Universitat Autònoma de Barcelona

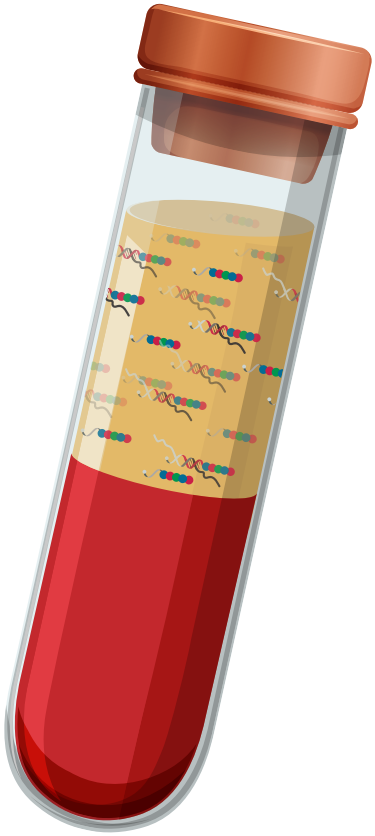
Validation of liquid biopsy-based analysis on the NanoString nCounter platform

Jillian W.P. Bracht-Loman

Validation of liquid biopsy-based analysis on the NanoString nCounter platform

"Finding your car's engine somewhere other than under the bonnet."

Dennis Lo



Jillian W.P. Bracht-Loman

Cover design: Pauline van Putten
Layout and design: Jillian W.P. Bracht
Printed by: Ipskamp printing, proefschriften.net

The work described in this thesis was part of the European Liquid Biopsies Academy (ELBA, https://elbaproject.eu/) which has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 765492.

Validation of liquid biopsy-based analysis on the NanoString nCounter platform

A dissertation submitted in fulfillment of the requirements for the PhD degree in Biochemistry, Molecular Biology and Biomedicine with the Mention of International Doctoral Research

by

Jillian Wilhelmina Paulina Bracht-Loman

Born on the 21st of January 1994
in Huizen, the Netherlands

Barcelona, February 2021

Universitat Autònoma
de Barcelona

The present doctoral thesis entitled Validation of liquid biopsy-based analysis on the NanoString nCounter platform, by Jillian W.P. Bracht-Loman, was carried out at the medical services company Pangaea Oncology, under the supervision of Dr. Miguel Angel Molina Vila, technical director of the Laboratory of Oncology, and Dr. Rafael Rosell from the Cancer Biology and Precision Medicine Program at the Catalan Institute of Oncology, Germans Trias i Pujol Health Sciences Institute and Hospital (IGTP), and the tutoring of Dr. Anna Maria Bassols Teixidó, from the Department of Biochemistry and Molecular Biology at the Universitat Autònoma de Barcelona (UAB).

Miguel Angel Molina Vila Rafael Rosell Anna Maria Bassols Teixidó

Barcelona, February 2021

What we lack in natural ability, we can make up for in discipline.

To my amazing Mother and Brother,

Hasta la luna y vuelta.

Acknowledgements

Usually the acknowledgements are saved for the end of the thesis. However, without the support of my colleagues, friends and family, it would have been impossible to perform the work presented in this thesis. Therefore I would like to start by thanking all of you. Anything is possible when you have the right people there to support you.

Dear **Miguel-Angel**, **Dr. Rosell**, **Niki** and **Anna Bassols**. I would like to thank you for your supervision, advice and for answering the millions of questions I asked in the past years. You gave me the freedom to design and manage my projects but provided support whenever I needed it. Without you this work would not have been possible, thank you!!

To the **members of the yearly thesis review** and the **examination board**, thank you for your willingness, your feedback and your help.

I also owe a big thank you to **all the amazing colleagues in Pangaea**. The friendly atmosphere in the lab, combined with the endless knowledge and Spanish lessons have made my Barcelona experience complete. I always loved going to work, and you are all the main reason for that. A lot has happened in the last four years, but it never affected the strong bond you have with each other. Specifically I want to thank **Jordi Berenguer** and **Ana Gimenez**. In the past years you have taught me everything I know now, and I cannot thank you enough for that! **Jordi**, the best lab buddy I ever had, although I still don´t know how to calculate dilutions your way (ups.. busted), I finally found my own way thanks to your perseverance. When I heard you were leaving Pangaea, I sort of panicked but you taught me well and I managed to find my way. We kept in touch over (Beata) beers and of course to discuss science and life in general, and I hope we can keep that up. **Ana**, you are like the RNA angel of the lab. Whatever I needed, you always knew the answer or where to find the answer. We have worked very closely in the last years, and I will miss you a lot! **Erika** and **Sonia**, thank you for all. You are my Sol o Sombra and Callejón o Paki amigas, and I really enjoyed this break time. As the pathology team, you also did a massive amount work for this thesis and other projects by cutting all the samples. And last but not least, **Clara** and **Miguel-Angel**. **Clara**, la madre Cubana del laboratorio, you are always ready to solve everyone's problems and never lose your smile while doing so. You´re one of the sweetest people I know, and I hope to see you sometime in Amsterdam (not on a barco ;)). **Miguel-Angel**, you stepped in and became my supervisor halfway through this PhD. Your knowledge, advice and perseverance have helped me to finish this thesis in a very short time but always keeping the quality in mind. Thank you!

The biggest learning school during my PhD was the close contact with the clinic. To all the oncologists I worked with, especially: **Dr. Rosell**, **Niki**, **Manolo**, **Maria**, **Juanjo**, **Andres**, **Carlos**, **Santi**, **Jano**, **Katy**, **Irene** & the **clinical trials team**. You have taught me to look at my research from a clinical- and even from a patient´s perspective. Everything we do in the lab is interesting, but it only becomes valuable when we can improve either the quality of life or survival of cancer patients. I will keep that in mind forever, and use it to make a real difference.

Without the financial support of the **European Union´s Horizon 2020** research and innovation programme under the Marie Skłodowska-Curie grant, the ELBA project would not exist. To all the ELBA supervisors and partners, thank you for all! **Tom**, thank you for coordinating ELBA and for your help and feedback in the last years. **Dani**, I hope you know how much the ELBA students appreciate all your help and advice. You are an amazing person with a very big heart, thank you for everything! Then a shoutout to all my ELBAnian colleagues who have become close friends in the last years. From the first meeting in Prague we connected in a very special way, and I cannot thank you enough for your support, positive distractions and scientific input. Especially **Martyna**, **Silvia**, **Mafalda**, **Nicolas**, **Stavros, Julia** and **Diogo**, thank you for the great collaborations and the calls when they were most needed. **Carlos**, I don´t even know where to start. The few sentences here will not cover all the things I´m grateful for, but you are a very important part of this PhD. You know how to deal with my stubbornness (yes, I´m aware of that..), and you pulled me through when something (or every single thing) did not go as I wanted. Besides all the advice you gave me during this rollercoaster that is called a PhD, you were also there for coffee breaks with a lot of chocolate, costa brava weekends, Spanish translations and just to talk about nothing. You have no idea how much I´ll miss you!

I would have never been able to perform the work presented in this thesis without the help of NanoString. Dear **Alan** and **Sarah**, thank you for your endless support and the opportunities you have given me. I believe in the use of the nCounter technologie for tumor tissue and for liquid biopsies, and I hope this thesis will convince others to believe in it as well.

Lieve **Mama** en **Jamie**. De dag voor ik naar Barcelona vloog kreeg ik het lichtelijk benauwd. Waar was ik nou weer aan begonnen? Ik heb er geen seconde spijt van gehad, mede dankzij jullie support! Als ik advies nodig heb of gewoon iets kwijt moet zijn jullie er altijd, over facetime of live. Het niet in Nederland zijn heeft ook moeilijkere tijden gekend, maar ook toen stonden jullie op Schiphol klaar om mij op te vangen en naar **Opa** te gaan. Bedankt voor alle steun, kansen, Barca bezoekjes, facetime gesprekken, motivatie en opvoeding. Jullie zijn mijn voorbeelden en ik zou nergens zijn zonder jullie! Ook mijn bonus-familie, de **Familie Jansen en aanhang**, jullie zijn er altijd voor mij en voor ons. Bedankt voor het samen lachen, huilen, reizen en alle mooie herinneringen en toekomstige plannen. Lieve **Deb**, mijn beste en liefste vriendinnetje, dankjewel dat je er altijd voor me bent. Eindelijk zijn je twee lullen weer in het land! Elke donderdagavond en voor iedere vlucht belde jij me even, en tijdens Sant Juan zat jij meestal bij mij in Barcelona. Zelfs op afstand blijven we onafscheidelijk, en ik hoop dat dat altijd zo blijft!

To my **friends** and **Barca roomies**, thank you for the visits, for the dinners (and drinks ;)), for the advice, the distractions and the laughing. To the ones abroad, we will visit each other when things are back to normal. For the once in the Netherlands, see you soon!

Table of contents

Acronyms and abbreviations

ABC:	activated B-cell
ADC:	adenocarcinoma
ALK:	anaplastic lymphoma receptor tyrosine kinase
ARDS:	acute respiratory distress syndrome
AUC:	area under the curve
BAL:	bronchoalveolar lavage
C-EXT:	extraction control
cDNA:	complementary DNA
cfDNA:	circulating cell free DNA
cfRNA:	cell-free RNA
CIP:	checkpoint inhibitor pneumonitis
COPD:	chronic obstructive pulmonary disease
CR:	complete response
CRC:	colorectal cancer
Cryo-EM:	cryogenic electron microscopy
CSF:	cerebrospinal fluid
CT:	computed tomography
CTC:	circulating tumor cell
ctDNA:	circulating tumor DNA
CTLA-4:	cytotoxic T-lymphocyte-associated protein 4
Cyc:	cycles
DE:	differential expression
DLBCL:	diffuse large B-cell lymphoma
EV:	extracellular vesicle
FC:	fold change
FDA:	US food and drug administration
FFPE:	formalin-fixed paraffin embedded
FISH:	fluorescence in-situ hybridization
gbm:	stochastic gradient boosting
GCB:	germinal center B-cell
gDNA:	genomic DNA
GEP:	gene expression profile
h:	hour
HIV-1:	human immunodeficiency virus 1
HRP:	horseradish peroxidase
ICI:	immune checkpoint inhibitors
IHC:	immunohistochemistry
irAE:	immune-related adverse event
knn:	k-nearest neighbors
LOOCV:	leave-one-out cross validation
m:	month

MET:	mesenchymal epithelial transition
MSI/dMMR:	microsatellite instability/defective mismatch repair
MTE:	multiplexed target enrichment
NC:	normalized counts
NGS:	next-generation sequencing
NK:	natural killer
NPV:	negative predictive value
NS:	not significant
NSCLC:	non-small cell lung cancer
OS:	overall survival
PC:	positive control
PCR:	polymerase chain reaction
PD-1:	programmed cell death protein
PD-L1:	programmed death-ligand 1
PD:	progressive disease
PE:	pleural effusion
PF4:	platelet factor 4
PFS:	progression-free survival
PPV:	positive predictive value
PR:	partial response
Q-PNA-PCR:	quantitative-peptide nucleic acid-mediated PCR
RET:	RET proto-oncogene
rf:	random forest
RFE:	recursive feature elimination
RNAseq:	RNA sequencing
ROC:	receiver operating characteristic
ROS1:	ROS proto-oncogene 1 receptor tyrosine kinase
RT-qPCR:	reverse transcription-quantitative-PCR
RT:	room temperature
SD:	standard deviation
SD:	stable disease
SLLIP:	Spanish lung liquid vs. invasive biopsy program
SNV:	single nucleotide variants
TCGA:	the cancer genome atlas
TEP:	tumor educated platelet
TIS:	tumor inflammation signature
TMB:	tumor mutational burden
TNBC:	triple negative breast cancer
Tregs:	regulatory T cells
WES:	whole exome sequencing
WGS:	whole genome sequencing
yr:	year

Summary

The assessment of predictive- and prognostic molecular markers in tumor tissue, also known as personalized treatment, has transformed clinical practice for many cancer types. This genotype-directed therapy was found to improve patient survival, and several technical platforms have been introduced in clinical laboratories since then. However, not all tumors can be biopsied and tissue quantities are often insufficient for tumor characterization. Liquid biopsies, such as membrane-encapsulated- or circulating free RNA, DNA and proteins, can be derived from body fluids and can replace or complement tissue biopsies. They have several advantages, such as repeated sampling, a minimally invasive character and heterogeneous profiling. Unfortunately, there is still a big gap between basic research and clinical implementation of liquid biopsies, mainly due to the lack of standardized methodologies. In addition, currently used technical platforms are not always suitable to analyze the low quantity and quality of tumor-derived material that can be found in a liquid biopsy. In consequence, large-scale validation and clinical implementation of liquid biopsy-based biomarker assays requires a sensitive, quick, easy-to-use, relatively cheap, flexible and standardized technical platform with low input requirements.

The nCounter platform can be used to analyze all types of molecules, including RNA, DNA and proteins. Binding of color coded barcodes to targets of interest allows for either a direct read-out of gene- or protein expression levels or the detection of mutations. Tissue-based biomarker assay development on nCounter led to the FDA approval of the Prosigna™ assay for clinical use in breast cancer subtyping. Previous efforts have also highlighted the potential of this platform to analyze amplified liquid biopsy-derived molecules, although validation studies in the clinical setting are needed. In this thesis we validated the use of the NanoString nCounter platform to analyze material from liquid biopsies and develop clinically relevant biomarker assays. Two manuscripts have been accepted for publication on this topic (**chapter 2 and 3**), and two additional manuscripts are in preparation for publication (**chapter 4 and 5**).

In **chapter 2**, we focused on one of the most advanced liquid biopsies in terms of clinical implementation; circulating tumor DNA. We extracted cell-free DNA (cfDNA) from body fluids of cancer patients, including 70 retrospective- and 91 prospective samples, and used the nCounter platform with the 3D Single Nucleotide Variant (SNV) Solid Tumor Panel for routine mutation detection. We obtained highly concordant results when comparing mutation detection in five clinically relevant genes by nCounter with other routinely used technical platforms (98.9%). In addition, follow-up analysis showed that this workflow can be used to detect resistance mutations. Importantly, the nCounter platform was able to detect these mutations with lower cfDNA input, in a shorter time frame, with less hands-on time and using a straightforward data analysis.

We then continued with a less clinically advanced liquid biopsy; extracellular vesicles (EVs). The nCounter platform had never been used to analyze gene expression in plasma-derived EVs. To this end, in **chapter 3**, we performed a proof-of-concept study where we optimized a workflow for EV enrichment, EV-mRNA purification and subsequent gene expression analysis

on nCounter to develop biomarker assays. Our final workflow, including a DNase treatment, was highly reproducible and could be performed in only three days. We then used this workflow to assess gene expression in plasma EVs from cancer and control individuals, using the Human Immune V2 594-gene panel. Although our sample cohort did not qualify for the development of a clinically relevant signature, we developed a machine learning algorithm as a proof-of-concept. We found that an eight gene signature could be used to accurately distinguish between the EVs of cancer patients and controls.

Biomarker assays are most valuable if they meet a clinical need. Therefore, we implemented the EV-mRNA workflow in another project, performing a side-by-side comparison of tumor tissue- and EV-mRNA in **chapter 4**. Treatment with immune checkpoint inhibitors (ICIs) can improve patient outcome, but it can also cause an overactivation of the immune system. This can result in the development of immune related adverse events (irAEs), such as checkpoint inhibitor pneumonitis (CIP). CIP occurs in about 20% of ICI-treated patients, and is an inflammation of the lung tissue that can be lethal if not detected and treated in time. In this study, we used pre-treatment tumor tissue and EVs to develop a predictive CIP signature based on the expression of 770 genes included in the IO360 panel. We initially obtained an 8-gene tissue CIP signature, that was able to predict CIP with an accuracy of 86.6%. Interestingly, this signature did not work in EVs. Considering our previous EV proof-of-concept results, and based on the advantages of an EV-based biomarker assay, we then developed a new EV-based signature. This 4-gene EV CIP signature had a high accuracy (87.0%) and negative predictive value (92.7%). Moreover, the CIP signature score could not only be used to predict CIP before starting ICI treatment, the EV CIP-score also increased during actual CIP development.

Lung cancer is the leading cause of cancer-related death worldwide and new tools for early detection are an unmet clinical need. In **chapter 5**, we have used the EV-mRNA workflow for the development and validation of a signature for early lung cancer detection from plasma EV samples using the Human Immune V2 panel. A 16-gene signature was able to classify early- and late stage lung cancer versus controls with an accuracy of 81-84% in the training cohort. Equally high accuracy values were obtained in a separately collected validation cohort.

In conclusion, this thesis has provided evidence that the nCounter platform can be used for the analysis of liquid biopsies in the clinical setting, especially for ctDNA and EV-mRNA. The low quantity and quality requirements, accuracy, short turn-around time, flexibility, high reproducibility and straightforward data analysis are important advantages of this platform. To this end, clinically relevant biomarker assays can be developed and validated on this standardized platform and can aid in the transition of liquid-derived biomarker assays from bench to bedside. The results presented in this thesis have also generated new possible lines of research. Many more liquids (e.g. urine, saliva), biosources (e.g. tumor educated platelets, circulating tumor cells, white blood cells) and circulating molecules (e.g. cfRNA, circular RNA, long non-coding RNA, proteins) require similar validation studies. Combining several types of circulating molecules, or even biosources, in a multi-omics approach is also a possibility on this platform and may provide a more robust source of biomarker assay development.

Resumen

La evaluación de los marcadores moleculares en tejido tumoral para el pronóstico del cáncer y la predicción de respuesta al tratamiento (lo que habitualmente se conoce como tratamiento personalizado) ha transformado la práctica clínica a la hora de tratar muchos tipos de cáncer. Son numerosos los trabajos que desde hace tiempo respaldan el efecto que esta terapia dirigida por genotipo tiene sobre los pacientes oncológicos mejorando la supervivencia del paciente; consecuentemente, un amplio rango de plataformas técnicas han sido implementadas en los laboratorios clínicos en los últimos años. Sin embargo, no todos los tumores se pueden biopsiar y, a menudo, las cantidades de tejido son insuficientes para la caracterización del tumor. Las biopsias líquidas, como el ARN, el ADN o las proteínas circulantes tanto libres como encapsuladas en una membrana, pueden extraerse de los fluidos corporales reemplazando o complementando de este modo las tradicionales biopsias de tejido. Las biopsias líquidas tienen varias ventajas: ofrecen la posibilidad de realizar estudios seriados, son mínimamente invasivas y permiten analizar la heterogeneidad tumoral. Desafortunadamente, todavía existe una gran brecha entre la investigación básica y la implementación clínica de las biopsias líquidas, principalmente debido a la falta de metodologías estandarizadas. Además, las plataformas técnicas que se utilizan actualmente no siempre son adecuadas para analizar la baja cantidad y calidad de material del tumor procedente de una biopsia líquida. En consecuencia, la validación e implementación de los ensayos de biomarcadores en biopsias líquidas en los laboratorios clínicos requieren una plataforma técnica estandarizada que sea sensible, rápida, fácil de usar, viable económicamente, flexible y que requiera un aporte inicial de ácidos nucleicos bajo, debido a la baja concentración que normalmente se obtiene en las biopsias líquidas.

