



Universiteit  
Leiden

The Netherlands

## **Biomarkers for the response to immunotherapy in patients with non-small cell lung cancer**

Muller, M.

### **Citation**

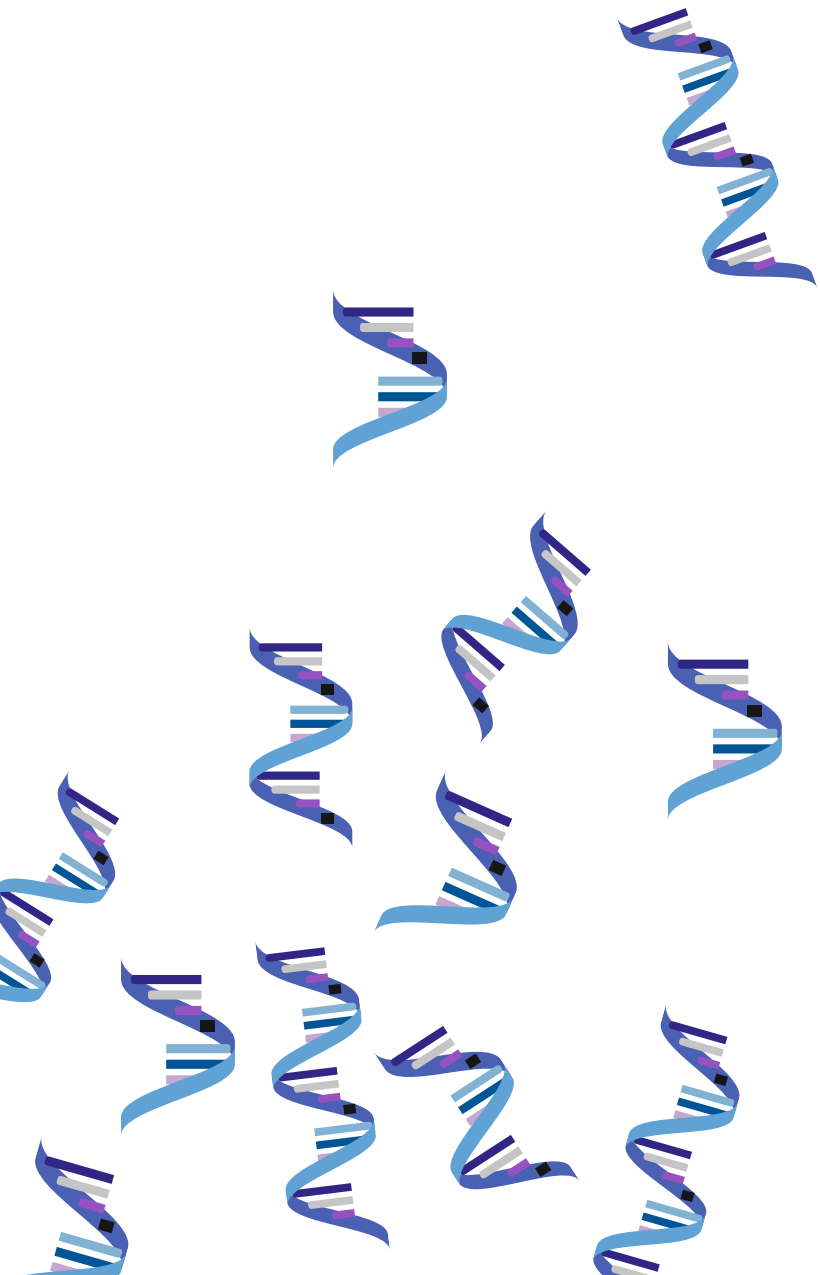
Muller, M. (2024, May 29). *Biomarkers for the response to immunotherapy in patients with non-small cell lung cancer*. Retrieved from <https://hdl.handle.net/1887/3754842>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3754842>

**Note:** To cite this publication please use the final published version (if applicable).



# 3

## Blood platelet RNA profiles do not enable for nivolumab response prediction at baseline in patients with non-small-cell lung cancer

Mirte Muller\*, Myron G. Best\*, Vincent van der Noort, T. Jeroen N. Hiltermann, Anna-Larissa N. Niemeijer, Edward Post, Nik Sol, Sjors G.J.G. In 't Veld, Tineke Nogaede, Lianne Visser, Robert Schouten, Daan van den Broek, Karlijn Hummelink, Kim Monkhorst, Adrianus J. de Langen, Ed Schuurin, Egbert Smit, Harry J.M. Groen, Thomas Wurdinger, Michel M. van den Heuvel

*Tumor Biology* 46 (2024) S327-S340

## Abstract

**Background:** Anti-PD-(L)1 immunotherapy has emerged as a promising treatment approach for non-small-cell lung cancer (NSCLC), though the response rates remain low. Pre-treatment response prediction may improve patient allocation for immunotherapy. Blood platelets act as active immune-like cells, thereby constraining T-cell activity, propagating cancer metastasis, and adjusting their spliced mRNA content.

