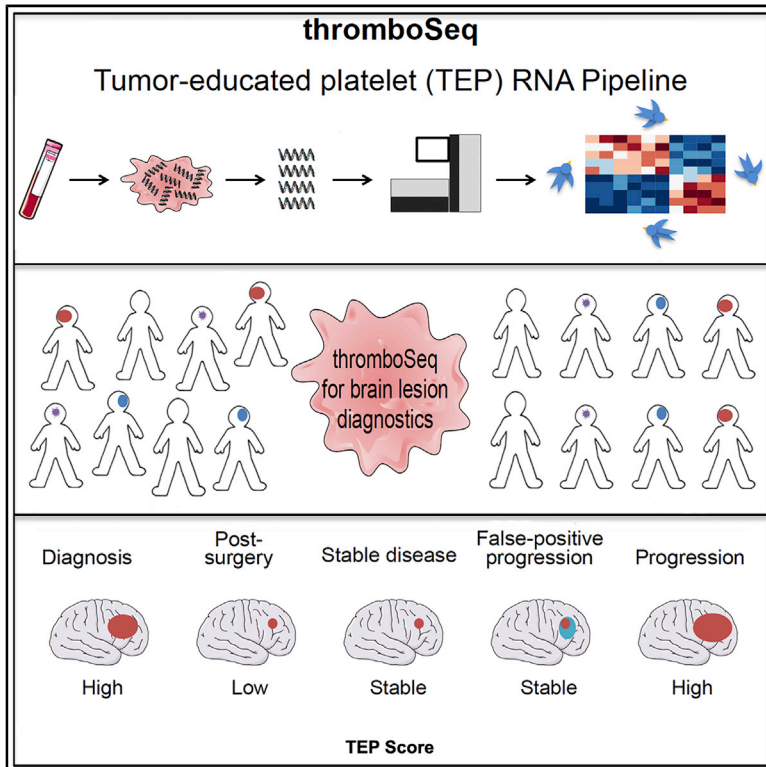


# Tumor-Educated Platelet RNA for the Detection and (Pseudo)progression Monitoring of Glioblastoma

## Graphical Abstract



## Authors

Nik Sol, Sjors G.J.G. in 't Veld, Adrienne Vancura, ..., Jaap C. Reijneveld, Myron G. Best, Thomas Wurdinger

## Correspondence

ni.sol@amsterdamumc.nl (N.S.),  
t.wurdinger@amsterdamumc.nl (T.W.)

## In Brief

Sol et al. show that tumor-educated platelets (TEPs) can be employed for detection of glioblastoma. Glioblastoma TEP RNA profiles can be differentiated from patients with metastatic brain cancer or MS. Furthermore, the tumor signals in TEPs are dynamic, indicating that TEPs can be employed for blood-based therapy monitoring.

## Highlights

- TEP RNA enables blood-based brain tumor diagnostics
- TEP RNA is dynamic throughout anti-tumor treatment
- TEP RNA may be employed for therapy monitoring



## Article

# Tumor-Educated Platelet RNA for the Detection and (Pseudo)progression Monitoring of Glioblastoma

Nik Sol,<sup>1,2,\*</sup> Sjors G.J.G. in 't Veld,<sup>1,3,13</sup> Adrienne Vancura,<sup>1,3,13</sup> Maud Tjerkstra,<sup>1,3</sup> Cyra Leurs,<sup>2,4</sup> François Rustenburg,<sup>1,3</sup> Pepijn Schellen,<sup>1,3</sup> Heleen Verschueren,<sup>1,3</sup> Edward Post,<sup>1,3</sup> Kenn Zwaan,<sup>1,3</sup> Jip Ramaker,<sup>1,3</sup> Laurine E. Wedekind,<sup>1,3</sup> Jihane Tannous,<sup>5</sup> Bauke Ylstra,<sup>6</sup> Joep Killestein,<sup>2,4</sup> Farrah Mateen,<sup>5</sup> Sander Idema,<sup>3</sup> Philip C. de Witt Hamer,<sup>3</sup> Anna C. Navis,<sup>7</sup> William P.J. Leenders,<sup>8</sup> Ann Hoeben,<sup>9</sup> Bastiaan Moraal,<sup>10</sup> David P. Noske,<sup>1,3</sup> W. Peter Vandertop,<sup>1,3</sup> R. Jonas A. Nilsson,<sup>3,11</sup> Bakhos A. Tannous,<sup>5</sup> Pieter Wesseling,<sup>1,6,12</sup> Jaap C. Reijneveld,<sup>1,2</sup> Myron G. Best,<sup>1,3,6,14</sup> and Thomas Wurdinger<sup>1,3,14,15,\*</sup>

<sup>1</sup>Brain Tumor Center Amsterdam, Cancer Center Amsterdam, Amsterdam UMC, VU University Medical Center, Amsterdam, the Netherlands

<sup>2</sup>Department of Neurology, Cancer Center Amsterdam, Amsterdam UMC, VU University Medical Center, Amsterdam, the Netherlands

<sup>3</sup>Department of Neurosurgery, Cancer Center Amsterdam, Amsterdam UMC, VU University Medical Center, Amsterdam, the Netherlands

<sup>4</sup>MS Center Amsterdam, Amsterdam UMC, VU University Medical Center, Amsterdam, the Netherlands

<sup>5</sup>Department of Neurology, Massachusetts General Hospital and Neuroscience Program, Harvard Medical School, Boston, MA, USA

<sup>6</sup>Department of Pathology, Cancer Center Amsterdam, Amsterdam UMC, VU University Medical Center, Amsterdam, the Netherlands

<sup>7</sup>Department of Pathology, Radboud University Medical Center, Nijmegen, the Netherlands

<sup>8</sup>Department of Biochemistry, Radboud Institute for Molecular Life Sciences, Nijmegen, the Netherlands

<sup>9</sup>Department of Medical Oncology, Maastricht Academic Medical Center, Maastricht, the Netherlands

<sup>10</sup>Department of Radiology and Nuclear Medicine, Amsterdam UMC, VU University Medical Center, Amsterdam, the Netherlands

<sup>11</sup>Department of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden

<sup>12</sup>Princess Máxima Center for Pediatric Oncology, Utrecht, the Netherlands

<sup>13</sup>These authors contributed equally

<sup>14</sup>Senior author

<sup>15</sup>Lead Contact

\*Correspondence: [ni.sol@amsterdamumc.nl](mailto:ni.sol@amsterdamumc.nl) (N.S.), [t.wurdinger@amsterdamumc.nl](mailto:t.wurdinger@amsterdamumc.nl) (T.W.)

<https://doi.org/10.1016/j.xcrm.2020.100101>

## SUMMARY

Tumor-educated platelets (TEPs) are potential biomarkers for cancer diagnostics. We employ TEP-derived RNA panels, determined by swarm intelligence, to detect and monitor glioblastoma. We assessed specificity by comparing the spliced RNA profile of TEPs from glioblastoma patients with multiple sclerosis and brain metastasis patients (validation series,  $n = 157$ ; accuracy, 80%; AUC, 0.81 [95% CI, 0.74–0.89;  $p < 0.001$ ]). Second, analysis of patients with glioblastoma versus asymptomatic healthy controls in an independent validation series ( $n = 347$ ) provided a detection accuracy of 95% and AUC of 0.97 (95% CI, 0.95–0.99;  $p < 0.001$ ). Finally, we developed the digitalSWARM algorithm to improve monitoring of glioblastoma progression and demonstrate that the TEP tumor scores of individual glioblastoma patients represent tumor behavior and could be used to distinguish false positive progression from true progression (validation series,  $n = 20$ ; accuracy, 85%; AUC, 0.86 [95% CI, 0.70–1.00;  $p < 0.012$ ]). In conclusion, TEPs have potential as a minimally invasive biosource for blood-based diagnostics and monitoring of glioblastoma patients.

## INTRODUCTION

Several blood-based biosources, such as plasma, serum, plasma-derived extracellular vesicles, and circulating tumor cells, are currently being evaluated as liquid biopsies for many types of cancer.<sup>1,2</sup> However, brain tumors are notoriously difficult to detect in blood.<sup>3–6</sup> Analysis of cerebrospinal fluid collected from patients with diffuse glioma revealed the presence of multiple molecular biomarkers.<sup>7</sup> Alternatively, the presence of glioblastoma can only be identified in plasma DNA in less than 10% of patients while analyzing a plethora of point mutations and genomic rearrangements<sup>8</sup> and in 60% of IDH1 mutant glioma patients when specifically analyzing *IDH1* mutants.<sup>9</sup> In addition, circulating tumor cells were detected in

20%–73% of patients with glioblastoma, depending on the detection method applied.<sup>10–12</sup> Detection of brain cancer in blood might leverage the use of such “liquid biopsies” for analysis of tumor progression, tumor recurrence, therapy response prediction, and monitoring<sup>13</sup> and for differentiating glioblastoma tumor progression from false positive progression (pseudo-progression or radiation necrosis) after initial therapy.<sup>3</sup>

Blood platelets act as local and systemic responders during tumorigenesis and cancer metastasis<sup>14</sup> and are exposed to tumor-induced platelet education, resulting in altered platelet behavior.<sup>15–17</sup> We have shown that tumor-educated platelets (TEPs) are a potential biosource for blood-based cancer diagnostics<sup>18–20</sup> by using the highly multiplexed biomarker detection platform thromboSeq.<sup>21</sup> We previously detected glioblastoma



with an accuracy of 84% in a 32-sample validation series among age-unmatched healthy (asymptomatic) controls.<sup>18</sup> The TEP-derived RNA signatures allowed separation of patients with glioblastoma from those with metastases to the brain with 92% accuracy.<sup>18</sup>

In this study, we investigate the diagnostic power of thromboSeq for differentiation of glioblastoma from other brain lesions (i.e., brain metastases and active multiple sclerosis lesions), unifying unique TEP educational programs in neurological diseases. Furthermore, we show that TEP RNA profiles of patients with glioblastoma are different from those of asymptomatic healthy controls. Finally, we show that, in patients with glioblastoma, TEP RNA signatures correlated to tumor volume and tumor recurrence and may facilitate discrimination of true tumor progression from false positive progression.

