original report

Mutational Signature and Transcriptomic Classification Analyses as the Decisive Diagnostic Tools for a Cancer of Unknown Primary

Roger Olofsson Bagge
Akif Demir
Joakim Karlsson
Babak Alaei-Mahabadi
Berglind O. Einarsdottir
Henrik Jespersen
Mattias F. Lindberg
Andreas Muth
Lisa M. Nilsson
Marta Persson
Johanna B. Svensson
Elin M.V. Söderberg
Ronald R. de Krijger
Ola Nilsson
Erik Larsson
Göran Stenman
Jonas A. Nilsson

Author affiliations and support information (if applicable) appear at the end of this article.

R.O.B., A.D., and J.K. are co-first authors.

Corresponding author: Jonas A. Nilsson, PhD, Goteborgs Universitet Sahlgrenska Akademin, Sahlgrenska Cancer Center, Medicinaregatan (continued)

abstract

Purpose **Cancer of unknown primary is a group of metastatic tumors in which the standard diagnostic workup fails to identify the site of origin of the tumor. The potential impact of precision oncology on this group of patients is large, because actionable driver mutations and a correct diagnosis could provide treatment options otherwise not available for patients with these fatal cancers. This study investigated if comprehensive genomic analyses could provide information on the origin of the tumor.**

Patients and Methods **Here we describe a patient whose tumor was misdiagnosed at least three times. Next-generation sequencing, a patient-derived xenograft mouse model, and bioinformatics were used to identify an actionable mutation, predict resistance development to the targeted therapy, and correctly diagnose the origin of the tumor. Transcriptomic classification was benchmarked using The Cancer Genome Atlas (TCGA).**

Results **Despite the lack of a known primary tumor site and the absence of diagnostic immunohistochemical markers, the origin of the patient's tumor was established using the novel bioinformatic workflow. This included a mutational signature analysis of the sequenced metastases and comparison of their transcriptomic profiles to a pan-cancer panel of tumors from TCGA. We further discuss the strengths and limitations of the latter approaches in the context of three potentially incorrectly diagnosed TCGA lung tumors.**

Conclusion **Comprehensive genomic analyses can provide information on the origin of tumors in patients with cancer of unknown primary.**

INTRODUCTION

Massively parallel sequencing of tumor DNA and RNA offers new possibilities to identify actionable mutations and enables precision medicine in the clinic.[1,2] Identification of a high mutational load (eg, resulting from carcinogen exposure or DNA repair defects[3]) can also provide information on increased responsiveness to immunotherapies, such as anti–programmed death 1 antibodies.[4] Trials such as the NCI-MATCH (National Cancer Institute Molecular Analysis for Therapy Choice) trial, the TAPUR (Targeted Agent and Profiling Utilization Registry) study, and the NCT-MASTER (Molecularly Aided Stratification for Tumor Eradication) study[5,6] include next-generation sequencing (NGS) to identify actionable mutations and will generate new data to inform future trial designs or best clinical practice. For a few diseases, most notably breast cancer, malignant melanoma, and lung cancer, sequencing to identify therapeutically actionable driver alterations is a mainstay in many countries.

Cancer of unknown primary (CUP) accounts for 2% to 5% of all diagnosed malignancies and is a frequent cause of death resulting from

1G, Plan 6, Gothenburg, NA 41390, Sweden; e-mail: jonas.a.nilsson@surgery.gu.se.

cancer.[7] The clinical presentation of CUP is heterogeneous, ranging from a single metastasis to widespread metastatic disease. Therapy options include chemotherapy protocols based on the presumed anatomic site of origin or a taxane and platinum combination. Because their prognosis is poor,[7] patients may benefit significantly from extended genomic analysis not only for identification of targeted treatments but also for identification of the primary site of the tumor.[8,9] Targeted therapies directed against a particular driver mutation can act differently in different tumor types[10]; therefore, it may be beneficial to identify both the actionable mutation and the primary tumor site.

Here we describe an unusual case of a patient seeking a second opinion while receiving therapy with pembrolizumab for a presumed lung cancer. By employing comprehensive, targeted, whole-genome and RNA sequencing, we identified an actionable mutation and established the correct diagnosis and origin of the tumor. The bioinformatic workflow described, with emphasis on transcriptome-based classification and mutational signature analysis, could be useful as a template for establishing the diagnoses and origins of CUPs, which could enable new treatment options for patients with these otherwise fatal cancers, especially when combined with standard profiling of known actionable driver mutations.

PATIENTS AND METHODS

Case of a *BRAF*-Mutant Tumor With Neuroendocrine Differentiation

We describe a nonsmoking 67-year-old man who initially sought medical attention because of coughing in December 2015. A computed tomography (CT) scan revealed a central mass around the right pulmonary hilus, together with enlarged mediastinal lymph nodes. Endobronchial ultrasound-guided aspiration suggested that the patient had an ALK-positive non–small-cell lung cancer. However, a subsequent fluorescent in situ hybridization analysis failed to confirm the ALK positivity, and re-evaluation of the original specimen suggested instead a large-cell neuroendocrine carcinoma of the lung.

At the patient's request, the cell block sections were re-examined in May 2016 at the Sahlgrenska University Hospital in Sweden. Cytologic and immunohistochemical (IHC) examination of the original specimen suggested a malignant neuroendocrine neoplasm (NEN), possibly a malignant paraganglioma (PGL) with differential diagnoses being gangliocytic neuroblastoma, malignant melanoma, Merkel cell carcinoma, carcinoid, and small-cell lung cancer (Fig 1). Analysis of an extirpated subcutaneous metastasis revealed that the tumor was positive for the neuroendocrine markers tyrosine hydroxylase (TH), chromogranin A, synaptophysin (SYN), Neu-N, and CD99, as well as for the neural crest markers SOX10 and S100. In contrast, the tumor was negative for cytokeratins as well as the melanoma markers tyrosinase, HMB45 (PMEL), and melan-A (Fig 1; Appendix Figs A1 to A3). A patient-derived xenograft (PDX) mouse model was generated as previously described[11] from the subcutaneous metastasis, and the xenograft was also positive for TH, SYN, SOX10, and S100 (Fig 1). Taken together, these analyses argue for a NEN of unknown primary, which accounts for > 10% of these diagnoses.[12]

Urinary dopamine levels were transiently elevated (approximately three times the upper limit of normal), but repeated measurements of plasma metanephrines (including metoxytyramine) were negative, arguing against a primary PGL. NGS and multiplex ligation-dependent probe amplification analyses of genes previously associated with hereditary PGL were all negative,[13,14] and there was no family history of PGL. Array comparative genomic hybridization and whole-genome sequencing analyses of the metastasis revealed an aneuploid tumor with copy number alterations preferentially affecting whole chromosomes and chromosome arms, including gain of 1q, 2, 7, 8, 9p24.3-p23, 9q, 13, 15, 16, 20, and 21 and loss of 1p, 9p23-p13.2 (homozygous loss of *CDKN2A*), 10, and 18 (Fig 2A and 2B; Appendix Table A1; Data Supplement). *MAML3* gene fusions found in PGL[15] were not detected (Data Supplement). Finally, and atypically for PGL[15,16] or other NENs, allele-specific PCR and NGS demonstrated that the biopsies harbored a *BRAF* V600E mutation, which was confirmed by whole-genome and RNA sequencing (Fig 2C; Data Supplement).

Treatment of the Patient

Upon initial diagnosis, medical treatment with crizotinib (because of presumed ALK positivity) and platinum-based chemotherapy were

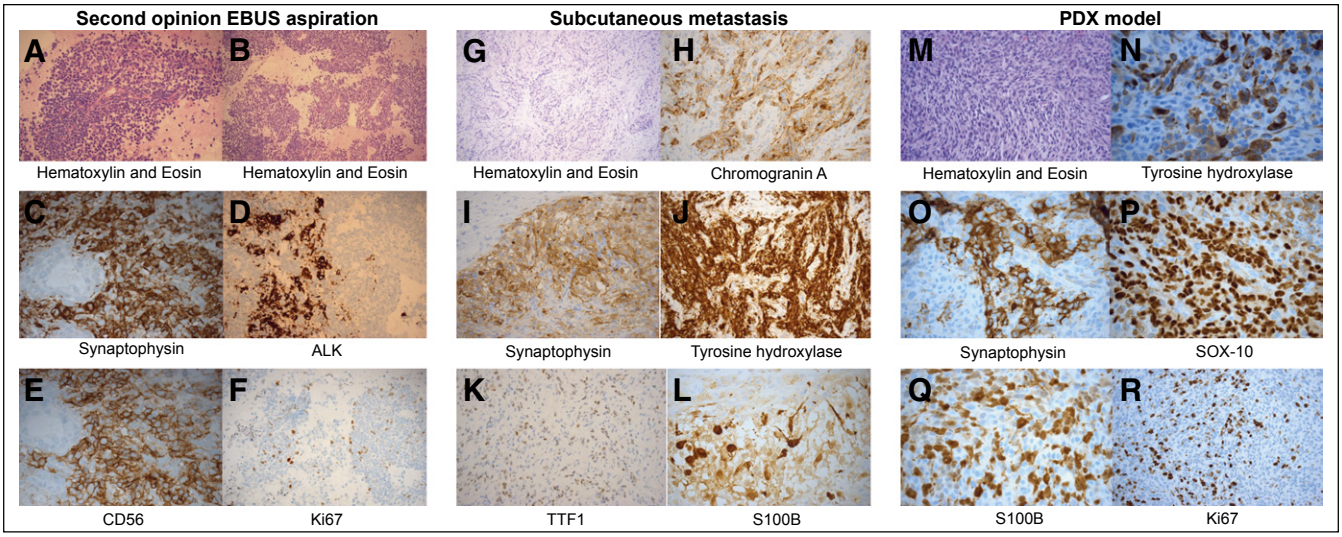


Fig 1. Morphologic and immunohistochemical characteristics of the cell block from the (A-F) second opinion endobronchial ultrasound-guided aspiration, (G-L) subcutaneous metastatic lesion, and (M-R) patient-derived xenograft model. Sections were stained with (A, B, G, M) hematoxylin and eosin or with antibodies directed against the indicated antigens as follows: (C, I, O) synaptophysin, (D) ALK, (E) CD56, (F, R) Ki67, (H) chromogranin A, (J, N) tyrosine hydroxylase, (K) TTF1, (L, Q) S100B, and (P) SOX10.

tried, but the patient experienced progression. The patient was therefore switched to immunotherapy with pembrolizumab. [^{18}F]Fluorodeoxyglucose–positron emission tomography/CT examination at Sahlgrenska University Hospital (Fig 3A) showed progression, and the patient had clinical symptoms of stridor and shortness of breath. He also had a problem with hypotension. The pembrolizumab treatment was stopped, and iodine-123–metaiodobenzylguanidine scintigraphy and somatostatin receptor scintigraphy were performed. Both were negative, ruling out the use of metaiodobenzylguanidine- or peptide receptor–mediated radionuclide therapy.

After a multidisciplinary conference in August 2016, the patient started with the BRAF inhibitor (BRAFi) dabrafenib (based on the V600E mutation). A rapid partial response was observed, with shrinkage of the subcutaneous metastases, disappearance of clinical symptoms including the stridor, and metabolic response on [^{18}F]fluorodeoxyglucose–positron emission tomography/CT after evaluation at 8 weeks (Fig 3A). The clinical response lasted nearly 4 months, after which a general performance decrease occurred. Brain magnetic resonance imaging in late November revealed three brain metastases, of which the largest measured 5 cm and was located superficially in the frontal lobe. The patient was prescribed the MEK inhibitor (MEKi) trametinib in combination with dabrafenib. He also underwent surgery in December 2016 to remove the frontal lobe metastasis. By January 2017, the patient had experienced further progression. Treatment with BRAFi and MEKi was stopped, and he received temozolomide 250 mg daily. Unfortunately, he rapidly deteriorated; a CT brain scan revealed massive progression of the brain metastases, and the patient died shortly thereafter.