**Objective:** We investigated whether platelet RNA profiles before start of nivolumab anti-PD1 immunotherapy may predict treatment responses.

**Methods:** We performed RNA-sequencing of platelet RNA samples isolated from stage III-IV NSCLC patients before treatment with nivolumab. Treatment response was scored by the RECIST-criteria. Data were analyzed using a predefined thromboSeq analysis including a particle-swarm-enhanced support vector machine (PSO/SVM) classification algorithm.

**Results:** We collected and processed a 286-samples cohort, separated into a training/evaluation and validation series and subjected those to training of the PSO/SVM-classification algorithm. We observed only low classification accuracy in the 107-samples validation series (area under the curve (AUC) training series: 0.73 (95%-CI: 0.63-0.84, n=88 samples), AUC evaluation series: 0.64 (95%-CI: 0.51-0.76, n=91 samples), AUC validation series: 0.58 (95%-CI: 0.45-0.70, n=107 samples)), employing a five-RNA biomarker panel.

**Conclusions:** We concluded that platelet RNA may have minimally discriminative capacity for anti-PD1 nivolumab response prediction, with which the current methodology is insufficient for diagnostic application.

**Keywords:** Blood platelets, RNA-sequencing, nivolumab, immunotherapy

## Introduction

Lung cancer, especially non-small cell lung cancer (NSCLC), is the most common cause of cancer-related mortality in the world, causing over one million deaths annually [66]. Interestingly, PD-(L)1 checkpoint inhibitors have shown remarkable effectiveness, durability of response and tolerability [158], although overall response rates remain low (~20%) [3, 5]. Numerous studies have identified correlations between smoking, tumor tissue mutational load, infiltration of CD8-positive T-cells and response to anti-PD-(L)1 immunotherapy [159, 160]. At the moment, current stratification of patients for anti-PD-(L)1 targeted therapy is hampered by limited accuracy and concordance of available biomarkers, including the currently used tumor tissue PD-L1-protein expression marker. Identification of patients with a low likelihood of response to anti-PD-(L)1 immunotherapy, while still correctly identifying individuals who most likely benefit from this therapy, may prevent unnecessary treatment and concomitant costs, and potential exposure of patients to serious immunological adverse events.

Blood platelets are cell fragments involved in hemostasis, initiation of wound healing and metastasis of cancer [161, 162]. They originate from the megakaryocytes that reside in the bone marrow and lung parenchyma [161, 163], and are loaded with premature RNA molecules [164-166]. Upon environmental cues, platelets splice their mRNA molecules resulting in platelet RNA repertoires of both spliced and unspliced mRNAs [59, 164, 167]. In addition, platelets have the capability to sequester RNAs [168] and load platelet microparticles with RNAs that can be transferred to other cells, such as endothelial and tumor cells [169, 170]. Functionally, platelet-derived TGF $\beta$  and direct physical platelet-tumor cell interactions can induce tumor cell epithelial-mesenchymal transition [171], a process that has previously also been associated with resistance to anti-PD1 immunotherapy [63]. In addition, platelets can behave as immunomodulators in inflammatory conditions [172, 173], regulate immune cell extravasation [174, 175], and cross-communicate with lymphocytes [176] and neutrophils [177]. It was shown that platelets armed with immunotherapeutic antibodies home to lung tumors in mice [178], exhibiting profound anti-tumor activity. Also, previously pretreatment increased platelet-to-lymphocyte ratio in patients with metastatic NSCLC treated with nivolumab was associated with worse overall survival and a reduced response rate [179]. Platelets are thus potentially involved in the immune response towards tumor cells, and might thereby alter their RNA composition.

We recently provided evidence that platelets may serve as blood-based liquid biopsies for the detection and subclassification of cancer [59, 61, 168], as well as for therapy response monitoring [180, 181]. In this prospective study, we investigated whether platelet RNA signatures may provide classification power for nivolumab immunotherapy response prediction before start of treatment.