## RESULTS

### Unique Spliced RNA Profiles Can Be Identified in Platelets of Patients with Several Neurological Diseases

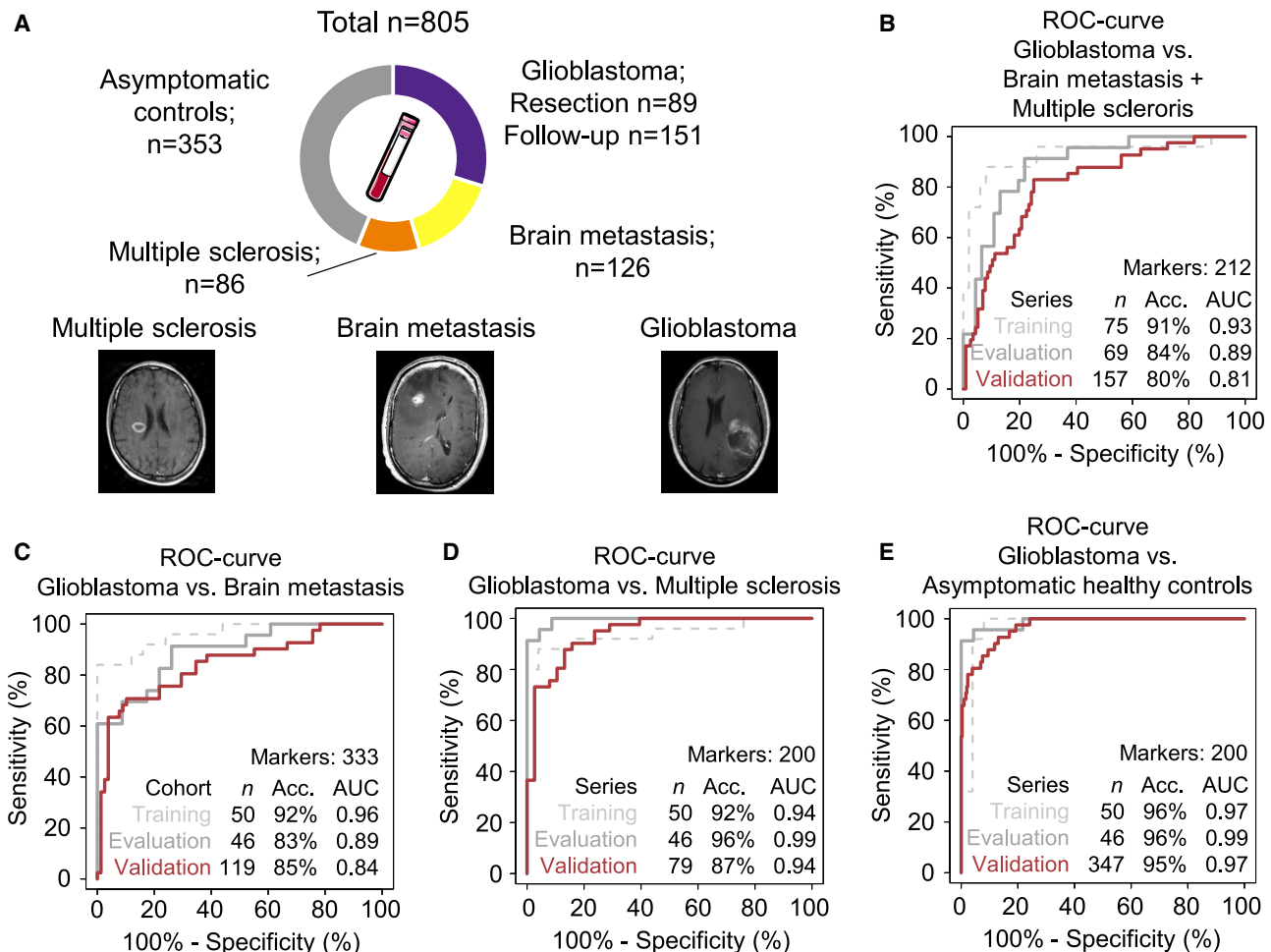
We first investigated whether several neurological diseases may differentially educate platelets, resulting in different spliced RNA profiles. For this, we prospectively collected and isolated platelet pellets from whole blood by differential centrifugation from 89 patients with primary glioblastoma collected on the day of first tumor resection (glioblastoma resection). From 52 of these 89 patients, follow-up blood samples were collected during postoperative chemo- and radiotherapy treatment (glioblastoma follow-up; 151 samples in total, 2–9 samples per patient). In addition, we collected blood from 126 patients with one or multiple brain metastases and 86 patients with relapsing-remitting multiple sclerosis (Figure 1A). Finally, we included platelet samples from 353 asymptomatic healthy controls without self-reported symptoms of neurological disorders or cancer, resulting in a total series size of 805 samples (Figure 1A; Table S1). Patients with brain metastases were diagnosed with different primary tumors: non-small cell lung carcinoma (NSCLC;  $n = 85$ ), breast cancer ( $n = 15$ ), melanoma ( $n = 15$ ), renal cell cancer ( $n = 7$ ), colorectal cancer (CRC;  $n = 1$ ), esophagus cancer ( $n = 1$ ), pancreatic cancer ( $n = 1$ ), and an unknown primary tumor ( $n = 1$ ). The majority of patients with brain metastases also had metastatic disease in other organs, including the liver, lungs, bones, and adrenal tissue (Table S1). Patients with multiple sclerosis were part of an early inception cohort in which patients were included at diagnosis and followed annually. These patients were diagnosed at least 10 years ago; all had clinically stable disease, and a subset of patients used disease-modifying drugs. Patient characteristics are provided in Table S1. From a small subset of 21 glioma patient samples, platelet counts were determined. No correlation was observed between platelet count and RNA concentration (Figure S1A). All samples were subjected to the thromboSeq platelet RNA sequencing pipeline (Figures S1B–S1D).<sup>18,19,21</sup>

To minimize potential confounding effect of previously identified confounding factors, we next selected, from this blood sample series, a sample set matched for age and whole-blood storage time, resulting in a matched series of 48 patients with glioblastoma at the time of first tumor resection, patients with brain metastases, patients with multiple sclerosis, and asymp-

tomatically healthy controls. The median age was comparable among the four groups (Table 1), but it should be noted that the average age of patients with multiple sclerosis is lower because of a difference in time of disease onset compared with those with glioblastoma or brain metastases. Hence, we accepted the suboptimal results of age matching for patients with multiple sclerosis. We ensured that comparable numbers of samples were isolated within 24 or 48 h after blood withdrawal (Table 1), circumventing potential effects of blood storage time on the platelet-spliced RNA profiles.<sup>19</sup>

We first investigated the specificity of TEPs from glioblastoma patients by comparing the RNA profiles from patients with glioblastoma at the time of the first tumor resection with platelets from brain metastasis patients and patients with multiple sclerosis. For this, we randomly separated the total matched sample series into training, evaluation, and validation series (Table 1). The training samples were employed to select a spliced RNA biomarker panel and to build a machine learning support vector machine (SVM) algorithm, whereas the evaluation series was employed to further optimize the spliced RNA biomarker panel and performance of the SVM algorithm by swarm intelligence (STAR Methods). First, an algorithm biomarker panel of 212 platelet spliced RNAs was calculated to distinguish glioblastoma versus brain metastasis and multiple sclerosis. This resulted in a training series ( $n = 75$ ) with an area under the curve (AUC) of 0.93 (95% confidence interval [CI], 0.85–1.00), an evaluation series ( $n = 69$ ) with an AUC of 0.89 (95% CI, 0.81–0.97), and an independent validation series ( $n = 157$ ) with an AUC of 0.81 (95% CI, 0.74–0.89;  $p < 0.001$ ; Figure 1B; Table S2A). Second, another biomarker panel of 341 platelet-spliced RNAs was used in a multiclass comparison of glioblastoma versus multiple sclerosis versus brain metastasis. This resulted in a training series ( $n = 75$ ) with an accuracy of 73%, an evaluation series ( $n = 69$ ) with an accuracy of 77%, and a validation series ( $n = 157$ ) with an accuracy of 75% ( $p < 0.001$ ; Figure S1E; Table S2B). The patient groups were also analyzed separately and compared directly. For glioblastoma versus brain metastasis, a biomarker panel of 333 platelet-spliced RNAs was selected. This resulted in a training series ( $n = 50$ ) with an AUC of 0.96 (95% CI, 0.92–1.00), an evaluation series ( $n = 46$ ) with an AUC of 0.89 (95% CI, 0.81–0.98), and an independent validation series ( $n = 119$ ) with an AUC of 0.84 (95% CI, 0.76–0.92;  $p < 0.001$ ; Figure 1C; Table S2C). For glioblastoma versus multiple sclerosis, a biomarker panel of 200 platelet-spliced RNAs was selected. This resulted in a training series ( $n = 50$ ) with an AUC of 0.94 (95% CI, 0.87–1.00), an evaluation series ( $n = 46$ ) with an AUC of 0.99 (95% CI, 0.98–1.00), and an independent validation series ( $n = 79$ ) with an AUC of 0.94 (95% CI, 0.89–0.99;  $p < 0.001$ ; Figure 1D; Table S2D).

Because our data indicate that TEPs from glioblastoma patients have a unique spliced RNA profile, we next investigated whether patients with glioblastoma at the time of tumor resection can be differentiated from asymptomatic healthy controls. Following training and optimization of the thromboSeq classification algorithm with a biomarker panel of 200 platelet-spliced RNAs, a 347-sample validation series reached an accuracy of 68% for detection of glioblastoma at a specificity for detection of asymptomatic healthy controls of more than 98%. In this



**Figure 1. TEP RNA Profiling for Brain Tumor Diagnostics**

(A) Schematic overview of TEPs as biosource for liquid biopsies and number of groups and sample series sizes included.

(B and C) receiver operating characteristics (ROC) curves of (B) “glioblastoma versus brain metastasis plus multiple sclerosis” algorithm, (C) “glioblastoma versus brain metastasis” algorithm, (D) “glioblastoma versus multiple sclerosis” algorithm, and (E) “glioblastoma versus asymptomatic healthy controls” algorithm, including the training series (dashed gray), evaluation series (gray), and validation series (red). Indicated are sample series sizes, best accuracy, and AUC value. Sample HC0068 and sample Maas-GBM-NICT-035G were inadvertently duplicated in the training or validation series because of a randomization code error identified after the validation process.