La plataforma nCounter se puede utilizar para analizar todo tipo de moléculas, incluyendo ARN, ADN y proteínas. La hibridación de diferentes códigos formados por moléculas de colores siguiendo patrones específicos con secuencias de interés permite una lectura directa de los niveles de expresión de genes y proteínas o la detección de mutaciones. El desarrollo de ensayos de biomarcadores en tejidos usando nCounter condujo a la aprobación por la administración de fármacos y alimentos de los Estados Unidos (FDA) del ensayo Prosigna ™ para su uso clínico en la tipificación del cáncer de mama. Numerosos estudios han destacado el potencial de esta plataforma para analizar moléculas derivadas y amplificadas de biopsias líquidas, aunque estudios de validación en el entorno clínico aun son necesarios. El objeto de esta tesis es la validación del uso de la plataforma NanoString nCounter para analizar material de biopsias líquidas y desarrollar ensayos de biomarcadores clínicamente relevantes. Se han aceptado dos manuscritos para su publicación sobre este tema (**capítulos 2 y 3**), y dos manuscritos adicionales estan actualmente en preparación para su incipiente publicación (**capítulo 4 y 5**).

En el **capítulo 2**, nos enfocamos en una de las biopsias líquidas más avanzadas en términos de implementación en la clínica; ADN tumoral circulante (ctDNA). Extrajimos ADN libre circulante (cfDNA) de fluidos corporales de pacientes con cáncer, incluyendo 70 muestras retrospectivas y 91 muestras prospectivas. Durante este estudio hemos utilizado la plataforma

nCounter con el panel de tumores sólidos que muestran una variante única de nucleótido (SNV) para la detección de mutaciones clínicamente relevantes. Obtuvimos resultados de la detección de mutaciones en cinco genes mediante nCounter los cuales mostraron una alta correlación con los resultados obtenidos mediante otras plataformas técnicas de uso habitual (98,9%). Además, el análisis de seguimiento mostró que este flujo de trabajo se puede utilizar para detectar mutaciones en oncogenes asociados con la resistencia a terapias oncológicas. Sobre todo, la plataforma nCounter pudo detectar estas mutaciones más rápido, con una menor cantidad de cfDNA, con una necesidad mínima de personal concebido y utilizando un análisis de datos sencillo.

El segundo estudio se centra en unos biomarcadores menos avanzados clínicamente dentro del campo de las biopsias líquidas, como es el caso de las vesículas extracelulares (EVs). La plataforma nCounter nunca se había utilizado para analizar la expresión génica en EVs derivados de plasma. A tal efecto, en el **capítulo 3**, realizamos un estudio piloto en el que optimizamos un flujo de trabajo para el enriquecimiento de las EVs, la purificación del ARN contenido en estas vesículas y el posterior análisis de expresión génica mediante nCounter evaluando el potencial que alberga esta plataforma para desarrollar ensayos de biomarcadores. Como resultado, nuestro flujo de trabajo final, incluyendo un tratamiento con DNasa, fue altamente reproducible con una duración total de solo tres días. Una vez optimizado el flujo de trabajo, este mismo protocolo fue utilizado para evaluar la expresión génica en EVs de plasma de los pacientes con cáncer e individuos controles, utilizando el "Human Immune V2 panel" de Nanostring el cual alberga 594 genes. Aunque nuestra cohorte de muestras no calificó para el desarrollo de una firma clínicamente relevante, desarrollamos un algoritmo de aprendizaje automático como prueba de concepto. Descubrimos que se podría usar una firma de ocho genes para distinguir con precisión entre las EVs derivadas de los pacientes con cáncer y los controles.

Los ensayos de biomarcadores son más valiosos si satisfacen una necesidad clínica. Por esta razón, implementamos el flujo de trabajo de EV ARN en otro proyecto, realizando una comparación directa entre ARN de tejido tumoral y ARN derivado de EVs, tal y como se detalla en el **capítulo 4**. El tratamiento con inhibidores de puntos de control inmunitarios (ICIs) puede mejorar el resultado del paciente, pero también puede provocar una sobreactivación del sistema inmunológico. Esto puede resultar en el desarrollo de eventos adversos relacionados con el sistema inmunológico (irAE), como la neumonitis por inhibidores de puntos de control inmunitarios (CIP). La CIP ocurre en aproximadamente el 20% de los pacientes tratados con ICIs y es una inflamación del tejido pulmonar que puede ser letal si no se detecta y trata a tiempo. En este estudio, utilizamos tejido tumoral y EVs basales para desarrollar una firma predictiva de CIP basada en la expresión de 770 genes incluidos en el panel IO360. Inicialmente obtuvimos una firma CIP de tejido de 8 genes, que fue capaz de predecir CIP con una exactitud del 86.6%. Curiosamente, esta firma no funcionó en EVs. Teniendo en cuenta nuestros resultados anteriores de prueba de concepto de EVs, y basándonos en las ventajas de un ensayo de biomarcadores basado en EVs, desarrollamos una nueva firma en EVs. Esta firma de 4 genes tuvo una alta exactitud (87.0%) y un alto valor predictivo negativo (92.7%). Además, la puntuación de la firma CIP no solo se puede utilizar para predecir CIP antes de

comenzar el tratamiento con ICIs, sino que la puntuación de la firma CIP también aumentó durante el desarrollo real de la CIP.

El cáncer de pulmón es la principal causa de muerte relacionada con el cáncer en todo el mundo y las nuevas herramientas para la detección temprana son una necesidad clínica aun no alcanzada. En el **capítulo 5**, hemos utilizado el flujo de trabajo de ARN de las EVs para el desarrollo y la validación de una firma genética para la detección temprana de cáncer de pulmón a partir de muestras de EV de plasma utilizando el panel Human Immune V2 de NanoString. Una firma de 16 genes fue capaz de clasificar el cáncer de pulmón en etapa temprana y tardía frente a los controles con una precisión del 81-84% en la cohorte de entrenamiento. Se obtuvieron valores de precisión igualmente altos en una cohorte de validación recopilada por separado.

En conclusión, esta tesis ha proporcionado evidencia de que la plataforma nCounter se puede utilizar para el análisis de biopsias líquidas en el ambiente clínico, especialmente para el análisis de ctADN y EV-mARN. Los requisitos de baja cantidad y calidad de la muestra, la sensibilidad, el análisis rápido, la flexibilidad, la alta reproducibilidad y análisis de datos sencillo son algunas de las ventajas importantes de esta plataforma. Por lo tanto, los ensayos de biomarcadores clínicamente relevantes se pueden desarrollar y validar en esta plataforma y esto puede ayudar en la transición de biopsias líquidas del laboratorio a la práctica. Los resultados que hemos presentado en esta tesis también han generado nuevas posibles líneas de investigación. Muchos más fluidos (p. ej. orina, saliva), fuentes biológicas (p. ej. plaquetas educadas sobre el tumor, células tumorales circulantes, glóbulos blancos) y moléculas circulantes (p. ej., ARN libre de células, ARN circular, ARN largo no codificante, proteínas) requieren estudios de validación adicionales. La combinación de varios tipos de moléculas circulantes, o de fuentes biológicas, en un enfoque multiómico también es una posibilidad en esta plataforma y puede producir un ensayo de biomarcadores con mayor precisión.

Resum

L'avaluació dels marcadors moleculars en teixit tumoral per predir el pronòstic del càncer i la resposta al tractament, també coneguda com a tractament personalitzat, ha transformat la pràctica clínica per a molts tipus de càncer. Es va descobrir que aquesta teràpia dirigida per genotipar, millora la supervivència del pacient i per tant s'han introduït diverses plataformes tècniques en els laboratoris clínics. No obstant això, no tots els tumors es poden biopsiar i, sovint, les quantitats de teixit són insuficients per a la caracterització del tumor. L'ARN, l'ADN i les proteïnes lliures circulants de biòpsies líquides, es poden extreure dels fluids corporals i poden reemplaçar o complementar les biòpsies de teixit. Les biòpsies líquides tenen diversos avantatges: la possibilitat de realització d' estudis de serie o mínimament invasiva i permet analitzar l'heterogeneïtat tumoral. Malauradament, encara hi ha una gran bretxa entre la recerca bàsica i la implementació clínica de biòpsies líquides, principalment a causa de la manca de metodologies estandarditzades. A més, les plataformes tècniques que s'utilitzen actualment, no sempre són adequades per analitzar la baixa quantitat i qualitat de material del tumor derivat d'una biòpsia líquida. En conseqüència, la validació i implementació dels assajos de biomarcadors en biòpsies líquides en els laboratoris clínics, requereixen una plataforma tècnica estandarditzada que sigui sensible, ràpida, fàcil d'utilitzar, relativament econòmica, flexible i amb poca quantitat de mostra.

La plataforma nCounter es pot utilitzar per analitzar tota classe de molècules, inclosos ARN, ADN i proteïnes. La hibridació de codis de barres codificats per colors pels objectius d'interès permet una lectura directa dels nivells d'expressió de gens i proteïnes o la detecció de mutacions. El desenvolupament d'assaigs de biomarcadors en teixits, usant el nCounter, va conduir a l'aprovació per la FDA de l'assaig Prosigna™ per a ús clínic en la tipificació del càncer de mama. Els esforços anteriors també han destacat el potencial d'aquesta plataforma per analitzar molècules derivades i amplificades de biòpsies líquides, encara que es necessiten estudis de validació en l'entorn clínic. En aquesta tesi validem l'ús de la plataforma NanoString nCounter per analitzar material de biòpsies líquides i desenvolupar assajos de biomarcadors clínicament rellevants. S'han acceptat dos manuscrits per a la seva publicació sobre aquest tema (**capítols 2 i 3**), i estem preparant dos manuscrits addicionals per a la seva publicació (**capítol 4 i 5**).

En el **capítol 2**, ens enfoquem en una de les biòpsies líquides més avançades en termes d'implementació a la clínica; ADN tumoral circulant (ctDNA). Es va extreure ADN lliure de cèl·lules (cfDNA) de fluids corporals de pacients amb càncer, incloses 70 mostres retrospectives i 91 mostres prospectives. Hem utilitzat la plataforma nCounter, amb el panell de tumors sòlids de variant de nucleòtid únic 3D (SNV), per a la detecció de mutacions clínicament rellevants. Vam obtenir resultats de la detecció de mutacions en cinc gens molt concordants mitjançant nCounter al comparar amb altres plataformes tècniques d'ús habitual (98.9%). A més, l'anàlisi de seguiment va mostrar que aquest flux de treball es pot utilitzar per detectar mutacions de resistència. Sobretot, la plataforma nCounter va poder detectar aquestes

mutacions de forma ràpida, amb menor quantitat de cfDNA, amb una necessitat mínima de personal concebut i mitjançant una anàlisi de dades senzill.

Després continuem amb una biòpsia líquida clínicament menys avançada; vesícules extracel·lulars (EVs). La plataforma nCounter mai s'havia utilitzat per analitzar l'expressió gènica en EVs derivats del plasma. A aquest efecte, en el **capítol 3**, vam realitzar un estudi de prova de concepte en el qual optimitzem un flux de treball per a l'enriquiment de les EVs, la purificació d'ARN de les EVs i la posterior anàlisi d'expressió gènica en nCounter per desenvolupar assajos de biomarcadors. El nostre flux de treball final, que inclou un tractament amb DNasa, va ser altament reproduïble i es va poder realitzar en només tres dies. Després vam fer servir aquest flux de treball per avaluar l'expressió gènica en EVs de plasma dels pacients amb càncer i individus controls, utilitzant el "Human Immune V2 panel" amb 594 gens. Encara que la nostra cohort de mostres no va qualificar per al desenvolupament d'una signatura clínicament rellevant. Vam desenvolupar un algoritme d'aprenentatge automàtic com a prova de concepte. Vam descobrir que es podria fer servir una signatura de vuit gens per distingir amb precisió entre els EVs derivat dels pacients amb càncer i els controls.

Els assajos de biomarcadors són més valuosos si satisfan una necessitat clínica. Per tant, vam implementar el flux de treball d'EV ARN en un altre projecte, fent una comparació de teixit tumoral i EV ARN en el **capítol 4**. El tractament amb inhibidors de punts de control immunitari (ICIs) pot millorar el resultat del pacient, però també pot provocar una sobre activació del sistema immunitari. Això pot resultar en el desenvolupament d'esdeveniments adversos relacionats amb el sistema immunitari (Irae), com la pneumonitis per inhibidors de punts de control immunitaris (CIP). Els CIP ocorren en aproximadament el 20% dels pacients tractats amb ICIs i és una inflamació del teixit pulmonar que pot ser letal si no es detecta i tracta a temps. En aquest estudi, vam utilitzar teixit tumoral i EVs basal per desenvolupar una signatura predictiva de CIP basada en l'expressió de 770 gens inclosos en el panel IO360. Inicialment vam obtenir una signatura CIP de teixit de 8 gens, que va ser capaç de predir amb una exactitud del 86.6%. Curiosament, aquesta firma no va funcionar en EVs. Tenint en compte els nostres resultats anteriors de prova de concepte de EVs, i basant-nos en els avantatges d'un assaig de biomarcadors, vam desenvolupar una nova firma en EVs. Aquesta signatura de 4 gens va tenir una alta exactitud (87.0%) i un alt valor predictiu negatiu (92.7%). A més, la puntuació de la signatura CIP no només es pot utilitzar per predir CIP abans de començar el tractament sino que la puntuació de la signatura CIP també va augmentar durant el desenvolupament real de la CIP.

El càncer de pulmó és la principal causa de mort relacionada amb el càncer a tot el món i les noves eines per a la detecció primerenca són una necessitat clínica encara no assolida. En el **capítol 5**, hem utilitzat el flux de treball d'EV ARN per al desenvolupament i la validació d'una signatura genètica per a la detecció primerenca del càncer de pulmó a partir de mostres d´EV de plasma utilitzant el panell Human Immune V2 de NanoString. Una firma de 16 gens va ser capaç de classificar el càncer de pulmó en fase inicial i tardana davant els controls amb una precisió de l'81-84% en la cohort d'entrenament. Es van obtenir valors de precisió igualment alts en una cohort de validació recopilada per separat.

En conclusió, aquesta tesi ha proporcionat evidència que la plataforma nCounter es pot utilitzar per a l'anàlisi de biòpsies líquides en l'ambient clínic, especialment per a l'anàlisi de ctDNA i EV-mRNA. Els requisits de baixa quantitat i qualitat de la mostra, la sensibilitat, l'anàlisi ràpid, la flexibilitat, l'alta reproductibilitat i anàlisi de dades senzilles són alguns dels avantatges importants d'aquesta plataforma. Per tant, els assajos de biomarcadors clínicament rellevants es poden desenvolupar i validar en aquesta plataforma i això pot ajudar en la transició de biòpsies líquides de laboratori a la pràctica. Els resultats que hem presentat en aquesta tesi també han generat noves possibles línies d'investigació. Els líquids (p. ex. orina, saliva), mostres d'origen biològic (p. ex. plaquetes educades sobre el tumor, cèl·lules tumorals circulants, glòbuls blancs) i molècules circulants (p. ex., ARN lliure de cèl·lules, ARN circular, ARN no codificant llarg, proteïnes) requereixen estudis de validació addicionals. La combinació de diversos tipus de molècules circulants o mostres d'origen biològic, en un enfocament multiómic també és una possibilitat en aquesta plataforma i pot produir un assaig de biomarcadors amb més precisió.

Samenvatting

Het meten en evalueren van voorspellende en prognostische biomarkers in tumorweefsel, beter bekend als gepersonaliseerde behandeling, heeft de zorg voor kankerpatiënten drastisch veranderd. Doordat overlevingskansen aanzienlijk toenemen met deze gerichte vorm van behandeling, zijn nieuwe technieken geïntroduceerd in klinische laboratoria om dit soort metingen uit te voeren. Echter, het verkrijgen van een biopt is niet altijd mogelijk en de hoeveelheid weefsel is vaak onvoldoende voor een nauwkeurige en volledige analyse van de tumor. Alle cellen in het lichaam, waaronder tumor cellen, scheiden informatie uit in het bloed. Deze informatie bestaat uit RNA, DNA en eiwitten. Dit soort moleculen kunnen vrij door het bloed circuleren of beschermd zijn door een membraan zoals een blaasje of bloedplaatje. Dit soort vloeibare biopten noemen we ook wel "liquid biopsies", en kunnen weefselbiopten vervangen of aanvullende informatie geven. Liquid biopsies kennen verschillende voordelen, waaronder de mogelijkheid tot herhaalde afname, een minimaal invasief karakter en een beter beeld van de aanwezige tumor. Toch is er nog een grote kloof tussen fundamenteel onderzoek en klinische implementatie van liquid biopsies, voornamelijk vanwege het ontbreken van een standaard methode van analyseren. Bovendien zijn de momenteel gebruikte technische systemen niet altijd geschikt om de minimale hoeveelheid en lage kwaliteit van tumor-afgeleid materiaal uit een liquid biopsy te analyseren. Grootschalige validatie en klinische implementatie van biomarker testen uit liquid biopsies vereist een gestandaardiseerde techniek die nauwkeurig, snel, gebruiksvriendelijk, relatief goedkoop en flexibel is, ondanks de minimale hoeveelheid en lage kwaliteit van het materiaal.

Het NanoString nCounter systeem kan worden gebruikt om RNA, DNA en eiwitten te analyseren. Het binden van kleurgecodeerde streepjescodes aan deze moleculen maakt het mogelijk de expressie van genen en eiwitten direct uit te lezen of mutaties te detecteren. Een voorbeeld van een FDA-goedgekeurde biomarker test in weefsel die is ontwikkeld op dit systeem, is de Prosigna™ test. Deze test kan prognostische informatie geven voor borstkanker patiënten. Naast het direct uitlezen van weefsels, kan een minimale hoeveelheid materiaal ook eerst gekopieerd worden voordat de analyse plaatsvindt. Hierdoor kan de kleine hoeveelheid materiaal dat uit bloed te verkrijgen is toch geanalyseerd worden. Echter zijn validatie studies noodzakelijk.

