analytical chemistry

pubs.acs.org/ac

Letter

Multivariate Analysis of RNA Chemistry Marks Uncovers Epitranscriptomics-Based Biomarker Signature for Adult Diffuse Glioma Diagnostics

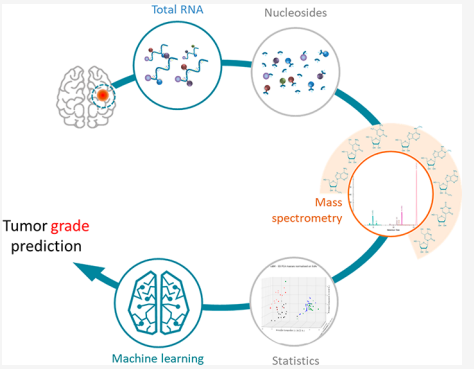
S. Relier, A. Amalric, A. Attina, I.B. Koumare, V. Rigau, F. Burel Vandenbos, D. Fontaine, M. Baroncini, J.P. Hugnot, H. Duffau, L. Bauchet, C. Hirtz,* E. Rivals,* and A. David*

Cite This: *Anal. Chem.* 2022, 94, 11967−11972

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

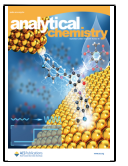
ABSTRACT: One of the main challenges in cancer management relates to the discovery of reliable biomarkers, which could guide decision-making and predict treatment outcome. In particular, the rise and democratization of high-throughput molecular profiling technologies bolstered the discovery of "biomarker signatures" that could maximize the prediction performance. Such an approach was largely employed from diverse OMICs data (i.e., genomics, transcriptomics, proteomics, metabolomics) but not from epitranscriptomics, which encompasses more than 100 biochemical modifications driving the post-transcriptional fate of RNA: stability, splicing, storage, and translation. We and others have studied chemical marks in isolation and associated them with cancer evolution, adaptation, as well as the response to conventional therapy. In this study, we have designed a unique pipeline combining multiplex analysis of the epitranscriptomic landscape by high-performance liquid chromatography coupled to tandem mass spectrometry with statistical multivariate analysis and machine learning approaches in order to identify biomarker signatures that could guide precision medicine and improve disease diagnosis. We applied this approach to analyze a cohort of adult diffuse glioma patients and demonstrate the existence of an "epitranscriptomics-based signature" that permits glioma grades to be discriminated and predicted with unmet accuracy. This study demonstrates that epitranscriptomics (co)evolves along cancer progression and opens new prospects in the field of omics molecular profiling and personalized medicine.



A significant fraction of cancer-related deaths could be prevented by improving biomarker-based early detection of cancer and molecular stratification. A biomarker is an objectively measured substance, structure, or process that serves as an indicator of normal or pathological state. As such, it can be exploited for patient management in various clinical settings such as disease risk estimation, cancer screening, and malignancy detection. A wide variety of biomarkers have been employed so far, including proteins, DNA, RNA, microRNA, antibodies, peptides, and metabolites.[1−4] They can either be measured from biological fluids (e.g., blood, urine) or directly from a tumor biopsy. Taken individually, current cancer biomarkers lack robustness and their variability weakens their predictive power. To address this challenge, recent advances in omics approaches combined with the emergence of machine learning applications[5] have spurred the identification of biomarker signatures, i.e., a set of biomarkers that maximize the prediction performance.[6] Discovery of novel accurate molecular tumor signature, related to known cellular pathways, opens the possibility of identifying drug-resistance pathways as well as new potential targets. Furthermore, the use of more than one biomarker may strengthen diagnosis accuracy and facilitate patient stratification. Recently, chemical modifications of RNA (a.k.a. epitranscriptomics) emerged as a novel layer of gene expression regulation in healthy tissues (e.g., brain), as well as in several pathologies such as neurodegenerative diseases[7] and cancer.[8,9] In particular, we and others associated several chemical marks with cancer evolution, adaptation, as well as the response to conventional therapy.[8,10,11] Building on these observations, we reasoned that several epitranscriptomics marks may vary along cancer progression and even play a role in this process. Epitranscriptomics may represent a novel source for biomarkers. Some studies correlate the level of a given nucleoside with disease onset or progression (reviewed in ref 9). Nevertheless, analysis of multiple epitranscriptomics marks has never been employed for diagnosis purpose.