See also [Figure S1](#) and [Tables S1](#) and [S2A–S2E](#).

case, the training series (n = 50) resulted in an AUC of 0.97 (95% CI, 0.92–1.00), the evaluation series (n = 46) in an AUC of 0.99 (95% CI, 0.95–1.00), and the validation series (n = 347) in an AUC of 0.97 (95% CI, 0.95–0.99; p < 0.001; [Figure 1E](#); [Table S2E](#)). We thus conclude that platelet-spliced RNA profiles may enable blood-based diagnosis of neuro-oncological and neuro-inflammatory conditions and are most likely not a response of aspecific platelet activation.

### TEP RNA Profiles of Patients with Glioblastoma Normalize toward Platelet RNA Profiles from Asymptomatic Healthy Controls following Tumor Resection

Because we observed unique spliced RNA repertoires in TEPs of patients with glioblastoma at the time of tumor resection, we

questioned whether these profiles might quantitatively decrease concordant with tumor volume. For this, following the blood sample at the time of tumor resection, we collected blood at several follow-up time points during the concurrent chemo-radiation phase together with magnetic resonance imaging (MRI) tumor visualization<sup>22</sup> (52 unique patients, 151 platelet samples in total, 2–9 samples per patient; [Table S1](#)). Of one glioblastoma patient included in the follow-up sample collection, one pre-operative sample failed the sample processing quality checks (VU429 T0), but the follow-up samples were included in the analysis. Collection of follow-up samples was completed when there were no further treatment options or when a patient was lost to follow-up for any reason ([STAR Methods](#)). The sample collection schedule took into account the naturally occurring renewal of the platelet pool in blood following tumor resection because

**Table 1. Series Layout**

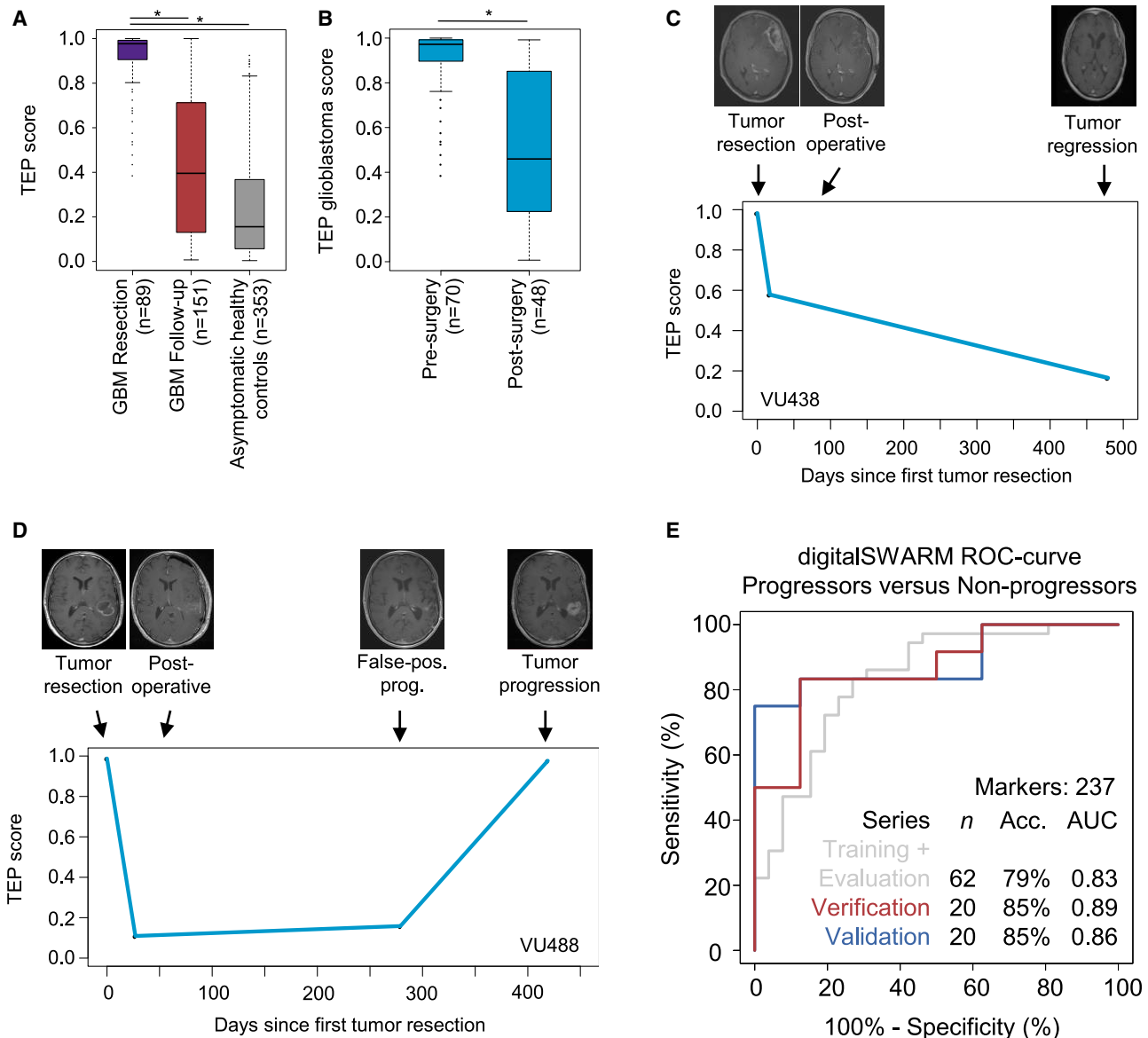
Series	Group	n	Median age in years (IQR)	Blood storage time (<24 h)	Gender (% male; % unknown)
Training series	glioblastoma baseline	25	65 (23)	18 (72%)	64%; 0%
	multiple sclerosis	25	49 (17)	22 (88%)	24%; 12%
	brain metastasis	25	62 (11)	20 (80%)	40%; 0%
	asymptomatic controls	25	63 (30)	16 (64%)	44%; 0%
Evaluation series	glioblastoma baseline	23	63 (15)	18 (78%)	70%; 0%
	multiple sclerosis	23	47 (14.5)	19 (83%)	26%; 17%
	brain metastasis	23	65 (16)	18 (78%)	35%; 0%
	asymptomatic controls	23	63 (13.5)	16 (70%)	39%; 0%
Validation series	glioblastoma baseline	34	55 (18)	35 (85%)	68%; 15%
	multiple sclerosis	38	39 (43)	26 (68%)	18%; 32%
	brain metastasis	78	58 (14)	38 (49%)	49%; 0%
	asymptomatic controls	306	47 (29)	282 (92%)	33%; 8%

platelets have a 7- to 10-day lifespan.<sup>23</sup> This follow-up sample series included glioblastoma patients with stable, progressive, and regressive disease as well as glioblastoma patients with false positive progression. We employed the “glioblastoma versus asymptomatic healthy controls” algorithm (Figure 1E) for classification of all glioblastoma baseline samples assigned to the validation series (n = 89) and glioblastoma follow-up samples (n = 151) and all asymptomatic healthy controls (n = 353). The “glioblastoma versus asymptomatic healthy controls” algorithm reports for each sample a binary classification result (“asymptomatic healthy control” or “glioblastoma”). However, in addition, the algorithm can also provide, for each sample, a quantitative score ranging from 0–1, representing the classification confidence score. The more pronounced the glioblastoma signal in a certain sample, the more this sample has a classification confidence score toward 1. We employed the classification confidence score to indicate the relative and quantitative signal of glioblastoma in platelet RNA profiles collected during therapy follow-up and called this classification confidence score the TEP score. First we observed that the TEP score of the blood samples collected from patients with glioblastoma just before tumor resection was significantly higher compared with those collected during the follow-up period (mean glioblastoma at tumor resection [n = 89], 0.91; mean glioblastoma follow-up [n = 151], 0.44;  $p < 2.2 \times 10^{-16}$ ; independent Student’s t test; Figure 2A), and asymptomatic healthy controls (mean asymptomatic healthy controls [n = 353], 0.24;  $p < 2.2 \times 10^{-16}$ ; independent Student’s t test; Figure 2A). To determine whether TEP-derived spliced RNA profiles do mirror the disease burden in the longitudinal sample collection, we analyzed the TEP score in 52 patients who were treated for their glioblastoma (Figures 2C and 2D; Data S1). When comparing pre-operative baseline blood samples with samples collected after tumor resection (median number of days after resection, 19; min-max, 10–33 days), we observed a mean decrease in TEP score after resection of 0.40 ( $p < 8.7 \times 10^{-11}$ , n = 70 versus n = 48 patients, unpaired independent Student’s t test; Figure 2B). These results indicate that the reduced tumor load after tumor resection significantly reduces the TEP score in blood. Moreover, we discovered a negative

correlation between the TEP score calculated from the first post-operative sample and overall survival ( $r = -0.33$ ,  $p = 0.03$ , Pearson’s correlation, n = 30 samples; Figure S2A). However, the contribution of potential additional prognostic factors, such as the age of the patient, Karnofsky performance score, and molecular tumor characteristics, cannot be excluded.

Subsequently, for each available time point, we determined the TEP score and matched the scores with the available MRI of the tumor (Figures 2C and 2D; Data S1). The patient with glioblastoma, VU438 (Figure 2C), had continuing lower volumes of gadolinium-enhancing tumor tissue in MRIs following tumor resection with a concomitant decrease in TEP score. Contrary, patient VU488, following a measure point with false positive progression, had true tumor progression with an increasing TEP score (Figure 2D). Interestingly, for some patients, the TEP score increased in blood prior to radiological tumor progression (Data S1), potentially indicating that a blood test may precede clinical and/or radiological tumor evolution. Although the correlation varied in individual cases, these results do indicate that the TEP score could possibly be employed for glioblastoma therapy monitoring.