In the PDX model generated from the subcutaneous metastasis, treatment with the BRAFi dabrafenib alone or in combination with the MEKi trametinib was tested. Notably, there was no effect of the BRAFi alone, whereas in two of three mice, the combination treatment led to markedly slower disease progression, which correlated with a reduction of phosphorylation of the MEK target ERK (Fig 3B). The lack of complete inhibition of ERK phosphorylation could be explained by extended analyses of the genome sequencing of the subcutaneous metastasis, which revealed P124L and F53L mutations in *MAP2K1* (encoding MEK1; Fig 2D). These mutations have previously been associated with resistance to BRAFi in melanoma.[17]

Comprehensive Immunoprofiling to Establish the Correct Diagnosis

To diagnose this unusual case of presumed NEN, additional morphologic and IHC analyses were performed (Appendix Table A2; Appendix Figs A1 to A3). Analyses of all specimens, including the brain metastasis, showed that the tumor had partly epithelioid and partly elongated cells, with moderate nuclear pleomorphism but no prominent nucleoli, squamous differentiation, or gland formation. On the basis of marker expression and the sites of the metastases, a differential diagnosis of neuroendocrine

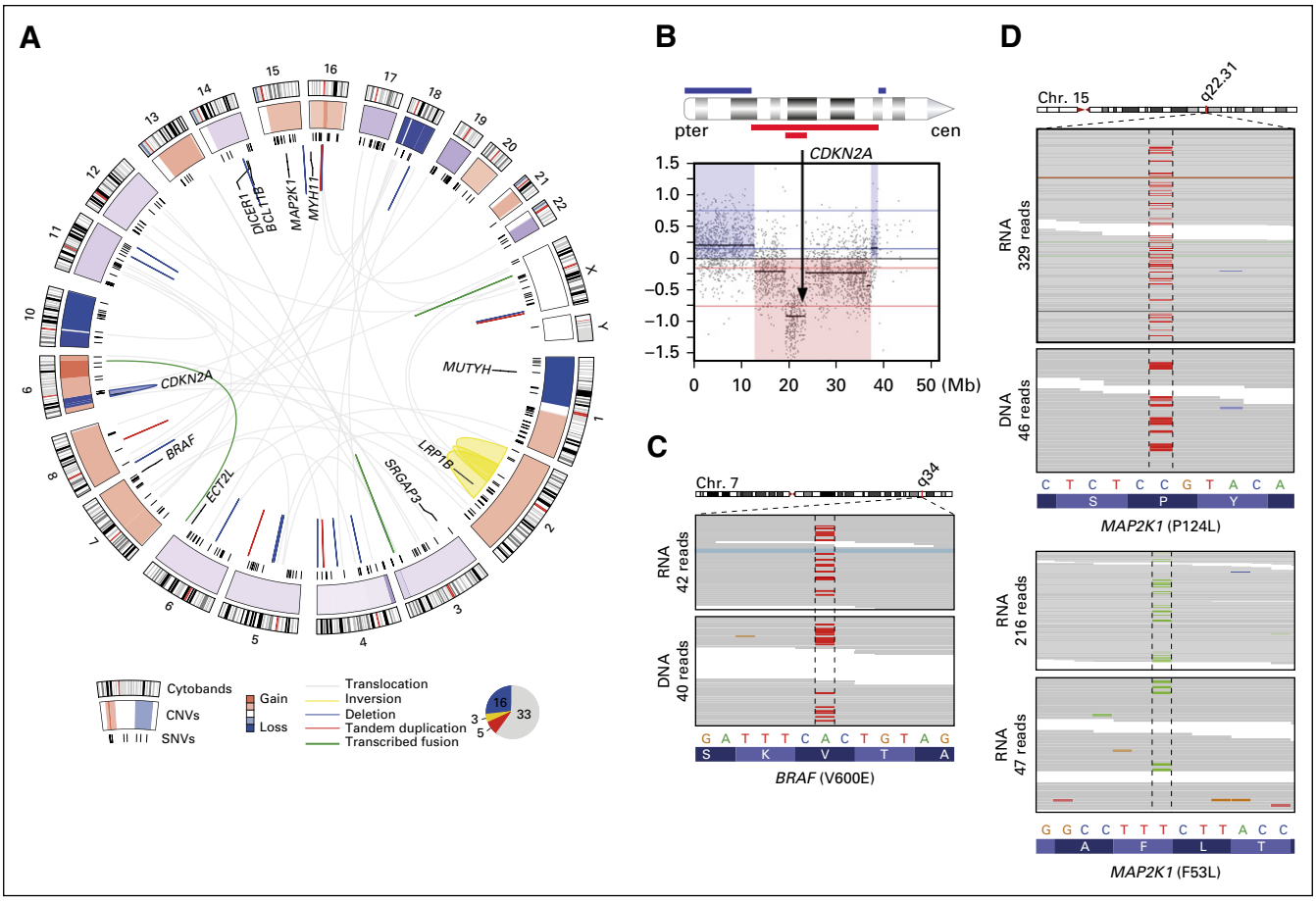


Fig 2. (A) Somatic copy number variants (CNVs), single-nucleotide variants (SNVs), expressed fusion genes (thick green lines), and somatic structural variants revealed by whole-genome sequencing and (B) array comparative genomic hybridization. Highlighted are mutations in genes included in the Catalogue of Somatic Mutations in Cancer Cancer Gene Census, as well as a homozygous deletion spanning *CDKN2A*. All mutations and copy number alterations can be found in Appendix Tables A1 to A3. (C) Somatic variants detected with whole-genome and RNA sequencing in *BRAF* (V600E) and (D) *MAP2K1* (P124L and F53L).

carcinoma, malignant PGL, neuroblastic tumor, or malignant melanoma could be entertained. However, neuroendocrine carcinomas were unlikely, because they express keratins and show more diffuse and stronger expression of SYN and chromogranin-A. PGLs are negative for keratins but have strong SYN and chromogranin-A immunostaining, even when they are malignant. In addition, they have a nested growth pattern with S100-positive sustentacular cells. Although such cells may be lost in malignant forms, and these may also have a more diffuse growth pattern, the extent of atypia and necrosis and the complete lack of nesting and sustentacular cell pattern would be unusual. A neuroblastic tumor would be unusual at the patient's age and with this dissemination pattern, and the lack of ganglion cell differentiation and NB84 expression is not compatible with this diagnosis. Lack of such differentiation may occur in undifferentiated neuroblastoma, but the patient's tumor did not show the typical small blue round-cell tumor pattern characteristic of this entity. Finally, the possibility of a malignant melanoma was entertained. The presence of combined epithelioid and spindle cell differentiation would be compatible with this diagnosis, although there were no prominent nucleoli, which are frequently observed in melanoma. In contrast, positive staining for SOX10 and S100 favored this diagnosis but was not definitive. Neural and neuroendocrine differentiation, as exemplified by focal reactivity for SYN, chromogranin-A, and TH, has been noted in melanoma[18] but has always been accompanied by at least one of the melanoma markers HMB45, tyrosinase, or melan-A. However, in our patient, the decisive melanoma markers were not expressed. Thus, our extended IHC profiling did not provide conclusive evidence in favor of any of the abovementioned diagnoses.

RESULTS

Mutational Signature and Transcriptome Analyses Reveal the Identity of the Primary Tumor

Parallel to the IHC analyses of the brain metastasis, we also took a genomic approach to this

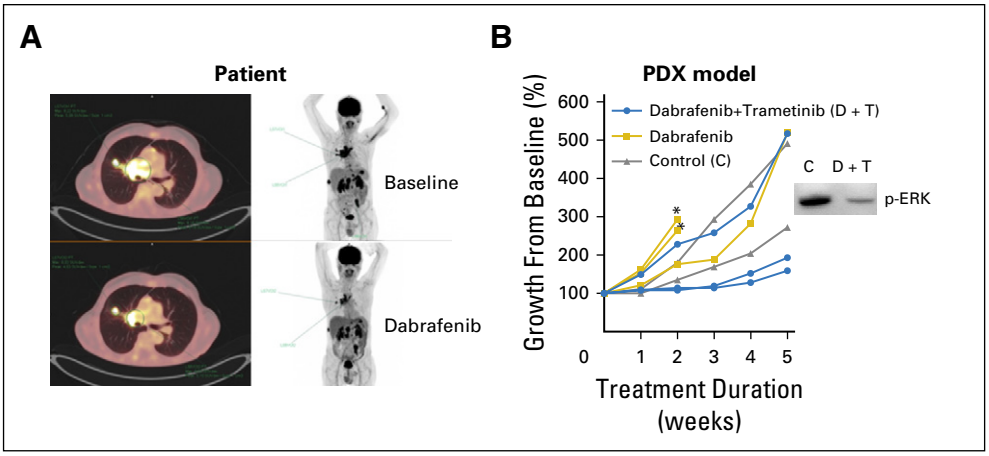


Fig 3. Tumor responses to the BRAF-targeted therapy. (A) Tumor load in patient visualized on [^{18}F] fluorodeoxyglucose ([^{18}F] FDG) positron emission tomography/computed tomography at baseline (top panels) and after 3 months of treatment with the BRAF inhibitor dabrafenib (bottom panels). Green ring and arrows highlighting main tumor mass and showing tumor regression with decreased [^{18}F]FDG uptake. (B) Patient-derived xenograft (PDX) model in nonobese severe combined immune-deficient interleukin-2 chain receptor γ knockout mice (NOG mice). Mice were treated with vehicle (control [C]), dabrafenib (D), or dabrafenib plus trametinib (D + T). Inset shows western blot analysis of activated and phosphoprylated ERK (pERK) in PDX tumors from mice treated with C or D + T. (*) Mice were euthanized because of rapid tumor growth or weight loss.

case. Mutational signatures have emerged as powerful tools to determine the somatic mutational processes underlying the development of tumors.[19,20] By calculating the frequencies of base substitutions across the genome, we observed a dominance of C>T transitions that preferentially occurred in dipyrimidine trinucleotide contexts (Fig 4A) and that were biased toward the untranscribed strand (Poisson test, false discovery rate < 0.05) in both the skin and brain metastases. This pattern closely matched a known mutational signature from the Catalogue of Somatic Mutations in Cancer (signature 7) associated with ultraviolet radiation (UV)-induced damage (Fig 4A), which is only found in cancers originating from sun-exposed areas.[19,20] A decomposition of the mutational spectrum into known mutational signatures, using the MutationalPatterns R package,[21] yielded an estimated 80% contribution from signature 7 in both samples. These results strongly indicate that the primary tumor had originated from the skin.

The classic melanoma markers TYR, PMEL, and MLANA and the Merkel cell carcinoma marker cytokeratin 20 were all negative, both by IHC analyses and by RNA sequencing (RNA-seq). Therefore, we instead compared the transcriptome of the present tumor with that of 9,583 other tumors from 32 cancer types sequenced by The Cancer Genome Atlas (TCGA) consortium. Notably, for both the skin and brain metastases, the strongest correlating samples were melanomas (Fig 4B). Moreover, RNA-seq demonstrated expression of the melanoblast markers *MITF*, *DCT*, and *SOX10* in the metastases as well as in the PDX tumors. *DCT* and *SOX10* were also elevated in xenografts from BRAFi-treated mice, which correlated with suppression of the neuroendocrine markers *CHGB*, *TH*, and *ENO2* (Appendix Fig A4; Data Supplement). Taken together, our findings of an UV-induced mutational signature, *BRAF/MAP2K1/CDKN2A* mutations, the dissemination pattern, the expression of S100 and SOX10, and the correlation with the transcriptome profiles of TCGA melanomas strongly indicated that the present tumors represented metastases from an unknown primary melanoma with neuroendocrine differentiation.

Pan-Cancer Transcriptome-Based Classification Can Predict the Cellular Origin of a Tumor

To investigate how transcriptome-based classification performance generalizes beyond the present case, we performed leave-one-out cross-validation of k nearest neighbor classification on the pan-cancer data set, varying the number of nearest neighbors (k) considered, in terms of Spearman correlation, from one to 50 (Appendix). The optimal choice (k = 6) yielded an overall accuracy of 95% (Appendix Figs A5A and A5B). However, certain cancers were more difficult to predict than others; for example, rectal and colon adenocarcinomas were difficult to distinguish, as were squamous carcinoma of the lung, squamous carcinoma of the head and neck, and bladder cancer, which share molecular characteristics[22] (Fig 5A).

