiol.* **19**, 731–744 (2009).
51. Beisenayeva, A. *et al.* The importance of immunohistochemical studies in the diagnosis of cancer of unknown primary origin. *Georgian Med. News* **249**, 73–80 (2015).
52. Biernacka, A., Linos, K. D., DeLong, P. A., Suriawinata, A. A., Padmanabhan, V., Liu, X. A case of S-100 negative melanoma: A diagnostic pitfall in the workup of a poorly differentiated metastatic tumor of unknown origin. *Cytojournal* **13**, 21 (2016).
53. Tothill, R. W. *et al.* An expression-based site of origin diagnostic method designed for clinical application to cancer of unknown origin. *Cancer Res.* **65**, 4031–4040 (2005).
54. Horlings, H. M. *et al.* Gene expression profiling to identify the histogenetic origin of metastatic adenocarcinomas of unknown primary. *J. Clin. Oncol.* **26**, 4435–4441 (2008).
55. Varadhachary, G. R. *et al.* Molecular profiling of carcinoma of unknown primary and correlation with clinical evaluation. *J. Clin. Oncol.* **26**, 4442–4448 (2008).
56. van Laar, R. K. *et al.* Implementation of a novel microarray-based diagnostic test for cancer

of unknown primary. *Int. J. Cancer* **125**, 1390–1397 (2009).

57. Handorf, C. R. *et al.* 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 (2013).
58. Yoon, H. H. *et al.* Gene expression profiling identifies responsive patients with cancer of unknown primary treated with carboplatin, paclitaxel, and everolimus: NCCTG N0871 (alliance). *Ann. Oncol.* **27**, 339–344 (2016).
59. Greco, F. A., Lennington, W. J., Spigel, D. R. & Hainsworth, J. D. Molecular profiling diagnosis in unknown primary cancer: Accuracy and ability to complement standard pathology. *J. Nat. Cancer Inst.* **105**, 782–790 (2013).
60. Varadhachary, G. R. *et al.* 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 (2011).
61. Ferracin, M. *et al.* MicroRNA profiling for the identification of cancers with unknown primary tissue-of-origin. *J. Pathol.* **225**, 43–53 (2011).
62. Esteller, M. Epigenetics in cancer. *New Engl. J. Med.* **358**, 1148–1159 (2008).
63. Heyn, H. & Esteller, M. DNA methylation profiling in the clinic: applications and challenges. *Nat. Rev. Genetics* **13**, 679–692 (2012).
64. Rando, O. J. Daddy issues: paternal effects on phenotype. *Cell* **151**, 702–708 (2012).
65. Schubeler, D. Epigenetic islands in a genetic ocean. *Science* **338**, 756–757 (2012).
66. Portela, A. & Esteller, M. Epigenetic modifications and human disease. *Nat. Biotechnol.* **28**, 1057–1068 (2010).
67. Vizoso, M. *et al.* Epigenetic activation of a cryptic TBC1D16 transcript enhances melanoma progression by targeting EGFR. *Nat. Med.* **21**, 741–750 (2015).
68. Sandoval, J. *et al.* A prognostic DNA methylation signature for stage I non-small-cell lung cancer. *J. Clin. Oncol.* **31**, 4140–4147 (2013).