False positive tumor progression remains a notorious challenge for therapy response monitoring using imaging modalities. To further investigate the potential value of the TEP-spliced RNA profiles following tumor resection and successful anti-tumor treatment, we questioned whether the TEP score would allow detection of false positive progression in glioblastoma treatment. Hence, we stratified all TEP samples from patients with glioblastoma collected during treatment follow-up into patients who had tumor progression at the time of blood collection (true progressive, n = 42), no tumor progression (stable disease or partial response [non-progressive], n = 40), or false positive progression (n = 20), scored according to the RANO criteria.<sup>24</sup> False positive tumor progression was determined based on longitudinal clinical observation and combined-modality MRI, including perfusion MRI, but was not confirmed by tissue biopsy. We observed a decrease in average TEP score in patients in whom MRI enhancement was suspected for false positive progression compared with those with suspected true progressive tumors (mean true progressive [n = 42], 0.55; mean false positive



**Figure 2. TEP RNA Signatures for Glioblastoma Therapy Monitoring and (False Positive) Progression Analysis**

(A) Boxplot of the TEP score (classification strength as output by the thromboSeq software) of glioblastoma at the moment of first tumor resection ( $n = 89$ ), glioblastoma follow-up ( $n = 151$ ), and asymptomatic healthy control ( $n = 353$ ) samples classified in the “glioblastoma versus asymptomatic healthy controls” algorithm. Classification of glioblastoma follow-up samples ( $n = 151$ ) results in significantly reduced TEP scores compared with glioblastoma samples collected at the moment of first tumor resection ( $n = 89$ ). Per boxplot, the median, IQR, and  $1.5 \times$  IQR (whiskers) are shown.

(B) Boxplot of the TEP score before and at the first time point after tumor resection (pre-surgery,  $n = 70$ ; post-surgery,  $n = 48$ ), indicating reduced TEP scores following (partial) tumor removal. Per boxplot, the median, IQR, and  $1.5 \times$  IQR (whiskers) are shown.

(C and D) TEP score plotted during the therapy course, indicated as days since primary tumor resection, for patients VU438 (C) and VU488 (D), connected by a straight line. The MRI images acquired at each time point are shown at the top of the graph. Radiological evaluation of tumor growth is indicated below each MRI image.

(E) ROC curve of the “progressive versus non-progressive” digitalSWARM classifier, including the combined training plus evaluation series (gray), verification series (red), and validation series (blue). Indicated are sample series sizes, best accuracy, and AUC value.

See also Figure S2 and S3 and Table S2F.

progression [ $n = 20$ ] 0.29;  $p = 0.001$ ; independent Student’s  $t$  test; Figure S2B), and comparable with those with a non-progressive tumor (mean non-progressive [ $n = 40$ ], 0.34;  $p = 0.003$ ; Figure S2B). We next determined whether a dedicated

classification algorithm capable of distinguishing progressors from non-progressors could be established; i.e., patients with stable disease, tumor regression, or false-positive progression. Unfortunately, for this particular analysis of tumor progression,

the conventional particle swarm optimization (PSO)-enhanced thromboSeq classifier performed suboptimally, with an AUC of 0.71 (Figure S2C), in contrast to the other classifications showing AUCs of 0.81–0.97 for detection of glioblastoma (Figures 1B–1E). Hence, we decided to develop a variant of the classical PSO-enhanced thromboSeq SVM algorithm, which we called the digitalSWARM classification algorithm (STAR Methods). Briefly, this algorithm iteratively improves the biomarker RNA panel by binary including or excluding RNAs highly ranked according to ANOVA statistics and summarizes the analysis as a TEP score for tumor progression monitoring (Figure S3A). In addition, it randomly selects multiple training and evaluation series, enabling the algorithm to select the most optimal combination of training samples for independent verification and validation (Figure S3B). Of note, this algorithm resulted in similar outcomes for the glioblastoma diagnostics classifiers (Figure S3C). Next the digitalSWARM algorithm was used to determine a biomarker panel of 267 platelet RNAs, reaching an AUC of 0.86 in the validation series of the “progressors versus non-progressors” algorithm. In summary, the combined training and evaluation series ( $n = 62$ ) reached an AUC of 0.83 (95% CI, 0.73–0.94), the verification series ( $n = 20$ ) an AUC of 0.89 (95% CI, 0.73–1.00), and the validation series ( $n = 20$ ) an AUC of 0.86 (95% CI, 0.70–1.00;  $p < 0.012$ ; Figure 2E; Figure S3D; Table S2F). Importantly, the four patients with false positive progression included in the validation series were classified correctly as non-progressors. Gene Ontology analysis of the RNAs included in this spliced RNA biomarker panel enhanced in patients with tumor progression was associated with negative regulation of B cell proliferation and microglial cell activation (Figure S3E). Although Gene Ontology analyses should be interpreted with caution regarding platelet RNA profiles, as opposed to nucleated cells, these results potentially indicate that the platelet-immune axis is involved in glioblastoma tumor progression. Hence, we conclude that patients with false positive progression might be efficiently discriminated from those with true tumor progression by employing TEP RNA signatures, although large-scale external validation remains warranted.

## DISCUSSION

Blood is a promising biosource to acquire molecular information regarding the presence and molecular make-up of a tumor.<sup>1</sup> So far, it remains difficult to identify traces from brain tumors in blood.<sup>8</sup> Our results show that patients with glioblastoma have markedly altered TEP-spliced RNA profiles that enable high-accuracy classification compared with TEP-spliced RNA profiles from asymptomatic healthy controls and patients with neuro-inflammatory or other (neuro)oncological conditions. This indicates that the measured platelet-spliced RNA profiles are not merely a result of aspecific platelet activation. We confirmed our previous observation<sup>18</sup> that platelets collected from patients with brain metastases do contain differential spliced RNA profiles compared with those from patients with glioblastoma. Finally, we provide evidence that the glioblastoma fingerprint in TEPs may gradually regress following tumor resection but recurs during tumor progression, but the number of samples analyzed so far is limited.

Previously, it has been shown that increased pre-operative and post-chemo-radiotherapy platelet counts in patients with glioblastoma are associated with poorer survival.<sup>25–27</sup> Also, glioblastoma cells induce platelet aggregation by expression of podoplanin, provoking venous thrombo-embolisms.<sup>28</sup> Furthermore, platelets contain many growth factors that can be released upon activation, providing a pro-tumoral microenvironment.<sup>14,29,30</sup> Hence, platelets may participate in progression of glioblastoma. Platelets are a potential anucleated source for RNA-based disease detection in blood.<sup>18–20,31–34</sup> They can be easily isolated in any clinical laboratory and contain high-quality RNA. Despite their anucleated state, platelets have the capability to inherit, splice, and translate their pre-mRNA content,<sup>35–37</sup> but from the data collected in this study it cannot be deduced what effect the platelet RNA alterations may have on the platelets’ proteome.

In the presence of a tumor, platelets appear to be “educated” by the tumor environment. This may be achieved by (1) transcriptional alterations in the megakaryocyte; (2) differential splicing of platelet RNA upon external queues<sup>36–39</sup> derived from tumor, stromal, (peri)vascular, and/or immune cells; (3) sequestration of spliced (tumor-derived) RNAs;<sup>20,31</sup> (4) platelet RNA alternative splicing events; or (5) the presence and evolution of differential platelet subpopulations in blood. Of particular interest here is tumor-bone marrow cross-talk, in which the glioblastoma tumor may influence bone marrow- or lung-resident megakaryocytes<sup>40</sup> (for instance, via cytokines or extracellular vesicles), which may alter the transcriptional profile of megakaryocytes;<sup>41</sup> however, direct biological evidence is lacking, as far as we are aware. In addition, megakaryocytes actively allocate RNA molecules into platelets,<sup>35</sup> and therefore these megakaryocytes may actively influence the platelet RNA profiles. Because the platelet RNA content is tightly maintained with input from external events,<sup>42</sup> local or systemic conditions might influence platelet RNA composition as well. Additional studies are required to determine which part of the observed platelet RNA profiles is subject to changes during thrombopoiesis and which part is acquired while in circulation.

Educated platelets may, in return, supply the tumor with pro-angiogenic factors<sup>14,43</sup> or regulatory microRNA (miRNA) molecules.<sup>44</sup> Because patients with a primary brain tumor could be differentiated from those with metastatic brain cancer, it may well be that the educational profiles are primarily derived from the primary non-central nervous system (CNS) tumor, but an effect of additional non-CNS metastases on the platelet-spliced RNA profiles cannot be excluded. Because we were able to distinguish patients with glioblastoma from those with the neuro-inflammatory condition multiple sclerosis, this dataset may indicate that the cancer signature is at least partly of tumorigenic origin and less likely due to peri-tumor inflammatory conditions only. Furthermore, platelets might be used to distinguish tumefactive MS from glioblastoma or primary CNS lymphoma. The gradual regression of the glioblastoma fingerprint in the platelet RNA samples collected following tumor resection and during the course of treatment supports this notion. It remains to be investigated why certain platelet RNA profiles from patients with glioblastoma included in the follow-up sample series show continuing TEP score reduction (Data S1), whereas their clinical course, supported by MRI, indicates tumor progression.

An advantage of the thromboSeq RNA sequencing (RNA-seq) platform is the multiplexed readout of ~4,500 potential biomarkers profiled at once. The software tool enables selection of most contributing genes by employing swarm intelligence, leveraging the synergistic data in these RNA repertoires. The thromboSeq algorithm has shown high-accuracy performance with multiple tumor types,<sup>18,19</sup> including neuro-oncological subtypes. It can be expected that, by including significantly more samples in the training process, the algorithms will be trained even more accurately, raising the overall classification performance. We provide evidence that the TEP-spliced RNA profiles are dynamic during tumor treatment. Glioblastomas cannot be fully eradicated; therefore, we cannot determine whether the platelet profile will completely normalize when the tumor is treated radically with chemoradiotherapy management and/or tumor resection. The exact “half-life” of the tumor signal in TEPs remains unknown and cannot be assessed in this tumor type. We observed a significant reduction in TEP score following tumor resection; on average, these samples were collected 19 (min-max, 10–33) days after surgery. This suggests that after the platelet pool is replaced, the TEP signal is also reduced. Like other liquid biopsies, these dynamic profiles possibly identify treatment failure before standard imaging techniques do.<sup>45–47</sup> This potentially enables TEP RNA analysis to be used for therapy monitoring, perhaps also with other tumor types and therapy conditions.