In this study, we established an experimental procedure for mass spectrometry analysis of RNA extracted from patient

Received: April 7, 2022
Accepted: July 25, 2022
Published: August 23, 2022



https://doi.org/10.1021/acs.analchem.2c01526
Anal. Chem. 2022, 94, 11967−11972

biopsies. Resulting raw data was processed by the means of a bioinformatics pipeline dedicated to multivariate analysis of RNA marks. We applied this setting for analyzing a previously established cohort of adult diffuse glioma patients.[12,13] Glioma is the most common malignant tumor of the central nervous system, characterized by a significant variability in age of onset, histological and biological characteristics, and prognosis. It can be classified histologically according to the phenotype of tumoral cells: oligodendrogliomas and astrocytomas showing morphology reminiscent of oligodendrocyte and astrocyte lineage cells, respectively.[14] The latter two can be further ranked based on the consensus World Health Organization (WHO) classification, which assigns a degree of malignancy, graded from II to IV. Tumor grading, which describes the level of differentiation and/or how "normal" the cells look under a microscope, is the simplest way to determine the potential severity of the disease. Yet, diagnoses based upon biopsy have limitations for tumor grading and diagnosis of glioma. Most particularly, grades II and III cannot be easily distinguished since intratumor heterogeneity of the tumor grade is not uncommon in patients treated with extensive surgical resection.[15,16] Grade II designates a low grade, while grade-III represents a transition toward glioblastoma multiform (GBM), the most aggressive state. Despite a growing interest in the field,[17−19] not much is known about RNA marks dynamic throughout glioma progression.

Our study establishes the proof of concept that RNA marks evolve alongside grading tumor progression. Using machine learning, we demonstrate the existence of an "epitranscriptomics-based signature" that permits to discriminate glioma grades with remarkable ease. Finally, we show that a machine learning procedure on epitranscriptomic profiles can also predict glioma grading with unmet accuracy.

EXPERIMENTAL SECTION

Patients and Tumor Samples. A total of 58 tumors were surgically resected from adult patients diagnosed with diffuse glioma. None of the patients received any kind of chemical or radiation therapy before surgery. All samples were used in accordance with French bioethics laws regarding patient information and consent. At the time of resection, for each tumor, an aliquot was immediately frozen and stored at −80 °C, and the remaining tissue was fixed in 4% formalin as well as embedded in paraffin, and 3 μm sections were cut and stained with hematoxylin and eosin. The histopathological type of the tumor was determined, according to the revised World Health Organization classification.[20] Tumors consisted of grade-II (n = 20), grade-III gliomas (n = 20), and grade-IV glioblastomas (n = 18). Control samples (n = 19) originated from nononcological brain surgeries (epilepsy, benign lesion, etc.).

A specific consent for the study and validated by the CCPPRB has been signed by the patient or attested to in writing if the volunteer was unable to sign (e.g., hemiplegic subjects), by a third party independent of the investigator and the sponsor (validated in the study protocol). The information leaflet given to the persons as well as the consent form were also submitted to the opinion of the CCPPRB, authorizing the use of the samples and data in the study. This was a Biomedical Study: This biomedical research was conducted in accordance with the Declaration of Helsinki, the French Law n° 94-653 of July 29, 1994 relating to the respect of the human body (article 16-10 of the Civil Code), the Law n° 94-654 of July 29, 1994 relating to the donation and use of elements and products of the human body, to medical assistance in procreation and prenatal diagnosis (articles L1131-1 and L1245-2 of the Public Health Code), and the Law n°88-1138 of December 20, 1988 modified (Huriet-Serusclat Law, articles L1121 to L1126 of the Public Health Code). It was defined as research without direct individual benefit. For all samples, two senior pathologists (V.R. and F.B.V.) performed a double pathological review independently. Total RNA was extracted from tumor specimens by using the acid-phenol guanidium method. The quality of the RNA samples was determined by electrophoresis through agarose gels and ethidium bromide staining, and the 18S and 28S RNA bands were visualized under UV light.

A list of samples is available in "Supporting Information 1: Cohort description".

Sample Preparation. *RNA Digestion.* A 100 ng sample of total RNA was diluted in a total volume of 20 μL of Milli-Q water. A 3 μL aliquot of acetic-acid-acidified 0.1 M ammonium acetate, pH 5.3, and 1U of Nuclease P1 (Sigma, N8630) were added. Incubation at 42 °C for 2 h was performed. Then, 3 μL of 1 M ammonium acetate and 1U of alkaline phosphatase (Sigma, P4252) were added. The mixture was incubated at 37 °C for 2 h. Next, the nucleoside solution was diluted (28 μL sample + 60 μL solvent A) and was filtrated with 0.22 μm filters (Millex-GV, Millipore, SLGVR04NL). A 5 μL aliquot of each sample was injected, and all samples were analyzed in duplicate by LC-MS/MS.