In dit proefschrift hebben we het gebruik van het nCounter systeem gevalideerd met materiaal uit liquid biopsies om zo klinisch relevante biomarker tests te ontwikkelen. Twee manuscripten zijn geaccepteerd voor publicatie over dit onderwerp (**hoofdstuk 2 en 3**), en twee extra manuscripten zijn in voorbereiding voor publicatie (**hoofdstuk 4 en 5**).

In **hoofdstuk 2** hebben we ons gericht op een van de meest geavanceerde liquid biopsies wat betreft klinische implementatie; circulerend tumor-DNA. We hebben celvrij DNA (cfDNA) uit lichaamsvloeistoffen van kankerpatiënten geïsoleerd, waaronder 70 retrospectieve en 91 prospectieve samples en het nCounter systeem gebruikt voor routine mutatiedetectie. Onze resultaten in vijf klinisch relevante genen waren voor 98.9% in overeenstemming met

technieken die momenteel worden gebruikt in laboratoria. Ook tonen we aan dat het analyseren van vervolgsamples therapie resistentie kan detecteren. Het nCounter systeem kon bovendien mutaties vinden in samples met minder materiaal, in een korter tijdsbestek, met minder hands-on tijd en met behulp van een eenvoudige data analyse.

Vervolgens zijn we ons gaan richten op een minder klinisch geavanceerde liquid biopsy; (kanker)blaasjes of extracellulaire vesicles (EVs). Het nCounter systeem was nooit eerder gebruikt om genexpressie in EVs uit bloedplasma te analyseren. Om die reden hebben we in **hoofdstuk 3** een proof-of-concept studie uitgevoerd waarin we een workflow optimaliseerden voor het isoleren van EVs uit plasma, het verkrijgen van mRNA uit EVs en de genexpressie analyse op het nCounter systeem om vervolgens biomarker testen te kunnen ontwikkelen. Onze uiteindelijke workflow was in hoge mate reproduceerbaar en kon in slechts drie dagen worden uitgevoerd. Vervolgens hebben we deze workflow gebruikt om genexpressie in plasma-EVs van kanker- en controlepersonen te analyseren, met behulp van een panel dat tegelijkertijd kan kijken naar 594 verschillende genen. Hoewel de geïncludeerde patiënten te uiteenlopend waren om een klinisch relevante biomarker test te ontwikkelen, hebben we als proof-of-concept toch een zelflerend algoritme ontwikkeld. Dit algoritme kon, door het combineren van de expressie in slechts acht genen, worden gebruikt om nauwkeurig onderscheid te maken tussen de EVs van kankerpatiënten en controles.

Biomarker testen zijn het meest waardevol als ze voorzien in een klinische behoefte. Daarom hebben we de EV-mRNA workflow geïmplementeerd in een ander project in **hoofdstuk 4**. Hierin hebben we tumorweefsel- en EV-mRNA met elkaar vergeleken. Behandeling met immuuntherapie (ICIs) kan de overlevingskansen van patiënten verbeteren, maar het kan ook een overactivering van het immuunsysteem veroorzaken. Dit kan resulteren in de ontwikkeling van immuungerelateerde bijwerkingen (irAEs), zoals auto-immuun pneumonitis (CIP). CIP komt voor bij ongeveer 20% van de patiënten die behandeld zijn met immuuntherapie en is een ontsteking van het longweefsel die levensbedreigend kan zijn als deze niet op tijd wordt ontdekt en behandeld. In deze studie hebben we tumorweefsel en EVs van vóór de behandeling gebruikt om een voorspellende CIP-test te ontwikkelen op basis van een panel die de expressie meet van 770 verschillende genen. We hebben aanvankelijk een CIP-test ontwikkeld op basis van tumor weefsel, waar een zelflerend algoritme - met gebruik van slechts acht genen - nauwkeurig (86.6%) onderscheid kon maken tussen patiënten die wel en geen CIP zouden ontwikkelen. Deze CIP-score kon niet gebruikt worden in de EV samples van dezelfde patiënten. Gezien onze eerdere proof-of-concept resultaten en gebaseerd op de voordelen van een biomarker test in EVs, hebben we vervolgens een nieuwe test ontwikkeld op basis van EVs. De nieuwe EV CIP-test is gebaseerd op de expressie van slechts vier genen, en had zowel een hoge nauwkeurigheid (87.0%) als een goede negatief voorspellende waarde (92.7%). Bovendien kon de test niet alleen worden gebruikt om CIP te voorspellen voordat de ICI-behandeling werd gestart. De EV CIP-score nam ook toe tijdens de daadwerkelijke ontwikkeling van CIP.

Longkanker is wereldwijd de kankersoort met de hoogste sterfte, en nieuwe technieken voor vroege detectie zijn cruciaal voor betere overlevingskansen. In **hoofdstuk 5** hebben we de EV-

mRNA workflow gebruikt voor de ontwikkeling en validatie van een test voor de detectie van longkanker in een vroeg stadium met behulp van een panel die de expressie meet van 592 verschillende genen. De uiteindelijke test is gebaseerd op de expressie van 16 genen, en kon onderscheid maken tussen patiënten met longkanker (zowel in een vroeg als laat stadium) en controles met een nauwkeurigheid van 81-84%. Validatie van deze test in een onafhankelijk cohort van longkanker patiënten en controles bevestigde deze nauwkeurigheid.

Concluderend heeft dit proefschrift bewijs geleverd dat het nCounter systeem kan worden gebruikt voor het analyseren van liquid biopsies in de klinische setting, met name voor ctDNA en EV-mRNA. De lage eisen voor de hoeveelheid en de kwaliteit van het te analyseren materiaal, de hoge nauwkeurigheid, snelheid, flexibiliteit, hoge reproduceerbaarheid en eenvoudige data-analyse zijn belangrijke voordelen van dit systeem. Biomarker tests kunnen worden ontwikkeld en gevalideerd op dit gestandaardiseerde systeem en kunnen helpen bij de implementatie van liquid biopsy tests in de klinische praktijk. De resultaten die in dit proefschrift worden gepresenteerd, hebben ook nieuwe mogelijke onderzoekslijnen opgeleverd. Voor andere lichaamsvloeistoffen (bv. urine, speeksel), biologische bronnen binnen die vloeistoffen (bv. bloedplaatjes gemodificeerd door tumoren, circulerende tumorcellen, witte bloedcellen) en circulerende moleculen (bv. celvrij RNA, circulair RNA, lang niet-coderend RNA, eiwitten) zijn vergelijkbare validatiestudies vereist. Het combineren van verschillende soorten circulerende moleculen, of zelfs biologische bronnen, in een multi-omics benadering is ook een mogelijkheid op dit systeem en kan een robuustere bron zijn voor de ontwikkeling van biomarker testen.

Chapter 1. Introduction

The information in Chapter 1 is adapted from the following publications:

Bracht JWP, Gimenez-Capitan A, Huang CY et al. Analysis of extracellular vesicle mRNA derived from plasma using the nCounter platform. *Sci. Rep. 2021.* 11:3712.

Gimenez-Capitan A, Bracht JWP, García-Mosquera JJ et al. Multiplex detection of clinically relevant mutations in liquid biopsies of cancer patients using a hybridization-based platform. *Clinical Chemistry 2021. [published Online First: 2021/01 /14]*

Bracht JWP, Mayo-de-las-Casas C, Berenguer J et al. The Present and Future of Liquid Biopsies in Non-Small Cell Lung Cancer: Combining Four Biosources for Diagnosis, Prognosis, Prediction, and Disease Monitoring. *Curr. Oncol. Rep 2018.* 20;20(9):70.

Gonzalez-Cao M, Morán T, Dalmau J et al. Assessment of the Feasibility and Safety of Durvalumab for Treatment of Solid Tumors in Patients With HIV-1 Infection. The Phase 2 DURVAST Study. *JAMA Oncol. 2020*. 1;6(7):1063-1067.

Bracht JWP, Berenguer J, Karachaliou N et al. Combining plasma-based biosources to predict treatment response in NSCLC patients. *Ann. Of Oncol. 2018.* 1;29(9):2022.

Bracht JWP, Karachaliou N, Bivona T et al. BRAF Mutations Classes I, II, and III in NSCLC Patients Included in the SLLIP Trial: The Need for a New Pre-Clinical Treatment Rationale. *Cancers 2019.* 11(9). pii: E1381.

Bracht JWP, Aguilar A, Viteri S et al. A four-gene EV-mRNA signature to predict checkpoint inhibitor pneumonitis in lung cancer patients. *Manuscript in preparation.*

Cancer genotyping and patient stratification

Despite progress in all fields of cancer research, the global cancer burden is still growing. With 18 million cases, and 9 million deaths in 2018, cancer is still one of the biggest health problems(1). Late stage diagnosis, ineffective treatment and acquired therapy resistance are the main causes for this high mortality rate. In addition, cancer accounts for a major financial impact on health care costs, with a total amount of 103 billion euros spent in Europe in 2018. This includes direct cost of pharmaceuticals, facilities, medical equipment and medical staff, in addition to indirect costs such as productivity loss due to early mortality(2). Earlier cancer detection, effective molecular profiling and improved therapy response prediction are indispensable to enhance patient survival and lower cancer health care costs.

Technical advances in tumor profiling, including genomic, transcriptomic and proteomic-based analyses of tumor tissues, can be used to identify targetable gene abnormalities that drive tumor growth(3). These abnormalities include mutations, rearrangements, copy number variations, splice variants and protein overexpression. Importantly, genotype-directed therapy was shown to significantly improve overall survival of cancer patients(4). In addition, tumor characterization may be used to identify predictive biomarkers, such as programmed death-ligand 1 (PD-L1) expression and tumor mutational burden (TMB) analyses in tumor tissues. Although both approaches still lack robust sensitvity and specificity, both PD-L1 and TMB measurements are routinely used for immunotherapy response prediction and patient inclusion in clinical trials. Effective predictive biomarkers can ultimately improve patient survival and quality of life, and lower health care costs(5).

Most commonly used approaches for tumor characterization include fluorescence in-situ hybridization (FISH), immunohistochemistry (IHC), real-time polymerase chain reaction (PCR), end-point PCR, sanger sequencing and microarrays to detect alterations such as mutations, fusions, splice variants, amplifications and expression levels(6) (**Figure 1**). Next-generation sequencing (NGS) is a more recently emerged technique and can be used for sequencing of DNA or RNA. DNA sequencing can be employed to detect a wide array of alterations in multiple genes, avoiding the need to perform multiple tests and sparing tissue samples. Three different types of NGS analyses can be performed; a targeted gene panel, all genome coding regions (whole exome sequencing; WES) or the entire genome (whole genome sequencing; WGS)(7). For clinical implementation, targeted gene panels including actionable alterations, are often used due to the lower cost, easier interpretation and shorter turnaround time.

RNA sequencing (RNAseq) is used to detect and quantify RNA in order to uncover gene fusions, mutations, alternative splicing, post-transcriptional modifications and differences in gene expression. Gene expression studies can also be conducted using reverse transcription-quantitative-PCR (RT-qPCR) and microarrays, each with their own (dis)advantages. While RT-qPCR provides high sensitivity and specificity at high speed, it only allows for low-throughput transcriptomic analysis. Microarrays are a medium-throughput platform but have a limited dynamic range of detection and are therefore not suitable to detect genes that are

expressed at either low or high levels. RNAseq is a very accurate, high-throughput platform with a wide dynamic range, but limitations include turn-around-time, high cost and complex data analysis(8, 9).

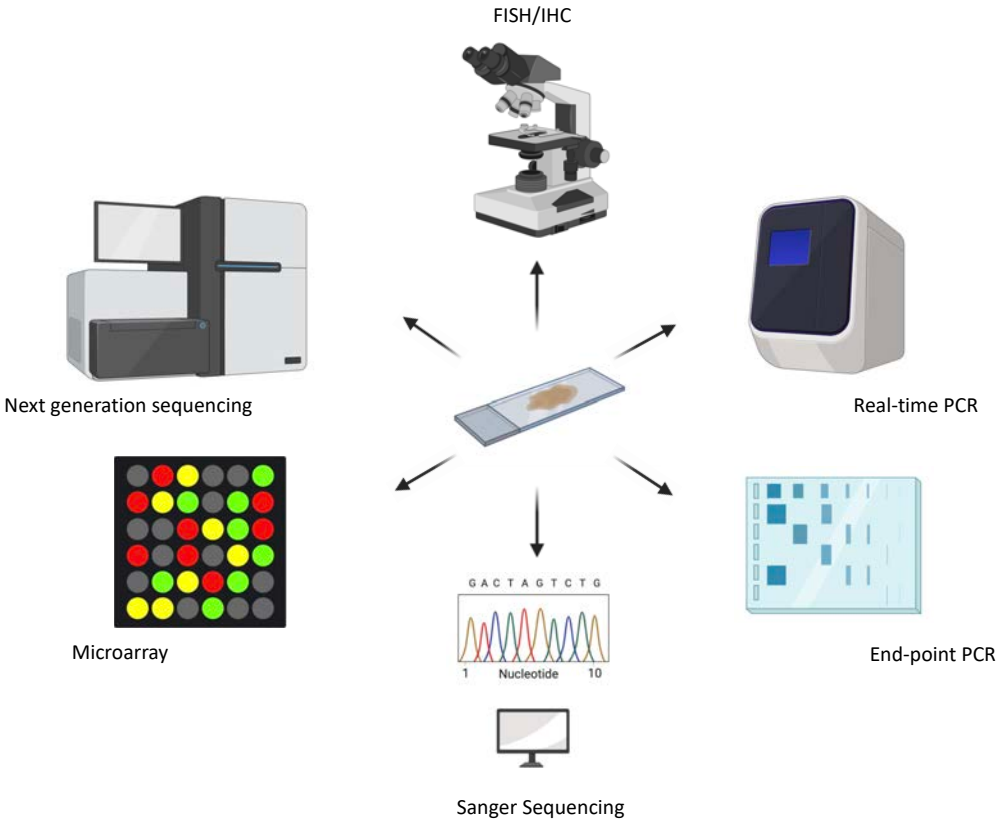


Figure 1. Commonly used laboratory approaches for tumor tissue characterization, including fluorescence in-situ hybridization (FISH), immunohistochemistry (IHC), real-time polymerase chain reaction (PCR), end-point PCR, sanger sequencing, microarrays and next generation DNA/RNA sequencing (NGS).

The NanoString nCounter platform

The NanoString nCounter platform has gained popularity in translational research in recent years. This platform provides multiplexed analysis of up to 800 RNA, DNA, or protein targets from low quality and quantity formalin-fixed paraffin embedded (FFPE) tissue samples. Besides commercially available panels, they can also be customized. The multiplexed CodeSets provide a direct capture and readout of individual targets in the sample, based on unique color coded barcodes for each target of interest. The capture and reporter probes are mixed with the sample for overnight hybridization (**Figure 2**). Excess capture and reporter probes are washed out in the Preparation Station the next day, and target-probe complexes are immobilized onto a cartridge. The color-coded barcodes are then counted on the Digital Analyzer, to determine the abundance of each target in the sample. Depending on the nature of the target (DNA, RNA or protein) this output can then be used to detect mutations or measure gene- and protein expression levels.

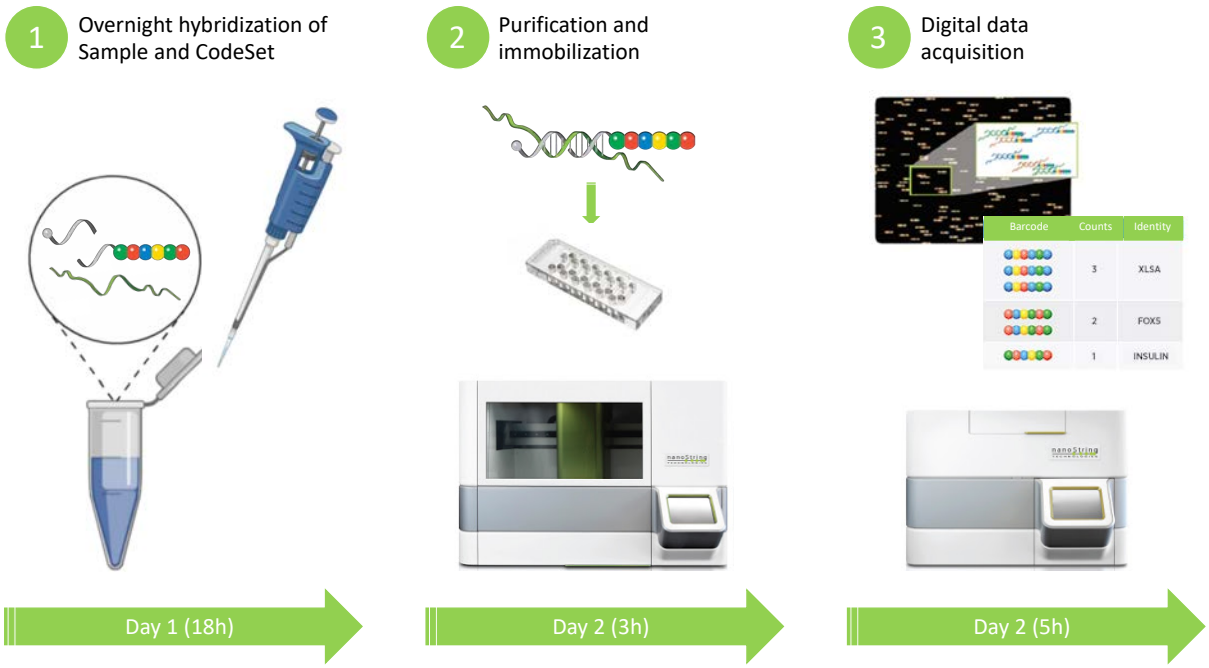


Figure 2. NanoString nCounter workflow. The CodeSet, consisting of capture and reporter probes, is mixed with the sample for overnight target-probe hybridization. Excess CodeSet and sample are washed out in the Preparation Station the next day, and target-probe complexes are immobilized onto a cartridge. Hereafter, the color-coded barcodes are counted on the Digital Analyzer, to determine the abundance of each target in the sample. Depending on the type of target (DNA/RNA/protein) mutation detection or gene/protein expression analysis can be carried out.