Although the platelet pellets are isolated using a standardized protocol based on differential centrifugation, we cannot exclude that a part of the platelet RNA-derived signal is caused by co-isolation of other cell fragments of similar size. Contamination of remaining nucleated blood cells (e.g., leukocytes) in the platelet isolations cannot be ruled out and may at least partially influence the observed RNA profiles. Also, despite the fact that we aimed to minimize the effect of potential confounding factors by age and whole-blood storage time matching, residual (at present unknown) variables may still confound the identified classification accuracy. Follow-up studies should investigate the biological mechanism underlying the educational process of TEPs in patients with brain tumors. Also, the potential role of glioblastoma—platelet—megakaryocyte signaling should be investigated. As discussed previously, in patients with glioblastoma, platelet counts are associated with prognosis during adjuvant chemo-radiotherapy treatment.<sup>48</sup> Also, the differential splicing patterns in platelets collected from patients with lower-grade glioma compared with those with glioblastoma and brain metastasis, is of interest, potentially providing blood-based opportunities to predict tumor progression of lower-grade toward higher-grade glioma. In addition, molecular tumor characteristics (e.g., IDH status, EGFR amplification, and other molecular characteristics) may be represented in TEP-spliced RNA profiles and deserve further investigation. Especially IDH status has shown effects on the coagulation status of glioma patients.<sup>49</sup> Unfortunately, the number of samples with a known genetic tissue profile to classify patients into a glioma molecular subgroup were too low to draw any statistically valid conclusions, which is a limitation of this study, even though 4 patients with confirmed mutant IDH status were mostly classified correctly in the platelet-based predictions. Follow-up

analysis should include an integrated morphological and molecular diagnosis.

Finally, integration of the TEP score together with MRI analysis may enhance detection and clinical management of patients with false positive glioblastoma progression. Prior to introduction of the TEP-based blood test in neuro-oncological clinics, thorough evaluation of pre-analytical variables and sample processing standardization as well as large-scale validation of TEP-spliced RNA profiles for brain tumor diagnostics and therapy monitoring are required.

### Limitations of Study

One limitations of this study is the low sample numbers, especially for detection of false positive progression; more samples would lead to a more robust prediction algorithm. Second, a head-to-head comparison with MRI modalities used to make false positive progression more or less likely is missing. Furthermore, more imaging and blood collection time points are necessary to determine whether treatment failure can be identified in TEPs before standard imaging techniques do. Another limitation is the unknown genetic profile of most GBM samples; especially IDH status could influence the RNA profile of platelets. Although several hypotheses of platelet education are discussed, the exact mechanism of platelet education remains unclear. All of these are important aspects that will need to be addressed in future studies.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Study design and sample selection
  - Clinical data annotation
- **METHOD DETAILS**
  - Blood processing and platelet isolation
  - Total RNA isolation, SMARTer amplification, and Tru-seq library preparation
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - Processing of raw RNA-sequencing data
  - Support Vector Machine (SVM)-based algorithm development and particle swarm optimization (PSO)-driven SVM-parameter optimization
  - Performance measurement of the PSO-enhanced thromboSeq algorithm
  - digitalSWARM classification algorithm
  - Gene ontology analysis

### SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.xcrm.2020.100101>.



## ACKNOWLEDGMENTS

Financial support was provided by European Research Council E8626 (to R.J.A.N. and T.W.) and 336540 and 713727 (to T.W.), Dutch Organisation of Scientific Research 93612003 and 91711366 (to T.W.), the Dutch Cancer Society (to T.W. and J.C.R.), Stichting STOPherentumoren.nl (to N.S., P.W., J.C.R., M.G.B., and T.W.), NIH/NCI CA176359 and CA069246 (to B.A.T.), CFF Norrland (to R.J.A.N.), and the Swedish Research Council (to R.J.A.N.). We are thankful to Dirk van Essen, Paul Eijk, Reno Bladergroen, Mirte Muller, Jillian Bracht, Michelle Esenkbrink, Leon Wils, Francesca Favaro, Jillian D. Schoonhoven, Michal Hegar, Saskia C. de Jager, Rolf T. Urbanus, Imo E. Hoefler, Gerard Pasterkamp, Daan van den Broek, Adrianus J. de Langen, Egbert F. Smit, Michel M. van den Heuvel, M.C.M. Kouwenhoven, and Henk M. Verheul for support and Sebastiaan van de Sand (SIT B.V.) for computational resources.

## AUTHOR CONTRIBUTIONS

N.S., M.G.B., J.C.R., and T.W. designed the study and wrote the manuscript. N.S., M.T., C.L., J.T., B.Y., J.K., F.M., S.I., A.C.N., W.P.J.L., A.H., B.M., D.P.N., W.P.V., R.J.A.N., B.A.T., P.W., and J.C.R. provided clinical samples and data. N.S., A.V., F.R., P.S., H.V., E.P., L.E.W., J.T., M.G.B., and T.W. performed sample processing for RNA-seq. N.S., S.G.J.G.i.t.V., A.V., P.W., B.A.T., M.G.B., and T.W. performed data analyses and interpretation. All authors provided critical comments on the manuscript.

## DECLARATION OF INTERESTS

M.G.B., R.J.A.N., and T.W. are inventors on relevant patent applications. T.W. and R.J.A.N. received funding from Illumina and are shareholders of GRAIL, Inc.

Received: November 14, 2019

Revised: April 23, 2020

Accepted: September 10, 2020

Published: October 1, 2020

## REFERENCES

- Crowley, E., Di Nicolantonio, F., Loupakis, F., and Bardelli, A. (2013). Liquid biopsy: monitoring cancer-genetics in the blood. *Nat. Rev. Clin. Oncol.* *10*, 472–484.
- Loo, H.K., Mathen, P., Lee, J., and Camphausen, K. (2019). Circulating biomarkers for high-grade glioma. *Biomark Med.* *13*, 161–165.
- Best, M.G., Sol, N., Zijl, S., Reijneveld, J.C., Wesseling, P., and Wurdinger, T. (2015). Liquid biopsies in patients with diffuse glioma. *Acta Neuropathol.* *129*, 849–865.
- Kros, J.M., Mustafa, D.M., Dekker, L.J.M., Sillevius Smitt, P.A.E., Luiders, T.M., and Zheng, P.-P. (2015). Circulating glioma biomarkers. *Neuro-oncol.* *17*, 343–360.
- Weller, M., van den Bent, M., Hopkins, K., Tonn, J.C., Stupp, R., Falini, A., Cohen-Jonathan-Moyal, E., Frappaz, D., Henriksson, R., Balana, C., et al. (2014). EANO guideline for the diagnosis and treatment of anaplastic gliomas and glioblastoma. *Lancet Oncol.* *15*, e395–e403.
- Westphal, M., and Lamszus, K. (2015). Circulating biomarkers for gliomas. *Nat. Rev. Neurol.* *11*, 556–566.
- Miller, A.M., Shah, R.H., Pentsova, E.I., Pourmaleki, M., Briggs, S., Distefano, N., Zheng, Y., Skakodub, A., Mehta, S.A., Campos, C., et al. (2019). Tracking tumour evolution in glioma through liquid biopsies of cerebrospinal fluid. *Nature* *565*, 654–658.
- Bettgowda, C., Sausen, M., Leary, R.J., Kinde, I., Wang, Y., Agrawal, N., Bartlett, B.R., Wang, H., Luber, B., Alani, R.M., et al. (2014). Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci. Transl. Med.* *6*, 224ra24.
- Boisselier, B., Gállego Pérez-Larraya, J., Rossetto, M., Labussière, M., Ciccarino, P., Marie, Y., Delattre, J.-Y., and Sanson, M. (2012). Detection of IDH1 mutation in the plasma of patients with glioma. *Neurology* *79*, 1693–1708.
- Macarthur, K.M., Kao, G.D., Chandrasekaran, S., Alonso-Basanta, M., Chapman, C., Lustig, R.A., Wileyto, E.P., Hahn, S.M., and Dorsey, J.F. (2014). Detection of brain tumor cells in the peripheral blood by a telomerase promoter-based assay. *Cancer Res.* *74*, 2152–2159.
- Sullivan, J.P., Nahed, B.V., Madden, M.W., Oliveira, S.M., Springer, S., Bhere, D., Chi, A.S., Wakimoto, H., Rothenberg, S.M., Sequist, L.V., et al. (2014). Brain Tumor Cells in Circulation are Enriched for Mesenchymal Gene Expression. *Cancer Discov.* *4*, 1299–1309.
- Muller, C., Holtschmidt, J., Auer, M., Heitzer, E., Lamszus, K., Schulte, A., Matschke, J., Langer-Freitag, S., Gasch, C., Stoupiec, M., et al. (2014). Hematogenous dissemination of glioblastoma multiforme. *Sci. Transl. Med.* *6*, 247ra101.
- Weller, M., van den Bent, M., Tonn, J.C., Stupp, R., Preusser, M., Cohen-Jonathan-Moyal, E., Henriksson, R., Le Rhun, E., Balana, C., Chinot, O., et al.; European Association for Neuro-Oncology (EANO) Task Force on Gliomas (2017). European Association for Neuro-Oncology (EANO) guideline on the diagnosis and treatment of adult astrocytic and oligodendroglial gliomas. *Lancet Oncol.* *18*, e315–e329.
- McAllister, S.S., and Weinberg, R.A. (2014). The tumour-induced systemic environment as a critical regulator of cancer progression and metastasis. *Nat. Cell Biol.* *16*, 717–727.
- Labelle, M., Begum, S., and Hynes, R.O. (2011). Direct signaling between platelets and cancer cells induces an epithelial-mesenchymal-like transition and promotes metastasis. *Cancer Cell* *20*, 576–590.
- Schumacher, D., Strilic, B., Sivaraj, K.K., Wettschureck, N., and Offermanns, S. (2013). Platelet-derived nucleotides promote tumor-cell trans-endothelial migration and metastasis via P2Y2 receptor. *Cancer Cell* *24*, 130–137.
- Kerr, B.A., McCabe, N.P., Feng, W., and Byzova, T.V. (2013). Platelets govern pre-metastatic tumor communication to bone. *Oncogene* *32*, 4319–4324.
- Best, M.G., Sol, N., Kooi, I., Tannous, J., Westerman, B.A., Rustenburg, F., Schellen, P., Verschueren, H., Post, E., Koster, J., et al. (2015). RNA-Seq of Tumor-Educated Platelets Enables Blood-Based Pan-Cancer, Multiclass, and Molecular Pathway Cancer Diagnostics. *Cancer Cell* *28*, 666–676.
- Best, M.G., Sol, N., In 't Veld, S.G.J.G., Vancura, A., Muller, M., Niemeijer, A.L.N., Fejes, A.V., Tjon Kon Fat, L.A., Huis In 't Veld, A.E., Leurs, C., et al. (2017). Swarm Intelligence-Enhanced Detection of Non-Small-Cell Lung Cancer Using Tumor-Educated Platelets. *Cancer Cell* *32*, 238–252.e9.
- Nilsson, R.J.A., Balaj, L., Hulleman, E., van Rijn, S., Pegtel, D.M., Walraven, M., Widmark, A., Gerritsen, W.R., Verheul, H.M., Vandertop, W.P., et al. (2011). Blood platelets contain tumor-derived RNA biomarkers. *Blood* *118*, 3680–3683.
- Best, M.G., In 't Veld, S.G.J.G., Sol, N., and Wurdinger, T. (2019). RNA sequencing and swarm intelligence-enhanced classification algorithm development for blood-based disease diagnostics using spliced blood platelet RNA. *Nat. Protoc.* *14*, 1206–1234.
- Stupp, R., Hegi, M.E., Mason, W.P., van den Bent, M.J., Taphoorn, M.J., Janzer, R.C., Ludwin, S.K., Allgeier, A., Fisher, B., Belanger, K., et al. (2009). Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol.* *10*, 459–466.
- George, J.N. (2000). Platelets. *Lancet* *355*, 1531–1539.
- Wen, P.Y., Chang, S.M., Van den Bent, M.J., Vogelbaum, M.A., Macdonald, D.R., and Lee, E.Q. (2017). Response Assessment in Neuro-Oncology Clinical Trials. *J. Clin. Oncol.* *35*, 2439, 1449.
- Brockmann, M.A., Giese, A., Mueller, K., Kaba, F.J., Lohr, F., Weiss, C., Gottschalk, S., Nolte, I., Leppert, J., Tuettenberg, J., et al. (2007).