LC-MS/MS. A mass spectrometer was calibrated to identify and quantitate accurately 25 modified nucleosides and 4 unmodified nucleosides (A, U, G, C) (Table S2.2). The nucleosides were separated by Nexera LC-40 systems (Shimadzu) using a Synergi Fusion-RP C18 column (4 μm particle size, 250 × 2 mm, 80 Å) (Phenomenex, 00G-4424-B0). The mobile phase consisted of 5 mM ammonium acetate adjusted to pH 5.3 with glacial acetic acid (solvent A) and pure acetonitrile (solvent B). The 30 min elution gradient started with 100% phase A followed by a linear gradient to 8% solvent B at 13 min. Solvent B was increased further to 40% over 10 min. After 2 min, solvent B was decreased back to 0% at 25.5 min. Initial conditions were regenerated by rinsing with 100% solvent A for an additional 4.5 min. The flow rate was 0.4 mL/min, and the column temperature was 35 °C. The detection was performed by a Shimadzu TripleQuad 8060 in positive ion mode. MS was operated in dynamic MRM mode with a retention time window of 3 min and a maximum cycle time set at 258 ms. The peak areas were determined using Skyline 4.1 software.

A thorough description of LC-MS/MS calibration is available in "Supporting Information 2: LC-MS/MS calibration".

Data Analysis. *Preprocessing and Normalization.* Each sample was measured thrice; we obtained thus three technical replicates: for each nucleoside, we checked that the corresponding retention times were homogeneous (less than 6% divergence), and if not, the measures were discarded (both mcm5U and hm5C were discarded). Then, we computed the average quantity for each nucleoside and transposed the table to get all nucleoside measures on one line for each biological sample. Since these quantities are relative and not absolute numbers of nucleosides, we normalized them. Because the quantity of an unmodified nucleoside also varies, this normalization was not ideal. We reasoned that the sum of the quantities of A, C, G, and U may be less variable than the individual quantities. We propose a new normalization (from

https://doi.org/10.1021/acs.analchem.2c01526
Anal. Chem. 2022, 94, 11967−11972

now on denoted by SUM): to divide the quantity of any nucleoside by the above-mentioned sum.

Exploratory Data Analysis and Machine Learning. The table of MS data was merged with that containing the grade information and healthy status (in one column of the dataframe). First, we computed the pairwise Pearson correlations between nucleoside quantitites using the corr function of the Pandas library[35] and plotted the heatmap using the Seaborn library.[36] Second, we explored the variability of any chosen nucleoside quantity with the cancer grade (which takes four possible values: Normal, II, III, IV) using boxplots (from Seaborn library). Third, we performed principal component analysis (PCA) on the vector of nucleoside quantities relative to those four classes defined by the cancer grades and the healthy status. We used the Seaborn and Matplotlib[37] libraries for all visualizations; the PCA analysis and other machine learning were programmed in Python using the Scikit-Learn library.[38]

The same dataframe was used for predicting the grades or healthy status with a support vector machine (SVM) equipped with a linear kernel and default parameters that were trained for multiclass predictions. In the evaluation procedure, the samples of data set were randomly splitted into the training set (75%) and the test set (25%).

Each step of data analysis is described in detail in "Supporting Information 3: Data processing".

RESULTS

Experimental Pipeline Description. In order to investigate whether the RNA modification pattern may evolve throughout glioma evolution, we employed RNA samples extracted from a previously established and characterized cohort including 58 samples from tumor biopsies (grade-II, -III, -IV gliomas) as well as 19 control samples (C) (n = 77 samples total, Sup. Table 1). We designed an experimental pipeline dedicated to feed a bioinformatics process with both experimental and clinical data (Figure 1A).

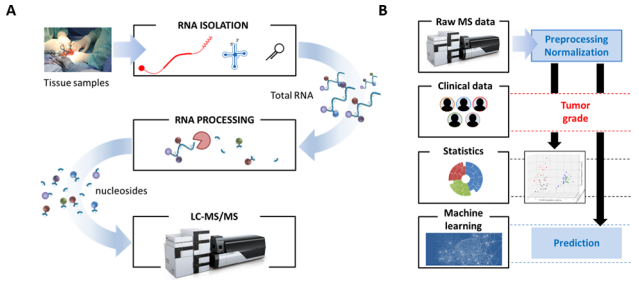


Figure 1. Overview of the method. (A) Experimental pipeline. This part is broken down into three steps: (1) RNA isolation from biological sample (tissue or plasma); (2) enzymatic processing of RNA into nucleosides; (3) injection and analysis by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS). (B) Data processing pipeline.