An additional evaluation was performed on an independent set of previously published data[11] and unpublished RNA-seq data of melanoma PDXs, uveal melanoma metastases, and one lung adenocarcinoma, yielding an overall accuracy of 0.84 (Appendix Fig A5C). Some metastases of

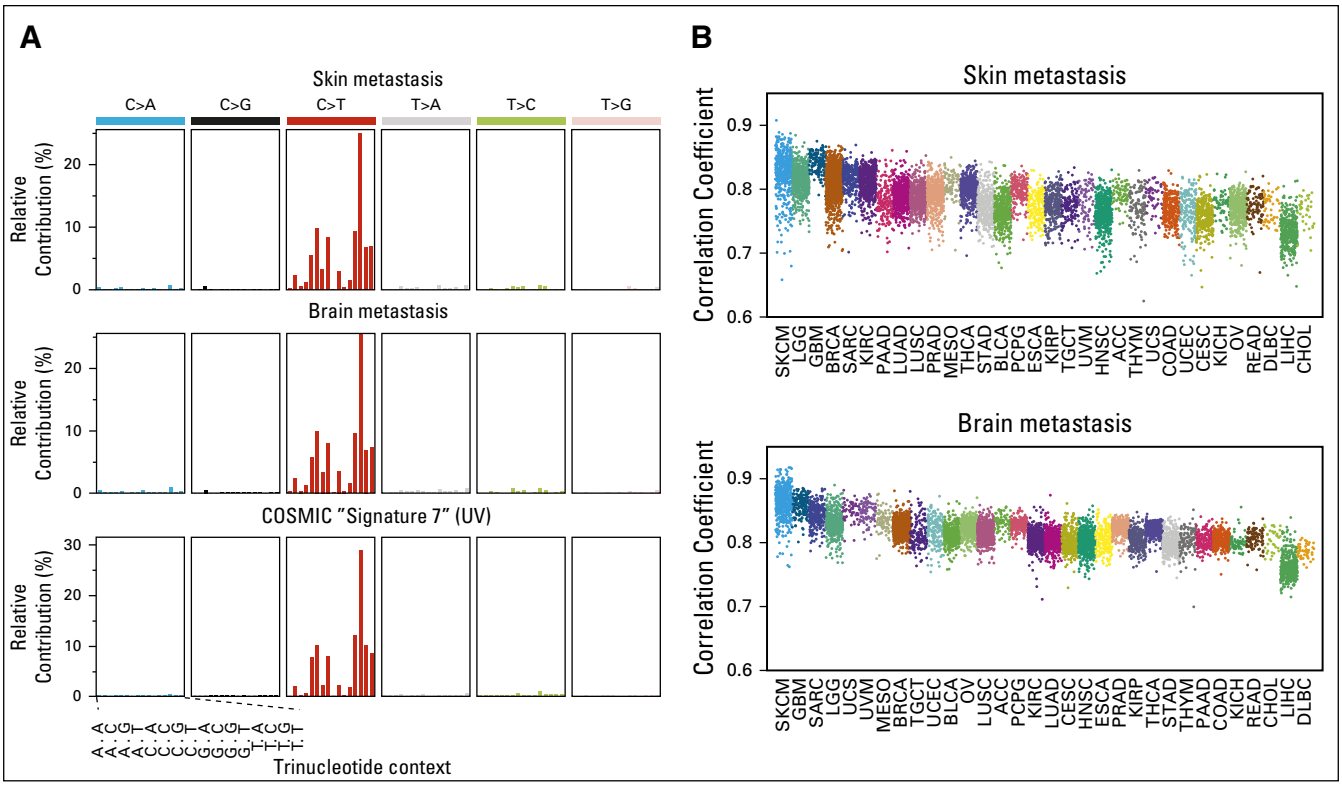


Fig 4. (A) Trinucleotide substitution frequencies observed in exonic regions of whole-genome sequencing data from the skin (top) and brain metastases (middle), as well as the ultraviolet radiation (UV)-associated reference signature from Catalogue of Somatic Mutations in Cancer (bottom). Substitution classes are defined on the basis of the pyrimidine of the altered base pair. (B) Correlations of the transcriptomes of the skin (top) and brain metastases (bottom) with 9,583 tumors from The Cancer Genome Atlas. The boxplots were ordered on the basis of the mean of the top six samples with the largest correlation coefficients within each cohort. ACC, adrenocortical carcinoma; BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL, cholangiocarcinoma; COAD, colon adenocarcinoma; DLBC, lymphoid neoplasm diffuse large B-cell lymphoma; ESCA, esophageal carcinoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; KICH, kidney chromophobe; KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LGG, brain lower-grade glioma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; MESO, mesothelioma; OV, ovarian serous cystadenocarcinoma; PAAD, pancreatic adenocarcinoma; PCPG, pheochromocytoma and paraganglioma; PRAD, prostate adenocarcinoma; READ, rectum adenocarcinoma; SARC, sarcoma; SKCM, skin cutaneous melanoma; STAD, stomach adenocarcinoma; TGCT, testicular germ cell tumor; THCA, thyroid carcinoma; THYM, thymoma; UCS, uterine carcinosarcoma; UCEC, uterine corpus endometrial carcinoma; UVM, uveal melanoma.

uveal melanomas were misclassified according to the organs where metastases were sampled, suggesting that normal tissue contamination could influence transcriptome-based predictions. Notably, however, both the skin and brain metastases of the present case were predicted to be melanoma by this classifier.

Ultraviolet Radiation Signature Identifies Incorrectly Labeled TCGA Lung Tumors

To further study the utility of UV signature analysis in detecting misdiagnosed tumors, we investigated all squamous lung cancer samples in TCGA. Three tumors were detected with > 50% UV signature contribution[23,24] (Figs 5B and 5C). Transcriptome-based analyses predicted one to be head and neck squamous cell carcinoma and the others to be lung squamous carcinomas (Fig 5A; Appendix Fig A6A). Melanoma could be ruled out by a lack of decisive melanoma markers and common driver mutations, as well as by high expression of the squamous/basal marker TP63[25] (Appendix Fig A6B). Although these samples still shared many Catalogue of Somatic Mutations in Cancer mutations with skin melanomas (Appendix Fig A6C), this is likely explained by spuriously shared mutations with highly mutated tumors in general. Some of these could be UV induced and selected for, as suggested for *RAC1* P29S and *PPP6C* R264C mutations.[26] However, given the close transcriptomic relationships between squamous cancers of the lung and head and neck, as well as other tumors with squamous differentiation, and their lower classification accuracies (Fig 5A; Appendix

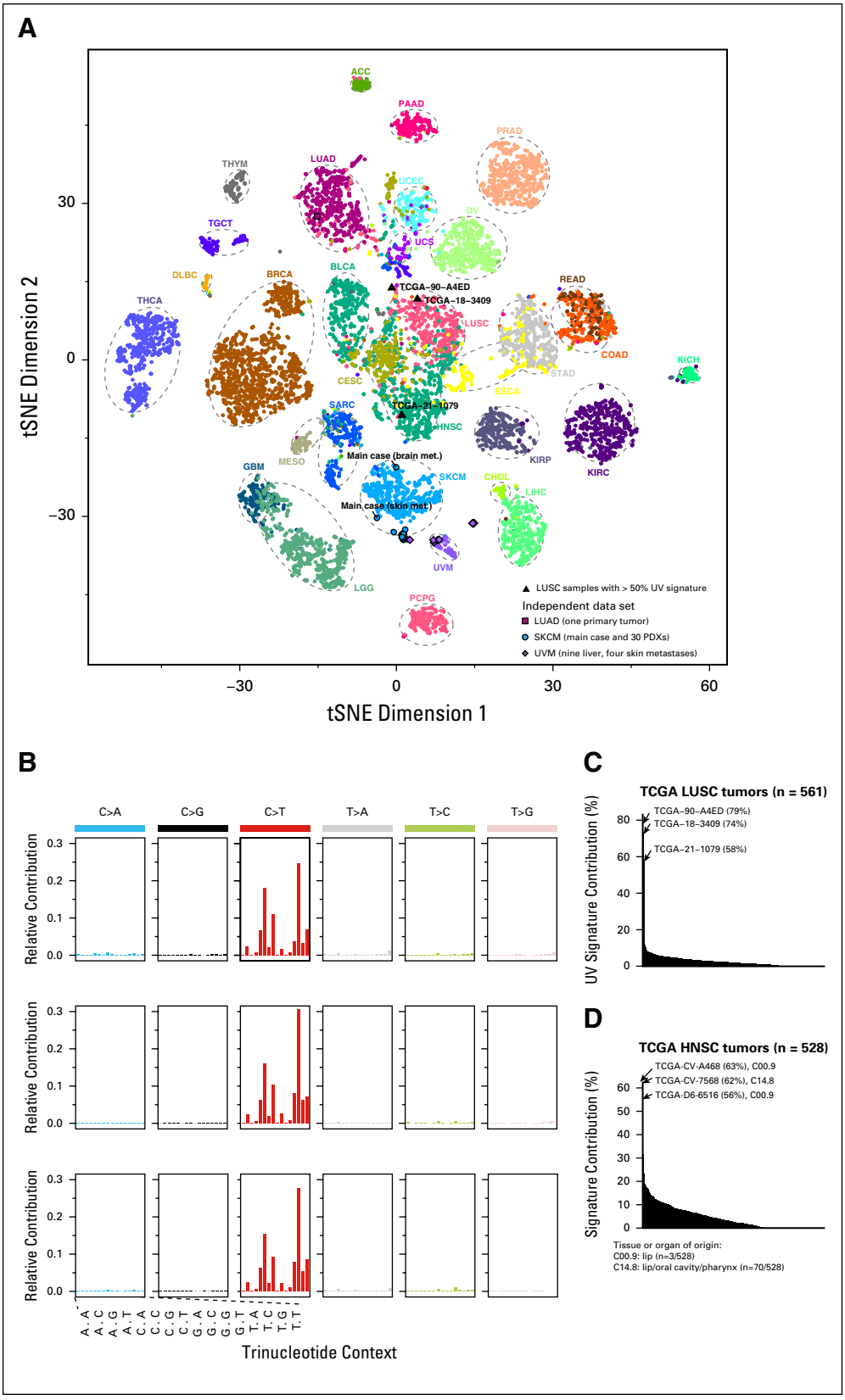


Fig 5. Mutational signature analysis and pan-cancer transcriptome-based comparison of lung squamous cell carcinoma (LUSC) samples from The Cancer Genome Atlas (TCGA). (A) t-distributed stochastic neighbor embedding (tSNE) projection of the pan-cancer data set, together with the main case skin and brain metastasis samples, as well as the independent data set used for additional validation of the transcriptome-based classifier. Three ultraviolet radiation (UV) signature–associated TCGA

Fig A5B), the predictions of lung squamous cell carcinoma for two of the UV-associated tumors are more likely to be a general indication of a squamous phenotype. Although primary lung tumors cannot be sun exposed, a subset of head and neck squamous cell cancers can be. Out of 528 TCGA HNSC tumors, three had over 50% UV signature, all of which were from the lip region (Fig 5D). Notably, skin squamous and basal cell carcinomas are absent from TCGA and therefore cannot be excluded as possible diagnoses for either of the three lung squamous cancers in TCGA with an UV signature. In conclusion, transcriptome-based classification is capable of indicating the cellular origin, in certain cancers with high accuracy, but not always the primary site. The UV signature is a biomarker that can indicate an origin from a sun-exposed site.

DISCUSSION

Here we demonstrate the application of mutational signature analysis and transcriptomic profiling as an adjunct in the diagnosis of a CUP with neuroendocrine differentiation. Although the tumor had a *BRAF* mutation, the morphology was ambiguous and atypical for malignant melanoma.[27] At least three hospitals had misdiagnosed the tumor, because the following features did not support a melanoma diagnosis: the tumor in the thoracic region was believed to be the primary, and the patient had never had a skin lesion removed; classic melanocyte markers were negative, whereas the tumor was positive for several neuroendocrine markers; and S100 and SOX10 were expressed, but these markers may also be expressed in NEN. Therefore, it was only after comprehensive genomic, transcriptomic, and IHC analyses that we obtained sufficient evidence to confidently classify the tumor as a malignant melanoma. In particular, the findings of a characteristic mutational signature associated with UV-induced damage and a transcriptome profile consistent with malignant melanoma strongly support this diagnosis. Indeed, the possibility of a diagnosis of melanoma was confirmed by a specialist in NEN pathology (R.R.d.K.) who was blinded to the genome analyses.