When analyzing RNA expression, the lack of cDNA synthesis and enzymatic reactions while using this platform significantly lowers the amplification bias, false positives and technical variation. In terms of sensitivity, the nCounter platform was found to be more sensitive than microarrays and have similar sensitivity to RT-qPCR. Moreover, reproducibility experiments revealed a replicate correlation coefficient of 0.999(9-12). A strong correlation has also been reported when comparing gene expression patterns obtained through nCounter and RNA-seq platforms, demonstrating how expression analysis can be performed without the required infrastructure, turn-around-time, RNA quality and bioinformatic pipeline for RNA-seq technologies(13). Finally, implementation of tumor characterization platforms in routine clinical practice requires an accurate, simple and cost-effective medium-throughput solution, such as the nCounter platform(9).

The nCounter platform for discovery of prognostic- and predictive biomarkers in tissue

Biomarkers are defined as: “objectively measured characteristics, describing an (ab)normal biological state in an organism by analysis of DNA, RNA, protein, peptide or chemical modifications”(14, 15). Biomarkers can be classified into 1) diagnostic; to identify a specific disease in a patient, 2) prognostic; used to inform on the risk of clinical outcome such as disease progression or recurrence, regardless of treatment, and 3) predictive; used to identify patient cohorts that are more likely to respond to clinical interventions(14, 16) (**Figure 3**). Due to its suitability for clinical implementation, the nCounter platform has been extensively used for

FFPE biomarker discovery, consequent assay development, analytical and clinical assay validation and to study clinical utility of the developed assays.

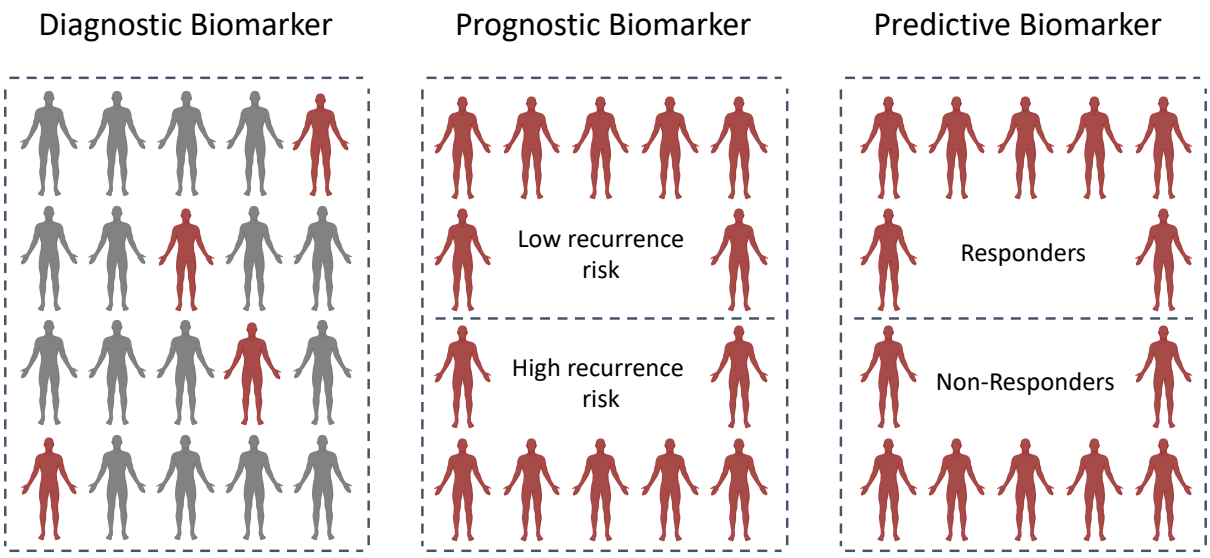


Figure 3. Classification of existing biomarker assays. Diagnostic biomarkers can be used to detect a specific disease in a patient. Prognostic biomarkers can determine patient outcome, independent of treatment. Predictive biomarkers can be used to predict which patients are more likely to experience a certain event (e.g. response to treatment and/or development of adverse events).

Prognostic biomarker assays on nCounter

One of the first studies that made use of the nCounter platform for prognostic biomarker assays focused on diffuse large B-cell lymphoma (DLBCL)(17). Molecular subtypes of DLBCL have significant prognostic value and include activated B-cell (ABC) and germinal center B-cell (GCB)-like DLBCL. Previously, patient stratification was based on gene expression and IHC, with limitations including low quality RNA from FFPE tissue samples and non-quantitative analysis. Gene expression profiling with the nCounter platform was used to analyze the expression of 145 genes and accurately classified patients into DLBCL subtypes, which could be used to guide treatment selection(17).

Breast cancer can also be classified into distinct subtypes; Basal-like, HER2-enriched, Luminal A and Luminal B, resulting in different prognostic outcomes. A gene expression study revealed that the nCounter platform could be used to identify these subtypes using the PAM50 assay; a molecular signature derived from the expression of 50 genes(18-20). This has led to the development of the Prosigna™ assay (NanoString Technologies, Seattle, WA), which can not only distinguish between the four intrinsic subtypes, but can also define a risk category and risk of recurrence. Prosigna™ is currently the only nCounter-based assay that has been validated in the clinical practice(21-23), having received US Food and Drug Administration (FDA) approval in 2013. The simplified workflow and shorter turnaround time supported the use of Prosigna™ in daily clinical practice.

Predictive biomarker assays on nCounter for targeted therapy selection

It is now known that accurate detection of driver alterations can guide cancer treatment and improve patient survival. In advanced non-small-cell lung cancer (NSCLC), 5-7% of tumors harbor anaplastic lymphoma receptor tyrosine kinase (ALK) rearrangements, ROS proto-oncogene 1 receptor tyrosine kinase (ROS1) fusions, mesenchymal epithelial transition (MET) exon 14 skipping mutations and ret proto-oncogene (RET) fusions, which can be targeted with specific inhibitors(3). FISH and IHC are currently the gold standard for the detection of these alterations. However, these techniques have several disadvantages, such as the quantity of tissue that is needed for single assay fusion detection, turn-over-time and cost. A recent, retrospective validation study used a multiplexed NanoString nCounter Assay for fusion detection in mRNA derived from FFPE tissue samples(24). The assay was able to detect ALK, ROS1 and RET fusions in patient samples with high concordance compared to IHC or FISH.

Another nCounter panel that can be used to detect targetable tumor driver alterations is the Vantage 3D™ DNA SNV Solid Tumor Panel, which targets 97 different single nucleotide variants (SNV) and small InDels in 24 genes. Two validation studies reported that mutations that were previously detected by NGS in FFPE samples, could also be detected by nCounter with a concordance of 99.9%(25) and >95%(26). In summary, these studies have shown that rapid, robust and cost-effective multiplexed profiling of tumor driver alterations is feasible with nCounter. In addition, this platform allows for easier clinical implementation compared to other NGS-based platforms.

Predictive biomarker assays on nCounter for immunotherapy response

Effective use of immune checkpoint inhibitors (ICIs) has revolutionized the treatment of some tumor types. Antibodies against programmed cell death protein (PD-1), its ligand PD-L1 and cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) can be used to inhibit tumor growth and potentially lead to curability in several advanced and metastatic cancer types(27). Despite clinical success and FDA approval of six different checkpoint inhibitors in the last nine years, only a subset of patients experience durable responses(28).

Research on predictive biomarkers for ICI response is currently a burning issue, and evolves mainly around PD-L1, TMB, microsatellite instability/defective mismatch repair (MSI/dMMR), specific mutations and gene signatures(29). Although PD-L1 and MSI/dMMR assays have already been approved for clinical use in different tumor types, both assays lack method standardization and require optimization to improve their predictive values(30). Other, less advanced, predictive biomarkers include analysis of immune cell infiltration and gene expression signatures. Due to the complex involvement of both the tumor and the microenvironment in triggering immune responses, it has been proposed that a single biomarker assay may not provide enough information to predict immunotherapy response. In this respect, gene panels could provide more wide-ranging information on both tumor biology, the microenvironment and immune system activation(16).

Efforts have been made to use gene expression signatures to distinguish between so-called "hot" and "cold" tumors. Hot tumors have a T cell-inflamed tumor microenvironment, which suggests a pre-existing immune response indicating tumor infiltration of T cells. On the other hand, cold tumors do not have a T cell-inflamed tumor microenvironment, and lack a pre-existing immune response(31). A signature has been developed to measure immune infiltration and the functional status of T cells and consequent responses to the anti-PD-1 antibody pembrolizumab(32, 33). This Tumor Inflammation Signature (TIS) is based on the expression of 18 genes and was developed as a clinical trial assay for multiple tumor types on the nCounter platform in 2017(33). TIS score was found to outperform IHC and TMB analyses in predicting clinical response to pembrolizumab. The 18 genes included in the TIS score are mainly involved in chemokine expression, cytotoxic activity, IFN-y activity and antigen presentation. The TIS score has been extensively validated in pre-immunotherapy patient tissues, including 10 different tumor types, where patients with improved response to anti-PD-1 blockade tended to have higher pre-treatment TIS scores(34-38). Although highly promising, TIS has not yet been implemented in the clinical practice to stratify patients that will benefit from ICIs.

In addition to the TIS score, other studies have aimed to develop nCounter-based gene expression signatures that can predict response to anti-PD-1 antibodies in specific tumor types, such as melanoma(39, 40), triple negative breast cancer (TNBC)(41), gastric cancer(42) or glioblastoma(43, 44). Further efforts focused on a pan-cancer predictive signature of immunotherapy response(37, 45), including our recent study in which we developed a predictive signature of clinical benefit from immunotherapy in HIV-1 infected cancer patients(46). Combined, these studies highlight the opportunity of using the nCounter platform in prospective validation studies to stratify patients that are more likely to respond to immunotherapy.

Due to increased use and new combinations with conventional standard-of-care treatment, immunotherapy can trigger auto-inflammatory immune-related adverse events (irAE). Development of irAEs occurs after inappropriate immune activation and could be related to environmental circumstances, genetic predisposition or tumor-host interactions(47). One of these irAEs is checkpoint inhibitor pneumonitis (CIP), a non-infection induced inflammation of the lung tissue which is reported to occur in about 1-5% of immunotherapy treated patients(48, 49). CIP arises with diverse clinical and radiographic manifestations, complicating the diagnosis in the clinical setting. If not detected in time, CIP can lead to heart- and/or respiratory failure and death, independent of therapy response. Early management consists of discontinuation of immunotherapy and start of corticosteroid treatment. Currently, in our clinic, almost 19% of patients developed CIP under ICI treatment, and this number is expected to increase in the near future with novel treatment combinations. Pre-treatment gene expression analysis can be employed to predict which patients have higher risk of developing CIP, and should therefore be monitored more extensively, or even receive a different treatment regimen. This type of stratification could improve patient survival and quality of life. One major advantage of using nCounter is the option to create custom panels with multiple gene signatures that can simultaneously predict response to treatment and risk of toxicities from immunotherapy.

Liquid biopsies

Even though tumor tissue profiling can provide significant prognostic and predictive information, advanced lung cancer and other tumors are sometimes difficult or even impossible to biopsy and tissue quantity is often insufficient for downstream profiling. Moreover, intra-tumoral heterogeneity may only yield a snapshot of the actual tumor alterations and repeated sampling to monitor treatment response is not possible(50-52). A recent approach to overcome or complement invasive tissue biopsies is the use of liquid biopsies(50, 52-57). Different body fluids can be used as liquid biopsies, including blood, urine and saliva. Circulating molecules, such as DNA, RNA or proteins, can be freely present within these liquids, or can be extracted and analyzed from circulating tumor cells (CTCs)(58-60), extracellular vesicles (EVs)(61-64) and tumor educated platelets (TEPs)(65, 66) (**Figure 4**). These liquid biosources, including their (dis)advantages and their potential application for biomarker development in NSCLC have been extensively reviewed by our group (**Annex 1**)(50).

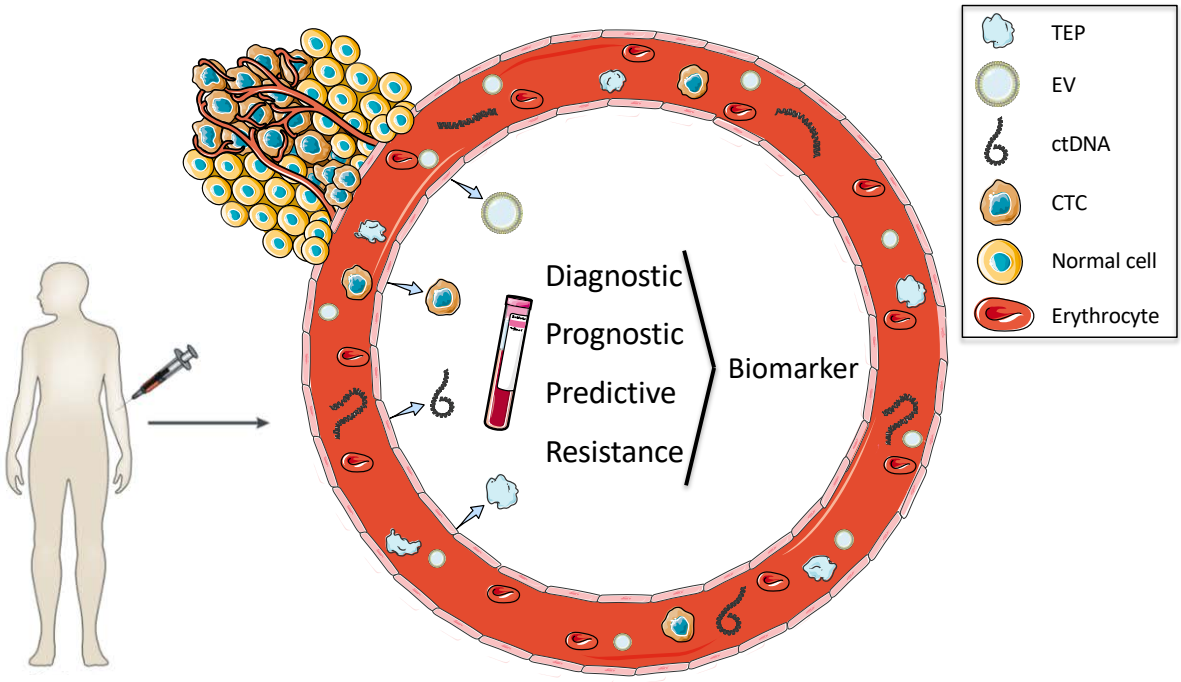


Figure 4. Liquid biopsies, including blood and other types of body fluids, contain circulating molecules that can be isolated and used to analyze ongoing biological processes in the body. These molecules can be freely present (e.g. ctDNA, RNA, proteins) or can be found within other liquid biosources (CTCs, EVs and TEPs). Characterization of these processes on a personalized (follow-up) basis or in large cohorts of individuals can then be used to create diagnostic, prognostic, predictive and/or resistance biomarker assays. *ctDNA: circulating tumor DNA; CTCs: circulating tumor cells; EVs: extracellular vesicles; TEPs: tumor-educated platelets.*

Genetic and transcriptomic characterization of tumors, based on liquid biopsy-derived materials, is currently a hot topic in translational research. Importantly, liquid biopsies could also provide a source for early diagnostic- or screening biomarker assays. Circulating cell free DNA (cfDNA) is currently the most commonly used liquid biosource. A variable fraction of cfDNA consists of circulating tumor DNA (ctDNA) in cancer patients, which can be used to

detect targetable genetic alterations. The introduction of ctDNA analysis led to enormous accelerations in clinical practice, and studies have shown that liquid biopsy based cfDNA analysis is as effective as standard-of-care tissue testing, with concordant results in >90% of patients and a shorter turnaround time(67). However, there is still room for improvement. Around 30% of NGS-based analysis using cfDNA fail, due to high sample input and quality requirements of these platforms(68). Other technical pitfalls related to NGS include assay turnaround time, high cost and complicated analysis due to the huge amount of data output. In addition, protocol standardization and thus technical reproducibility are crucial for validation studies and to establish clinical utility.

Apart from clinical implementation of cfDNA, RNA expression analysis from liquid biopsies can yield clinically relevant information about both the tumor and the immune system. Changes in gene expression that occur downstream of DNA alterations can be used to detect signatures indicative of the presence of a tumor. In addition, differential expression of RNA can be used to define signatures of treatment response and localize a tumor within the body. RNA can be freely circulating in the blood (cfRNA), but is much more stable when conserved in a membrane-enclosed particle, such as EVs, CTCs or TEPs. EVs are heterogeneously sized vesicles (10 nm – 1 μm) that are released from several cell types, and can be isolated from almost all body fluids. Their double-layered phospholipid membrane protects their nucleic acid cargo from degradation(69-71), and allows for intercellular communication by transferring their contents to target cells(69, 72, 73). The quantity and cargo of EVs is regulated by the producing cell, and EVs may therefore be tumor-derived. Consequently, RNA profiles within EVs provide attractive biomarkers for the development and progression of cancer(74). Gene expression studies using EV-RNA from cancer patients is an active area of research(63, 64), but the lack of standardized protocols for the extraction and analysis of EVs and their RNA cargo have led to inconsistent results and hamper clinical implementation.

The NanoString nCounter platform for the analysis of liquid biopsies

The nCounter platform could provide a solution to the technical drawbacks that the liquid biopsy field is currently facing. This platform shows high accuracy and reproducibility, lower cost compared to other multiplexed methodologies, and a shorter and simpler workflow and data analysis. In short, standardization of liquid biopsy-based analysis on the nCounter platform could aid in the transition from bench to bedside. Some attempts have already been made to replace or complement tissue biopsies by using liquid biopsy-based analysis on the nCounter platform, including the profiling of cfRNA(58, 75, 76) and (mi)RNA derived from CTCs(58, 77), leukocytes(78) and EVs(79, 80). Genetic characterization of DNA derived from EVs has also been performed(61).

One of the advantages of tissue-based nCounter analysis is the fact that cDNA conversion and consequent amplification are not necessary for gene expression analysis. In the case of liquid biopsies, the quantity of nucleic acids that can be extracted is much lower when compared to tissue samples. To this end, NanoString has developed a Low Input Kit, which enables gene

expression profiling or mutation detection from samples with low RNA or DNA input. A PCR amplification step is added to the workflow and used to perform multiplexed target enrichment (MTE) prior to panel hybridization. However, amplification bias is a concern and one study has reported on the detection of false-positive results after performing a pre-amplification step(77). Of note, the authors used a different pre-amplification protocol from the one that was developed by NanoString. In contrast, other studies found that the Low Input Kit yielded gene expression results that highly correlated with a non-amplified sample input(81).