69. Esteller, M. *et al.* Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *New Engl. J. Med.* **343**, 1350–1354 (2000).
70. Hegi, M. E. *et al.* MGMT gene silencing and benefit from temozolomide in glioblastoma. *New Engl. J. Med.* **352**, 997–1003 (2005).
71. Veeck, J. *et al.* BRCA1 CpG island hypermethylation predicts sensitivity to poly(adenosine diphosphate)-ribose polymerase inhibitors. *J. Clin. Oncol.* **28**, e563-564 (2010).
72. Ter Brugge, P. *et al.* Mechanisms of Therapy Resistance in Patient-Derived Xenograft Models of BRCA1-Deficient Breast Cancer. *J. Nat. Cancer Inst.* **108**, djw148 (2016).
73. Lopez-Serra, P. *et al.* A DERL3-associated defect in the degradation of SLC2A1 mediates the Warburg effect. *Nat. Comm.* **5**, 3608 (2014).
74. Nogales, V. *et al.* Epigenetic inactivation of the putative DNA/RNA helicase SLFN11 in human cancer confers resistance to platinum drugs. *Oncotarget* **7**, 3084-3097 (2016).
75. Moutinho, C. *et al.* Epigenetic inactivation of the BRCA1 interactor SRBC and resistance to oxaliplatin in colorectal cancer. *J. Nat. Cancer Inst.* **106**, djt322 (2014).
76. Roll, J. D. *et al.* Dysregulation of the epigenome in triple-negative breast cancers: Basal-like and claudin-low breast cancers express aberrant DNA hypermethylation. *Exp. Mol. Pathol.* **95**, 276–287 (2013).
77. Stefansson, O. A. *et al.* CpG island hypermethylation of BRCA1 and loss of pRb as co-occurring events in basal/triple-negative breast cancer. *Epigenetics* **6**, 638–649 (2011).
78. Alvarez, S. *et al.* DNA methylation profiles and their relationship with cytogenetic status in adult acute myeloid leukemia. *PloS One* **5**, e12197 (2010).
79. Wilop, S. *et al.* Array-based DNA methylation profiling in acute myeloid leukaemia. *British Journal of Haematology* **155**, 65–72 (2011).
80. Ibragimova, I. *et al.* Genome-wide promoter methylome of small renal masses. *PloS One* **8**, e77309 (2013).
81. Lenhard, K. *et al.* Analysis of promoter methylation in stool: a novel method for the

detection of colorectal cancer. *Clin. Gastroenterol. Hepatol.* **3**, 142–149 (2005).

82. Sabbioni, S. *et al.* Multigene methylation analysis of gastrointestinal tumors: TPEF emerges as a frequent tumor-specific aberrantly methylated marker that can be detected in peripheral blood. *Mol. Diag.* **7**, 201–207 (2003).
83. Sameer, A. S. & Nissar, S. Epigenetics in diagnosis of colorectal cancer. *Mol. Biol. Res. Comm.* **5**, 49–57 (2016).
84. Lee, B. Bin *et al.* Aberrant methylation of APC, MGMT, RASSF2A, and Wif-1 genes in plasma as a biomarker for early detection of colorectal cancer. *Clin. Cancer Res.* **15**, 6185–6191 (2009).
85. Lee, W. H., Isaacs, W. B., Bova, G. S. & Nelson, W. G. CG island methylation changes near the GSTP1 gene in prostatic carcinoma cells detected using the polymerase chain reaction: a new prostate cancer biomarker. *Cancer Epidemiol. Biom. & Prevent.* **6**, 443–450 (1997).
86. Goessl, C. *et al.* Fluorescent methylation-specific polymerase chain reaction for DNA-based detection of prostate cancer in bodily fluids. *Cancer Res.* **1**, 5941–5945 (2000).
87. Peng, X. *et al.* Accumulated promoter methylation as a potential biomarker for esophageal cancer. *Oncotarget* **8**, 679-691 (2017)
88. Su, Y., Fang, H. & Jiang, F. Integrating DNA methylation and microRNA biomarkers in sputum for lung cancer detection. *Clin. Epigenet.* **8**, 109 (2016).
89. Wei, J. *et al.* Integrated analysis of genome-wide DNA methylation and gene expression profiles identifies potential novel biomarkers of rectal cancer. *Oncotarget* **7**, 62547-62558 (2016).
90. Toyota, M. *et al.* CpG island methylator phenotype in colorectal cancer. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 8681-8686 (1999).
91. Kanai, Y., Ushijima, S., Kondo, Y., Nakanishi, Y., Hirohashi, S. DNA methyltransferase expression and DNA methylation of CPG islands and peri-centromeric satellite regions in human colorectal and stomach cancers. *Int. J. Cancer* **91**, 205-212 (2001).