We benchmarked the performance of transcriptome-based cancer type classification according to our method using TCGA data and found high accuracy in general. However, some cancer types were more difficult to distinguish because of similar transcriptome profiles. Another limitation was the absence of some cancer types in TCGA. In addition, normal tissue contamination can influence correlation analyses. Our findings indicate that that the cellular origin, but not always the primary site, could be determined accurately on the basis of transcriptome correlations. Mutational signature analysis can, if UV damage is indicated, help further distinguish between internal and sun-exposed sites. Although beyond the scope of our study to investigate, we expect that other signatures associated with site-specific carcinogen exposure, such as tobacco smoke, could also be found informative when deciphering the primary origin of tumors.

Current knowledge indicates that skin melanocytes may have two developmental origins: directly from the neural crest or via Schwann cell progenitors (SCPs) migrating along nerves in the developing embryo.[28] Interestingly, SCPs also generate the chromaffin cells of the adrenal gland.[29] It is therefore tempting to speculate that coexpression of the neuroendocrine markers TH, chromogranin A, and SYN in the present tumor was the result of a transformation event involving an SCP-generated SOX10-positive/MITF-positive/DCT-positive melanoblast. An interesting future avenue of research would be to investigate how many patients with CUP or

Fig 5. (Continued).
LUSC samples are also highlighted. (B) Overall trinucleotide substitution frequencies for the three TCGA LUSC samples with UV signature: TCGA-21-1079 (top), TCGA-90-A4ED (middle), and TCGA-18-3409 (bottom). (C) Relative estimated UV signature contributions for all samples in the TCGA LUSC cohort. Samples with UV signature contribution > 50% are highlighted. (D) Estimated contributions of the UV signature in samples from TCGA head and neck squamous cell carcinoma (HNSC) cohort, with site of origin noted according to available clinical metadata. ACC, adrenocortical carcinoma; BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL, cholangiocarcinoma; COAD, colon adenocarcinoma; DLBC, lymphoid neoplasm diffuse large B-cell lymphoma; GBM, glioblastoma multiforme; KICH, kidney chromophobe; KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LGG, brain lower-grade glioma; LIHC, liver hepatocellular carcinoma;LUAD, lung adenocarcinoma; MESO, mesothelioma; met, metastasis; OV, ovarian serous cystadenocarcinoma; PAAD, pancreatic adenocarcinoma; PCPG, pheochromocytoma and paraganglioma; PRAD, prostate adenocarcinoma; READ, rectum adenocarcinoma; SARC, sarcoma; SKCM, skin cutaneous melanoma; STAD, stomach adenocarcinoma; TGCT, testicular germ cell tumor; THCA, thyroid carcinoma; THYM, thymoma; UCS, uterine carcinosarcoma; UCEC, uterine corpus endometrial carcinoma; UVM, uveal melanoma.

NEN with an unknown primary actually may have melanoma or other diseases with approved therapies. Given the approval of targeted therapies and immunotherapies for melanoma and lung cancer, a change of diagnosis could have real clinical impact within a short timeframe.

Despite the incorrect diagnosis, we would argue that the treatments our patient received included those used in the clinical management of malignant melanoma, such as dabrafenib and pembrolizumab. One might argue that he should have received the dabrafenib and trametinib combination treatment upfront given the *MAPK1* mutation,[17] but at that time, this mutation was unknown to the clinician, and the diagnosis was PGL, meaning that dabrafenib was already off label. Moreover, although two of three PDX mice displayed stable disease with the combination treatment, none exhibited tumor regression. Possibly, a recently postulated triple combination targeting BRAF, MEK, and ERK would have been beneficial in this case.[30]

In summary, this is to our knowledge the first case study where mutational signature analysis has been used as a decisive adjunct to routine histopathology to ascertain the diagnosis of metastatic deposits from a CUP. Our study also demonstrates that therapy targeting driver mutations can yield clinical response, even though the cancer lacks a definitive diagnosis. We conclude that mutational signature analysis and transcriptome-based comparison with a pan-cancer panel are useful adjuncts in the diagnosis and determination of origin of CUPs. An accurate diagnosis, in turn, can enable new treatment options for patients with these otherwise fatal cancers.

DOI: https://doi.org/10.1200/PO.18.00002
Published online on ascopubs.org/journal/po on July 30, 2018.

AUTHOR CONTRIBUTIONS

Conception and design: Roger Olofsson Bagge, Akif Demir, Joakim Karlsson, Andreas Muth, Johanna B. Svensson, Erik Larsson, Göran Stenman, Jonas A. Nilsson

Financial support: Erik Larsson, Göran Stenman, Jonas A. Nilsson

Provision of study material or patients: Roger Olofsson Bagge, Andreas Muth, Ola Nilsson, Jonas A. Nilsson

Collection and assembly of data: Roger Olofsson Bagge, Akif Demir, Joakim Karlsson, Berglind O. Einarsdottir, Henrik Jespersen, Mattias F. Lindberg, Andreas Muth, Lisa M. Nilsson, Marta Persson, Johanna B. Svensson, Elin M.V. Söderberg, Ola Nilsson, Jonas A. Nilsson

Data analysis and interpretation: Roger Olofsson Bagge, Akif Demir, Joakim Karlsson, Babak Alaei-Mahabadi, Andreas Muth, Johanna B. Svensson, Ronald R. de Krijger, Ola Nilsson, Erik Larsson, Göran Stenman, Jonas A. Nilsson

Manuscript writing: All authors

Final approval of manuscript: All authors

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/po/author-center.

Roger Olofsson Bagge
Consulting or Advisory Role: Amgen
Consulting or Advisory Role: Bristol-Myers Squibb
Speakers' Bureau: Roche

Akif Demir
Consulting or Advisory Role: Cook Medical

Joakim Karlsson
No relationship to disclose

Babak Alaei-Mahabadi
No relationship to disclose

Berglind O. Einarsdottir
No relationship to disclose

Henrik Jespersen
No relationship to disclose

Mattias F. Lindberg
No relationship to disclose

Andreas Muth
No relationship to disclose

Lisa M. Nilsson
Stock and Other Ownership Interests: Exocure Ventures (I), Oxcia (I)

Travel, Accommodations, Expenses: Merck Sharp & Dohme Sweden (I)

Marta Persson
No relationship to disclose

Johanna B. Svensson
No relationship to disclose

Elin M.V. Söderberg
No relationship to disclose

Ronald R. de Krijger
No relationship to disclose

Ola Nilsson
No relationship to disclose

Erik Larsson
Stock and Other Ownership Interests: Kite Pharma, Juno Therapeutics

Göran Stenman
No relationship to disclose

Jonas A. Nilsson
Stock and Other Ownership Interests: Exocure Ventures, Oxcia

Travel, Accommodations, Expenses: Merck Sharp & Dohme Sweden

ACKNOWLEDGMENT

This research would not have been possible without the co-operation of the patient and his close relatives, to whom we extend our sincere gratitude. We also thank Sofia Stenqvist and Gülay Altiparmak for animal care and histology, respectively. The results published here are in part based upon data generated by the TCGA Research Network (cancergenome.nih.gov/). Sequence data generated for this study has been deposited at the European Genome-Phenome Archive, which is hosted by the EBI and the CRG under accession number EGAS00001003026. Additional information about EGA can be found at ega-archive.org.

Affiliations

Roger Olofsson Bagge, **Joakim Karlsson**, **Babak Alaei-Mahabadi**, **Berglind O. Einarsdottir**, **Mattias F. Lindberg**, **Andreas Muth**, **Lisa M. Nilsson**, **Johanna B. Svensson**, **Elin M.V. Söderberg**, **Erik Larsson**, and **Jonas A. Nilsson**, Sahlgrenska Academy, University of Gothenburg; **Roger Olofsson Bagge**, **Berglind O. Einarsdottir**, **Henrik Jespersen**, **Mattias F. Lindberg**, **Lisa M. Nilsson**, **Marta Persson**, **Elin M.V. Söderberg**, **Ola Nilsson**, **Göran Stenman**, and **Jonas A. Nilsson**, Sahlgrenska Cancer Center, University of Gothenburg; **Roger Olofsson Bagge**, **Akif Demir**, **Andreas Muth**, **Johanna B. Svensson**, Sahlgrenska University Hospital, Gothenburg, Sweden; and **Ronald R. de Krijger**, Reinier de Graaf Hospital, Delft; and University Medical Center Utrecht/Princess Maxima Center for Pediatric Oncology, Utrecht, the Netherlands.

Support

Supported by the Swedish Cancer Society; Region Västra Götaland Avtal om Läkarutbildning och Forskning grants; the Knut and Alice Wallenberg Foundation; the Erling-Persson Foundation; Lion's Cancer Foundation West; the Sjöberg Foundation; and BioCARE, a National Strategic Research Program at the University of Gothenburg.

REFERENCES

1. Frampton GM, Fichtenholtz A, Otto GA, et al: Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. Nat Biotechnol 31:1023-1031, 2013
2. Wagle N, Berger MF, Davis MJ, et al: High-throughput detection of actionable genomic alterations in clinical tumor samples by targeted, massively parallel sequencing. Cancer Discov 2:82-93, 2012
3. Stadler ZK, Battaglin F, Middha S, et al: Reliable detection of mismatch repair deficiency in colorectal cancers using mutational load in next-generation sequencing panels. J Clin Oncol 34:2141-2147, 2016
4. Le DT, Uram JN, Wang H, et al: PD-1 blockade in tumors with mismatch-repair deficiency. N Engl J Med 372:2509-2520, 2015
5. Sholl LM, Do K, Shivdasani P, et al: Institutional implementation of clinical tumor profiling on an unselected cancer population. JCI Insight 1:e87062, 2016
6. Zehir A, Benayed R, Shah RH, et al: Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat Med 23:703-713, 2017
7. Golfinopoulos V, Pentheroudakis G, Salanti G, et al: Comparative survival with diverse chemotherapy regimens for cancer of unknown primary site: Multiple-treatments meta-analysis. Cancer Treat Rev 35:570-573, 2009
8. Ross JS, Wang K, Gay L, et al: Comprehensive genomic profiling of carcinoma of unknown primary site: New routes to targeted therapies. JAMA Oncol 1:40-49, 2015
9. Tothill RW, Li J, Mileshkin L, et al: Massively-parallel sequencing assists the diagnosis and guided treatment of cancers of unknown primary. J Pathol 231:413-423, 2013
10. Kopetz S, Desai J, Chan E, et al: Phase II pilot study of vemurafenib in patients with metastatic BRAF-mutated colorectal cancer. J Clin Oncol 33:4032-4038, 2015