Several studies have reported on the analysis of blood-derived liquid biosources using nCounter (**Table 1**), including miRNAs. Khodadadi-Jamayran et al(76) used a 800-miRNA panel, and found that plasma-miR-24-3P expression differed between patients with- and without metastatic breast cancer. High miR-24-3P expression correlated with lower survival, and results were validated in an online cancer genome atlas (TCGA) tissue cohort. Garcia-Contreras et al(79) developed a plasma EV-miRNA signature, consisting of 7 miRNAs, to diagnose type 1 diabetes melitus. Finally, Vicentini et al(80) reported on the use of plasma EV-derived miRNAs to differentiate between four different pancreatic cancer types. Importantly, the nCounter miRNA panel did not require any pre-amplification.

Table 1. Previous performed studies using liquid biopsies for nCounter-based analysis

Study	Liquid biosource	Analyte	Pre-amplification
Khodadadi-Jamayran et al.	Plasma	Cell-free miRNA	No
Garcia Contreras et al.	Plasma EVs	miRNA	No
Vicetini et al.	Plasma EVs	miRNA	No
Porras et al.	CTCs	mRNA	Yes
Beck et al.	Plasma	Cell-free mRNA	Yes
	CTCs	mRNA	Yes
Wu et al.	Leukocytes	mRNA	No
Kossenkov et al.	Whole blood	mRNA	No
Kamyabi et al.	Plasma EVs	DNA	Yes

Other studies have described the use of CTC-mRNA for nCounter-based transcriptomic analysis. Porras et al(77) spiked CTCs, isolated from breast cancer cell lines, into the blood of healthy controls to determine if an nCounter panel could distinguish between gene expression patterns of spiked and unspiked samples. The authors found that pre-amplification of the CTC-mRNA was necessary to obtain gene expression values above background levels, but yielded a high fraction of false positive results when the gene expression signature was used to diagnose breast cancer with the PAM50 gene panel. This was likely due to the leukocyte-derived background RNA expression. Beck et al(58) used a combination of CTC-mRNA and plasma cfRNA to investigate prognostic and predictive biomarkers in lung cancer patients. They found a set of transcripts related to platelet factor 4 (PF4), present in patients with metastatic lung cancer, that was associated with patient survival. The pre-amplification protocol, used by Porras et al, was also implemented in this study and the authors did not report on amplification bias.

Besides EVs and CTCs as RNA carriers, leukocyte mRNA can also be used on the nCounter platform. Wu et al(78) focused on inflammation markers in leukocyte-derived mRNA, isolated from metastatic breast cancer patients, before and after targeted treatment. They defined a set of genes, related to IL-1 associated inflammation, that was downregulated after treatment initiation. These profile changes indicated an enhanced antitumor activity, and may therefore be exploited to predict therapy response. No pre-amplification step was reported by the authors. Instead of using encapsulated RNA from liquid biosources, cfRNA can also be extracted from whole blood. In a recently published study, Kossenkov et al(75) developed an nCounter-based classifier that was able to differentiate between low-dose computed tomography (CT) detected, malignant and non-malignant pulmonary nodes. The final classifier made use of only 41 transcripts to distinguish between malignant pulmonary nodes that require follow-up and non-malignant nodes. The use of whole blood samples led to higher RNA quantities, and prevented the need for a pre-amplification step.

Finally, EV-DNA has been analyzed with nCounter to detect mutations. Although only one sample was analyzed using this methodology, Kamyabi et al(61) reported on the use of plasma EV-DNA to detect a KRAS mutation in a pancreatic cancer patient(61). Since the quantity of DNA that can be extracted from EVs is expected to be much lower than the commonly used total plasma cfDNA extraction, the nCounter platform may provide new perspectives to standardize the detection of targetable genetic alterations in liquid biopsy samples. These studies combined indicate that, although the evidence is still preliminary, analysis of liquid biopsies on the nCounter platform could be a valuable tool to detect tumor specific alterations and transcriptional profiles. These alterations and patterns can be developed into biomarker assays and used for diagnostic, predictive and response evaluation purposes on the clinically employable nCounter analysis platform.

Central motivation and objectives

Taking into account the information that has already been published on liquid biopsies and the nCounter platform, this PhD thesis aimed to validate the use of the NanoString nCounter platform for the analysis of liquid biopsies. In particular:

- Chapter 2 contains a validation study on liquid biopsy-derived ctDNA for mutation detection
- In Chapter 3 we optimized existing protocols to analyze EV-mRNA
- We then compared FFPE tissue and matched plasma EV-mRNA analysis to predict risk of checkpoint inhibitor pneumonitis in Chapter 4
- In Chapter 5 we explored the use of EV-mRNA analysis for early lung cancer detection
- Finally, Chapter 6 contains a thorough discussion regarding the obtained results, followed by an overview of future perspectives.

Accepted manuscripts:

- Gimenez-Capitan A, Bracht JWP, García-Mosquera JJ et al. Multiplex detection of clinically relevant mutations in liquid biopsies of cancer patients using a hybridization-based platform. *Clinical Chemistry 2021. DOI: 10.1093/clinchem/hvaa248 [published Online First: 2021/01/14].*
- Bracht JWP, Gimenez-Capitan A, Huang CY et al. Analysis of extracellular vesicle mRNA derived from plasma using the nCounter platform. *Sci. Rep. 2021.* 11:3712.

Chapter 2. Multiplex detection of clinically relevant mutations in liquid biopsies of cancer patients using a hybridization-based platform.

Ana Giménez-Capitán, Jillian Bracht, Juan José García, Núria Jordana-Ariza, Beatriz García, Mónica Garzón, Clara Mayo-de-las-Casas, Santiago Viteri-Ramirez, Alejandro Martinez-Bueno, Andrés Aguilar, Ivana Gabriela Sullivan, Eric Johnson, Chung-Ying Huang, Jay L Gerlach, Sarah Warren, Joseph M. Beechem, Cristina Teixidó, Rafael Rosell, Noemí Reguart, Miguel A. Molina-Vila.

Clinical Chemistry 2021. [published Online First: 2021/01/14].

DOI: https://doi.org/10.1093/clinchem/hvaa248

Abstract

Background: With the advent of Precision Oncology, liquid biopsies are quickly gaining acceptance in the clinical setting. However, in some cases, the amount of DNA isolated is insufficient for Next Generation Sequencing (NGS) analysis. The nCounter platform could be an alternative, but it has never been explored for detection of clinically relevant alterations in fluids.

Methods: Circulating-free DNA (cfDNA) was purified from blood, cerebrospinal fluid and ascites of cancer patients and analyzed with the nCounter 3D Single Nucleotide Variant (SNV) Solid Tumor Panel, which allows for detection of 97 driver mutations in 24 genes.

Results: Validation experiments revealed that the nCounter SNV panel could detect mutations at allelic fractions of 0.02%-2% in samples with ≥5 pg mutant DNA/μL. In a retrospective analysis of 70 cfDNAs from cancer patients, the panel successfully detected *EGFR*, *KRAS*, *BRAF*, *PIK3CA* and *NRAS* mutations when compared with previous genotyping in the same liquid biopsies and paired tumor tissues [Cohen kappa of 0.96 (CI=0.92-1.00) and 0.90 (CI=0.74-1.00), respectively]. In a prospective study including 91 liquid biopsies from patients with different malignancies, 90 yielded valid results with the SNV panel and mutations in *EGFR, KRAS, BRAF, PIK3CA, TP53*, *NFE2L2*, *CTNNB1*, *ALK*, *FBXW7* and *PTEN* were found. Finally, serial liquid biopsies from a NSCLC patient revealed that the semi-quantitative results of the mutation analysis by the SNV panel correlated with the evolution of the disease.

Conclusions: The nCounter platform requires less DNA than NGS and can be employed for routine mutation testing in liquid biopsies of cancer patients.

Introduction

Although genetic analysis of tumor tissue provides useful information for prognosis and treatment decision making, a significant percentage of advanced-stage cancer patients cannot be biopsied or the amount of tumor tissue is insufficient for genetic analyses. In addition, repeated sampling for monitoring the course of the disease and detecting the emergence of mechanisms of resistance is frequently not feasible. Liquid biopsies constitute a minimally invasive, safe, and sensitive alternative in these cases; and are quickly gaining acceptance in the clinical setting(50, 82-87).

Circulating free DNA (cfDNA) purified from blood or other body fluids(88) is the most commonly used type of liquid biopsy. In cancer patients, cfDNA contains a variable fraction of DNA originating in the tumor cells (circulating tumor DNA, or ctDNA) and can be used to identify clinically relevant mutations, amplifications and gene fusions(85, 88-90), with polymerase chain reaction (PCR) based methods and targeted next-generation sequencing (NGS) being the most frequently used techniques. The nCounter platform (NanoString Technologies) is a relatively novel technology initially developed for multiplex analysis of RNA molecules, and has been successfully applied for the detection of clinically relevant fusion transcripts and gene signatures in tumor tissues(18, 24, 91). In addition, a new hybridization probe chemistry has been developed for the detection of hotspot somatic variants in tumor tissue samples(92). However, despite the growing number of laboratories using nCounter, the platform has never been tested for the routine analysis of liquid biopsies.

In this study, we performed a retrospective validation on collection-stored cfDNA samples that revealed an excellent correlation of nCounter with other methodologies for mutation detection. Then, we prospectively analyzed fluids derived from cancer patients and were able to detect a substantial number of relevant mutations, demonstrating that nCounter can be implemented in the clinical setting for the routine testing of liquid biopsies.

Materials and methods

Patients and cell lines

Fifteen liquid biopsy samples from healthy donors and seventy from cancer patients were used for the retrospective validation of the nCounter Vantage 3D DNA Single Nucleotide Variant (SNV) Solid Tumor Panel (NanoString Technologies) (**Table 1**). All of them had been stored in a sample collection approved by the Spanish Ministry of Health (reference number C.0005039). Then, 91 liquid biopsies (**Table 2**) from 83 patients collected in six Spanish hospitals were analyzed with the same panel (online **Supplemental Tables 1 and 2**). The study was carried out in accordance with the principles of the Declaration of Helsinki under an approved protocol of the institutional review board of Quirón Hospitals. Written informed consent was obtained from all patients and documented; samples were de-identified for patient confidentiality. Clinical information collected from each patient was limited to stage, gender, smoking status and tumor histology. Cell lines with EGFR, KRAS, PIK3CA, BRAF, and

NRAS mutations were used for analytical validation purposes and also as positive and negative controls (online **Supplemental Table 3**).

Table 1. Characteristics of the liquid biopsy samples included in the retrospective cohort.

Characteristics	**N =70**	**(%)**
Type of tumor		
Lung Cancer	49	70%
Colorectal Cancer	11	15.7%
Breast Cancer	1	1.4%
Melanoma	5	7.3%
Leukemia	1	1.4%
Pancreatic	1	1.4%
Endometrial	1	1.4%
Ovarian	1	1.4%
Type of fluid		
Plasma	62	88.6%
Ascites	3	4.3%
Serum	3	4.3%
Pleural fluid	1	1.4%
Cerebrospinal fluid	1	1.4%
Mutations previously detected by NGS or Q-PNA-PCR		
EGFR	19	27.1%
KRAS	9	12.%
BRAF	11	15.7%
NRAS	1	1.4%
PIK3CA	3	4.3%
EGFR and *PIK3CA*	2	2.8%
KRAS and *PIK3CA*	1	1.4%
BRAF and *PIK3CA*	1	1.4%
NRAS and *KRAS*	2	2.8%
No mutations	21	30%

DNA isolation

Plasma samples (10 mL) were collected in sterile Vacutainer tubes (BD) and cerebrospinal-, pleural- and ascitic fluid samples (3 to 500 mL) in sterile containers. After two consecutive centrifugation steps (500 g, 10 min), cfDNA was purified using the QIAsymphony® DSP Virus/Pathogen Midi Kit and a QIAsymphony robot (Qiagen), following the manufacturer's instructions. Initial volume was 1.2 mL, final elution volume was 30 µL. DNA concentration was estimated using Qubit®. Finally, DNA from the cell lines was purified using the DNA Easy® extraction kit (Qiagen,), according to the manufacturer's instructions.

Mutation detection by nCounter

The nCounter Vantage 3D DNA SNV Solid Tumor Panel enables detection of 97 driver mutations in 24 clinically relevant genes (online **Supplemental Table 4**). For mutation

Table 2. Characteristics of the samples prospectively evaluated in the study.

Characteristics	N = 90	(%)
Tumor type and histology		
Lung Cancer	51	56.6%
Adenocarcinoma	42	46.6%
Squamous	2	2.2%
Others	7	7.7%
Colorectal Cancer	20	22.2%
Adenocarcinoma	19	21.1%
Others	1	1.1%
Breast Cancer	4	4.4%
Melanoma	4	4.4%
Others	11	12.2%
Collection time		
Basal	62	68.8%
Progression	8	8.8%
Follow up	18	20%
Unknown	2	2.2%
Type of fluid		
Plasma	87	96.7%
Pleural fluid	1	1.1%
Cerebrospinal fluid	2	2.2%

detection using nCounter, 5 μL of purified cfDNA and a reference DNA (NanoString Technologies, provided with the panel) were subjected to a 21-cycle preamplification step in a Verity thermal cycler (Applied Biosystems), according to the manufacturer's instructions. Amplified DNA was denatured at 95ºC for 10 minutes and hybridized at 65ºC for 18-24 hours with the SNV pool, which contains mutation-specific and exon-specific probes that bind to DNA independently of the presence of mutations. Capture, cleanup and digital data acquisition were performed using the nCounter Prep Station™ and Digital Analyzer™ (NanoString Technologies) (online **Supplemental Fig. 1**).

Data analysis

Count values were exported to Excel 2016 (Microsoft) using nSolver software v4.0 (NanoString Technologies). For each mutation, samples with count values lower than the reference DNA were automatically considered negative and excluded from further analysis. The reference consists of a wild-type DNA that does not harbor any mutation and allows estimating the "background noise" counts for every mutation targeted by the kit. The remaining mutation-specific counts were normalized using the geometric mean of the exon counts of the corresponding gene in the same sample. The same procedure was applied to the count values derived from the reference DNA. Finally, the mutation-specific normalized counts of the samples were divided by the corresponding normalized counts of the reference DNA and the

result was subjected to a base-2 logarithmic transformation to obtain the log MUT values. For every mutation in the SNV panel, the average log MUT of all the negative samples for this particular mutation included in the retrospective study plus three standard deviations (SD) was established as the cut-off value for positivity. The only exceptions were KRAS mutations, where the mean plus two SDs was used.

Mutation testing by PNA-Q-PCR and NGS
Samples used in the retrospective validation had been previously genotyped by PNA-Q-PCR or NGS(85, 88, 89, 93, 94). For details about these two techniques, see the online Supplemental Methods.

Results

Analytical validation

First, we analyzed 15 cfDNA samples purified from the blood of healthy donors. All of them tested negative for the 97 mutations targeted by the Vantage 3D DNA SNV panel. Next, using DNA from two mutant cell lines, we found that 5 pg of mutated genomes per µL were sufficient to detect *EGFR* 15-bp deletions and *KRAS* G12C mutations (online **Supplemental Table 5**). Serial dilutions of a mixture of DNAs from 13 mutant cell lines spiked into a pan-negative cfDNA were employed to determine the limit of detection of the panel. Mutations in *EGFR*, *KRAS*, *NRAS* and *PIK3CA* were detected at allelic fractions between 0.02% and 2% (online **Supplemental Table 6**). Finally, spiked samples of two mutant cell lines and a cfDNA purified from the blood of a *KRAS* G12D positive patient were tested on different days by different operators, and the reproducibility of the SNV panel was found to be 100% (online **Supplemental Table 7**).

Retrospective validation in clinical samples

A total of 70 liquid biopsies from cancer patients were used in the retrospective validation of the nCounter SNV panel. Most of them corresponded to plasma samples (n=62), but sera (n = 3), ascites (n = 3), PE and CSF (n=1 each) were also represented (**Table 1**). Regarding tumor types, the majority of samples were from lung cancer patients (n=49), followed by colorectal cancer (n=11), melanoma (n=5) and other malignancies (n=5). All liquid biopsies in the retrospective cohort had been previously genotyped for *EGFR*, *KRAS*, *BRAF*, *NRAS* and *PIK3CA* hotspot mutations by NGS or PNA-Q-PCR. The cfDNAs were re-analyzed using the nCounter SNV panel, the counts for each mutation were normalized and positive and negative samples identified as explained in Methods (see also online **Supplemental Fig. 2**). The results obtained for three representative hotspot mutations are shown in **Figure 1**. In all cases, the distribution of the normalized counts was bimodal, with the mutant samples representing a different subpopulation.

The results of the previous cfDNA genotyping were used for comparison purposes (**Fig. 2**). For *EGFR*, *KRAS*, *BRAF*, *NRAS* and *PIK3CA* mutation detection, nCounter and NGS/PNA-Q-PCR showed concordance rates ranging from 97.1% to 100% and Cohen kappas from 0.91 to 1.00 (**Table 3**). If the five genes were considered together, mutation status by nCounter showed

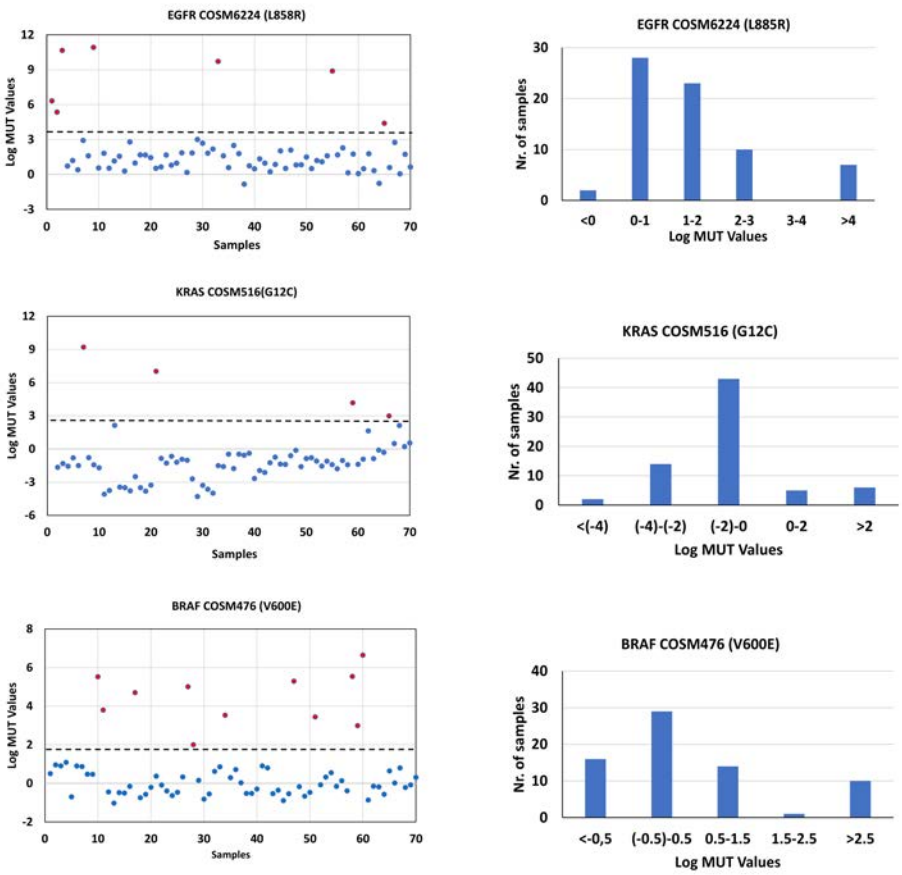


Figure 1. Results of the mutation detection by nCounter in the retrospective cohort (n=70) for three representative mutations, L858R in EGFR **(A)**, G12C in KRAS **(B)** and V600E in BRAF **(C)**. Left, individual normalized counts in the liquid biopsy samples, expressed as log MUT values. The dotted lines indicate the threshold. Right, distribution of the log MUT values.

an almost perfect agreement with previous genotyping, with only four discordant cases, 0.96 Cohen kappa (CI=0.92-1.00) and 98.9% concordance (CI=97.1-99.7%).