- Preoperative thrombocytosis predicts poor survival in patients with glioblastoma. *Neuro-oncol.* 9, 335–342.
26. Nolte, I., Przibylla, H., Bostel, T., Groden, C., and Brockmann, M.A. (2008). Tumor-platelet interactions: glioblastoma growth is accompanied by increasing platelet counts. *Clin. Neurol. Neurosurg.* 110, 339–342.
  27. Boonyawan, K., Hess, K.R., Yang, J., Long, L., Wang, Q., Ezhilarasan, R., Aulia, A., Alfaro-Munoz, K.D., de Groot, J.F., Bhat, K.P., and Sulman, E.P. (2017). A relative increase in circulating platelets following chemoradiation predicts for poor survival of patients with glioblastoma. *Oncotarget* 8, 90488–90495.
  28. Riedl, J., Preusser, M., Nazari, P.M.S., Posch, F., Panzer, S., Marosi, C., Birner, P., Thaler, J., Brostjan, C., Lötsch, D., et al. (2017). Podoplanin expression in primary brain tumors induces platelet aggregation and increases risk of venous thromboembolism. *Blood* 129, 1831–1839.
  29. Menter, D.G., Tucker, S.C., Kopetz, S., Sood, A.K., Crissman, J.D., and Honn, K.V. (2014). Platelets and cancer: a casual or causal relationship: revisited. *Cancer Metastasis Rev.* 33, 231–369.
  30. Sierko, E., and Wojtukiewicz, M.Z. (2004). Platelets and Angiogenesis in Malignancy. *Semin. Thromb. Hemost.* 30, 95–108.
  31. Nilsson, R.J.A., Karachaliou, N., Berenguer, J., Gimenez-Capitan, A., Schellen, P., Teixido, C., Tannous, J., Kuiper, J.L., Drees, E., Grabowska, M., et al. (2016). Rearranged EML4-ALK fusion transcripts sequester in circulating blood platelets and enable blood-based crizotinib response monitoring in non-small-cell lung cancer. *Oncotarget* 7, 1066–1075.
  32. Tjon-Kon-Fat, L.A., Lundholm, M., Schröder, M., Wurdinger, T., Thellenberg-Karlsson, C., Widmark, A., Wikström, P., and Nilsson, R.J.A. (2018). Platelets harbor prostate cancer biomarkers and the ability to predict therapeutic response to abiraterone in castration resistant patients. *Prostate* 78, 48–53.
  33. Calverley, D.C., Phang, T.L., Choudhury, Q.G., Gao, B., Oton, A.B., Weyant, M.J., and Geraci, M.W. (2010). Significant downregulation of platelet gene expression in metastatic lung cancer. *Clin. Transl. Sci.* 3, 227–232.
  34. Takagi, S., Tsukamoto, S., Park, J., Johnson, K.E., Kawano, Y., Moschetta, M., Liu, C.-J., Mishima, Y., Kokubun, K., Manier, S., et al. (2018). Platelets enhance Multiple Myeloma progression via IL-1 $\beta$  upregulation. *Clin. Cancer Res.* 24, 2430–2439.
  35. Cecchetti, L., Tolley, N.D., Michetti, N., Bury, L., Weyrich, A.S., and Gressle, P. (2011). Megakaryocytes differentially sort mRNAs for matrix metalloproteinases and their inhibitors into platelets: a mechanism for regulating synthetic events. *Blood* 118, 1903–1911.
  36. Denis, M.M., Tolley, N.D., Bunting, M., Schwartz, H., Jiang, H., Lindemann, S., Yost, C.C., Rubner, F.J., Albertine, K.H., Swoboda, K.J., et al. (2005). Escaping the nuclear confines: signal-dependent pre-mRNA splicing in anucleate platelets. *Cell* 122, 379–391.
  37. Nassa, G., Giurato, G., Cimmino, G., Rizzo, F., Ravo, M., Salvati, A., Nyman, T.A., Zhu, Y., Vesterlund, M., Lehtiö, J., et al. (2018). Splicing of platelet resident pre-mRNAs upon activation by physiological stimuli results in functionally relevant proteome modifications. *Sci. Rep.* 8, 498.
  38. Rondina, M.T., Schwartz, H., Harris, E.S., Kraemer, B.F., Campbell, R.A., Mackman, N., Grissom, C.K., Weyrich, A.S., and Zimmerman, G.A. (2011). The septic milieu triggers expression of spliced tissue factor mRNA in human platelets. *J. Thromb. Haemost.* 9, 748–758.
  39. Schwartz, H., Tolley, N.D., Foulks, J.M., Denis, M.M., Risenmay, B.W., Buerke, M., Tilley, R.E., Rondina, M.T., Harris, E.M., Kraiss, L.W., et al. (2006). Signal-dependent splicing of tissue factor pre-mRNA modulates the thrombogenicity of human platelets. *J. Exp. Med.* 203, 2433–2440.
  40. Lefrançois, E., Ortiz-Muñoz, G., Caudrillier, A., Mallavia, B., Liu, F., Sayah, D.M., Thornton, E.E., Headley, M.B., David, T., Coughlin, S.R., et al. (2017). The lung is a site of platelet biogenesis and a reservoir for haematopoietic progenitors. *Nature* 544, 105–109.
  41. Pang, L., Weiss, M.J., and Poncz, M. (2005). Megakaryocyte biology and related disorders. *J. Clin. Invest.* 115, 3332–3338.
  42. Alhasan, A.A., Izuogu, O.G., Al-Balool, H.H., Steyn, J.S., Evans, A., Colzani, M., Ghevaert, C., Mountford, J.C., Marenah, L., Elliott, D.J., et al. (2016). Circular RNA enrichment in platelets is a signature of transcriptome degradation. *Blood* 127, e1–e11.
  43. Klement, G.L., Yip, T.-T., Cassiola, F., Kikuchi, L., Cervi, D., Podust, V., Italiano, J.E., Wheatley, E., Abou-Slaybi, A., Bender, E., et al. (2009). Platelets actively sequester angiogenesis regulators. *Blood* 113, 2835–2842.
  44. Michael, J.V., Wurtzel, J.G.T., Mao, G.F., Rao, A.K., Kolpakov, M.A., Sabri, A., Hoffman, N.E., Rajan, S., Tomar, D., Madesh, M., et al. (2017). Platelet microparticles infiltrating solid tumors transfer miRNAs that suppress tumor growth. *Blood* 130, 567–580.
  45. Kallergi, G., Konstantinidis, G., Markomanolaki, H., Papadaki, M.A., Mavroudis, D., Stoumaras, C., Georgoulas, V., and Agelaki, S. (2013). Apoptotic circulating tumor cells in early and metastatic breast cancer patients. *Mol. Cancer Ther.* 12, 1886–1895.
  46. Wu, X.L., Tu, Q., Faure, G., Gallet, P., Kohler, C., and Bittencourt, Mde.C. (2016). Diagnostic and Prognostic Value of Circulating Tumor Cells in Head and Neck Squamous Cell Carcinoma: a systematic review and meta-analysis. *Sci. Rep.* 6, 20210.
  47. Wong, R., Tie, J., Lee, M., Cohen, J., Wang, Y., Li, L., Ma, S., Christie, M., Kosmider, S., Tomasetti, C., et al. (2019). The potential role of circulating tumor DNA (ctDNA) in the further investigation of colorectal cancer patients with nonspecific findings on standard investigations. *Int. J. Cancer* 145, 540–547.
  48. Williams, M., Liu, Z.W., Woolf, D., Hargreaves, S., Michalarea, V., Menashy, R., Kooner, I., and Wilson, E. (2012). Change in platelet levels during radiotherapy with concurrent and adjuvant temozolomide for the treatment of glioblastoma: a novel prognostic factor for survival. *J. Cancer Res. Clin. Oncol.* 138, 1683–1688.
  49. Unruh, D., Schwarze, S.R., Khoury, L., Thomas, C., Wu, M., Chen, L., Chen, R., Liu, Y., Schwartz, M.A., Amidei, C., et al. (2016). Mutant IDH1 and thrombosis in gliomas. *Acta Neuropathol.* 132, 917–930.
  50. Wen, P.Y., Macdonald, D.R., Reardon, D.A., Cloughesy, T.F., Sorensen, A.G., Galanis, E., Degroot, J., Wick, W., Gilbert, M.R., Lassman, A.B., et al. (2010). Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. *J. Clin. Oncol.* 28, 1963–1972.
  51. Lin, N.U., Lee, E.Q., Aoyama, H., Barani, I.J., Barboriak, D.P., Baumert, B.G., Bendszus, M., Brown, P.D., Camidge, D.R., Chang, S.M., et al. (2015). Response assessment criteria for brain metastases: proposal from the RANO group. *Lancet Oncol.* 16, e270–e278.
  52. van den Bent, M., Wefel, J., Schiff, D., Taphoorn, M., Jaeckle, K., Junck, L., Armstrong, T., Choucair, A., Waldman, A., Gorlia, T., et al. (2011). Response assessment in neuro-oncology (a report of the RANO group): assessment of outcome in trials of diffuse low-grade gliomas. *Lancet Oncol.* 12, 583–593.
  53. Ellingson, B.M., Wen, P.Y., and Cloughesy, T.F. (2017). Modified Criteria for Radiographic Response Assessment in Glioblastoma Clinical Trials. *Neurotherapeutics* 14, 307–320.
  54. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120.
  55. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21.
  56. Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological Samples		
805 blood platelet samples	This study	Table S1
Chemicals, Peptides, and Recombinant Proteins		
RNALater solution	Thermo Scientific	Cat# AM7020
Critical Commercial Assays		
mirVana miRNA isolation kit	Ambion, Thermo Scientific	Cat# AM1560
SMARTer Ultra Low RNA Kit for Illumina Sequencing version 3	Clontech	Cat# 634853
Truseq Nano DNA Sample Prep Kit	Illumina	Cat# FC-121-4001
RNA picochip and reagents, Bioanalyzer 2100	Agilent	Cat# 5067-1513
DNA 7500 chip and reagents, Bioanalyzer 2100	Agilent	Cat# 5067-1506
DNA High Sensitivity chip and reagents, Bioanalyzer 2100	Agilent	Cat# 5067-4626
Deposited Data		
Raw and processed RNA-seq data	This study	GEO: 156902
Software and Algorithms		
Trimmomatic (version 0.22)	Bolger et al. <sup>54</sup>	<a href="http://www.usadellab.org/cms/?page=trimmomatic">http://www.usadellab.org/cms/?page=trimmomatic</a>
STAR (version 2.3.0)	Dobin et al. <sup>55</sup>	<a href="https://github.com/alexdobin/STAR">https://github.com/alexdobin/STAR</a>
HTSeq (version 0.6.1)	Anders et al. <sup>56</sup>	<a href="https://www-huber.embl.de/HTSeq/doc/overview.html">https://www-huber.embl.de/HTSeq/doc/overview.html</a>
Picardtools (version 1.115)	Broad Institute, USA	<a href="https://broadinstitute.github.io/picard/">https://broadinstitute.github.io/picard/</a>
Samtools (version 1.115)		<a href="http://samtools.sourceforge.net">http://samtools.sourceforge.net</a>
MATLAB (version R2015b)	The MathWorks Inc., USA	<a href="https://nl.mathworks.com/products/matlab.html">https://nl.mathworks.com/products/matlab.html</a>
R (version 3.3.0)		<a href="https://www.r-project.org">https://www.r-project.org</a>
R-studio (version 0.99.902)		<a href="https://rstudio.com/">https://rstudio.com/</a>
Bioconductor package edgeR (version 3.12.1)		<a href="https://bioconductor.org/packages/release/bioc/html/edgeR.html">https://bioconductor.org/packages/release/bioc/html/edgeR.html</a>
Bioconductor package EDASeq (version 2.4.1)		<a href="http://bioconductor.org/packages/release/bioc/html/EDASeq.html">http://bioconductor.org/packages/release/bioc/html/EDASeq.html</a>
Bioconductor package PPSO (version 0.9-9991)		<a href="https://github.com/TillF/ppso">https://github.com/TillF/ppso</a>
Bioconductor package RUVSeq (version 1.4.0)		<a href="http://bioconductor.org/packages/release/bioc/html/RUVSeq.html">http://bioconductor.org/packages/release/bioc/html/RUVSeq.html</a>
R-package e1071 (version 1.6-7)	CRAN	<a href="https://cran.r-project.org/web/packages/e1071/index.html">https://cran.r-project.org/web/packages/e1071/index.html</a>
R-package Optunity (version 1.0)	STADIUS lab	<a href="https://optunity.readthedocs.io/en/latest/">https://optunity.readthedocs.io/en/latest/</a>
R-package pROC (version 1.8)	CRAN	<a href="https://cran.r-project.org/web/packages/pROC/index.html">https://cran.r-project.org/web/packages/pROC/index.html</a>
R-package ROCR (version 1.0-7)	CRAN	<a href="https://cran.r-project.org/web/packages/ROCR/index.html">https://cran.r-project.org/web/packages/ROCR/index.html</a>
ThromboSeq algorithm v1.4		<a href="https://github.com/MyronBest/thromboSeq_source_code">https://github.com/MyronBest/thromboSeq_source_code</a>