11. Einarsdottir BO, Bagge RO, Bhadury J, et al: Melanoma patient-derived xenografts accurately model the disease and develop fast enough to guide treatment decisions. Oncotarget 5:9609-9618, 2014
12. Dasari A, Shen C, Halperin D, et al: Trends in the incidence, prevalence, and survival outcomes in patients with neuroendocrine tumors in the United States. JAMA Oncol 3:1335-1342, 2017
13. Muth A, Abel F, Jansson S, et al: Prevalence of germline mutations in patients with pheochromocytoma or abdominal paraganglioma and sporadic presentation: A population-based study in western Sweden. World J Surg 36:1389-1394, 2012
14. Wilzén A, Rehammar A, Muth A, et al: Malignant pheochromocytomas/paragangliomas harbor mutations in transport and cell adhesion genes. Int J Cancer 138:2201-2211, 2016
15. Fishbein L, Leshchiner I, Walter V, et al: Comprehensive molecular characterization of pheochromocytoma and paraganglioma. Cancer Cell 31:181-193, 2017
16. Castro-Vega LJ, Letouzé E, Burnichon N, et al: Multi-omics analysis defines core genomic alterations in pheochromocytomas and paragangliomas. Nat Commun 6:6044, 2015
17. Nikolaev SI, Rimoldi D, Iseli C, et al: Exome sequencing identifies recurrent somatic MAP2K1 and MAP2K2 mutations in melanoma. Nat Genet 44:133-139, 2011
18. Romano RC, Carter JM, Folpe AL: Aberrant intermediate filament and synaptophysin expression is a frequent event in malignant melanoma: An immunohistochemical study of 73 cases. Mod Pathol 28:1033-1042, 2015
19. Alexandrov LB, Nik-Zainal S, Wedge DC, et al: Signatures of mutational processes in human cancer. Nature 500:415-421, 2013
20. Nik-Zainal S, Alexandrov LB, Wedge DC, et al: Mutational processes molding the genomes of 21 breast cancers. Cell 149:979-993, 2012
21. Blokzijl F, Janssen R, van Boxtel R, et al: MutationalPatterns: Comprehensive genome-wide analysis of mutational processes. Genome Med 10:33, 2018
22. Hoadley KA, Yau C, Wolf DM, et al: Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. Cell 158:929-944, 2014
23. Campbell BB, Light N, Fabrizio D, et al: Comprehensive analysis of hypermutation in human cancer. Cell 171:1042-1056.e10, 2017
24. Campbell JD, Alexandrov A, Kim J, et al: Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. Nat Genet 48:607-616, 2016
25. Di Como CJ, Urist MJ, Babayan I, et al: p63 expression profiles in human normal and tumor tissues. Clin Cancer Res 8:494-501, 2002
26. Hodis E, Watson IR, Kryukov GV, et al: A landscape of driver mutations in melanoma. Cell 150:251-263, 2012
27. Banerjee SS, Harris M: Morphological and immunophenotypic variations in malignant melanoma. Histopathology 36:387-402, 2000
28. Adameyko I, Lallemend F, Aquino JB, et al: Schwann cell precursors from nerve innervation are a cellular origin of melanocytes in skin. Cell 139:366-379, 2009
29. Furlan A, Dyachuk V, Kastriti ME, et al: Multipotent peripheral glial cells generate neuroendocrine cells of the adrenal medulla. Science 357:eaal3753, 2017
30. Xue Y, Martelotto L, Baslan T, et al: An approach to suppress the evolution of resistance in BRAFV600E-mutant cancer. Nat Med 23:929-937, 2017

Appendix

Patient

The patient received oral and written information and signed the informed consent according to the ethical approval at the regional ethical review board (#289-12). The patient and his relatives have approved the publication of this case report.

Immunohistochemical Staining of Tumor Tissues

Biopsies from patient tumor and patient-derived xenograft (PDX) tumor were fixed in neutral buffered formalin and embedded in paraffin. Sections (3 to 5 microns) were placed on positively charged glass slides and subjected to antigen retrieval using EnVision FLEX Target Retrieval Solution (high power of hydrogen; Agilent Technologies, Santa Clara, CA). Staining was performed in a Dako Autostainer Link using EnVision FLEX with EnVision FLEX+ (LINKER; DakoCytomation, Glostrup, Denmark). Positive and negative controls were included in each run. The following primary antibodies were used: anti-ALK (clone ALK1; Dako), anti–beta-catenin (BD610154; Becton Dickinson, Franklin Lakes, NJ), anti-CD56 (clone 123C3; Dako), anti-CD99 (clone EPR3097Y; Epitomics, Burlingame, CA), anti–chromogranin A (clone LK2H10; Millipore, Burlington, MA), anti-CDX2 (ab76541; Abcam, Cambridge, United Kingdom), anti–cytokeratin 8/18 (NCL-5D3; Leica, Wetzlar, Germany), anti–cytokeratin 20 (clone Ks20.8; Dako), anti–high molecular weight cytokeratin (clone AE1/3; Dako), anti-HMB45 (M0634; Dako), anti-Ki67 (clone MIB-1; Dako), anti–melan A (clone A103; Dako), anti-NB84 (NCL-NB84; Leica), anti–Neu-N (MAB377; Millipore), anti-NFP (M0762; Dako), anti-PAX8 (10336-1-AP; Proteintech, Chicago, IL), anti-S100 (catalog no IR504; Dako), anti-SOX10 (clone EP268; Epitomics), antisynaptophysin (clone DAK-SYNAP; Dako), anti-TTF1 (clone SPT24; Novocastra, Wetzlar, Germany), and anti–tyrosine hydroxylase (clone 1B5; Novocastra).

Extraction of DNA and RNA From Tumor

Snap-frozen tumor pieces were homogenized using Bullet Blender (Next Advance, Troy, NY). DNA and RNA were extracted using the AllPrep DNA/RNA kit (Qiagen, Hilden, Germany).

DNA Sequencing–Based Variant Calling

Raw DNA sequencing reads were aligned to the hg19 reference genome with bwa (version 0.7.12; options "mem" and "-M"; Li H, et al: Bioinformatics 25:1754-1760, 2009). Duplicates were marked with Picard (version 1.109; https://broadinstitute.github.io/picard). The resulting BAM files were recalibrated and indels were jointly realigned for the tumor and normal samples with GATK (version 3.3.0; McKenna A, et al: Genome Res 20:1297-1303, 2010). Variant calling was performed with GATK HaplotypeCaller on the tumor and normal sample separately and with Mutect (version 1.1.5; Cibulskis K, et al: Nat Biotechnol 31:213-219, 2013) on both samples paired.

RNA Sequencing–Based Variant Calling

Raw RNA sequencing reads were aligned to the hg19 reference genome with STAR (verison 2.5.2b; Dobin A, et al: Bioinformatics 29:15-21, 2013) in two passes, with the parameter "–sjdbOverhang 125." Read groups were added and duplicates marked with Picard "AddOrReplaceReadGroups" and "MarkDuplicates," respectively. Reads were then split into exon segments with GATK "SplitNCigarReads," and base recalibration was performed using GATK "BaseRecalibrator," for which the reference variant lists "dbsnp_132_b37.leftAligned.vcf," "1000G_phase1.indels.b37.vcf," and "Mills_and_1000G_gold_standard.indels.b37.nochr.vcf" (obtained from the GATK resource bundle; https://software.broadinstitute.org/gatk/download/bundle) were used. Variant calling was subsequently performed using GATK HaplotypeCaller with the options "-dontUseSoftClippedBases," "-stand_call_conf 20.0," and "-stand_emit_conf 20.0."

Variant Filtering

To obtain a sensitive set of whole-genome sequencing variant calls that included indels, the union of the GATK tumor and Mutect somatic calls was first constructed. This set was then further compared with the RNA sequencing calls by determining the intersection and complement sets, using the bcftools "isec" command. From the above sets, variants present in the GATK calls on the normal sample were then removed. Before the set operations, the variant vcf files were left normalized using bcftools (two steps: options "norm *-m*-both" in the first, and "norm –f human_g1k_v37.fasta" in the second; the reference fasta file was obtained from the GATK resource bundle). Subsequently, the sets were annotated with ANNOVAR (Wang K, et al: Nucleic Acids Res 38:e164, 2010) to identify and discard all variants present with ≥ 1% frequency in either the 1000 Genomes database (version August 2015; "1000g2015aug_all") or ESP6500 ("esp6500siv2_all") or listed in a version of dbSNP 138, from which flagged somatic and clinically associated variants had been removed ("snp138NonFlagged"). Synonymous variants were additionally discarded.

Mutational Signature Analysis

Variants detected by whole-genome sequencing, as described under DNA Sequencing–Based Variant Calling, were annotated using ANNOVAR (as described), after which all variants that were either nonexonic or present in the snp138NonFlagged, esp6500siv2_all, or 1000g2015aug_all databases were removed. Samples from the skin and brain metastases were processed identically. To determine the mutation spectrum of these variants, the VCF files were converted into a 96-trinucleotide mutation frequency matrix using the function "mut_matrix" (parameter: ref_genome = "Bsgenome.Hsapiens.UCSC.hg38", excluding the sex chromosomes) from the R package MutationalPatterns.[21] A set of known mutational signature trinucleotide frequencies, obtained from the Catalogue of Somatic Mutations in Cancer (http://cancer.sanger.ac.uk/cancergenome/assets/signatures_probabilities.txt (accessed October 27, 2017), was then fit to the observed mutations using the function "fit_to_signatures" from the same R package. The principle behind the fitting algorithm is to find the nonnegative linear combination of the predefined mutational signatures that best explains all mutations in a given sample, which is done by solving a nonnegative least squares optimization problem.[21] This allowed the estimation of the relative contributions of the known mutational signatures in a given sample. The function "plot_96_profile" was used to visualize the overall trinucleotide frequencies in a given sample. Strand bias was tested using the "strand_bias_test" function of the same package (Poisson test). *P* values were adjusted using the Benjamini-Hochberg method.

For mutational signature analysis of The Cancer Genoma Atlas (TCGA) tumors, annotated somatic mutation calls (MuTect2) in VCF format on the basis if whole-exome sequencing data were downloaded for 561 lung squamous cell carcinoma (LUSC) and 528 head and neck squamous cell carcinoma tumors from the Genomic Data Commons portal (https://portal.gdc.cancer.gov/; accessed on November 24, 2017, and February 26, 2018, respectively). From these, all variants not marked as passing the quality control filters were discarded. The filtered mutation calls of all samples were then analyzed with respect to the relative contributions of known mutational signatures as described. Clinical metadata that included the site of origin for the head and neck squamous cell carcinoma tumors were also downloaded from the Genomic Data Commons portal (accessed on February 26, 2018).

RNA Sequencing–Based Fusion Gene Detection

Expressed fusion genes were detected using FusionCatcher (version 0.99.3a beta; Nicorici D, et al: bioRxiv doi:https://doi.org/10.1101/011650) against the Ensembl (version 75) reference annotation, on the basis of the RNA sequencing fastq files of the tumor sample.

Transcriptome Comparison With Tumors From TCGA

RNA sequencing data for 9,583 tumors from 32 cancer types were downloaded from the cgHub repository on December 18, 2015, and aligned to the hg19 human genome assembly, excluding alternative haplotype regions, with hisat 0.1.6-beta (parameters: "–no-mixed–no-discordant–no-unal–known-splicesite-infile"; Kim D, et al: Nat Methods 12:357-360, 2015), using splice junctions defined in the GENCODE (version 19) reference human genome annotation. Gene read counts were derived with htseq-count (parameters: "-m intersection-strict -s no"; Anders S, et al: Bioinformatics 31:166-169, 2015). Reads per kilobase per million normalized values were calculated, taking into account the max mature transcript length of each gene and using robust size factors as previously described for the DESeq method (Anders S, et al: Genome Biol 11:R106, 2010). For the correlation analysis, reads from our own samples were realigned and read counts requantified and normalized using the same methods described for TCGA data. However, standard read depth–based size factors were used for the reads per kilobase million normalization of these samples. Pairwise Spearman correlation coefficients were then calculated between our samples and each TCGA sample, with respect to all coding genes (using the function "corr" in MATLAB R2017b).

For t-distributed stochastic neighbor embedding analysis of all samples, $\log_2$-transformed (pseudocount of 1 added) expression values of all coding genes were used, together with the "Rtsne" function from the "Rtsne" R package (default parameters; van der Maaten LJP, et al: J Mach Learn Res 9:2579-2605, 2008).

Implementation and Cross-Validation of Transcriptome-Based Cancer Type Classification

To classify a sample using the *k*-nearest neighbor approach, Spearman correlation coefficients were first calculated between the sample of interest and all samples in the TCGA pan-cancer data set. A prediction was then made using the majority vote of the top *k* most strongly correlated samples. In case of ties, the value of *k* was decreased by 1 until a majority vote was achieved. Leave-one-out cross-validation was performed by in turn excluding each sample from the TCGA data set and classifying it using correlations to the remaining samples. Values of *k* from one to 50 were evaluated, and the *k* that gave the highest accuracy was chosen for the final classifier. Overall accuracy statistics were calculated using the "confusionMatrix" function of the "caret" R package. The final classifier was then subjected to additional validation on an independent data set composed of 30 skin melanoma PDXs, 13 uveal melanoma metastases (four skin and nine liver), and one lung adenocarcinoma primary tumor sample, all preprocessed and normalized as described for the main case samples. Classification of the ultraviolet radiation (UV)-associated TCGA LUSC tumors and our main case samples were carried out with the same classifier ($k = 6$).