## Methods

### Study design

In this prospective, observational, multi-center study in patients with advanced (stage III-IV) NSCLC blood samples at moment of baseline for nivolumab treatment were collected. The aim was to optimally identify patients who respond to nivolumab, for which clinical imaging data was available as a gold standard for response evaluation. Sample size calculation was performed using the method of Dobbin and Simon [182], assuming an overall response rate of 17% [2], employing the results from a pilot study (data not shown). We estimated a required samples size for the training series of 180 patients (i.e. ~31 responders and ~149 non-responders) under assumption that the gene with maximal mean fold difference between groups will exhibit a fold change of 2.5. From a clinical perspective we expected a successful biomarker to have 95% sensitivity and at least 50% specificity, i.e. halving the number of unnecessarily treated patients while keeping the number of false negatives low. We define the new biomarker to be successful if the lower boundary of the 95%-confidence interval for the specificity lies above 38% at the threshold/cut-off where its specificity is 95%. The rationale for this is that at the 95% sensitivity level a specificity of 38% will yield an overall response rate (ORR) in the positively tested group of 24% which is the rate achieved by the tissue PD-L1-test [2] at the much lower sensitivity of 46%. A validation series of 90 patients (75 non-responders and 15 responders) has 90% power to achieve this assuming a true specificity of the test of 55%.

### Patient and sample collection

All patients with advanced (stage III/IV) NSCLC who were scheduled to start treatment with nivolumab were recruited for voluntary inclusion in this study. Patients were included in the Amsterdam UMC, Amsterdam, the Netherlands Cancer Institute (NKI/AvL), Amsterdam and the University Medical Center Groningen, Groningen, all from The Netherlands. Patients were included from December 2013 till June 2017, with follow-up until June 2018. The median time of follow-up was 195 days (Interquartile range (IQR) 97 – 399.5 days). All patients received treatment in accordance with recent literature [5, 183] and local guidelines. Platelet samples were collected up to one month before start of treatment, but not after start of treatment. Other clinical data, including computed tomography (CT) scans, blood tests and relevant clinical data such as the use of medication used in this study were collected for routine clinical practice. These were handled according to the Dutch Personal Data Protection Act (WvPb). Response assessment of patients treated with nivolumab was performed by CT-imaging at baseline, 6-8 weeks, 3 months and 6 months after start of treatment. Treatment response was assessed according to the updated RECIST version 1.1 criteria and was scored as progressive disease (PD), stable disease (SD), partial response (PR), or complete response (CR) [105, 184]. A responder was defined as those patients with PE or CR at three months, despite possible PD at six months. A patient was also regarded a responder with SD at three and six months.

Tumor tissue was stained using either the PD-L1 22C3 (Agilent) or PD-L1 28-8 (Abcam) antibodies and scored according to the manufacturer's instructions. This study was

conducted in accordance with the principles of the Declaration of Helsinki, and all participant provided written informed consent. Our study was approved by the local medical ethical committee (PTC NKI-AvL, Amsterdam, NL45524.031.13; medical ethics committee, UMCG, Groningen, 2010/109 [185]; Dutch trial register NL7839) and the patient privacy committee, and was performed according to the institutional patient privacy protocols.

### **Platelet isolation**

Peripheral whole blood was drawn by regular venipuncture in EDTA-coated Vacutainer tubes. Whole blood samples in 6- or 10-mL EDTA-coated Vacutainer tubes were processed using standardized protocols within 12 hours as described previously [168, 186, 187] at room temperature (21°C). Here, using a 20-minute 120xg centrifugation step platelet rich plasma (PRP) was separated from nucleated blood cells. Following, 9/10<sup>th</sup> of the PRP needs to be collected carefully without disturbing the buffy coat, with the risk of contaminating the platelet preparation with nucleated cells. Then the platelets were pelleted from the PRP by a 20-minute 360xg centrifugation step. Platelet pellets were carefully resuspended in RNAlater (Life Technologies), followed by an overnight incubation at 4°C frozen at -80°C.