The four discordant samples were further investigated (**Fig. 2**). One corresponded to a plasma sample with a T790M in *EGFR* detected by PNA-Q-PCR at an extremely low allelic fraction (0.004%), well below the limit of detection of the nCounter SNV panel. The only discordant sample for *BRAF* had been positive for a V600K mutation by PNA-Q-PCR, with a 2.1% allelic fraction. The plasma sample had been stored for four years at the moment of the nCounter analysis and showed very low exonic counts, suggesting cfDNA degradation. Regarding the two samples discordant for *KRAS*, one corresponded to a plasma positive for a Q61L mutation by nCounter but negative by PNA-Q-PCR and one to a serum sample harboring a G12S mutation by PNA-Q-PCR at 0.12% allelic fraction, not detected by nCounter.

Paired tissue samples with complete genotyping results were available for 30 liquid biopsies included in the retrospective study. For *EGFR*, *KRAS*, *BRAF*, *NRAS* and *PIK3CA* mutation detection, nCounter in liquid biopsy showed 71-100% sensitivity and 100% specificity vs. paired tissue (online **Supplemental Table 8**). When the five genes were considered together,

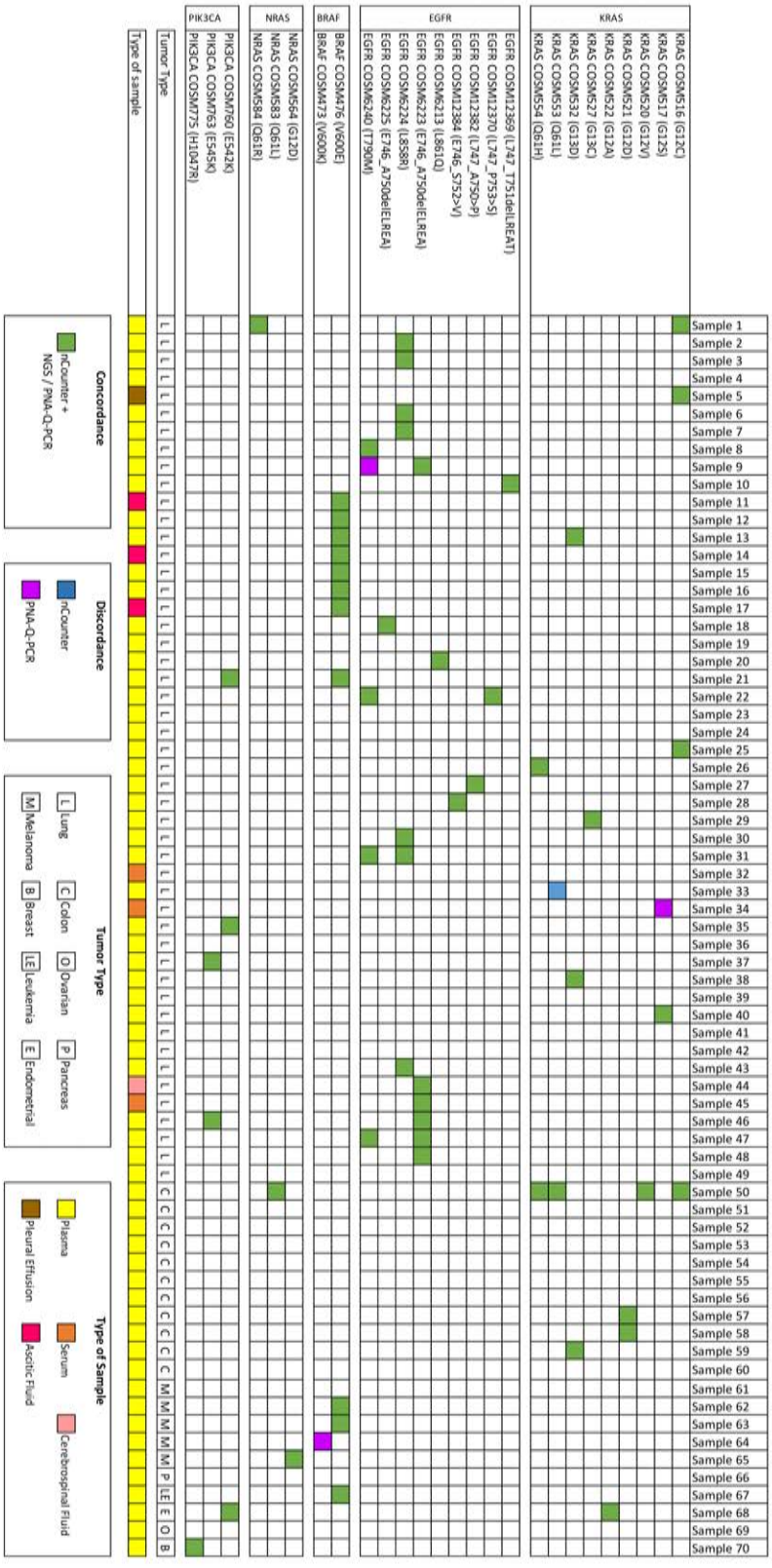


Figure 2. Heatmap of the liquid biopsies included in the retrospective cohort (n=70). All samples were analyzed by nCounter and an alternative technique, either NGS or PNA-Q-PCR. Green, mutations detected by nCounter and the alternative technique. Blue, mutations detected only by nCounter. Yellow, mutations detected only by the alternative technique.

Table 3. Concordance of mutation detection by nCounter with NGS and PNA-Q-PCR in liquid biopsy samples. The 95% confidence intervals are indicated for the overall results.

	nCounter vs. NGS/PNA-Q-PCR					
Genes*	***EGFR***	***KRAS***	***BRAF***	***NRAS***	***PIK3CA***	**Overall**
Nº concordant results	69	68	69	70	70	346
Nº discordant results	1	2	1	0	0	4
Sensitivity	100%	93.3%	100%	100%	100.0%	94.7% (CI=85.4-98.9%)
Specificity	98.0%	98.2%	98.3%	100%	100.0%	99.7% (CI=98.1-100%)
Concordance	98.6%	97.1%	98.6%	100%	100.0%	98.9% (CI=97.1-99.7%)
Cohen kappa	0.96	0.91	0.95	1.00	1.00	0.96 (CI=0.92-1.00)

*Samples not carrying a mutation in a particular gene were used as negatives for this gene, independently of the mutational status of other genes.

Abbreviations: NGS—next generation sequencing; PNA-Q-PCR--quantitative PCR in presence of a quencher-labeled peptide nucleic acid

mutation status in liquid biopsy by nCounter showed an almost perfect agreement with previous genotyping in tissue biopsies, with a Cohen kappa of 0.90 (CI=0.74-1.00) and a 97.3% concordance (CI=93.1-99.2%).

Finally, we compared the log MUT values obtained by nCounter with the allelic fractions previously found by NGS or PNA-Q-PCR in the same samples. For this analysis, we selected the *KRAS*-positive liquid biopsies and we found a linear correlation between the log MUT *KRAS* values and the log2 of the *KRAS* mutant allelic fractions derived from NGS or PNA-Q-PCR (R^2=0.63; Pearson r=0.80; online **Supplemental Figure 3**).

Prospective analysis of liquid biopsies

During a 6-month period (December 2018 to June 2019), liquid biopsy samples from 83 cancer patients were collected, submitted to DNA extraction and prospectively analyzed using the SNV nCounter panel. Six patients had two or more fluid samples available (online **Supplemental Table 1**) bringing the total number of liquid biopsies tested to 91. In all cases, the concentration of purified cfDNA was less than 1 ng/mL, as measured by Qubit. Despite these low concentrations, only one of the 91 liquid biopsies showed very low exonic counts and was considered as not evaluable.

The characteristics of the 90 liquid biopsy samples finally included in the prospective study are presented in **Table 2**. The majority of them corresponded to plasma samples (n=87), although two CSF and a PE were also included. Regarding the type of malignancy, most liquid biopsies were obtained from lung cancer patients (n=51); followed by colorectal (CRC) (n=20), breast (n=4), melanoma (n=4), prostate (n=3) and other tumors (n=8), including thyroid, ovarian, pancreatic, sarcoma and kidney cancer.

The results of the mutation analysis by nCounter are presented in **Fig. 3A**. Among the 51 fluid samples from lung cancer patients, mutations in *EGFR* were found in seven samples; three harbored exon 19 deletions, three exon 21 point mutations and one a G719A mutation in exon

18. Regarding *KRAS* mutations, seven samples were positive for the G12 (n=4), Q61 (n=2) or G13 (n=1) positions. Hotspot mutations in *PIK3CA* were found in four lung cancer liquid biopsies, three of them coming from the same patient (see below). Finally, eleven samples harbored mutations in a variety of genes, including *PTEN*, *CTNNB1*, *GNAQ*, *NFE2L2*, *FGFR2*, *FBXW7*, *TP53* and *ALK*. The two liquid biopsies positive for *ALK* mutations corresponded to patients in progression to ALK targeted therapies. In the case of the CRC patients, 2/20 liquid biopsies were positive for *KRAS* mutations and one each for *BRAF*, *NRAS* and *CTNNB1*. Finally, among samples collected from patients with other malignancies, a G12D mutation in *NRAS* was found in two serial samples from a melanoma patient in progression to BRAF/MEK inhibitors, while two consecutive liquid biopsies from a pancreatic cancer patient harbored the I195T mutation in *TP53* and a thyroid sample a Q61R mutation in *NRAS*.

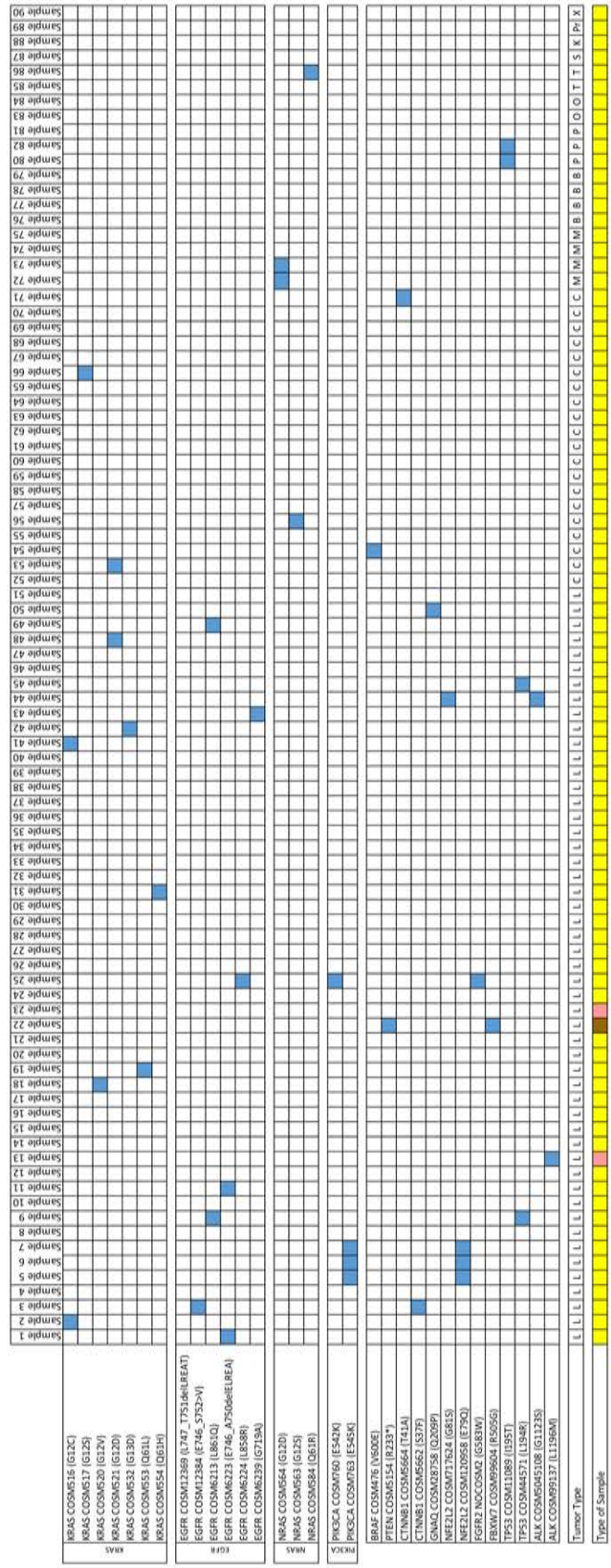
Eighty-four of the 90 liquid biopsy samples in the prospective cohort had not been previously submitted to any kind of testing. The remaining six samples (four blood and two CSFs) had been analyzed using liquid-biopsy NGS panels, yielding invalid results. Interestingly, all of them were evaluable by nCounter, which detected drivers in two, a *KRAS* G12D mutation in a blood sample and a L1196M *ALK* resistance mutation in a CSF.

Validation of results of the prospective testing

A subset of 16 cfDNA samples from the prospective cohort with remaining material was subsequently submitted to NGS for validation purposes. The subset included samples from the 11 patients carrying mutations in genes not validated in the retrospective part of the study; such as *TP53* or *ALK* (**Figure 3A**). The NGS panel employed did not cover *CTNNB1*, *FBXW7*, *FGFR2* or *GNAQ* and mutations in these genes could not be confirmed. For the rest of genes, NGS showed concordant results with nCounter for *EGFR* (n=4), *TP53* (n=3), *PIK3CA* (n=2), *NRAS* (n=2), *ALK* (n=1) and *NFE2L2* (n=1) mutations. Sequencing only failed to detect a *PTEN* mutation, while two samples were not evaluable due to insufficient material (**Figure 3A** and online **Supplemental Table 9**).

Mutation analysis of serial samples

A clinical case where serial liquid biopsies were collected will be described in further detail (**Fig. 3B**). It corresponded to a lung cancer patient, diagnosed in February 2016, harboring an *EML4-ALK* fusion and wild-type for *EGFR*, *KRAS*, *BRAF* and *PIK3CA* in tumor tissue at presentation. The patient started ceritinib on February 2016, which was replaced by alectinib three months later due to hepatic toxicity. The patient was in remission for more than two years, and the four serial blood samples obtained from April 2016 to June 2017 were pan-negative by the nCounter SNV panel. In contrast, two mutations were detected in a fifth sample collected in April 2018, E545K in *PIK3CA* and E79Q in *NFE2L2*, at log MUT values of 2.1 and 3.2, respectively. The patient showed radiological progression in multiple sites four months later, in August 2018. Alectinib was then replaced by brigatinib, a blood sample obtained in October revealed a substantial increase in the log MUT values of both mutations, which rose to 5.6 and 7.8. A subsequent radiological evaluation demonstrated lack of response to brigatinib.


A
KRAS
KRAS COSM516 (G12C)
KRAS COSM517 (G12S)
KRAS COSM520 (G12V)
KRAS COSM521 (G12D)
KRAS COSM532 (G13D)
KRAS COSM553 (Q61L)
KRAS COSM554 (Q61H)
EGFR
EGFR COSM12369 (L747_T751delLREAT)
EGFR COSM12384 (E746_S752>V)
EGFR COSM6213 (L861Q)
EGFR COSM6223 (E746_A750delELREA)
EGFR COSM6224 (L858R)
EGFR COSM6239 (G719A)
NRAS
NRAS COSM564 (G12D)
NRAS COSM563 (G12S)
NRAS COSM584 (Q61R)
PIK3CA
PIK3CA COSM760 (E542K)
PIK3CA COSM763 (E545K)
BRAF COSM476 (V600E)
PTEN COSM5154 (R233*)
CTNNB1 COSM5664 (T41A)
CTNNB1 COSM5662 (S37F)
GNAQ COSM28758 (Q209P)
NFE2L2 COSM717624 (G81S)
NFE2L2 COSM120958 (E79Q)
FGFR2 NOCOSM2 (G583W)
FBXW7 COSM99604 (R505G)
TP53 COSM11089 (I195T)
TP53 COSM44571 (L194R)
ALK COSM45045108 (G1123S)
ALK COSM59137 (L1196M)
Tumor Type
Type of Sample
Tumor Type
L Lung
C Colon
O Ovarian
P Pancreas
M Melanoma
B Breast
T Thyroid
X Unknown
S Sarcoma
K Kidney
Pr Prostate
Type of Sample
Plasma
Cerebrospinal Fluid
Pleural Effusion

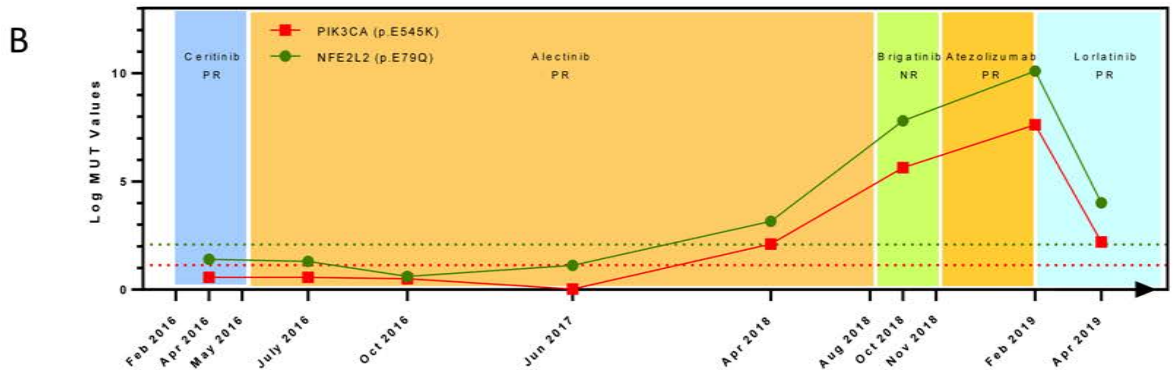


Figure 3. (A, previous page) Heatmap representing the results of the mutation detection by nCounter in the liquid biopsies included in the prospective cohort (n=90). **(B)** Case report of an EML4-ALK advanced NSCLC patient. Results of the mutation analysis of serial blood samples using the nCounter SNV panel are presented, together with the clinical evolution of the patient.