92. Ogino, S. *et al.* CpG island methylator phenotype, microsatellite instability, BRAF mutation and clinical outcome in colon cancer. *Gut* **58**, 90-96 (2009).
93. Abe, M. *et al.* CpG island methylator phenotype is a strong determinant of poor prognosis in neuroblastomas. *Cancer Res.* **1**, 828-834 (2005).
94. Asada, K. *et al.* Stronger prognostic power of the CpG island methylator phenotype than methylation of individual genes in neuroblastomas. *Jpn. J. Clin. Oncol.* **43**, 641-645 (2013).
95. Sandoval, J. *et al.* Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics* **6**, 692–702 (2011).
96. Moran, S., Arribas, C. & Esteller, M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. *Epigenomics* **8**, 389-399 (2016).
97. Heyn, H., Méndez-González, J. & Esteller, M. Epigenetic profiling joins personalized cancer medicine. *Exp. Rev. Mol. Diagn.* **13**, 473–479 (2013).
98. Moran, S. *et al.* Validation of DNA methylation profiling in formalin-fixed paraffin-embedded samples using the Infinium HumanMethylation450 Microarray. *Epigenetics* **9**, 829–833 (2014).
99. 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**, 596-599 (2008).
100. Hainsworth, J. D. *et al.* 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 (2013).
101. Greco, F. A. Improved diagnosis, therapy and outcomes for patients with CUP. *Nat. Rev. Clin. Oncol.* **14**, 5–6 (2016).
102. Varadhachary, G.R. *et al.* Carcinoma of unknown primary with gastrointestinal profile: immunohistochemistry and survival data for this favorable subset. *Int. J. Clin. Oncol.* **19**, 479-484 (2014).

103. Pavlidis, N., Petrakis, D., Golfopoulos, V. & Pentheroudakis, G. Long-term survivors among patients with cancer of unknown primary. *Critical Rev. in Oncol./Hematol.* **84**, 85–92 (2012).
104. Tothill, R.W., *et al.* Massively-parallel sequencing assists the diagnosis and guided treatment of cancers of unknown primary. *J. Pathol.* **231**, 413–423 (2013).
105. Ross, J. S. *et al.* Comprehensive Genomic Profiling of Carcinoma of Unknown Primary Site: New Routes to Targeted Therapies. *JAMA Oncol.* **1**, 40–49 (2015).
106. Stella, G. M. *et al.* MET mutations in cancers of unknown primary origin (CUPs). *Human Mut.* **32**, 44–50 (2011).
107. Gatalica, Z. *et al.* Comprehensive tumor profiling identifies numerous biomarkers of drug response in cancers of unknown primary site: analysis of 1806 cases. *Oncotarget* **5**, 12440–12447 (2014).
108. Varadhachary, G. Carcinoma of unknown primary site: the poster child for personalized medicine? *JAMA Oncol.* **1**, 19–21 (2015).
109. Amatu, A., *et al.* Promoter CpG island hypermethylation of the DNA repair enzyme MGMT predicts clinical response to dacarbazine in a phase II study for metastatic colorectal cancer. *Clin. Cancer Res.* **15**, 2265–2272 (2013).
110. Amatu A, *et al.* Tumor MGMT promoter hypermethylation changes over time limit temozolomide efficacy in a phase II trial for metastatic colorectal cancer. *Ann. Oncol.* **27**, 1062–1067 (2016).
111. Palma, N. A. *et al.* Durable Response to Crizotinib in a MET-Amplified, KRAS-Mutated Carcinoma of Unknown Primary. *Case Rep. Oncol.* **7**, 503–508 (2014).
112. Chung, J.H., *et al.* A Poorly Differentiated Malignant Neoplasm Lacking Lung Markers Harbors an EML4-ALK Rearrangement and Responds to Crizotinib. *Case Rep. Oncol.* **7**, 7628–7632 (2014).
113. Martinez-Cardús, A. *et al.* Epigenetic homogeneity within colorectal tumors predicts shorter relapse-free and overall survival times for patients with locoregional cancer. *Gastroenterol.* **151**, 961–972 (2016).