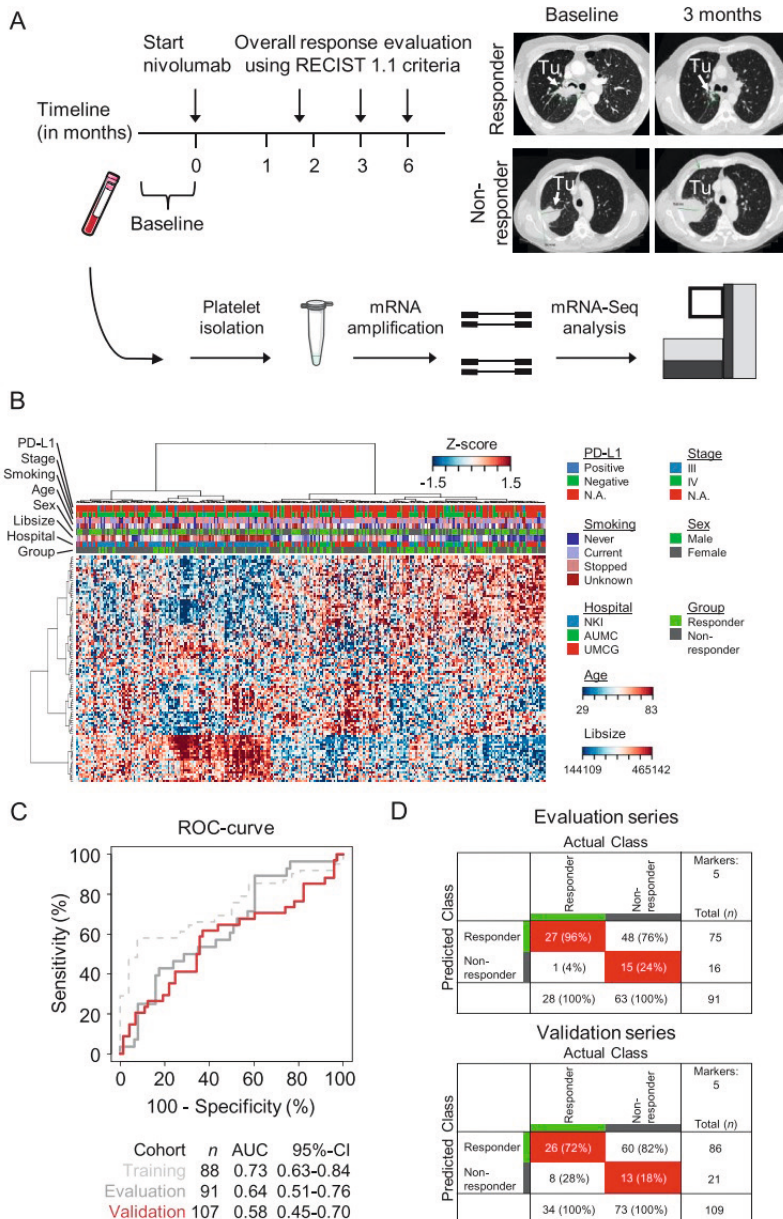
### **Platelet RNA-sequencing**

The thromboSeq wet-lab platelet mRNA-sequencing protocol and dry-lab software modules have recently been described in detail and are publicly available [186]. In brief, first platelet total RNA was isolated using the mirVana miRNA isolation kit (Ambion, Thermo Scientific, AM1560). All samples were subjected to cDNA synthesis and amplification using the SMARTer Ultra Low RNA Kit for Illumina Sequencing v3 (Clontech, cat. nr. 634853) with an input of 500 picogram of total RNA. All amplified platelet cDNA was 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). All bead clean-up steps were performed using a 15-minute bead-cDNA binding step and a 10-cycle enrichment PCR. Other steps were according to manufacturer's protocol. During this protocol, multiple quality control steps were performed using RNA and DNA chips available for the Bioanalyzer 2100 (Agilent). High-quality samples with product sizes between 300-500 bp were pooled (12-24 samples per pool) in equimolar concentrations for superficial thromboSeq and submitted for 100 bp Single Read sequencing on the Illumina HiSeq 2500 and 4000 platform. During whole blood, blood platelet and platelet RNA processing, the researchers were blinded to the clinical outcome measures.

### **Processing of raw RNA-sequencing data**

We employed Trimmomatic (v. 0.22) [188] for trimming and clipping of sequence adapters from the RNA-seq reads, of which the resulting reads were mapped to the human reference genome (hg19) using STAR (v. 2.3.0) [189]. The mapped reads were summarized using HTSeq (v. 0.6.1), guided by the Ensembl gene annotation version 75 [190]. All following statistical and analytical analyses were performed in R (version 3.3.0) and R-studio (version 0.99.902). Genes encoded on the mitochondrial DNA and the Y-chromosome were excluded from downstream analyses. To circumvent potential

contamination from cell-free DNA, only intron-spanning spliced RNA reads were employed for data analysis [60]. Sample filtering was performed by assessing the library complexity, which is partially associated with the intron-spanning reads library size. The raw sequencing data has been deposited at the GEO database under accession number GSE216297.