First, RNA samples (100 ng) are hydrolyzed using a stepwise enzymatic treatment. Second, the obtained mixture of nucleosides is analyzed using high-performance liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) (Figure 1 and Table S2.2). Third, a dedicated bioinformatics process takes over the raw mass spectrometry data (Figure 1B). This process is then divided into three blocks: (1) preprocessing and normalization; (2) joint analysis of LC-MS/MS data with cancer grade variable; (3) machine-learning-based analysis of epitranscriptomic profiles. Noteworthily, this process can be adapted to any type of clinical variable. In our case, we aim at distinguishing the cancer grades, in particular II and III.[15,16]

Epitranscriptome Modification Levels Vary with Cancer Grade. Establishing the epitranscriptome profile of cancer samples by LC-MS/MS is a novel approach to identify modified and unmodified nucleosides whose levels are altered following disease onset and progression. First, we investigate whether the variations of nucleoside quantities carry information about the cancer/healthy status and about the cancer grades. We observed the distribution of nucleoside quantities from samples belonging to either "control" tissues, to grade-II, to grade-III, or to grade-IV samples separately and in this order (i.e., Control < II < III < IV) (Figure 2B).

It appears that variations of quantities according to grades are similar between subsets of nucleosides: in one subset, the quantities augment from control toward grade-IV, in another subset, the quantities diminish from control toward grade-IV, and in a last subset, the quantities remain stable. This enables us to group nucleosides exhibiting similar variations relative to grade, in three groups (Figure 2B). Overall, this clearly suggests that variations of some individual nucleosides are an informative signal regarding the distinction of glioma grades.

Several pairs or groups of nucleosides exhibit highly correlated or anticorrelated measures, suggesting a lot of redundancy in these variations of quantities (Figure 2A). Noteworthily, N1-methylguanosine (m^1G), queuosine (Q), and Ac^4C behave similarly in a central cluster also containing N1-methyladenosine (m^1A), the four of them being particularly abundant in tRNA.[21] An alternative way of addressing our current question is to apply principal component analysis (PCA) on the entire vector of individual nucleoside quantities, which we term the *epitranscriptomic profile* or simply the *profile* of a sample. When applied to the full cohort, the PCA shows that 71% of the observed variations are accumulated on the first three dimensions (Figure 3 and Table S3.2): clearly, the percentage of variation observed on each component diminishes considerably until remaining below 8% after the fourth component and beyond. This means that we can project the points on the first three dimensions of this space and still carry over 70% of the variation (Figure 3). One clearly sees that control and grade-II samples are grouped together on the left (above value 0 on PC2), while grade-III and grade-IV samples are located on the right (above 0 on PC1). Moreover, on PC3, grade-III points tend to group below value 0, while grade-IV gliomas group above 0. The overall picture suggests that individual quantities combined in the epitranscriptomic profile carry sufficient information to "separate" cancer grades ("control" and grade-II samples appear to be globally mixed). Importantly, epitranscriptomics-based PCA analysis is not biased by age of patients (Figure S3.1).

Machine Learning on Epitranscriptomic Profiles Predicts Cancer Grade Accurately. Obviously, for patient management applications, it is natural to ask whether the information contained in epitrancriptomic profiles is sufficient to make a grade prediction, and with what accuracy. Next, we investigate this using machine learning approaches (a class of AI methods). A support vector machine (SVM) with a linear kernel and default parameters was trained for multiclass predictions using a randomly chosen subset of 70% of the samples (with the entire profiles). The resulting SVM is the