114. Stelzer, Y., Shivalila, C. S., Soldner, F., Markoulaki, S. & Jaenisch, R. Tracing dynamic changes of DNA methylation at single-cell resolution. *Cell* **163**, 218–229 (2015).
115. Kantlehner, M. *et al.* A high-throughput DNA methylation analysis of a single cell. *Nucleic Acids Res.* **39**, e44–e44 (2011).
116. Komine, K. *et al.* Utility of measuring circulating tumor cell counts to assess the efficacy of treatment for carcinomas of unknown primary origin. *Anticancer Res.* **34**, 3165–3168 (2014).
117. Fraga, M. F. *et al.* Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat Genet.* **37**, 391–400 (2005).
118. Seligson, D. B. *et al.* Global histone modification patterns predict risk of prostate cancer recurrence. *Nature* **435**, 1262–1266 (2005).
119. Schmidl, C., Rendeiro, A. F., Sheffield, N. C., Bock, C. ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors. *Nat. Methods* **12**, 963–965 (2015).
120. Rotem, A. *et al.* Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat Biotechnol.* **33**, 1165–1172 (2015).
121. Hyun, B. R., McElwee, J. L., Soloway, P. D. Single molecule and single cell epigenomics. *Methods* **72**, 41–50 (2015).
122. Buenrostro, J. D. *et al.* Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* **523**, 486–490 (2015).
123. Rendeiro, A. F. *et al.* Chromatin accessibility maps of chronic lymphocytic leukaemia identify subtype-specific epigenome signatures and transcription regulatory networks. *Nat. Commun.* **7**, 11938 (2016).
124. Esteller, M. *et al.* Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. *Cancer Res.* **59**, 67–70 (1999).
125. Wong, I. H. *et al.* Detection of aberrant p16 methylation in the plasma and serum of liver cancer patients. *Cancer Res.* **59**, 71–73 (1999).

126. Kristensen, L. S. *et al.* Aberrant methylation of cell-free circulating DNA in plasma predicts poor outcome in diffuse large B cell lymphoma. *Clin.Epigenet.* **8**, 95 (2016).
127. Vaca-Paniagua, F. *et al.* Targeted deep DNA methylation analysis of circulating cell-free DNA in plasma using massively parallel semiconductor sequencing. *Epigenomics* **7**, 353–362 (2015).
128. Lehmann-Werman, R, *et al.* Identification of tissue-specific cell death using methylation patterns of circulating DNA. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E1826-1834 (2016).
129. Guo, S. *et al.* Identification of methylation haplotype blocks aids in deconvolution of heterogeneous tissue samples and tumor tissue-of-origin mapping from plasma DNA. *Nat Genet.* (2017)
130. Tothill, R. W. *et al.* Development and validation of a gene expression tumour classifier for cancer of unknown primary. *Pathology* **47**, 7-12 (2015).

FIGURE LEGENDS

Figure 1. Biopsy sites and frequencies of tumor types predicted by EPICUP.

Frequency of sites (left) where orphan metastases were biopsied from the EPICUP study. Representation of the frequency of tumor type predicted by EPICUP (right). Those sites or tumor types with frequencies <1% are not shown. Ca.: carcinoma; CNS: central nervous system neoplasia; HCC: hepatocellular carcinoma; HNSC: head and neck squamous carcinoma; NSCLC: non-small cell lung cancer; Uroth.: urothelial.

Figure 2. Examples of histopathological validation of CUP cases whose origin was predicted by EPICUP.

a) CUP case predicted by EPICUP as sarcoma: histopathological validation of the tumor sample after necropsy, using hematoxylin–eosin (H&E) staining and immunohistochemistry, which was shown to be positive by the specific sarcoma marker vimentin. The S100 marker, unrelated to sarcoma, was negative. b) Predicted breast cancer CUP case by EPICUP, showing H&E staining, cytokeratin (CK) markers CK 7 (positive) and CK 20 (negative) and the specific breast marker mammoglobin that confirmed the breast cancer diagnosis. c) CUP case predicted by EPICUP as colon carcinoma, showing H&E staining, the cytokeratin markers CK 7 (negative) and CK 20 (positive), and the specific colon marker intestine-specific transcription factor (CDX2) positive, validating the colon cancer diagnosis.