**Figure 1 – Platelet RNA-based nivolumab response prediction.**



### ThromboSeq classification algorithm

We employed our previously standardized and optimized thromboSeq algorithm for data analysis [59-61], which is extensively described elsewhere. For this study, software-version v1.5 of the thromboSeq source code was employed, and is for research purposes publicly available via GitHub ([https://github.com/MyronBest/thromboSeq\\_source\\_code\\_v1.5](https://github.com/MyronBest/thromboSeq_source_code_v1.5)) [61]. In short, the particle-swarm-optimization (PSO-) enhanced support vector machine (SVM)-based thromboSeq algorithm employs an SVM-algorithm for RNA expression classification tasks, supplemented by a PSO algorithm that iteratively selects parameters for gene panel selection. The thromboSeq algorithm is developed with the use of a training, evaluation, and independent validation series. The training series is used for selection of stable genes that are used for remove unwanted variation (RUV)-mediated data normalization, gene panel selection and SVM-training. Here, RUV is employed to minimize variability among the samples introduced by potential confounding factors [60]. This correction measure is determined based on the training series to enable for independent validation of remaining samples. The gene panel is selected using ANOVA-statistics, supplemented by filtering for highly correlated RNAs as well as filtering for RNAs most contributive to the SVM-model. Based on this selected gene panel a SVM-model including class weight correction is trained. Next, the evaluation series is employed to evaluate the classification performance of the SVM-model with the particular gene panel and algorithm settings. Here, a PSO-algorithm is used to repeatedly provide more optimized thresholds to select the gene panel and set other algorithm settings, of which the optimization steps are evaluated using the classification scores in the evaluation series. Finally the algorithm and gene panel is locked following which the validation series is used for validation of the PSO-optimized SVM-classification algorithm.

### Figure 1 – Platelet RNA-based nivolumab response prediction. (continued)

(A) Schematic overview of the experimental setup. Blood of patients eligible for treatment with PD-1 inhibitor nivolumab was included from one month before till start of treatment (baseline, t=0). Tumor response read-out based on CT-imaging and according to the RECIST 1.1 criteria were performed at 6-8 weeks, 3 months, and 6 months, after start of nivolumab therapy. A responder was defined as PR, CR at three months and SD at three and six months. A non-responder was defined as PD and SD after three, but PD after six months.

(B) Heatmap and unsupervised hierarchical clustering analysis of RNAs with differential RNA levels between nivolumab responders (n=88) and non-responders (n=198). Columns indicate samples, rows indicate RNAs, and color intensity represents the Z-score transformed RNA expression values. Clustering of samples showed non-random partitioning ( $p < 0.003$ , Fisher's exact test). The clinical and technical variables 'hospital of sample location', 'library size', 'sex', 'age', 'smoking status', 'tumor stage', and 'PD-L1 tumor tissue status' are indicated on top of the heatmap with each owns color coding (legend on the right).

(C) Receiver operating characteristics curve of the thromboSeq nivolumab response prediction algorithm at moment of treatment baseline. Indicated are the training series (dashed gray line), evaluation series (gray line), and validation series (red line). Also indicated are sample numbers, AUC values, and the 95% confidence intervals (95%-CI).

(D) 2x2 tables summarizing the classification results for the evaluation series (upper table) and validation series (lower table) of the thromboSeq nivolumab response prediction algorithm at moment of treatment baseline geared towards a rule-out setting. Indicated are sample numbers and percentage of total samples.

For this study, the variables 'Age' and 'lib.size' were applied as potential confounding factors for RUV-correction with both a default threshold of 0.8. The swarm-parameters 'lib.size', 'FDR', 'correlatedTranscripts', and 'rankedTranscripts' were employed using the following boundaries, respectively: -0.1 – 1.0; 5 – number of genes with ANOVA FDR <0.5; 0.5 – 1.0; 5 – number of genes with ANOVA FDR <0.5. In total 100 particles were deployed with eight iterations. Class-weights and a rule-out optimization step forcing towards at least 95% sensitivity in the evaluation series during particle optimization were enabled. The algorithm employed has selected the following thresholds: 'lib.size': 0.131965039293101; 'FDR': 5; 'correlatedTranscripts': 0.989771223666639; 'rankedTranscripts': 12.2922697840516.

An ANOVA comparison was performed using the `thromboSeq.anova`-function in the thromboSeq software. The variables 'Age' and 'lib.size' were subjected to RUV-correction, with the thresholds '0.8' and '0.8', and with `k.variables=3`. The heatmap clustering was optimized by a PSO-function thereby adjusting the FDR-threshold evaluated by column-dendrogram clustering.

### Statistical analyses

SPSS (v25; SPSS, Chicago, USA) was used for the descriptive statistics. For a between-group-comparison the Mann-Whitney U test, two sample unpaired Students t-test, or chi-square test was used depending on the variable.