https://doi.org/10.1021/acs.analchem.2c01526
Anal. Chem. 2022, 94, 11967−11972

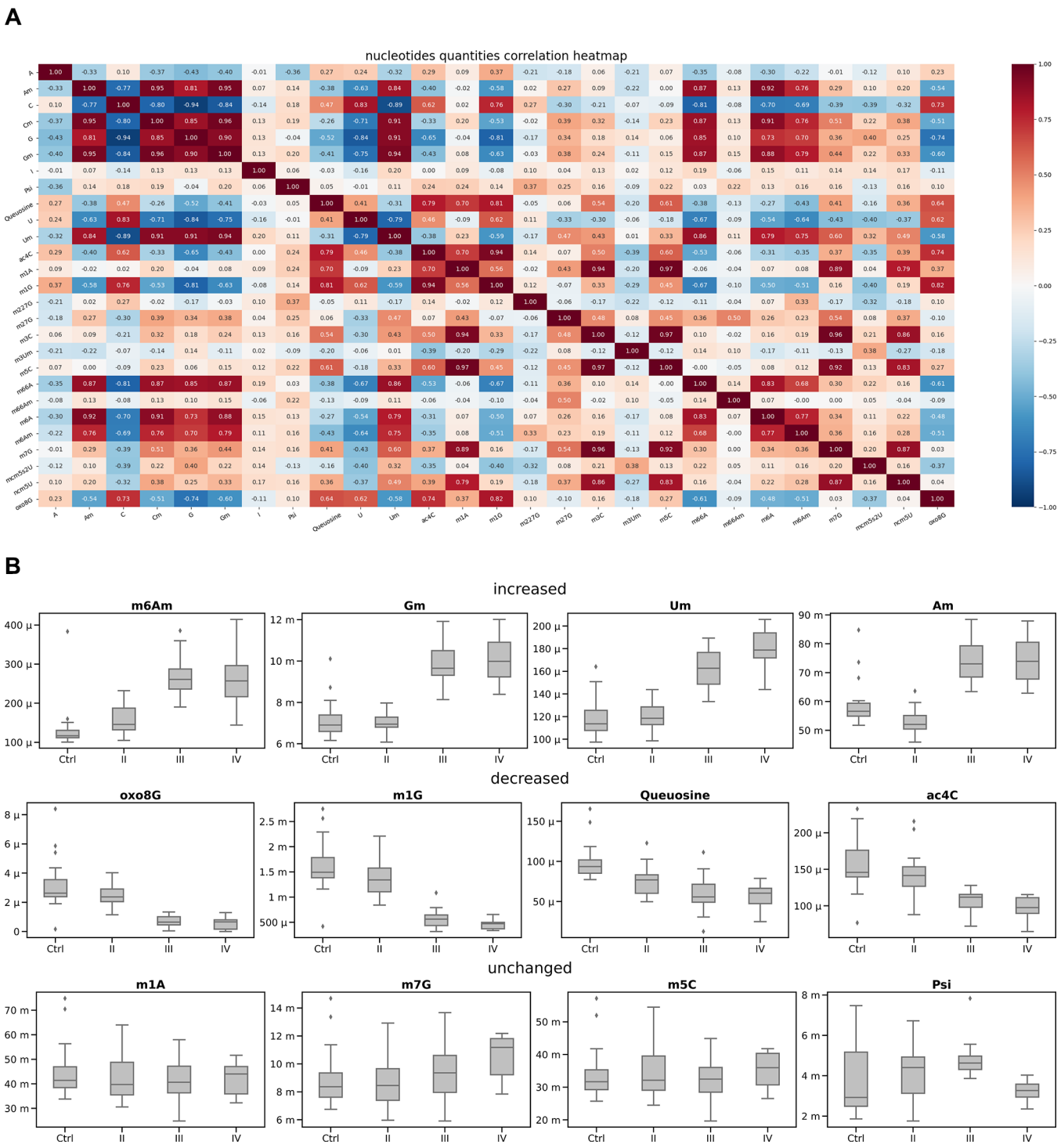


Figure 2. (A) Pairwise Pearson correlation of nucleoside levels. A central cluster displays several RNA modifications enriched in tRNA (dotted square). (B) Boxplots of selected nucleoside's levels throughout cancer grading. Three groups of nucleosides can be distinguished based on their level along cancer grading.

classifier: given an epitranscriptomic profile as input, it outputs a label corresponding to the predicted class (C, II, III, or IV). The SVM was then evaluated using the complementary subset of samples. The weighted average accuracy of prediction over all classes reaches 90% (with an ROC AUC score of 92.7%).

We tested the robustness of accuracy prediction with respect to the type of ML algorithms and obtained again an accuracy of 90% using linear discriminant analysis (see Sup Information 3). This confirmed that full epitranscriptomic profiles have sufficient predictive power to distinguish the "control" status and the three glioma grades.

Toward an Epitranscriptomic Signature for Glioma Grading. As illustrated earlier, distinct individual nucleoside measures behave differently with the cancer grade. Natural questions arise: are full profiles with 27 measures necessary to achieve prediction? In comparison to the full profile, nucleoside measures used *in isolation* are insufficient for prediction (Table S3.3). Then, among the 27 measures of the