Analysis of Driver Gene Mutations in TCGA Samples

To investigate how driver gene mutations found in the UV-associated LUSC samples were represented across other TCGA cancer types, MAF files of annotated somatic variants were downloaded for all 32 cancer types, using the "GDCquery_Maf" function from the "TCGAbiolinks" R package (parameters: pipelines="mutect2"; Colaprico A, et al: Nucleic Acids Res 44:e71, 2016). Only mutations in Cancer Gene Census genes listed in the Catalogue of Somatic Mutations in Cancer and annotated as "Missense_Mutation," "Nonsense_Mutation," "Nonstop_Mutation," "In_Frame_Ins," "In_Frame_Del," "Frame_Shift_Ins," or "Frame_Shift_Del" and found in any of the three LUSC samples were included. The Cancer Gene Census list was downloaded from http://cancer.sanger.ac.uk/cosmic (accessed on February 19, 2018). The number of times each mutation occurred in a given cohort was divided by the total number of samples in each cohort to adjust for differences in cohort size (the three UV-associated samples were removed from the LUSC cohort for these calculations).

DNA Sequencing–Based Copy Number Variant Calling

Copy number variant segments were determined using binocular (https://sourceforge.net/projects/binocular), with parameters "–delta=90," "–min-copy-ratio=1.1," "–min-maf-delta=0.05," and "–ai-cutoff=0.001" using the vcf file from Mutect as input together with the whole-genome sequencing BAM files of the tumor and normal samples.

Structural Variant Calling

Somatic structural variants were mapped using Meerkat (Yang L, et al: Cell 153:919-929, 2013). Only variants with at least three discordant read pairs and one read spanning the breakpoint junction were considered. The "somatic call" function with the "-n 1 -D 5 -Q 10 -u 1 -f 1 -e 1 -D 5 -z 1" parameters was used to identify somatic events. For additional removal of germ line variants, we first used the "somatic call" function with the "-F" parameter where the "discord" files were generated from seven normal genomes (unpublished data) and then removed events in which both breakpoints overlapped with known population variants the Database of Genomic Variants database (MacDonald JR, et al: Nucleic Acids Res 42:D986-D992, 2014).

Array-Based Comparative Genomic Hybridization

Genomic DNA was extracted from fresh-frozen tumor tissues of the subcutaneous and brain metastases as previously described (Persson F, et al: Oncogene 27:3072-3080, 2008). Array comparative genomic hybridization (CGH) analysis was performed using the Human Genome CGH Microarray 244K oligonucleotide arrays (G4411B; Agilent Technologies) as recommended by the manufacturer. Slides were scanned using an Agilent High-Resolution C Microarray Scanner, followed by data extraction and normalization using Feature Extraction (version 12.0.1.1; Agilent Technologies) with linear normalization (protocol CGH_1200_Jun14). Data analysis was carried out using NEXUS Copy Number (version 8.0) Discovery Edition (BioDiscovery, El Segundo, CA). The significance threshold for segmentation was $P < 1.0^{-6}$. The settings for aberration calls were 1.0 for amplification, 0.2 for gain, –0.2 for loss, and –1.0 for homozygous deletion. Each aberration was checked manually to confirm the accuracy of the call. Sex chromosomes were excluded from the analysis, as were regions partially or completely covered by a previously reported copy number variation (Database of Genomic Variants; http://dgvbeta.tcag.ca/dgv/app/news?ref=NCBI36/hg18; Iafrate AJ, et al: Nat Genet 36:949-951, 2004).

Generation of a PDX Model

A 3 × 3 × 3 mm piece of a biopsy of a subcutaneous metastasis of the patient was transplanted into the flank of immunocompromised, nonobese severe combined immune-deficient interleukin-2 chain receptor γ knockout mice (NOG mice; Taconic, Copenhagen, Denmark) to form xenografts. When the first PDX tumor reached 150 mm^3, it was passaged to new mice (passage 2), which subsequently were passaged to 10 mice, where tumor finally grew in eight mice (passage 3). These mice were treated 5 days per week for a minimum of 3 weeks with either vehicle or trametinib 0.3 mg/kg twice per day (Selleck Chem, Houston, TX). The mouse chow was either standard or supplied with dabrafenib 210 mg/kg (Tafinlar [Novartis, Basel, Switzerland] from the pharmacy). Once per week, tumor sizes were measured with a caliper. All animal experiments were performed in accordance with EU directive 2010/63 (Regional Animal Ethics Committee of Gothenburg approval #36-2014).

RNA Sequencing Analysis of Drug-Treated PDX Models

RNA sequencing reads were aligned to a combined human (hg19) and mouse (mm10) reference genome using STAR, with default settings. Reads aligning to human chromosomes were extracted. From these, reads that also mapped to mouse chromosomes and that did not have primary alignments to human chromosomes were removed, using the Picard Tools (version 1.109; http://broadinstitute.github.io/picard/) module FilterSamReads. Read counts were then binned to genes using the HTSeq (version 0.6.0; Zhao M, et al: Nucleic Acids Res 44:D1023-D1031, 2016) command htseq-count (parameters "-s reverse" and "-m intersection-strict"), as well as the human reference genome annotation GENCODE (version 17). Differential expression analysis was carried out using DESeq2 (Cancer Genome Atlas Network: Nature 517:576-582, 2015), with the parameter "alpha=0.05," considering vehicle-treated and untreated mice both as controls relative to the dabrafenib- and trametinib-treated samples. Enrichment analyses were carried out using MsidDB (Love MI, et al: Genome Biol 15:550, 2014), using Gene Ontology "biological process" gene sets (Anders S, et al: Bioinformatics 31:166-169, 2015), considering genes with an absolute $\log_2$ fold change > 2 and adjusted P values < .01. Terms with $q < 0.05$ were considered significantly enriched.

Table A1. Summary of Copy Number Alterations Detected by Array CGH in the Subcutaneous and Brain Metastases

Chromosome Region	Cytoband Location	Length (bp)	Event	No. of Genes
Subcutaneous metastasis				
chr1:0-120 474 160	pter-p12	120,474,160	Loss	1,436
chr1:143 682 903-247 138 725	q21.1-qter	103,455,823	Gain	1,247
chr2:0-242 717 069	pter-qter	242,717,069	Gain	1,718
chr7:0-158 620 978	pter-qter	158,620,978	Gain	1,279
chr8:0-146 274 826	pter-qter	146,274,826	Gain	997
chr9:153 131-13 016 679	p24.3-p23	12,863,549	Gain	53
chr9:13 016 679-37 375 731	p23-p13.2	24,359,052	Loss	188
chr9:19 407 697-23 668 840	p22.1-p21.3	4,261,144	Homozygous loss of *CDKN2A*	37
chr9:37 376 563-140 273 252	p13.2-qter	71,611,300	Gain	759
chr10:115 544-135 374 737	pter-qter	135,374,737	Loss	1,057
chr13:18 123 649-114 142 980	q11-qter	96,019,332	Gain	614
chr15:21 195 208-100 233 406	q11.2-qter	79,038,199	Gain	938
chr16:0-88 827 254	pter-qter	88,827,254	Gain	1,155
chr18:0-76 117 153	pter-qter	76,117,153	Loss	401
chr20:0-62 435 964	pter-qter	62,435,964	Gain	751
chr21:14 202 423-46 944 323	q11.2-qter	32,741,901	Gain	326
Brain metastasis				
chr1:0-120 474 160	pter-p12	120,474,160	Loss	1,436
chr1:143 682 903-247 138 725	q21.1-qter	103,455,823	Gain	1,247
chr2:0-242 717 069	pter-qter	242,717,069	Gain	1,718
chr3:0-199 379 625	pter-qter	199,379,625	Gain	1,452
chr7:0-158 620 978	pter-qter	158,620,978	Gain	1,279
chr8:0-146 274 826	pter-qter	146,274,826	Gain	997
chr9:153 131-13 016 679	p24.3-p23	12,863,549	Gain	53
chr9:13 016 679-37 355 243	p23-p13.2	24,338,564	Loss	188
chr9:19 406 154-23 669 793	p22.1-p21.3	4,263,639	Homozygous loss of *CDKN2A*	37
chr9:37 355 243-140 273 252	p13.2-qter	102,918,009	Gain	845
chr10:115 544-135 374 737	pter-qter	135,374,737	Loss	1,057
chr13:18 123 649-114 142 980	q11-qter	96,019,332	Gain	614
chr15:21 195 208-100 233 406	q11.2-qter	79,038,199	Gain	938
chr16:0-88 827 254	pter-qter	88,827,254	Gain	1,155
chr18:0-76 117 153	pter-qter	76,117,153	Loss	401
chr20:0-62 435 964	pter-qter	62,435,964	Gain	751
chr21:14 202 423-46 944 323	q11.2-qter	32,741,901	Gain	326

NOTE. Performed using Human Genome CGH Microarray 244K oligonucleotide arrays (Agilent Technologies, Santa Clara, CA).
Abbreviation: CGH, comparative genomic hybridization.

Table A2. Summary of IHC Analyses Performed on Tumor Biopsies

Antibody	Mediastinal Tumor	Skin Tumor	Brain Tumor
HMW cytokeratin	ND	Negative	Negative
Cytokeratin 8/18	ND	Negative	Negative
Cytokeratin 20	ND	Negative	Negative
CDX2	ND	Negative	Negative
TTF1	Negative	Focally positive	Negative
PAX8	ND	Focally positive	Negative
Beta-catenin	ND	Positive*	Focally positive*
Chromogranin A	ND	Positive	Focally positive
Synaptophysin	Focally positive	Positive	Focally positive
Tyrosine hydroxylase	ND	Positive	Focally positive
NB84	ND	Negative	Negative
CD99	ND	Focally positive†	Focally positive†
S100	ND	Positive	Positive
Melan A	ND	Negative	Negative
HMB45	ND	Negative	Negative
SOX10	ND	Focally positive	Positive
Neu-N	ND	Focally positive	Negative
NFP	ND	Positive	Focally positive
CD56	Focally positive	ND	ND

Abbreviations: IHC, immunohistochemical; ND, not determined.
*Membranous and cytoplasmic labeling; no nuclear labeling.
†Difficult to determine because of high background staining.

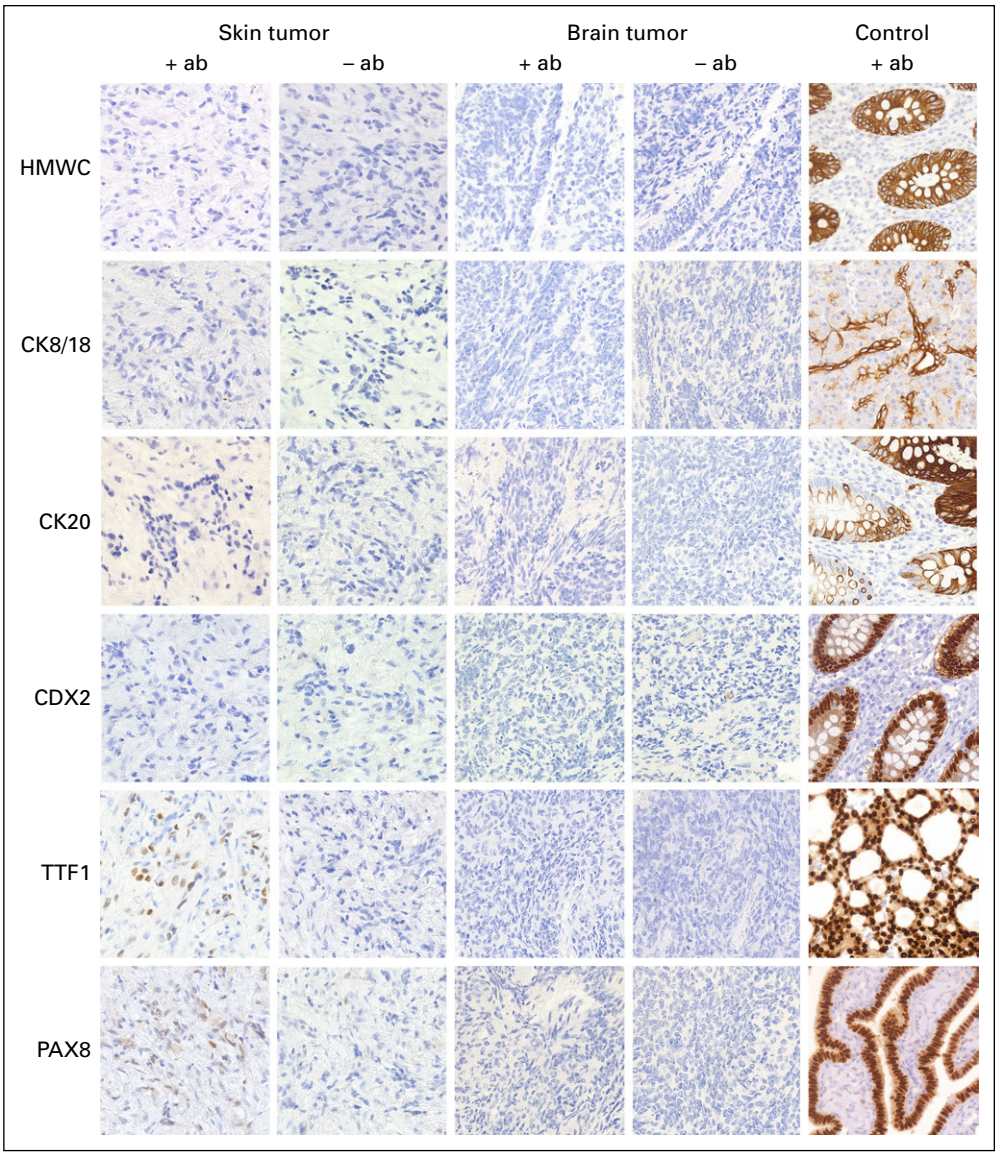


Fig A1. Morphologic and immunohistochemical characteristics of the cytology specimen from the subcutaneous metastatic lesion and the brain metastasis. Sections were stained with hematoxylin and eosin or with antibodies directed against the indicated antigens.