## Results

We prospectively collected 300 blood samples from patients with advanced NSCLC that were selected for treatment with the PD-1 monoclonal antibody nivolumab. These were subjected to the thromboSeq protocol (Figure 1A). 14 samples (4%) were excluded from follow-up analyses according to previously introduced thromboSeq quality measures (Supplemental Figure 1) [60, 186]. The remaining 286 samples were included for the analyses (Table 1). We observed no statistically significant differences between responders and non-responders regarding age, gender, and hospital of patient inclusion (Table 1). We also observed no statistical significant differences in total platelet counts between responders and non-responders of whom this data was available ( $p=0.660$ ,  $n=185$ , unpaired Student's t-test, Table 1). Tumor tissue was partially available for correlative PD-L1 immunohistochemistry analysis. Additional demographic variables are presented in Table 1 and Table S1. The full dataset was subdivided into training ( $n=88$ ), evaluation ( $n=91$ ), and validation series ( $n=107$ ; Supplemental Figure 1C), as previously performed according to our thromboSeq dry-lab protocol [186]. For comparative analysis, we forced the patients whom the tumor tissue PD-L1 status was available into the validation series. The hospital of origin (i.e. sample collection) was equally distributed among the three series (Table S2). The separation of samples among the three series was defined in the study protocol and based on preliminary data and power calculations (see Methods).

We first performed an ANOVA comparison on all platelet RNA samples, identifying platelet RNAs with discriminative power between nivolumab responders and non-responders. From the 3750 platelet RNAs that were detected in the dataset in total,

27 had statistically significant differential RNA levels (False Discovery Rate (FDR)<0.05, Table S3). From these RNAs, only one RNA from the GSEA inflammatory response gene data set [191] had a statistically significant differential expression level level (NAMPT, FDR = 0.0002).

Heatmap dendrogram clustering of the samples applying these 27 RNAs resulted in non-significant clustering between responders and non-responders (p-value: 0.24, Fisher's exact test, data not shown). Adjusting the ANOVA-statistics significance threshold employing a PSO-function (see Methods) resulted in dendrogram clustering discriminating responders from non-responders (p-value<0.003, Fisher's exact test, used FDR-threshold: 0.419696, Figure 1B). Improving discriminating power by adjusting the ANOVAs FDR-threshold implies that also statistically non-significant RNAs can contribute to the clustering accuracy. This provides the rationale for the PSO-enhanced classification algorithm development selecting an RNA panel based on most contributive power to a discriminative biomarker panel.

We next performed PSO-enhanced training of an SVM-classification algorithm, forcing the algorithm to identify optimal algorithm settings including the RNA biomarker panel for a rule-out nivolumab response prediction classifier. This would enable accurate identification of patients who do respond to nivolumab, while providing the opportunity to exclude patients that will not respond. This training process resulted into an area under the curve (AUC) for the training series of 0.73 (95%-CI: 0.63-0.84, n=88 samples, as determined by LOOCV), and for the evaluation series of 0.64 (95%-CI: 0.51-0.76, n=91 samples) for the evaluation series. Subsequent validation of the 107-samples validation series resulted into an AUC of 0.58 (95%-CI: 0.45-0.70, n=5 spliced RNA biomarker panel, Figure 1C). The five-gene biomarker panel as unbiasedly selected by the thromboSeq software was composed of the RNAs HPSE, HBD, PRSS50, CRYM, and LARP1 (Supplemental Figure 2A). Using the classification score readout according a 'rule-out' application, in a sensitivity of 96% (95%-CI: 82-100%, n=28) with a specificity of 24% (95%-CI: 14-36%, n=63) in the evaluation series, resulted in a concomitant sensitivity in the validation series of 76% (95%-CI: 59-89%, n=34) and a specificity of 18% (95%-CI: 10-29%, n=73, Figure 1D), with which our predefined threshold of a successful new biomarker is not reached.

## Discussion

Anti-PD-(L)1 therapies have revolutionized cancer treatment, although, (non-invasive) predictive biomarkers are highly desirable [192]. In this proof-of-concept study, we aimed to associate spliced platelet RNA profiles to nivolumab response prediction measures at moment of therapy baseline.

Platelets act as local systemic responders during tumorigenesis [162], thereby suffering from tumor-mediated platelet education, and altering its behavior [171, 174, 193]. In previous studies, we have demonstrated that platelet RNA can function as a biomarker to detect and classify cancer from blood [59, 61]. Using our previously developed thromboSeq platform, we could not find solid RNA repertoires associated with tumor

response at moment of start of therapy. Therefore, we were in this setting and study setup unable to successfully validate a nivolumab response prediction thromboSeq algorithm. Therefore at this point, interrogation of platelet RNA signatures before start of nivolumab treatment seems not to be useful for nivolumab therapy selection.