Atezolizumab was then administered, but the log MUT values further increased in blood, to 7.6 and 10.1; an evaluation of response on February 2019 revealed progression of the disease. Lorlatinib was finally started and the patient underwent a partial response that was accompanied by a substantial decrease in the log MUT values of the two mutations in plasma, which dropped to 2.2 and 4.0. Five serial samples of the patient had remaining material after nCounter, and *PIK3CA* mutations were tested by PNA-Q-PCR for validation purposes. The results showed a good agreement with those previously obtained by nCounter (online **Supplemental Table 10**).

Discussion

Precision oncology, based on the assessment of molecular markers predictive of treatment outcome, has transformed clinical practice for many types of cancer. Since tumor tissue is not always available or sufficient for genetic testing, liquid biopsies have quickly gained acceptance in the clinical setting(50, 82-87). Initially, blood and other fluids from cancer patients were mainly employed to detect clinically relevant mutations in *EGFR*, *KRAS*, *NRAS* or *BRAF* using PCR-derived techniques targeting a limited number of exons(95). However, in the last few years, several NGS platforms have been adapted to the requirements of liquid biopsies and are being employed by a growing number of laboratories.

The nCounter technology is a multiplex hybridization-based assay(11) that differs from NGS techniques, being based on direct counting of the RNA or DNA molecules(96). The technology has been adapted for the detection of mutations in DNA purified from tumor tissue(92) by the design of three types of probes (S, M and T). S probes have two binding regions, one detecting the presence of the mutation and the other binding to a nearby wild-type sequence; while M probes act like signal attenuators of the wild-type sequences, and T probes facilitate detection (online **Supplemental Figure 1**). The nCounter technology has been widely used in research studies to simultaneously determine mRNA expression levels of hundreds of genes in biological samples(97, 98), including liquid biopsies of cancer patients(99, 100). Some of these exploratory studies have led to the identification of expression-based signatures to discriminate

malignant lung nodules(75) and predict outcome to immunotherapy in solid tumors(101) or drug sensitivity in prostate cancer(102). However, nCounter has never been used for routine testing of liquid biopsies of cancer patients and the only signature in clinical use is the U.S. Food and Drug Administration-approved, tumor-tissue based Prosigna, which determines the risk of recurrence in breast cancer(18, 91, 103).

Here we have described the validation of the nCounter SNV panel, which can detect mutations and small indels in 27 genes, for the genotyping of liquid biopsy samples; followed by the implementation of the assay for the prospective testing of blood and other fluids of cancer patients. During the validation study, we found that 5 pg of mutant DNA, purified from 1.2 mL of blood or other body fluids, was sufficient for successful analysis. Regarding limits of detection, using spiking experiments with cell lines we found values of 0.02-2% allelic fraction. The concentration of cfDNA in the liquid biopsy samples used in our study ranged between 0.1 and 0.5 ng/µL, and 5 µL were loaded in the nCounter assay; meaning that the total cfDNA input was 500-2500 pg. Since 5 pg of mutant DNA are required, the minimum allelic fractions needed for mutation detection in liquid biopsy samples by nCounter would be 0.1%-0.02%, in coincidence with the values found in cell line experiments In contrast, using cfDNA inputs lower than 500 pg could lead to higher limits of detection.

The limits of detection of the nCounter SNV panel compare favorably with the requirements of liquid biopsy NGS assays (online **Supplemental Table 11**) and the low requirement of input material explains that, among the 91 liquid biopsies prospectively analyzed with the panel, only 1 (1.1%) sample was not evaluable. Remarkably, valid results could be obtained for two CSF samples, which are usually collected at small volumes and contain particularly low amounts of cfDNA(88, 104, 105). One of them corresponded to an *EML4-ALK* positive patient progressing to targeted therapies, where a L1196M resistant mutation was identified and used for the selection of subsequent therapies. Of note, NGS had been previously attempted with these two CSFs but yielded invalid results due to insufficient DNA concentration.

Finally, comparison with results obtained in tissue biopsies revealed diagnostic sensibility and specificity of 84.6% and 100%, respectively. All these values are in the range of those reported for liquid biopsy NGS platforms such as Guardant Health or Oncomine (online **Supplemental Table 11**). One of the limitations of our study was that the number of paired tissue samples was limited and the confidence interval calculated for the diagnostic sensitivity had a wide range, from 66.4-93.8%. However, we were able to compare the results obtained by nCounter in all the liquid biopsies in the retrospective cohort with the previous genotyping of the same samples by NGS or PNA-Q-PCR, showing an almost perfect agreement, with 99% concordance and a 0.96 Cohen kappa (CI=0.92-1.00).

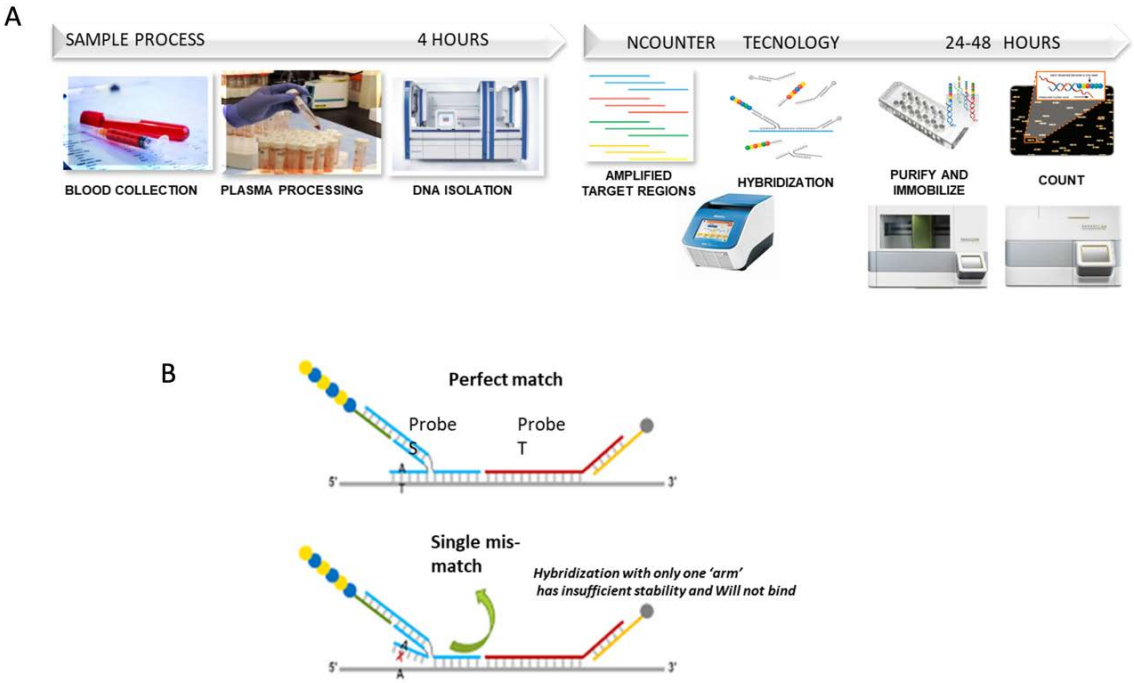
During prospective testing, *EGFR* mutations were found in 7/51 liquid biopsy samples from the NSCLC patients analyzed (13.7%)(3), a percentage in the range of the frequency described in European populations, while *KRAS* mutations were detected in another seven patients; *EGFR* and *KRAS* mutations were mutually exclusive, as expected. In the case of CRC, *KRAS*, *NRAS* or *BRAF* mutations were found in 5/21 (23.8%) of liquid biopsies. This relatively low

prevalence can be explained in two ways, (i) a considerable number of the CRC patients were stage I-IIIA, with less tumor burden than advanced patients; (ii) 3/21 samples corresponded to samples of patients in response to therapy. Finally, results obtained in serial liquid biopsies indicate that the nCounter SNV panel could be used to follow the evolution of cancer patients. Although allelic fractions as such cannot be estimated, the log MUT values were directly dependent on allelic fractions and could be easily calculated and monitored (**Figure 4**).

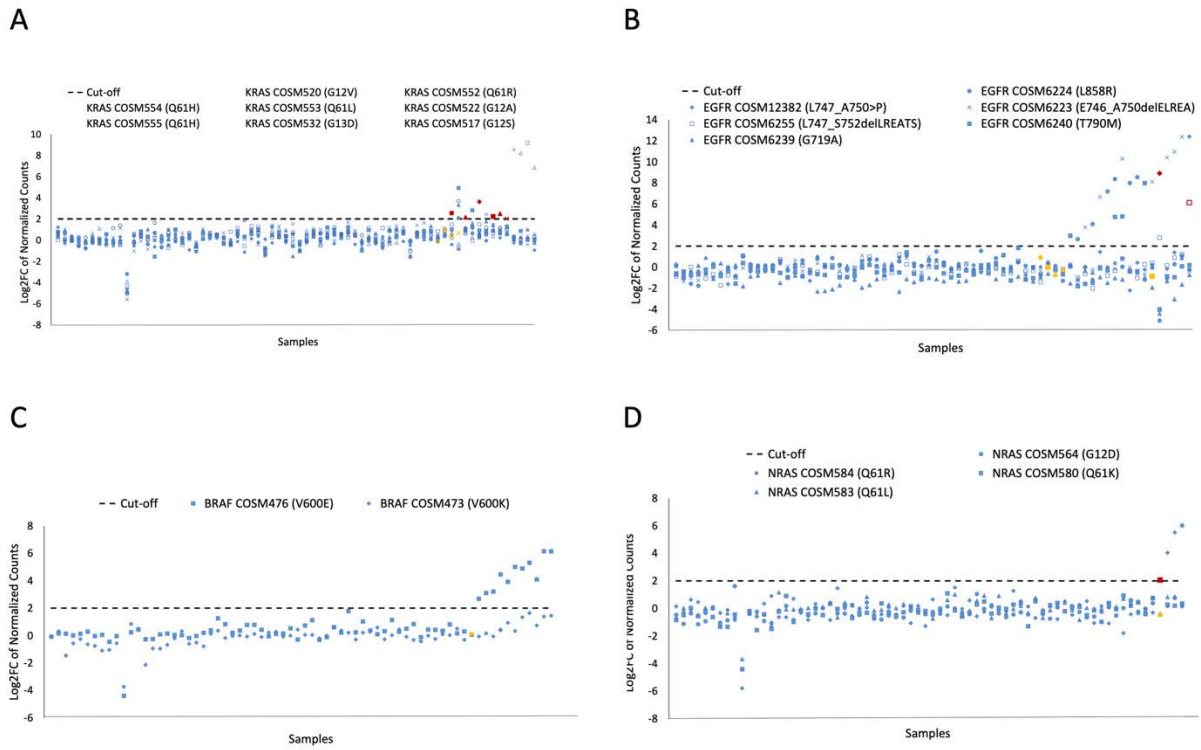
The nCounter platform confers several advantages over NGS techniques for mutation detection in liquid biopsies. It requires a substantially lower amount of material, has a 24-48h turnaround time with relatively short hands-on time, sample preparation is simple compared to NGS, data analysis is straightforward and does not require bioinformatics expertise (**Supplemental Figure 1**). The main disadvantage of the nCounter platform is that, not being a sequencing technique, it cannot detect mutations other than those contained in the SNV panel, although the panel can be customized. Detection of mutations by nCounter can be particularly useful in some settings. Examples include liquid biopsies with small volumes and/or low concentrations of cfDNA, such as CSF samples(88) or pleural and peritoneal lavages(94); liquid biopsies where NGS has failed; or cancer patients in urgent need of a genetic testing to determine if they are eligible for targeted therapies. Also, the nCounter SNV panel is well suited for monitoring patients in response to therapy where repeated NGS of liquid biopsies would not be cost-effective and can be spared until progression.

In summary, we have demonstrated that the nCounter SNV panel, initially developed for tumor tissue samples, shows an analytical performance similar to NGS in liquid biopsies, requires less material and can be implemented for multiplex detection of somatic mutations in the clinical setting. Our results also pave the way for testing the performance of nCounter for the detection of other relevant alterations in liquid biopsies from cancer patients, such as gene fusions or expression levels of genes predictive of response to immunotherapy.

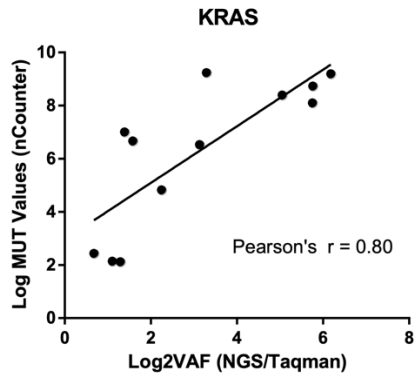
Supporting information



Supplemental Figure 1. (A) Workflow of the mutation analysis of liquid biopsies using nCounter. **(B)** Hybridization-based detection of mutations using S and T probes.



Supplemental Figure 2. Results of the mutation detection by nCounter in the retrospective cohort (n=70) for KRAS **(A)**, EGFR **(B)**, BRAF **(C)** and NRAS **(D)**. Individual normalized counts in the liquid biopsy samples, expressed as log MUT values, are represented. The dotted lines indicate the threshold for positivity.



Supplemental Figure 3. Comparison of the log2MUT values obtained by nCounter and the log2 of allelic fractions obtained by NGS or PNA-Q-PCR in 12 liquid biopsies positive for KRAS mutations.

The supplementary methods and tables can be found online at: https://academic.oup.com/clinchem/advance-article/doi/10.1093/clinchem/hvaa248/6095704?searchresult=1#supplementary-data

Conflicts of interest

CT has received fees for consultancy/advisory roles from Pfizer, Novartis, MSD, Roche, AstraZeneca and Takeda, and research funding from Pfizer and Novartis. NR has received fees for consultancy/advisory roles from MSD, BMS and Pfizer, and research funding from Pfizer, Novartis and Ministry of Health, Instituto de Salud Carlos III, Spain. C-YH, SW, RB, JMB and JG are full time employees of NanoString. The rest of the authors declare no conflicts of interest.

Research funding

The project received funding from a European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement ELBA No 765492. NR also received funding from the Ministry of Health, Instituto de Salud Carlos III, Spain (FIS PI16/00890), Some of the reagents used in the project were provided free-of-charge by NanoString Technologies.

Acknowledgements

We thank Ariadna Balada-Bel, María José Catalán, María González-Cao, Margarita Magem, Xavi Gonzalez, Veronica Pereira, Rich Boykin for their contributions. The investigators also wish to thank the patients for kindly accepting to donate samples to this study. We thank all the physicians who collaborated by providing clinical information.

Chapter 3. Analysis of extracellular vesicle mRNA derived from plasma using the nCounter platform.

Jillian WP Bracht, Ana Gimenez-Capitan, Chung-Ying Huang, Nicolas Potie, Carlos Pedraz-Valdunciel, Sarah Warren, Rafael Rosell, Miguel A Molina-Vila.

Scientific Reports 2021. 11:3712.

DOI: https://doi.org/10.1038/s41598-021-83132-0

Abstract

Extracellular vesicles (EVs) are double-layered phospholipid membrane vesicles that are released by most cells and can mediate intercellular communication through their RNA cargo. In this study, we tested if the NanoString nCounter platform can be used for the analysis of EV-mRNA. We developed and optimized a methodology for EV enrichment, EV-RNA extraction and nCounter analysis. Then, we demonstrated the validity of our workflow by analyzing EV-RNA profiles from the plasma of 19 cancer patients and 10 controls and developing a gene signature to differentiate cancer vs. control samples. TRI-reagent outperformed automated RNA extraction and, although lower plasma input is feasible, 500 μL provided highest total counts and number of transcripts detected. A 10-cycle pre-amplification followed by DNase treatment yielded reproducible mRNA target detection. However, appropriate probe design to prevent genomic DNA binding is preferred. A gene signature, created using a bioinformatic algorithm, was able to distinguish between control and cancer EV-mRNA profiles with an area under the ROC curve of 0.99. Hence, the nCounter platform can be used to detect mRNA targets and develop gene signatures from plasma-derived EVs.

Introduction

With the growing global cancer burden, estimated at 18 million cases in 2018(1), earlier cancer detection, enhanced disease monitoring and improved therapy selection are indispensable to enhance patient survival. Liquid biopsies provide a minimally invasive, safe and sensitive surrogate for tissue biopsies. Extracellular vesicles (EVs) are double-layered phospholipid membrane vesicles that are released by most cells, including cancer cells, immune cells and even blood platelets, and can be isolated from practically any body fluid. Cells liberate highly heterogeneous EVs in terms of size (10 nm – 1 um), cargo (nucleic acids, proteins and lipids), membrane composition, biogenesis and biological function(69-71). Several EV enrichment strategies have been described, including ultracentrifugation, size-exclusion chromatography and precipitation. Precipitation buffers capture water molecules and thereby decrease the hydration of particles, allowing their precipitation after a low-speed centrifugation.

Importantly, the active molecules that are found within EVs can be transported to local or distant target cells and execute biological functions, making EVs important mediators of intercellular communication(69, 72, 73). Since the quantity of released EVs and their specific cargo is regulated by the producing cells, the RNA profiles contained within EVs could potentially be used as biomarkers for development and progression of several diseases, including cancer(74). EVs also have the advantage of their lipid bilayer, which makes their cargo particularly stable and allows the use of biobank stored samples.