https://doi.org/10.1021/acs.analchem.2c01526
Anal. Chem. 2022, 94, 11967−11972

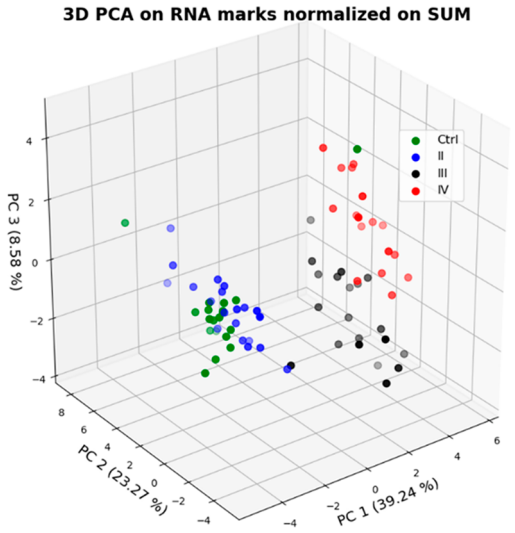


Figure 3. 3D PCA of epitranscriptomic profiles normalized on SUM. The number of dimensions was limited to three, since most of the variance can be attributed to three components (see Figure S3.1).

profile (termed features in ML terminology), which nucleosides contribute most important to prediction of grades? Given a classifier, recursive feature elimination (RFE) estimates the importance of each feature using cross validation and searches for an optimal number features for the prediction. When applied to the SVM trained with full profiles, RFE found a subset of nine features that optimizes prediction: {C_m, Psi, Q, U_m, m^1G, $m^{2,2,7}G$, m^5C, $m^{6,6}A$, m^6A_m} (Figure S3.2). Given this selection, we rebuilt the SVM classifier as above but using a profile restricted to these nine measures and obtained the same accuracy level of 90% for grade prediction with an improved ROC AUC score of 98.2%, showing that this selection is meaningful.

Given the redundancy observed in the full profile due to correlated measures, it is logical that RFE proposes a reduced subset. Because RFE (like other feature selection algorithms) is an iterative and heuristic procedure, the output selection likely is not unique: other subsets may achieve equally good or slightly suboptimal scores.

DISCUSSION

In this study, we implemented a novel experimental pipeline, coupling mass spectrometry with machine learning, to study the RNA marks landscape from patient samples. Applying this method to the glioma/GBM cohort, we uncovered the potential of analyzing epitranscriptomics for diagnostic/prognostic purpose. Several RNA marks have been previously considered as potential biomarkers, either in isolation or in small combinations,[22−25] but never at such a large scale.

While we established a proof of concept, further investigations will be required to solidify it for clinical application, in the context of brain cancer or other pathologies. This will require access to a larger cohort with refined clinic-biological data in order to get optimal patient stratification and an improved AI-based algorithm that ensures better prediction. For instance, diagnosis of gliomas can be further refined based on genetic characteristics (WHO 2016 and 2021 Classification[26]) such as isocitrate dehydrogenase 1 and 2 (IDH1 & 2) mutations,[27,28] 1P/19q codeletion,[29] and histone H3 K27 M mutation.[30] In particular, alterations on chromosomes 7 and 10 as well as mutations of TERT, EGFR, and CDKN2A were better developed in the last WHO 2021 classification. Depending on the status of the IDH gene mutation, GBM can be further classified into three subtypes: primary glioblastomas (wild type IDH), secondary glioblastomas (mutated IDH), and unclassified glioblastomas (NOS), which does not represent a GBM category per se due to its genetic heterogeneity.[31] Considering the emerging role of IDH1 and other metabolites in regulating RNA modifying enzymes,[32,33] we expect significant changes of RNA chemistry according to mutation status.

A strength of our approach resides in the use of total RNA, which facilitates RNA isolation and circumvents the issue of RNA degradation in frozen tissue.[34] The whole process could be easily automated and provide a prediction in matter of hours to guide clinical decision-making in support of regular histological and genomic procedures.

This study is in line with several reports showing the involvement of given RNA modifications throughout cancer progression (reviewed in ref 8). Epitranscriptomics-based signatures could be exploited to identify "essential" RNA modification whose deposition on specific RNA is a driver of cancer progression or resistance to conventional treatment. This may promote the development of novel therapeutic strategies as illustrated by the recent development of biotech companies in the field of RNA-chemistry-based therapeutics.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.2c01526.