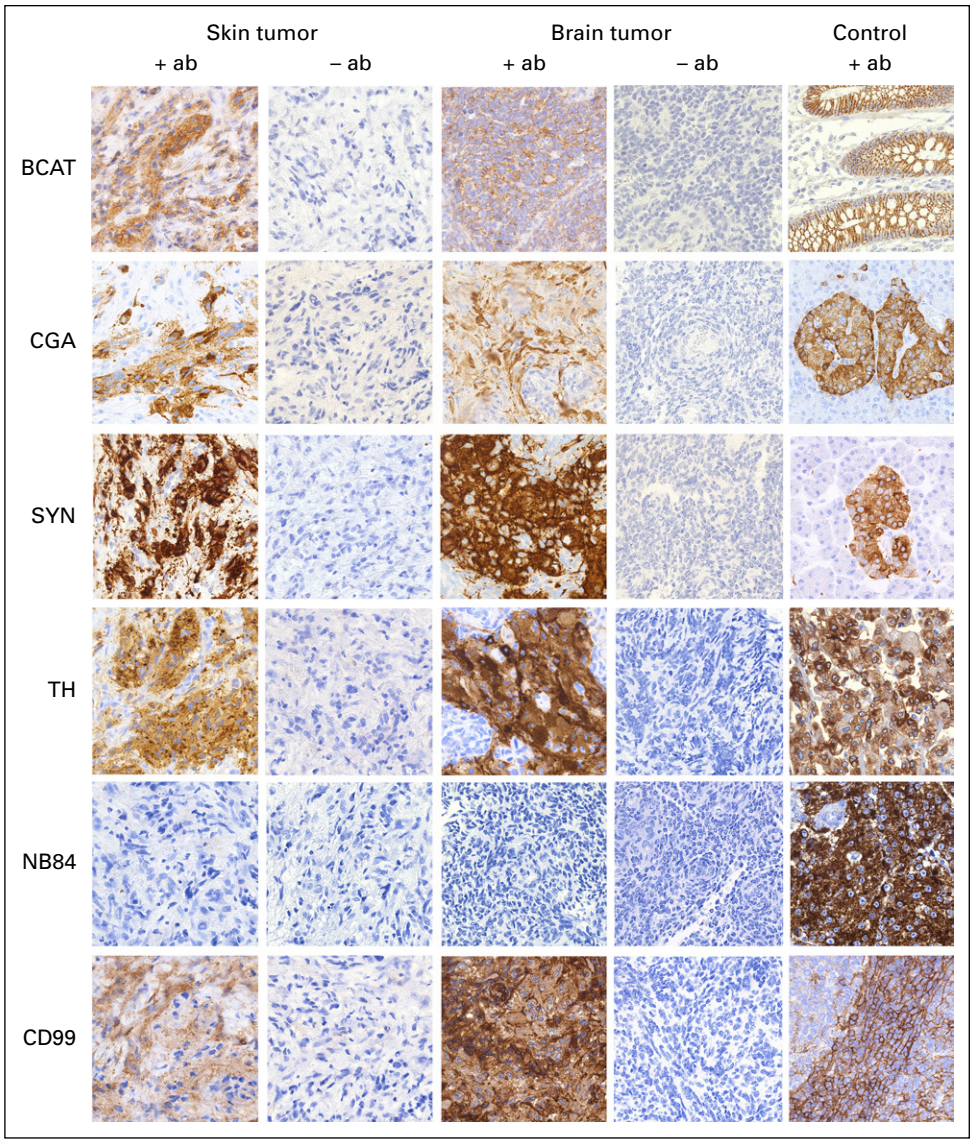
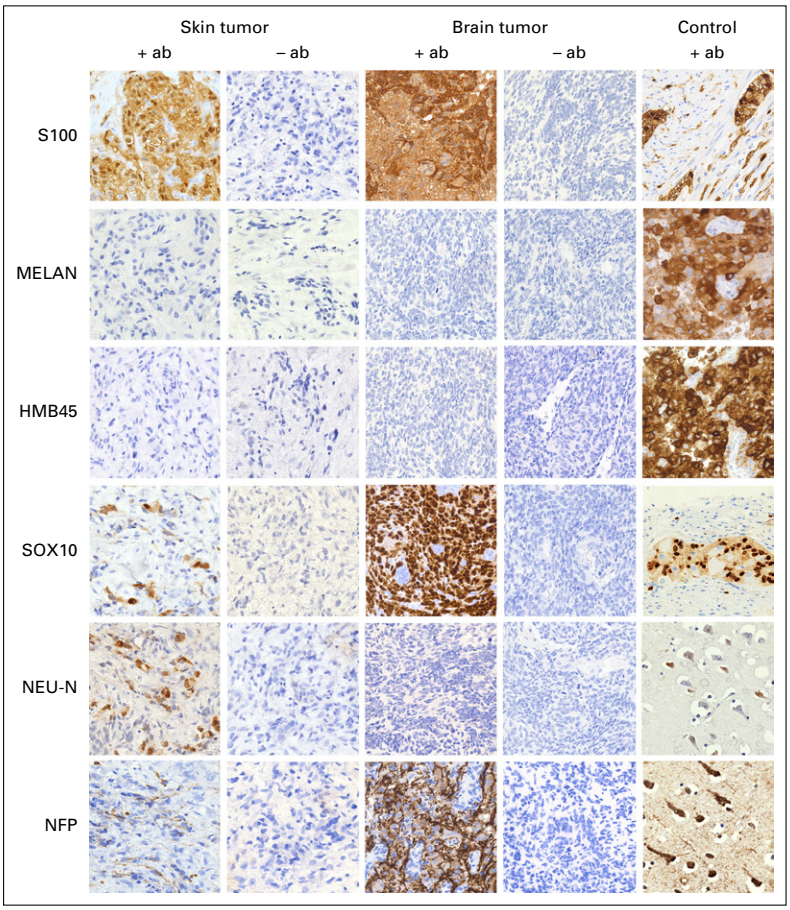


Fig A2. Morphologic and immunohistochemical characteristics of the cytology specimen from the subcutaneous metastatic lesion and the brain metastasis. Sections were stained with hematoxylin and eosin or with antibodies directed against the indicated antigens. SYN, synaptophysin; TH, tyrosine hydroxylase.

Fig A3. Morphologic and immunohistochemical characteristics of the cytology specimen from the subcutaneous metastatic lesion and the brain metastasis. Sections were stained with hematoxylin and eosin or with antibodies directed against the indicated antigens.



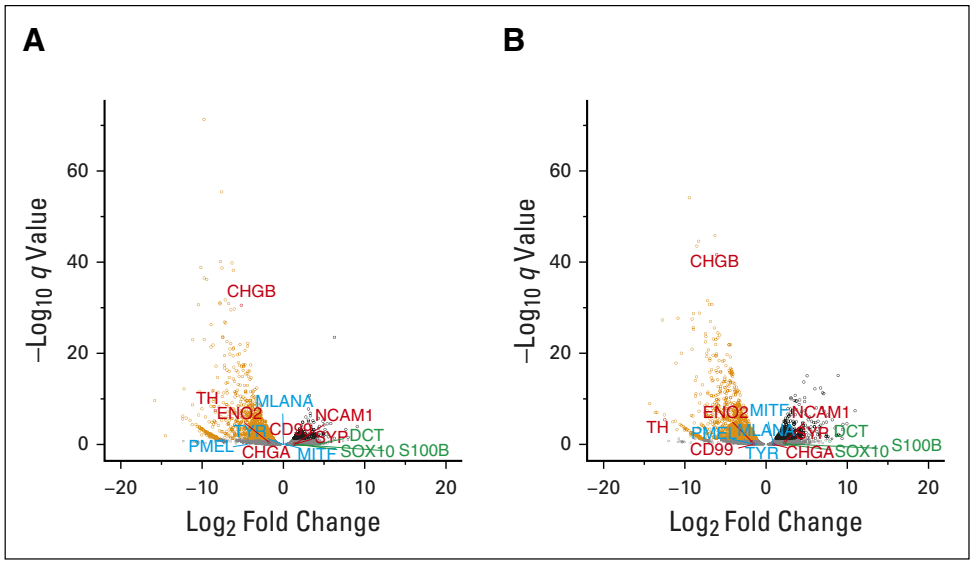


Fig A4. RNA sequencing analysis of dabrafenib- and trametinib-treated patient-derived xenograft models. Differential expression analysis of (A) dabrafenib-treated mice compared with vehicle and untreated controls and (B) dabrafenib and trametinib combination–treated mice compared with vehicle and untreated controls. Adjusted *P* values < .05 were considered statistically significant. Gray, nonsignificant genes; black, upregulated genes; orange, downregulated genes. Highlighted are selected marker genes of melanoblast (green), neuroendocrine (red), and melanocyte (blue) phenotypes. TH, tyrosine hydroxylase.