### RESOURCE AVAILABILITY

#### Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact Thomas Wurdinger ([t.wurdinger@amsterdamumc.nl](mailto:t.wurdinger@amsterdamumc.nl)).

## Materials Availability

This study did not generate new unique reagents.

## Data and Code Availability

The raw sequencing data reported in this paper has been deposited into the NCBI GEO database under accession number 156902. All employed bioinformatics software and code can be found online ([https://github.com/MyronBest; thromboSeq\\_source\\_code\\_v1.4](https://github.com/MyronBest; thromboSeq_source_code_v1.4) [to be released]).

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### Study design and sample selection

Peripheral whole blood was drawn by venipuncture from brain cancer patients, patients with multiple sclerosis, and asymptomatic healthy individuals at the various medical institutions in Europe and the USA. Whole blood was collected in 4-, 6-, or 10-mL EDTA-coated BD Vacutainers. Cancer patients were diagnosed by clinical, radiological and pathological examination, and were confirmed to have detectable tumor tissue load at the time of blood collection. Age- and whole blood storage time-matching was performed retrospectively, iteratively matching samples by excluding and including patients with glioblastoma at the time of first tumor resection, patients with brain metastasis, patients with multiple sclerosis, and asymptomatic healthy controls, aiming at a similar median age and age-range between groups. A detailed overview of the included samples, demographic characteristics, the hospital of origin, time between blood collection and platelet isolation (whole blood storage time), as well as an overview for which analyses and classifiers the samples were used is provided in [Table S1](#). Asymptomatic healthy controls were at the moment of blood collection, or previously, not diagnosed with cancer, but were not subjected to additional tests confirming the absence of cancer. None of the patients with multiple sclerosis had a malignancy at the moment of blood collection. This study was conducted in accordance with the principles of the Declaration of Helsinki. Approval for this study was obtained from the institutional review board and the ethics committee at each participating hospital. Each participant signed informed consent. Clinical follow-up of asymptomatic healthy controls is not available due to anonymization of these samples according to the ethical rules of the hospitals.

### Clinical data annotation

For collection and annotation of clinical data, patient records were manually queried for demographic and clinical variables, i.e., age, gender, type of tumor, metastases, details of current and prior treatments, survival rates, and co-morbidities. Blood samples from patients with glioblastoma were collected at the day of surgery (i.e., before surgery) and/or during treatment, respectively baseline and follow-up samples. During the follow-up period blood was preferably collected at the time of, or within weeks, of follow-up MRIs and always before a new treatment session. Sample collection was discontinued at the time of tumor progression, or because of any other reason through which a patient had become lost to follow-up (e.g., withdrawal of informed consent, discontinuation of treatment, continuation of treatment in another hospital, or death). Treatment response assessment of patients was done by MR-imaging and was performed by experienced neuroradiologists according to the updated RANO criteria, and scored as progressive disease (PD), stable disease (SD), partial response (PR), or complete response (CR).<sup>50–52</sup> False-positive progression was determined by subsequent MR-imaging,<sup>53</sup> including multiple MR modalities and perfusion MRI. A tissue biopsy is not required since tissue biopsies cover only a very limited area of the brain tumor and are therefore prone to sampling error. In combination with the additional surgical risk for the patient it is generally not used to confirm glioblastoma false-positive progression. All clinical data was anonymized and stored in a secured database.

## METHOD DETAILS

### Blood processing and platelet isolation

Whole blood samples in 4-, 6-, or 10-mL EDTA-coated BD Vacutainer tubes were processed using standardized protocols within 48 hours as described previously.<sup>18–21</sup> Whole blood collected at the Amsterdam UMC location VUmc, the Utrecht Medical Center, the Medical University of Vienna, the Radboud University Medical Center Nijmegen, the University Hospital of Maastricht, and the Netherlands Cancer Institute was subjected to platelet isolation within 12 hours after blood collection. Whole blood samples collected at Massachusetts General Hospital Boston and at the Amsterdam UMC location AMC were stored overnight and processed after 24 hours. Platelet pellets were isolated as described previously.<sup>18,19,21</sup> To isolate platelets, platelet rich plasma (PRP) was separated from nucleated blood cells by a 20-minute 120xg centrifugation step, after which the platelets were pelleted by a 20-minute 360xg centrifugation step. In order to reduce the risk of leukocyte contamination pelleted in the buffy coat, removal of 9/10<sup>th</sup> of the PRP has to be performed carefully. The remaining leukocyte yield with this isolation method is ~1 to 5 leukocytes per 1 million platelets.<sup>18</sup> Centrifugations were performed at room temperature. Finally, platelet pellets were carefully resuspended in RNAlater (Life Technologies) and after overnight incubation at 4°C frozen at –80°C. Platelet pellets in RNAlater can be stored at least for up to five years at –80°C while maintaining high-quality.