Figure 3. Illustration of how the determination of the primary site in CUP cases can be helpful for the right selection of the tailored-treatment.

Figure 4. CUP patient who received a site-specific treatment that matched the EPICUP prediction and a guided search for “actionable” molecular targets. CUP case that epigenetic profiling predicted to be derived from a non-small cell lung cancer (NSCLC) primary, a diagnosis validated by hematoxylin–eosin (H&E) staining and IHC studies with the cytokeratin (CK) markers CK 7 (positive) and CK 20 (negative) and the specific lung marker TTF-1 (thyroid transcription factor-1), in which a G719X EGFR (epidermal growth factor receptor) mutation was detected by real-time specific PCR amplification.

Figure 1

Biopsy site at diagnosis

EPICUP-predicted tumor type

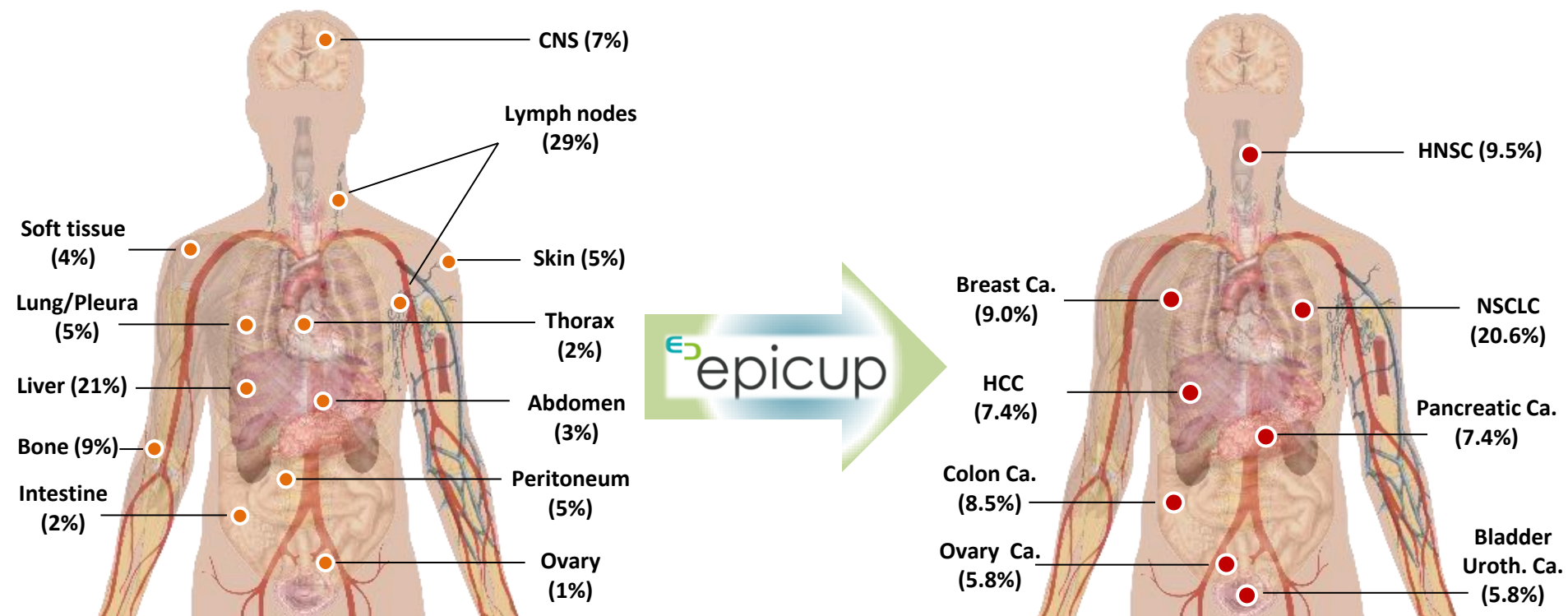
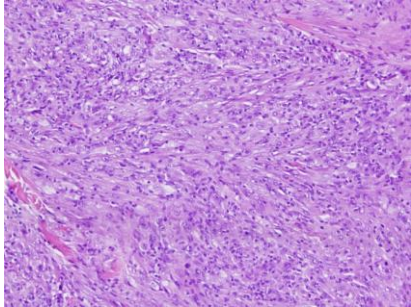