**Table 1 - Baseline characteristics of responders versus non-responders**

	<b>Non-responder N=198</b>	<b>Responder N=88</b>	<b>Total (N=286)</b>	<b>P-value</b>
<b>Patient</b>				
Male sex – No. (%)	115 (58.1)	43 (48.9)	158 (55.2)	0.15 <sup>A</sup>
Age (yr) – Mean (95% CI)	63.8 (62.5-65.1)	64.9 (62.8-67.1)	64.1 (63.0-65.3)	0.61 <sup>B</sup>
Smoking status – No. (%)				<b>0.05</b> <sup>C</sup>
Never	20 (10.1)	4 (4.5)	24 (8.4)	
Current	33 (16.7)	25 (28.4)	58 (20.3)	
Stop	132 (66.7)	57 (64.8)	189 (66.1)	
Pack Years – Mean (95% CI)	32.6 (29.6-35.6)	37.4 (33.8-41.1)	34.2 (31.9 – 36.5)	<b>0.01</b> <sup>B</sup>
WHO Performance status ≥ 2 – no / total known patients	31/182	2/84	33/266	<b>0.003</b> <sup>C</sup>
Platelet not normal range <sup>A*</sup> – no. / total known patients	27/128	7/57	34/185	0.15 <sup>A</sup>
Platelet Count Baseline – Mean (95% CI)	338.0 (314.5-361.6)	326.7 (295.4-357.9)	334.5(315.7-353.3)	0.66 <sup>B</sup>
Brain metastasis	43	17	60	0.65 <sup>A</sup>
<b>Pathology – no.</b>				
Adenocarcinoma	141	60	201	0.46 <sup>A</sup>
Squamous cell carcinoma	47	22	69	
Other	7	6	13	
Unknown	3	0	3	
<b>PD-L1 – no.</b>				
Negative	26	12	38	0.48 <sup>C</sup>
Positive 1 – 50%	10	5	15	
Positive >50%	3	4	7	
Unknown	159	67	226	
<b>Previous treatment – no.</b>				
CCRT	19	10	29	0.86 <sup>A</sup>
Platinum chemotherapy	109	47	156	
Unknown	67	29	96	
<b>Line of treatment – no.</b>				
1	2	1	3	0.58 <sup>A</sup>
2	101	43	144	
<2	25	14	39	
Unknown	70	30	100	
Previous RT / total known patients	87/127	34/58	121/185	0.24 <sup>A</sup>
Thoracic RT / total known patients	53/127	26/58	79/185	0.75 <sup>A</sup>

**Table 1 - Baseline characteristics of responders versus non-responders (Continued)**

	Non-responder N=198	Responder N=88	Total (N=286)	P-value
<b>Platelet related medication - no.</b>				
None	82	38	120	0.43 <sup>A</sup>
Acetylsalicylic acid	23	9	32	
Other	22	11	33	
Unknown	71	30	101	
<b>Hospital - no.</b>				
NKI	90	38	128	0.84 <sup>A</sup>
AUMC	39	20	59	
UMCG	69	30	99	
<b>Survival</b>				
OS – mean (95% CI)	165.3 (146.1- 184.55)	505.8 (465.7- 545.8)	270.4(244.7- 296.2)	<0.001 <sup>B</sup>
PFS – mean (95% CI)	72.6 (59.4-85.8)	455.8 (413.3- 498.4)	190.5 (164.5- 216.5)	<0.001 <sup>B</sup>

Abbreviations: N: Number of patients; no.: Number of patients; yr: years; 95% CI: 95% confidence interval; WHO performance-status score: World Health Organization performance status score. this is a score ranging from 0 to 5. where 0 indicates no symptom. 1 indicates mild symptoms and above 1 indicates greater disability; PD-L1: Programmed death ligand 1; CCRT: chemoradiation therapy; RT: radiation therapy; RECIST: Response evaluation criteria in Solid Tumors; NKI: Netherlands Cancer Institute. AUMC: Amsterdam University Medical Centers. UMCG: University Medical Center Groningen. OS: Overall Survival; PFS: progression free survival.\* Either below under limit of normal or above limit of normal. as defined by the clinical laboratory. In the Netherlands a normal platelet count is between 150-350 platelets x 10<sup>9</sup> / liter (150.000-350.000 platelets per microliter). Used statistical tests: <sup>A</sup>: Chi Square, <sup>B</sup>: One way ANOVA, <sup>C</sup>: Fisher-Freeman Halton exact test.

Subsequent analysis of classification accuracies according to PD-L1 tumor tissue status of whom the data was available showed on average increased classification scores for patients responding to nivolumab as compared to patients not responding to nivolumab (Supplemental Figure 2B).