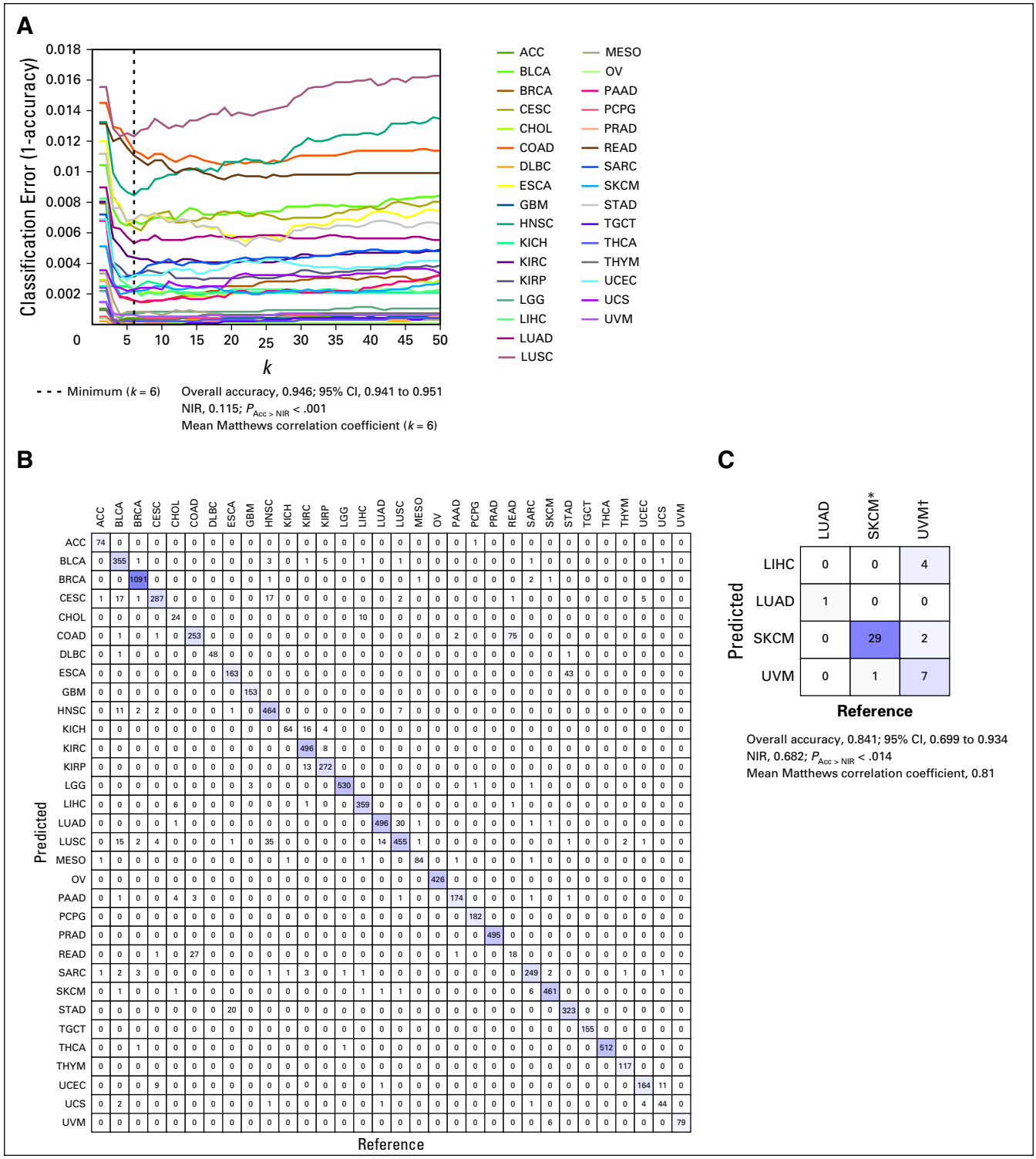


Fig A5. Benchmarking of pan-cancer transcriptome-based cancer type classification. (A) Classification error (defined as 1-accuracy) per cancer type obtained during leave-one-out cross-validation of the *k*-nearest neighbor classification approach on The Cancer Genome Atlas data set, with Spearman correlation coefficients calculated using all coding genes as the measure of similarity between samples. (B) Confusion matrix associated with the cross-validation. (C) Performance with $k = 6$ on an independent validation set, using unrelated skin melanoma patient-derived xenograft models, uveal melanoma skin and liver metastasis samples, and one lung adenocarcinoma primary tumor. (*) PDX:es. (†) Metastases (nine liver, four skin). Acc, accuracy; ACC, adrenocortical carcinoma; BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL, cholangiocarcinoma; COAD, colon adenocarcinoma; DLBC, lymphoid neoplasm diffuse large B-cell lymphoma; ESCA, esophageal carcinoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; KICH, kidney chromophobe; KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LGG, brain lower-grade glioma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; MESO, mesothelioma; NIR,

Fig A5. (Continued).
no information rate; OV, ovarian serous cystadenocarcinoma; PAAD, pancreatic adenocarcinoma; PCPG, pheochromocytoma and paraganglioma; PRAD, prostate adenocarcinoma; READ, rectum adenocarcinoma; SARC, sarcoma; SKCM, skin cutaneous melanoma; STAD, stomach adenocarcinoma; TGCT, testicular germ cell tumor; THCA, thyroid carcinoma; THYM, thymoma; UCS, uterine carcinosarcoma; UCEC, uterine corpus endometrial carcinoma; UVM, uveal melanoma.

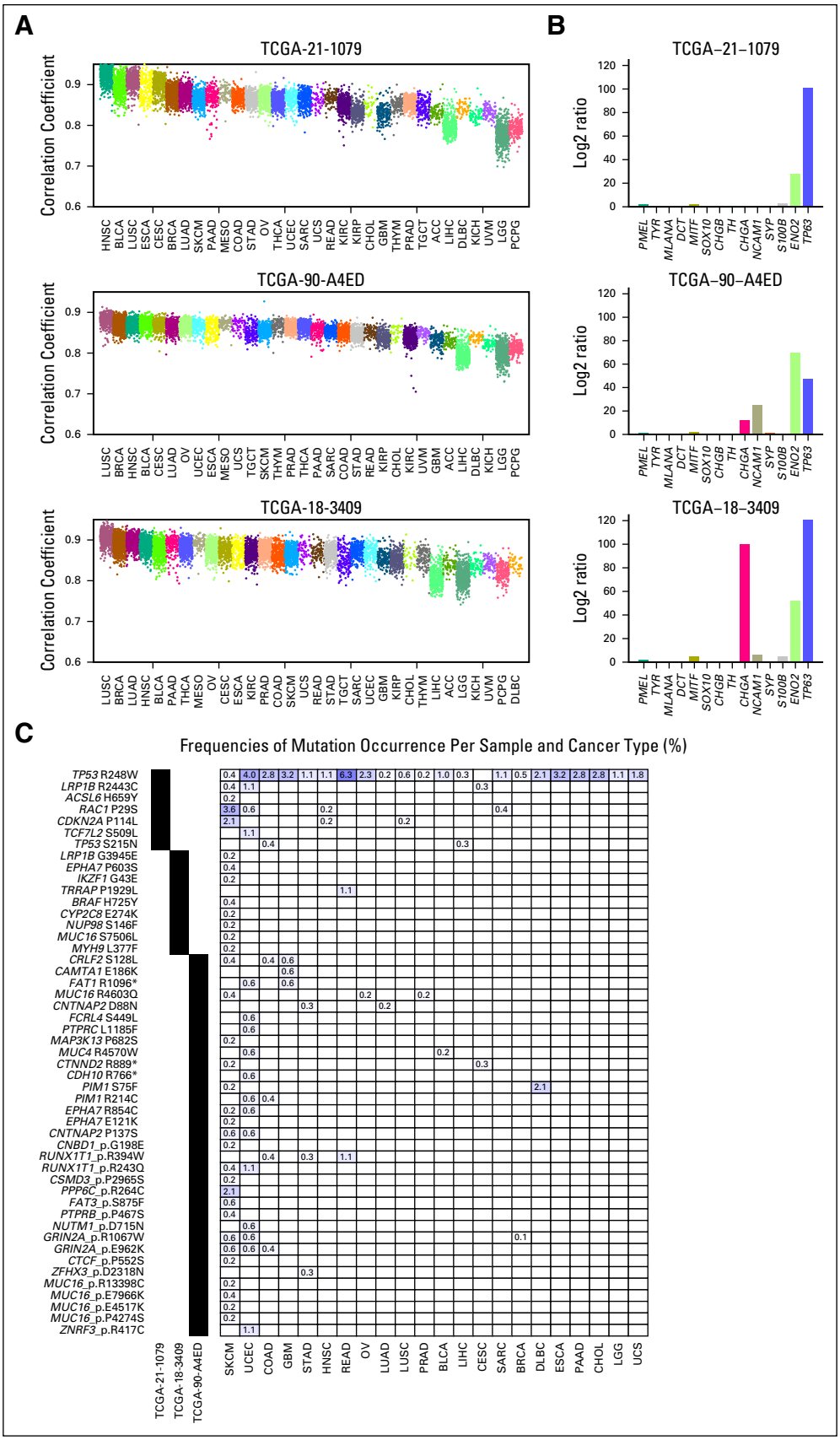

A
TCGA-21-1079
Correlation Coefficient
HNSC BLCA LUSC ESCA CESC BRCA LUAD SKCM PAAD MESO COAD STAD OV THCA UCEC SARC UCS READ KIRC KIRP CHOL GBM THYM PRAD TGCT ACC LIHC DLBC KICH UVM LGG PCPG
TCGA-90-A4ED
LUSC BRCA HNSC BLCA CESC LUAD OV UCEC ESCA MESO UCS TGCT SKCM THYM PRAD THCA PAAD SARC COAD STAD READ KIRP CHOL KIRC UVM GBM ACC LIHC DLBC KICH LGG PCPG
TCGA-18-3409
LUSC BRCA LUAD HNSC BLCA PAAD THCA MESO OV CESC ESCA KIRC PRAD COAD SKCM UCS READ STAD TGCT SARC UCEC GBM KIRP CHOL THYM LIHC ACC LGG KICH UVM PCPG DLBC
B
TCGA–21–1079
Log2 ratio
PMEL TYR MLANA DCT MITF SOX10 CHGB TH CHGA NCAM1 SYP S100B ENO2 TP63
TCGA–90–A4ED
TCGA–18–3409
C
Frequencies of Mutation Occurrence Per Sample and Cancer Type (%)
TP53 R248W
LRP1B R2443C
ACSL6 H659Y
RAC1 P29S
CDKN2A P114L
TCF7L2 S509L
TP53 S215N
LRP1B G3945E
EPHA7 P603S
IKZF1 G43E
TRRAP P1929L
BRAF H725Y
CYP2C8 E274K
NUP98 S146F
MUC16 S7506L
MYH9 L377F
CRLF2 S128L
CAMTA1 E186K
FAT1 R1096*
MUC16 R4603Q
CNTNAP2 D88N
FCRL4 S449L
PTPRC L1185F
MAP3K13 P682S
MUC4 R4570W
CTNND2 R889*
CDH10 R766*
PIM1 S75F
PIM1 R214C
EPHA7 R854C
EPHA7 E121K
CNTNAP2 P137S
CNBD1_p.G198E
RUNX1T1_p.R394W
RUNX1T1_p.R243Q
CSMD3_p.P2965S
PPP6C_p.R264C
FAT3_p.S875F
PTPRB_p.P467S
NUTM1_p.D715N
GRIN2A_p.R1067W
GRIN2A_p.E962K
CTCF_p.P552S
ZFHX3_p.D2318N
MUC16_p.R13398C
MUC16_p.E7966K
MUC16_p.E4517K
MUC16_p.P4274S
ZNRF3_p.R417C
TCGA-21-1079
TCGA-18-3409
TCGA-90-A4ED
SKCM UCEC COAD GBM STAD HNSC READ OV LUAD LUSC PRAD BLCA LIHC CESC SARC BRCA DLBC ESCA PAAD CHOL LGG UCS

Fig A6. Transcriptomic and mutational profiling of three ultraviolet signature–associated samples in The Cancer Genome Atlas (TCGA) lung squamous cell carcinoma (LUSC) cohort. (A) TCGA cancer types ranked by mean Spearman correlation of the top six most strongly correlating samples in each cohort: TCGA-21-1079 (top), TCGA-90-A4ED (middle), and TCGA-18-3409 (bottom). (B) Expression of selected marker genes relevant when considering a melanoma diagnosis: TCGA-21-1079 (top), TCGA-90-A4ED (middle), and TCGA-18-3409 (bottom). (C) Frequency of each driver gene mutation found in the ultraviolet radiation–associated LUSC tumors, in each of the different TCGA cancer types, relative to the total number of samples in each cohort. Driver gene mutations were defined as Catalogue of Somatic Mutations in Cancer mutations in Cancer Gene Census genes. Entries with zero frequencies in the matrix (after rounding to one digit) were left blank, and cancer types without these mutations represented were excluded. ACC, adrenocortical carcinoma; BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL, cholangiocarcinoma; COAD, colon adenocarcinoma; DLBC, lymphoid neoplasm diffuse large B-cell lymphoma; ESCA, esophageal carcinoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; KICH, kidney chromophobe; KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LGG, brain lower-grade glioma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; MESO, mesothelioma; OV, ovarian serous cystadenocarcinoma; PAAD, pancreatic adenocarcinoma; PCPG, pheochromocytoma and paraganglioma; PRAD, prostate adenocarcinoma; READ, rectum adenocarcinoma; RPKM, reads per kilobase million; SARC, sarcoma; SKCM, skin cutaneous melanoma; STAD, stomach adenocarcinoma; TGCT, testicular germ cell tumor; THCA, thyroid carcinoma; THYM, thymoma; UCS, uterine carcinosarcoma; UCEC, uterine corpus endometrial carcinoma; UVM, uveal melanoma.