Supporting Information 1: Cohort description; Supporting Information 2: LC-MS/MS calibration including the list of detected modifications, with detailed materials and method, chromatograms, and references; Supporting Information 3: Data processing including detailed description of preprocessing, normalization, data analysis and machine learning, and references (PDF)

AUTHOR INFORMATION

Corresponding Authors

C. Hirtz − *IRMB-PPC, INM, Univ Montpellier, CHU Montpellier, INSERM CNRS, Montpellier 34295, France;* orcid.org/0000-0002-7313-0629; Email: christophe.hirtz@umontpellier.fr

E. Rivals − *LIRMM, Univ. Montpellier, CNRS, Montpellier, Hérault 34095, France*; Email: rivals@lirmm.fr

A. David − *IGF, Univ. Montpellier, CNRS, INSERM, Montpellier, Hérault 34094, France; IRMB-PPC, INM, Univ Montpellier, CHU Montpellier, INSERM CNRS, Montpellier 34295, France;* orcid.org/0000-0003-3365-1339; Email: alexandre.david@igf.cnrs.fr

Authors

S. Relier − *IGF, Univ. Montpellier, CNRS, INSERM, Montpellier, Hérault 34094, France*

A. Amalric − *IGF, Univ. Montpellier, CNRS, INSERM, Montpellier, Hérault 34094, France; IRMB-PPC, INM, Univ Montpellier, CHU Montpellier, INSERM CNRS, Montpellier 34295, France*

A. Attina − *IRMB-PPC, INM, Univ Montpellier, CHU Montpellier, INSERM CNRS, Montpellier 34295, France*

https://doi.org/10.1021/acs.analchem.2c01526
Anal. Chem. 2022, 94, 11967−11972

I.B. Koumare − *Neurosurgery Department, Montpellier University Medical Center, Montpellier, Hérault 34295, France; Neurosurgery Department, CHU Gabriel Toure, Bamako, Mali*
V. Rigau − *Department of Pathology and Oncobiology, Montpellier University Medical Center, Montpellier, Hérault 34295, France*
F. Burel Vandenbos − *Central Laboratory of Pathology, Univ. Côte d'Azur, CHU Nice, CNRS, INSERM, Nice, Alpes-Maritimes 06000, France*
D. Fontaine − *Neurosurgery Department, Univ. Côte d'Azur, CHU Nice, Nice, Alpes-Maritimes 06000, France*
M. Baroncini − *Neurosurgery Department, CHU Lille, Univ. of Lille, Lille, Nord 59037, France*
J.P. Hugnot − *IGF, Univ. Montpellier, CNRS, INSERM, Montpellier, Hérault 34094, France*
H. Duffau − *IGF, Univ. Montpellier, CNRS, INSERM, Montpellier, Hérault 34094, France; Neurosurgery Department, Montpellier University Medical Center, Montpellier, Hérault 34295, France*
L. Bauchet − *IGF, Univ. Montpellier, CNRS, INSERM, Montpellier, Hérault 34094, France; Neurosurgery Department, Montpellier University Medical Center, Montpellier, Hérault 34295, France*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.2c01526

Author Contributions

A.D., E.R., and S.R. designed experiments and analyzed the results. J.P.H., I.B.K., M.B., H.D., D.F., and L.B. gathered clinical data and built the patient cohort. V.R. and F.B.V. performed histological analysis and review. S.R., A.A., and A.At performed experiments. S.R., E.R., A.D., C.H., A.A., and A.At, performed data analyses. S.R. and E.R. designed bioinformatics pipelines and performed bioinformatics analysis. E.R. devised and ran machine learning procedures. A.D., E.R., C.H, L.B., and V.R. wrote the manuscript. All the authors reviewed the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Our teams are generously supported by Occitanie Region/ FEDER (PPRi, SMART project), INCa DA N°2020-116 (AAP PLBIO 2020 - DAVID), Ligue contre le Cancer, SIRIC Montpellier Cancer (INCa-DGOS-Inserm 6045), CHU Montpellier, ARTC Sud, Des Etoiles dans la Mer, and Cancéropole GSO. Mass spectrometry experiments were carried out using the facilities of the Montpellier Proteomics Platform (PPM−PPC, BioCampus Montpellier).