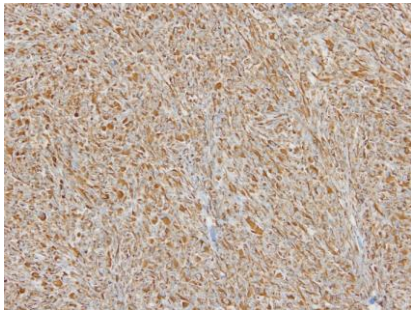
Figure 2

a

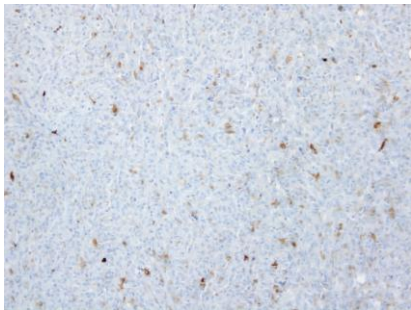
H&E



Vimentin

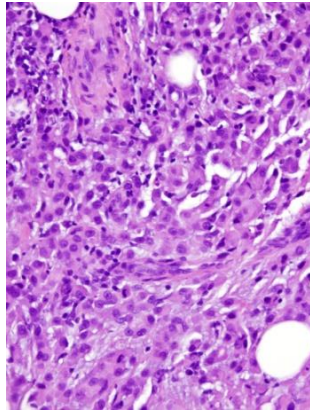


S100

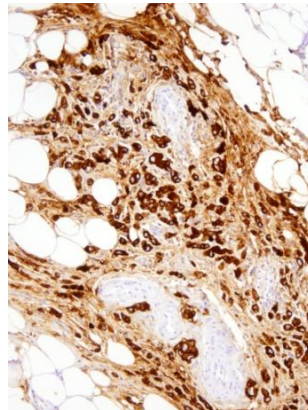


b

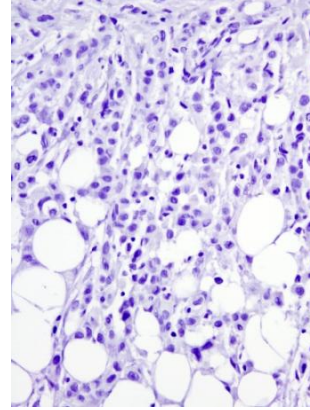
H&E



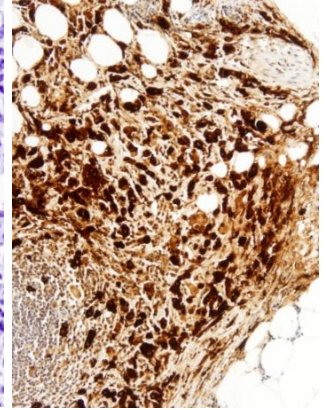
CK 7



CK 20