Because platelets are regarded as local immuno-like responders, the role of platelets in anti-tumor activity prior to and during anti-PD-(L)1 therapy, including direct and indirect physical interactions with lymphocytes, macrophages, and tumor cells should not be underestimated. In our five-gene platelet biomarker panel, both HPSE (heparanase) and LARP1 (La RibonucleoProtein 1) have lower expression values in nivolumab responders. HPSE is active in tumor invasion and inflammatory processes [194, 195], where-as LARP1 is associated with dengue virus and coronavirus disease-19 (COVID-19) [196, 197]. Higher expression levels in nivolumab non-responders may suggest either a different inflammatory process required for a anti tumor response, or a highly progressive tumor, which may explain the insensitivity to nivolumab treatment.

In this light, the current results were unexpected. We have several explanations for our findings; 1) platelet RNA is only very subtly altered in the presence of immunotherapy-sensitive as opposed to immunotherapy-resistant NSCLC tumors for which insufficient sample numbers were available to uncover such differences Alternatively, the employed

machine learning algorithms and gene panel selection approach had insufficient discriminatory power; 2) the alterations are masked by the presence of multiple platelet subpopulations; 3) the alterations may be best measured after start of treatment, e.g. a couple of weeks following the first nivolumab exposure; 4) insufficient samples were included into the training process; 5) despite matching of several confounding factors, remaining (unmeasured) factors and/or pre-analytical variables may increase the gene expression noise as opposed to the (low-level) signal and 6) in the power calculation the estimated RNA expression difference was overestimated, resulting into a too small anticipated cohorts sample size and therefore to a reduced algorithm's performance. Also, we cannot exclude that platelet RNA is of value as a predictive biomarker platform for other (immuno) therapeutics. Here, re-analysis of the data using other machine learning approaches is of interest and the data is publicly available for such purposes.

### **Acknowledgements**

We are thankful to the NKI-AVL Core Facility Molecular Pathology and Biobanking (CFMPB) for supplying NKI-AVL Biobank material and lab support and Sebastiaan van de Sand (SIT B.V.) for computational resources. We also thank Maarten Slagter for his support. Financial support was provided by European Research Council E8626 (R.J.A.N., E.F.S., T.W.) and 336540 (T.W.), the Dutch Organisation of Scientific Research 93612003 and 91711366 (T.W.), the Dutch Cancer Society (T.W.), BMS IION (M.M.v.d.H., T.W.), and Stichting STOPHersentumoren.nl.

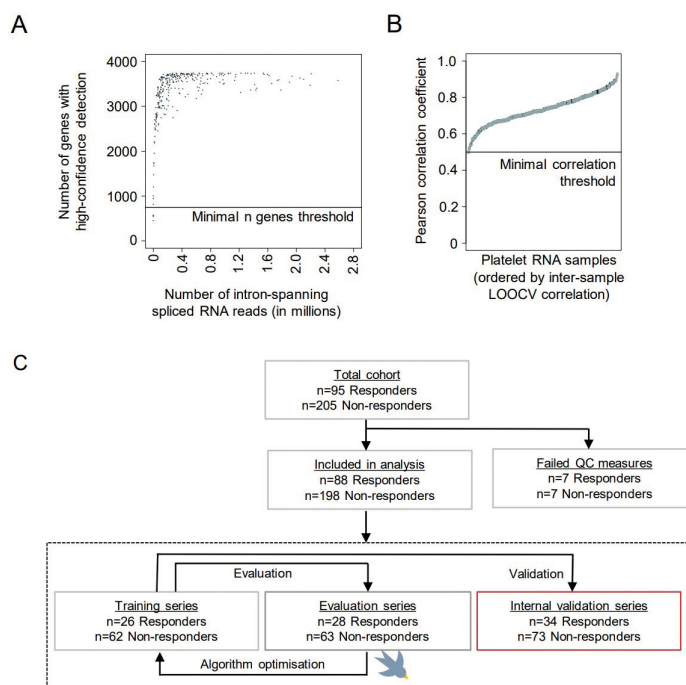
## Supplemental material

The supplemental material can be found at:



Included content in this thesis:

Figure S1 - ThromboSeq bioinformatics QCs



### Supplemental Figure 1 - thromboSeq bioinformatics QCs

(A) Number of intron-spanning spliced RNA reads versus number of genes detected. Samples with at least 1,500 different RNAs detected were included for further analyses.

(B) Leave-one-sample out cross-correlations, assessing the intersample comparability (Pearson correlation) based on all detected RNAs and filtering out samples with low inter-sample correlation (minimal correlation threshold >0.5).

(C) Schematic flow diagram indicating the number of samples at start of the analysis, following *in silico* QC-filtering, and available for algorithm training, evaluation and validation.