■ REFERENCES

(1) Ma, Y.; Zhang, P.; Wang, F.; et al. *Ann. Surg* **2012**, *255* (4), 720−30.
(2) Ma, Y.; Zhang, P.; Yang, J.; et al. *Int. J. Cancer* **2012**, *130* (9), 2077−87.
(3) Shi, H.; Li, X.; Zhang, Q.; et al. *Biomarkers* **2016**, *21* (7), 578−88.
(4) Tumas, J.; Kvederaviciute, K.; Petrulionis, M.; et al. *Med. Oncol* **2016**, *33* (12), 133.
(5) Pasolli, E.; Truong, D. T.; Malik, F.; et al. *PLoS Comput. Biol.* **2016**, *12* (7), e1004977.
(6) Lin, Y.; Qian, F.; Shen, L.; et al. *Brief Bioinform* **2019**, *20* (3), 952−975.
(7) Chatterjee, B.; Shen, C.J.; Majumder, P. *Int. J. Mol. Sci.* **2021**, *22* (21), 11870.
(8) Barbieri, I.; Kouzarides, T. *Nat. Rev. Cancer* **2020**, *20* (6), 303−322.
(9) Amalric, A.; Bastide, A.; Attina, A.; et al. *Crit Rev. Clin Lab Sci.* **2022**, *59* (1), 1−18.
(10) Relier, S.; Ripoll, J.; Guillorit, H.; et al. *Nat. Commun.* **2021**, *12* (1), 1716.
(11) Macari, F.; El-Houfi, Y.; Boldina, G.; et al. *Oncogene* **2016**, *35* (14), 1785−96.
(12) Reme, T.; Hugnot, J. P.; Bieche, I.; et al. *PLoS One* **2013**, *8* (6), e66574.
(13) Guichet, P. O.; Bieche, I.; Teigell, M.; et al. *Glia* **2013**, *61* (2), 225−39.
(14) Kim, H. J.; Park, J. W.; Lee, J. H. *Front Oncol* **2021**, *10*, 615400.
(15) Darlix, A.; Rigau, V.; Fraisse, J.; et al. *Neurology* **2020**, *94* (8), e830−e841.
(16) Pedeutour-Braccini, Z.; Burel-Vandenbos, F.; Goze, C.; et al. *Virchows Arch* **2015**, *466* (4), 433−44.
(17) Chai, R. C.; Wu, F.; Wang, Q. X.; et al. *Aging (Albany NY)* **2019**, *11* (4), 1204−1225.
(18) Pan, T.; Wu, F.; Li, L.; et al. *Mol. Brain* **2021**, *14* (1), 119.
(19) Dong, Z.; Cui, H. *Cancers (Basel)* **2020**, *12* (3), 736.
(20) Louis, D. N.; Ohgaki, H.; Wiestler, O. D.; et al. *Acta Neuropathol* **2007**, *114* (2), 97−109.
(21) Pan, T. *Cell Res.* **2018**, *28* (4), 395−404.
(22) Lo, W. Y.; Jeng, L. B.; Lai, C. C.; et al. *Clin. Chim. Acta* **2014**, *428*, 57−62.
(23) Opitz, P.; Herbarth, O.; Seidel, A.; et al. *Anticancer Res.* **2018**, *38* (11), 6113−6119.
(24) Hsu, W. Y.; Chen, C. J.; Huang, Y. C.; et al. *PLoS One* **2013**, *8* (12), e81701.
(25) Kumar, N.; Shahjaman, M.; Mollah, M. N. H.; et al. *Bioinformation* **2017**, *13* (6), 202−208.
(26) Louis, D. N.; Perry, A.; Wesseling, P.; et al. *Neuro Oncol* **2021**, *23* (8), 1231−1251.
(27) Zeng, A.; Hu, Q.; Liu, Y.; et al. *Oncotarget* **2015**, *6* (30), 30232−8.
(28) Hartmann, C.; Meyer, J.; Balss, J.; et al. *Acta Neuropathol* **2009**, *118* (4), 469−74.
(29) Griffin, C. A.; Burger, P.; Morsberger, L.; et al. *J. Neuropathol Exp Neurol* **2006**, *65* (10), 988−94.
(30) Lewis, P. W.; Muller, M. M.; Koletsky, M. S.; et al. *Science* **2013**, *340* (6134), 857−61.
(31) DeWitt, J. C.; Mock, A.; Louis, D. N. *Curr. Opin Neurol* **2017**, *30* (6), 643−649.
(32) Kim, J.; Lee, G. *Metabolites* **2021**, *11* (2), 80.
(33) Elkashef, S. M.; Lin, A. P.; Myers, J.; et al. *Cancer Cell* **2017**, *31* (5), 619−620.
(34) Richter, F.; Plehn, J. E.; Bessler, L.; et al. *Nucleic Acids Res.* **2022**, *50*, 4201.
(35) Reback, J.; McKinney, W.; Brockmendel, J., et al. pandas-dev/ pandas: Pandas 1.0.3 (Version v1.0.3); Zenodo, 2020. https://zenodo.org/record/3715232#.YuAMEXbMKUk.
(36) Waskom, M. L. *Journal of Open Source Software* **2021**, *6* (60), 3021.
(37) Hunter, J. D. *Computing in Science & Engineering* **2007**, *9* (3), 90−95.
(38) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; et al. *Journal of Machine Learning Research* **2011**, *12*, 2825−2830.