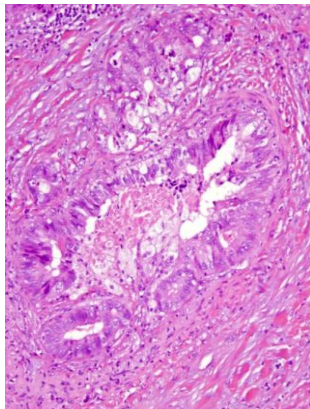


Mammoglobin

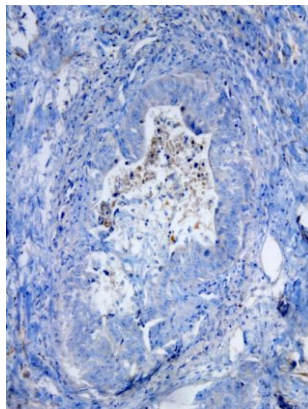


c

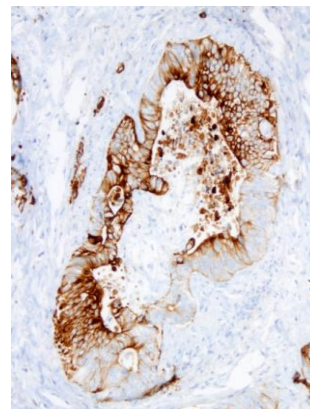
H&E



CK 7



CK 20



CDX2

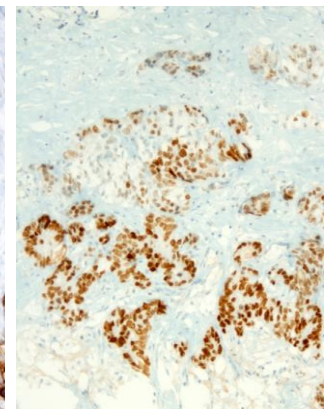


Figure 3

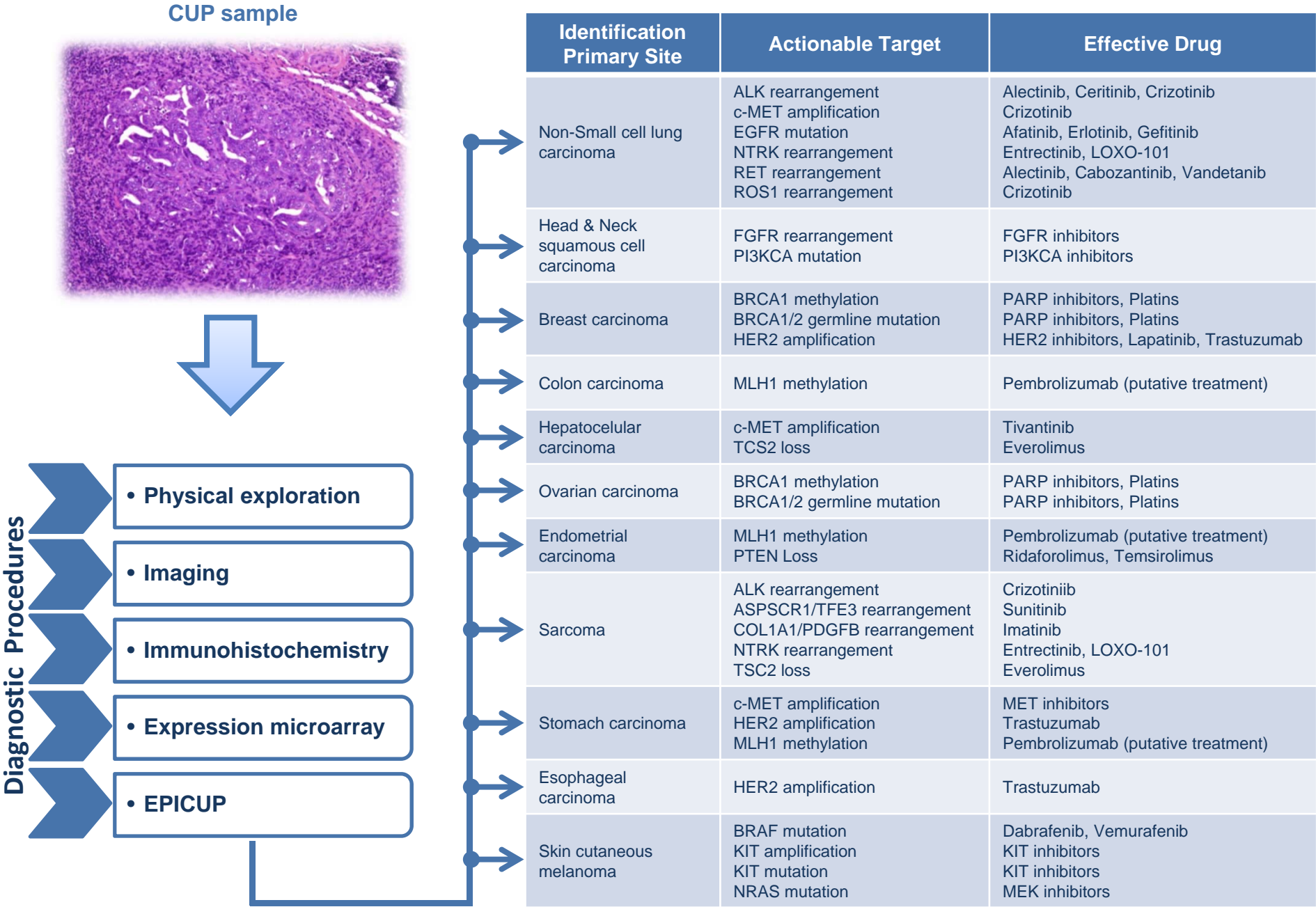
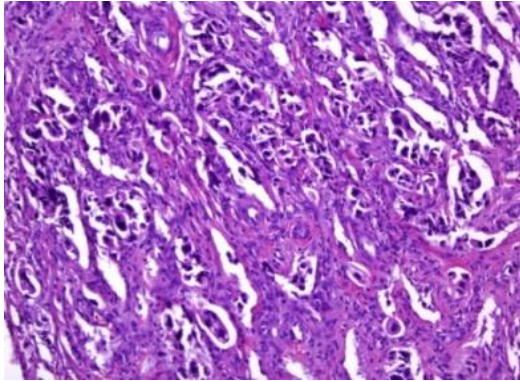
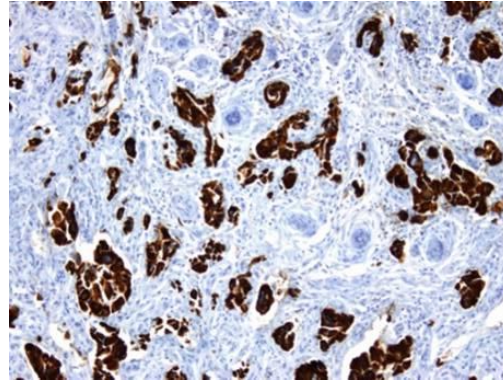