### Total RNA isolation, SMARTer amplification, and Truseq library preparation

Preparation of samples for sequencing was performed in batches as described previously,<sup>18,19,21</sup> and included a mixture of clinical conditions per batch. For platelet RNA isolation, frozen platelets were thawed on ice and total RNA was isolated using the mirVana miRNA isolation kit (Ambion, Thermo Scientific, AM1560). Platelet RNA was eluted in 30  $\mu$ L elution buffer. We evaluated the platelet RNA quality using the RNA 6000 Picochip (Bioanalyzer 2100, Agilent), and included, as a quality standard for subsequent experiments, only platelet RNA samples with a RIN-value > 7 and/or distinctive rRNA curves. All Bioanalyzer 2100 quality and quantity measures were collected from the automatically generated Bioanalyzer result reports using default settings, and after critical assessment of the reference ladder (quantity, appearance, and slope). High-quality samples were subjected to cDNA synthesis and amplification using the SMARTer Ultra Low RNA Kit for Illumina Sequencing v3 (Clontech, cat. nr. 634853). Prior to amplification, all samples were diluted to  $\sim$ 500 pg/ $\mu$ L total RNA and again the quality was determined and quantified using the Bioanalyzer Picochip. For samples with a stock yield below 400 pg/ $\mu$ L, a volume of two or more microliters of total RNA (up to  $\sim$ 500 pg total RNA) was used as input for the SMARTer amplification. Quality control of amplified cDNA was measured using the Bioanalyzer 2100 with DNA High Sensitivity chip (Agilent). All SMARTer cDNA synthesis and amplifications were performed together with a negative control, which was required to be negative by Bioanalyzer analysis. Samples with detectable fragments in the 300-7500 bp region were selected for further processing. For labeling of platelet cDNA for sequencing, all amplified platelet cDNA was first subjected to nucleic acid shearing by sonication (Covaris Inc) and subsequently labeled with single index barcodes for Illumina sequencing using the Truseq Nano DNA Sample Prep Kit (Illumina, cat nr. FC-121-4001). To account for the low platelet cDNA input concentration, all bead clean-up steps were performed using a 15-minute bead-cDNA binding step and a 10-cycle enrichment PCR. All other steps were according to manufacturer's protocol. Labeled platelet DNA library quality and quantity was measured using the DNA 7500 chip or DNA High Sensitivity chip (Agilent). High-quality samples with product sizes between 300-500 bp were pooled (12-19 samples per pool) in equimolar concentrations for thromboSeq and submitted for 100 bp Single Read sequencing on the Illumina HiSeq 2500 platform using version 4 sequencing reagents.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Processing of raw RNA-sequencing data

Raw RNA-seq data of platelets encoded in FASTQ-files were subjected to a standardized RNA-seq alignment pipeline, as described previously.<sup>18,19,21</sup> In summary, RNA-seq reads were subjected to trimming and clipping of sequence adapters by Trimmomatic (v. 0.22),<sup>54</sup> mapped to the human reference genome (hg19) using STAR (v. 2.3.0),<sup>55</sup> and summarized using HTSeq (v. 0.6.1), which was guided by the Ensembl gene annotation version 75.<sup>56</sup> All subsequent statistical and analytical analyses were performed in R (version 3.3.2) and R-studio (version 0.99.903). Of samples that yielded less than  $0.2 \times 10^6$  intron-spanning reads in total after sequencing, we again sequenced the original Truseq preparation of the sample and merged the read counts generated from the two individual FASTQ-files after HTSeq count summarization (performed for  $n = 31$  samples). Genes encoded on the mitochondrial DNA and the Y chromosome were excluded from downstream analyses. Sample filtering was performed by assessing the library complexity, which is partially associated with the intron-spanning reads library size. First, we excluded the genes that yielded < 30 intron-spanning reads in > 90% of the cohort for all platelet samples that were sequenced. This resulted in this platelet RNA-seq library in 4,487 different genes detected with sufficient coverage. For each sample, we quantified the number of genes for which at least one intron-spanning read was mapped, and excluded samples with < 750 detected genes. Following, we performed a leave-one-sample-out cross-correlation analysis to exclude platelet samples that show low intersample correlation (Pearson's correlation threshold: 0.4). Finally, we excluded platelet RNA samples with median logCPM < 3, resulting in a final data series of  $n = 805$  samples ( $n = 353$  asymptomatic healthy controls,  $n = 240$  glioblastoma,  $n = 126$  brain metastasis, and  $n = 86$  multiple sclerosis; [Figure 1A](#); [Table S1](#)). To prevent potential plasma DNA from contributing to our computational platelet RNA analyses, we only selected spliced intron-spanning RNA reads.

### Support Vector Machine (SVM)-based algorithm development and particle swarm optimization (PSO)-driven SVM-parameter optimization

The PSO-enhanced algorithm was extensively described previously.<sup>19,21</sup> Briefly, the algorithm employs a training and evaluation series for gene panel selection and algorithm development, of which selection parameters are optimized by PSO. For our classifiers we employed 100 particles with 6 iterations, and optimized four steps of the generic classification algorithms, i.e., (i) the iterative correction module threshold used for selection of genes identified as stable genes among the library size, (ii) the FDR-threshold included in the differential splicing filter applied to the results of the likelihood ANOVA test, (iii) the exclusion of highly correlated genes selected after the likelihood ANOVA test, and (iv) number of genes passing the recursive feature elimination (RFE)-algorithm. Predefined ranges were submitted to the PSO-algorithm for every classification task presented in this study. To identify a potential algorithm threshold at which the maximum number of patients with glioblastoma may be identified with a predefined specificity for asymptomatic healthy controls, we geared the algorithm readout in the evaluation series toward 95% specificity. The software code automatically applies the algorithm predictive strength threshold selected in the evaluation series to the validation series.

### Performance measurement of the PSO-enhanced thromboSeq algorithm

All classification experiments were performed with the PSO-enhanced thromboSeq algorithm, using parameters optimized by PSO. All random selection procedures were performed using the *sample*-function as implemented in R. For assignment of samples per series to the training and evaluation series, only the number of samples per clinical group was balanced, whereas other potentially contributing variables were not stratified at this stage (assuming random distribution among the groups). Performance of the training series was assessed by a leave-one-out cross validation approach (LOOCV, see also Best et al.<sup>18</sup>). The list of stable genes among the initial training series, determined RUV-factors for removal, and final gene panel determined by swarm-optimization of the training-evaluation series were used as input for the LOOCV procedure. As a control for internal reproducibility, we randomly sampled training and evaluation series, while maintaining the validation series and the swarm-guided gene panel of the original classifier, and performed 1000 training and classification procedures. As a control for random classification, class labels of the samples used by the SVM-algorithm for training of the support vectors were randomly permuted ( $n = 1000$  iterations), while maintaining the swarm-guided gene list of the original classifier. P values were calculated accordingly, as described previously.<sup>18</sup> Results were presented in receiver operating characteristics (ROC)-curves, and summarized using area under the curve (AUC)-values, as determined by the ROCR-package in R. AUC 95% confidence intervals were calculated according to the method of Delong using the *ci.auc*-function of the pROC-package in R.

### digitalSWARM classification algorithm

DigitalSWARM in essence iteratively and digitally (binary) selects RNAs to be included in the biomarker panel. In addition, digital-SWARM enables for inclusion of a verification series that allows for selection of multiple and ultimately the best combination of samples assigned to the training and evaluation series, a potential pitfall of the conventional PSO-enhanced thromboSeq classifier.<sup>21</sup> The algorithm performs the following steps of which the full dataset is employed as input data; first, random selection of  $n$  training and evaluation series ( $n = 200$  by default), with only one locked evaluation and one locked validation series. Then, the algorithm analyzes a grid of multiple potential confounding factor (e.g., library size) thresholds for RUV-correction. For each setting an ANOVA-comparison employing the training series is performed and the potential confounding factor setting of which the lowest ANOVA false discovery rate could be achieved is selected as the most optimal RUV-correction threshold. Next, a preliminary biomarker RNA panel is selected (by default  $p$  value  $< 0.1$ ,  $\log\text{CPM} > 3$ , no expression of RNAs encoded on the Y chromosome) separately for RNAs with increased or decreased numbers of reads mapping to splice junctions. Following, for both biomarker RNA panels, a particle swarm optimization (PSO)-algorithm iteratively enables for in- or exclusion ('1' or '0') of RNAs (default number of particles: 100, default number of iterations: 50). The performance of each biomarker RNA panel proposed by PSO is tested by calculating the median expression level of all samples in the evaluation series, progressive and non-progressive separately, and performing an independent Student's  $t$  test comparison. The ultimately best biomarker spliced RNA panel, either with in- or decreased RNAs, is employed for calculating a companion TEP score. This score is calculated by the median  $\log_2$  normalized counts of either biomarker RNA panel and for the progressive and non-progressive samples separately. For the training, evaluation, and verification series an area-under-the curve (AUC)-value is calculated, and for all iterations and each biomarker panel selected the best verification series AUC is selected. Employing this particular training and evaluation series and biomarker RNA panel the validation series is classified and summarized. 30% of the samples in the dataset were assigned to both the training and evaluation series, whereas 20% was assigned to both the verification and validation series. The specificity of the biomarker RNA panel is assessed by randomly shuffling the labels of the training and evaluation series samples before the procedure ( $n = 1000$  times), thereby requiring at least 20% of the sample labels to be different as compared to the original sample label.

### Gene ontology analysis

Gene ontology analysis was performed on the 5<sup>th</sup> of September 2019 with the PantherDB-database, performing gene ontology analysis for biological processes (<http://www.pantherdb.org>). PANTHER overrepresentation test was selected, employing the *Homo sapiens* reference list ( $n = 20,996$  genes), which were compared to RNAs with enhanced spliced junction reads in the 'non-progressive'-group ( $n = 126$ ) and RNA with enhanced spliced junction reads in the 'progressive'-group ( $n = 131$ ) in the 'progressive versus non-progressive'-algorithm (Figure 2E). Fisher's exact test was employed to calculate statistical significance and the top five hits sorted by enrichment score were plotted in barplots.