Figure 3

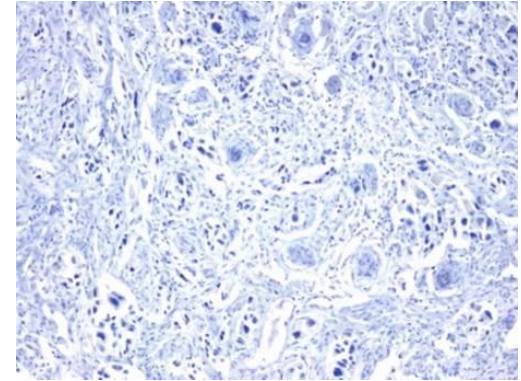
H&E



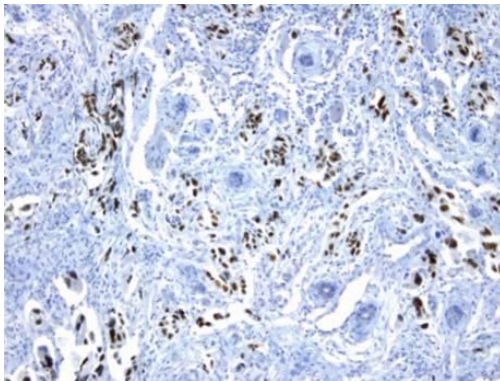
CK 7



CK 20



TTF-1



EGFR mutation in codon 18