Gene expression studies using EV-RNA from cancer patients are an active area of research(63, 64). Promising findings have been reported but the lack of standardized protocols for RNA extraction and analysis, leading to inconsistent results, is hampering clinical implementation. Transcriptomic analysis studies are often conducted using quantitative real time PCR (qRT-PCR), microarrays or RNA sequencing (RNAseq), each with their own (dis)advantages(106). While qRT-PCR provides high sensitivity and specificity with a short turnaround time, it only allows for low-throughput transcriptomic analysis of a limited number of genes. Microarrays represent a medium-throughput platform but have a narrower dynamic range of detection and are not suitable to detect genes expressed at either low or high levels. Finally, RNAseq is an accurate, high-throughput platform with a wide dynamic range, but limitations include longer turnaround time, high cost and complex data analysis(8, 9).

In recent years, the NanoString nCounter platform has gained popularity in translational research and clinical settings. The platform provides a simple, sensitive and cost-effective solution for multiplexed analysis of up to 800 RNA targets by direct capturing and counting of individual targets. In addition, it can be used with formalin-fixed paraffin embedded (FFPE) tumor tissue and allows for low quality and quantity tissue samples(11, 92). At this respect, the nCounter-based Prosigna assay, which differentiates breast cancer subtypes and predicts the risk of recurrence based on a 50-gene signature, has been validated in the clinical practice and received FDA approval in 2013(22, 23). Another assay developed using the nCounter platform is the 18-gene Tumor Inflammation Signature (TIS), which was able to predict clinical response to PD-1 blockade in an investigational clinical trial assay(33). These

two assays, which utilize tissue samples, emphasize the potential of the nCounter platform as biomarker assay development tool, especially in diagnostic laboratories. Regarding the analysis of liquid biopsies on nCounter, several studies have investigated the potential of some materials, including cf-(107) and EV-DNA(61), CTC-RNA(58, 77), leukocyte mRNA(78), cfRNA(75) and EV-miRNA(79, 80), with different success rates. However, nCounter has never been tested for the analysis of EV-derived mRNA.

Here, we present a proof-of-concept study where we optimized a workflow for EV enrichment from human blood samples, EV-mRNA purification and subsequent analysis by nCounter. Then, we used the workflow to develop an EV-mRNA based gene signature to differentiate cancer vs. control samples. Our work demonstrates that nCounter can be employed for biomarker discovery based on EV-mRNA.

Materials and methods

Patient Samples

This study was carried out in accordance with the principles of the Declaration of Helsinki under an approved protocol of the institutional review board of Quirón Hospitals. We selected for the study all blood samples from advanced stage cancer patients arriving to our institution in a two month period with sufficient volume for EV extraction after routine genetic testing (n=19). Blood samples from 10 healthy controls were also collected (**Table 1**). Written informed consent was obtained from all participants and documented; samples were de-identified for confidentiality. Clinical information collected from each participant was limited to gender, age, tumor type and stage.

Table 1. Clinical characteristics of the cancer patients and controls included in the study. *ADC: adenocarcinoma; SCC: squamous cell carcinoma; UPS: undifferentiated pleomorphic sarcoma.*

Characteristics	Cancer patients (n = 19)	Controls (n = 10)
Gender – no. (%)		
Male	11 (57.9)	5 (50.0)
Female	8 (42.1)	5 (50.0)
Age - yr		
Median	62	42
Range	45-78	24-53
Tumor type – no. (%)		
Lung (ADC)	10 (52.7)	-
Gastric (ADC)	3 (15.8)	-
Anal (SCC)	2 (10.5)	-
Rectal (ADC)	2 (10.5)	-
Sarcoma (UPS)	2 (10.5)	-
Stage – no. (%)		
Stage III	2 (10.5)	-
Stage IV	17 (89.5)	-

Extracellular Vesicle Enrichment

Whole blood samples (10 mL) were collected in sterile EDTA Vacutainer tubes (BD, Plymouth, UK) and centrifugated twice at 1000 x g for 10 min at room temperature (RT). Plasma samples were stored at -80°C until further processing. The miRCURY Exosome

Serum/Plasma Kit (Qiagen, Hilden, Germany) was used to enrich for EVs from 500 µL plasma, according to the manufacturer's instructions, unless otherwise specified. In short, dead cells and debris (including platelets and fibrin) were cleared with thrombin and centrifugation. Precipitation Buffer was added, samples were incubated overnight at 4°C and the EV fraction was pelleted by centrifugation. Supernatants were collected and stored for separate analysis. EV enriched pellets were resuspended for further processing.

EV Characterization by Western Blot

EV enriched pellets were resuspended in 300 µL ice-cold radioimmunoprecipitation assay buffer containing protease inhibitor mixture (Roche Applied Science, Penzberg, Germany), as previously described(108). Samples were incubated on ice for 30 min, homogenized and centrifugated 15 min, 12.000 x g at 4°C. Supernatants (lysates) were collected and 80 µg proteins were electrophoresed on 10% SDS-polyacrylamide gels (Life Technologies, Carlsbad, CA, USA) and transferred to PVDF membranes (Bio-Rad Laboratories Inc., Hercules, CA, USA). Membranes were blocked in Odyssey Blocking Buffer (Li-Cor Biosciences, Lincoln, NE, USA). All target proteins were immunoblotted with appropriate primary and horseradish peroxidase (HRP)-conjugated secondary antibodies (**Supplementary Table S5**). Chemiluminescent bands were detected in a ChemiDoc MP Imaging System (Bio-Rad Laboratories Inc.).

EV Characterization with Cryogenic Electron Microscopy

Cryogenic Electron Microscopy (Cryo-EM) was performed by the Microscopy Service of the Universitat Autonoma de Barcelona (UAB), and used for direct visualization of EVs using the TEM JEOL 2011 200 KV. As previously described(109), 2 µL volume of the resuspended EV sample was added to a carbon TEM grid. The grid was transferred onto the cyro-preparation chamber of a Leica electron microscope, containing a liquid ethane bath cooled to -180 °C. Using a piece of filter paper, the EV solution was taken off the grid and plunged into the liquid ethane. The orifice trapped frozen EV solution was assembled into a plunger (Leica EM GP) and blotted with Whatman No. 1 filter paper. The grid was placed in a liquid nitrogen bath, and then loaded into a liquid nitrogen-cooled TEM grid holder. The grid holder was placed into a JEOL 2011 TEM microscope. Imaging was performed using a Gatan UltraScan US1000 CCD camera and data was analyzed with Digital Micrograph 1.8.

RNA Extraction

EV-enriched pellets were treated with 4 µg/mL RNase A from bovine pancreas (Sigma-Aldrich, St. Louis, MO) for 1 h at 37°C, to remove extra-vesicular RNA not associated to EVs. For TRIzol LS Reagent (Thermo Fisher Scientific, Waltham, MA) and TRI-reagent (MRC, Cincinnati, OH) extraction, TRIzol solutions were added to a final volume of 1 mL and incubated at RT for 20 min to inactivate RNase A and lyse the EVs. Then, 200 µL Chloroform: Isoamyl Alcohol (24:1) (Panreac Química SLU, Barcelona, Spain) was added and samples were vigorously vortexed and centrifuged at 12,000 x g for 15 min at 4°C. The aqueous upper layer was kept and RNA was precipitated adding 2.5 µL Glycogen (Merck KGaA, Darmstadt, Germany) and 500 µL 2-propanol (Merck KGaA), incubating 10 min at RT and centrifugating

for 10 min, 12,000 x g, at 4°C. The final RNA pellet was washed with 75% ethanol, air dried and dissolved in 20 µL nuclease free water. The QIAsymphony DSP Virus/Pathogen Kit was also tested in the automated QIAsymphony SP System (Qiagen) for RNA extraction from EVs, according to the manufacturer's instructions.

Electrophoretic analysis of RNA
The approximate quantity and size distribution of the isolated EV-RNA was evaluated using the Bioanalyzer RNA 6000 Pico Assay (Agilent Technologies, Santa Clara, CA) according to manufacturer instructions.

DNase Treatment
In order to remove co-isolated DNA, the EV-RNA samples were treated with the DNA-free DNA Removal Kit (Thermo Fisher Scientific), according to manufacturer instructions. In short, 1 µL DNase buffer and 0.5 µL enzyme were added to 7.5 µL RNA sample, followed by incubation at 37°C for 30 min and DNase removal.

Gene Expression Analysis using nCounter
The nCounter Low RNA Input Amplification Kit (NanoString Technologies, Seattle, WA) was used to retrotranscribe and pre-amplify 4 µL EV-derived RNA using 10 cycles. Retrotranscription was carried out in 0.5 mL tubes while pre-amplification, using primers targeting the genes of the Human Immunology V2 Panel (NanoString Technologies), was performed in 384-well plates to prevent sample evaporation. In parallel, a Moloney Murine Leukema Virus (M-MLV) Reverse Transcriptase Enzyme (Thermo Fisher Scientific) was also tested for cDNA synthesis. The Human Immunology V2 Panel (NanoString Technologies) was used to analyze EV-derived, pre-amplified cDNA according to manufacturer instructions. This panel targets 594 general genes involved in the immune response such as cytokines, enzymes, interferons and their receptors. Samples were hybridized for 18 h at 65°C.

Data Normalization and Analysis
Raw nCounter counts of expressed genes were normalized in R and R studio v3.6.3 using the R package *NanoStringNorm.(110)* Normalization was performed following several steps: technical assay variability normalization using the geometric mean of the positive control probes, background correction using the mean plus two times standard deviation (SD) of the negative control probes, and sample content normalization using the total amount of counts for each sample. Normalized counts were log2-transformed, and used for differential expression (DE) analysis. Log2 fold change (FC) of each gene was calculated as the ratio of average log-2 transformed counts of the cancer patient cohort vs. the control cohort. Volcano plots were used to visualize log2 FC on the x-axis and nominal p-values on the y-axis. GraphPad Prism software (version 9.0.0; https://www.graphpad.com/scientific-software/prism/) was used for other statistical testing and to create figures.

Classifier Algorithm Development
Optimal gene selection was performed using recursive feature elimination (RFE). To this end, a leave-one-out cross validation (LOOCV) algorithm was used on the full Human Immunology

V2 gene Panel. The number of genes to select was set at 4, 8, 16 or 579 and the amount of genes that yielded optimal performance after cross-validation was automatically selected. Classification was performed with the selected gene signature using random forest (rf) and k-nearest neighbors (knn) classifiers with three iterations. The model with the highest accuracy was then selected as the final model. Signature scores for each sample were derived from the final model.

Gene Expression Analysis using qRT-PCR

Complementary DNA (cDNA) was synthesized using the M-MLV Reverse Transcriptase Enzyme (Thermo Fisher Scientific). Hereafter, cDNA was added to TaqMan Universal Master Mix (Applied Biosystems) in a 12.5 µL reaction with specific primers and probe designed for each gene. The primer and probe sets were designed using Primer Express Software (version 3.0.1; https://www.thermofisher.com/order/catalog/product/4363993#/4363993) (Applied Biosystems) according to their Ref Seq (http://www.ncbi.nlm.nih.gov/LocusLink) Gene-specific primers were designed as follows: GAPDH, forward: 5′-TGACC TCAACTACATGGTTTACATGTT-3′ and reverse: 5′-TGACGGTGCCATGGAATTT-3′; Caspase 8, forward: 5′-CAGGGCTCAAATTTCTGCCTAC-3′ and reverse: 5′-GAAGAA GTGAGCAGATCAGAATTGAG-3′; CCL5, forward: 5′ CATCT GCCTCCCCATATTCCT 3′ and reverse: 5′ AGTGGGCGGGCAATGTAG 3′. Quantification of gene expression was performed using the QuantStudio 7 Flex Real-Time PCR System (Thermo Fisher Scientific). Expression levels of mRNA were expressed as arbitrary units based on Ct values. Commercial RNAs were used as controls (liver and lung; Stratagene, La Jolla, CA). In all quantitative experiments, a sample was considered not evaluable when the standard deviation of the Ct values was > 0.30 in 2 independent analyses.

Data Availability

Supplementary information is available for this paper

Results

Optimization of plasma EV enrichment and EV-RNA extraction methodologies

EVs were enriched from 500 µL plasma of control samples using the miRCURY Exosome Serum/Plasma Kit (**Fig. 1a**). Final miRCURY sediments were submitted to western blotting, revealing enrichment in the exosome markers Flotillin and CD63, which were absent or detected at low levels in miRCURY supernatants and whole plasma samples. Sediments, supernatants and plasmas were negative for the cell-specific marker calnexin (**Fig. 1b, Supplementary Fig. S1**). Cryogenic electron microscopy (cryo-EM), a commonly used technique for EV characterization(111, 112), was used to visualize the miRCURY sediments, revealing EVs with the classical morphology and a diameter of 100-300 nm, in agreement with the reported 10 nm - 1 µm size range (**Fig. 1c**)(69-71).

TRIzol LS and TRI-reagent are mixtures of phenol, guanidine isothiocyanate and other components routinely used for nucleic acid extractions. We found that the quantity of RNA

that could be isolated from EV-enriched sediments using TRI-reagent was too low to be determined by the

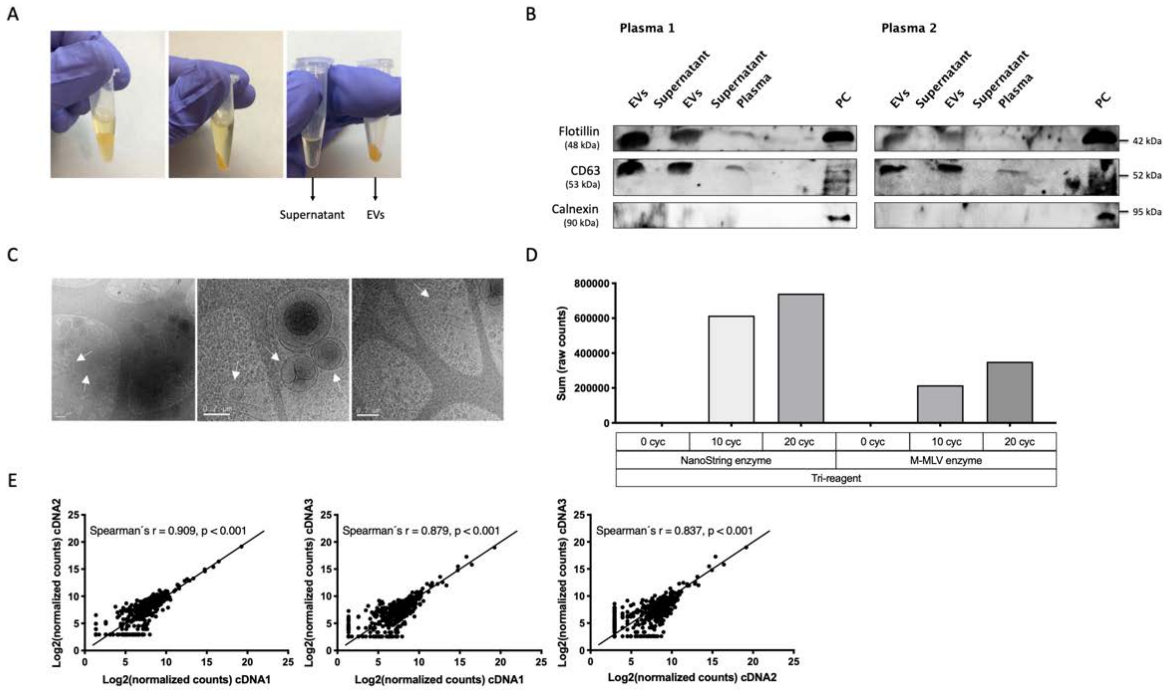


Figure 1. EV enrichment and characterization, assay reproducibility. **(A)** EV enrichment based on precipitation using the miRCURY Exosome Serum/Plasma Kit, including separation of supernatant and EV-enriched pellet. **(B)** Immunoblots showing expression of Flotillin, CD63 (both exosome markers) and Calnexin (cell specific marker) in EV-enriched pellets, supernatants, full plasma and a positive control sample. Experiments were performed in duplicates. Membranes were cut and incubated with specific antibodies for Flotillin, CD63 and Calnexin. Images were cropped for clarity purposes and full membranes can be found in **Supplementary Figure S1**. **(C)** Cryogenic Electron Microscopy (cryo-EM) of EV pellets. Arrows point to extracellular vesicles with different size ranges. Scale bars are 200 nm. **(D)** Total counts by nCounter after EV-RNA extraction using TRI reagent. Two different retrotranscriptases (NanoString versus M-MLV) and three pre-amplification conditions (0, 10 and 20 cycles) were tested. **(E)** Reproducibility experiment comparing the Log2 normalized counts by nCounter from three independent cDNAs derived from a single EV-RNA sample. Spearman´s correlation coefficient is indicated. *EVs: extracellular vesicles; PC: positive control; Cyc: cycles.*

Qubit RNA High Sensitivity Assay Kit (Thermo Fisher Scientific). Bioanalyzer profiles of two representative samples revealed RNA concentrations <150 pg/µL, insufficient for nCounter analysis using the Human Immunology V2 Panel, which we had selected for our study (**Supplementary Fig. S2**). Therefore, we tested retrotranscription and pre-amplification of the EV-mRNA with the nCounter Low RNA Input Amplification Kit, using primers targeting the genes of the Panel. Two reverse transcriptases were compared, the M-MLV and the enzyme provided by the kit, together with 10 vs. 20-cycles for the pre-amplification step. Results indicated that the retrotranscriptase provided by the kit was more efficient in terms of final counts and that both 10 and 20 cycles yielded sufficient raw counts for successful nCounter analysis with the Human Immunology V2 Panel (**Fig. 1d**). However, 20-cycles of pre-amplification led to saturation for some genes with higher expression levels (**Supplementary Fig. S2**). In consequence, 10 cycles were selected for the final workflow.

To test the reproducibility of the steps described above, we retrotranscribed and pre-amplified the same EV-RNA sample on three independent reactions. Then, we compared the results

obtained when submitting the three resulting cDNA samples to nCounter analysis. A strong correlation was found between the normalized counts for each individual gene obtained in the different cDNAs, represented by a Spearman´s r of 0.84 to 0.91, $p < 0.01$ (**Fig. 1e**).

All the experiments so far described had been performed with 500 µL plasma samples and TRI-reagent. Two additional RNA extraction methods were tested on EV-enriched preparations from control samples, the automated QiaSymphony and the manual TRIzol LS Reagent based isolation (**Fig. 2a**). When considering the total number of counts by nCounter, TRI-reagent was found to outperform both QiaSymphony and TRIzol LS, independently of the retrotranscriptase or the number of cycles used for the pre-amplification step (**Fig. 2b**). Finally, we also tested the effect of plasma input volume on downstream analysis of EV-RNA on the nCounter platform. Both the total counts and number of transcripts detected were higher with an initial plasma volume of 500 µL (**Fig. 2c**).

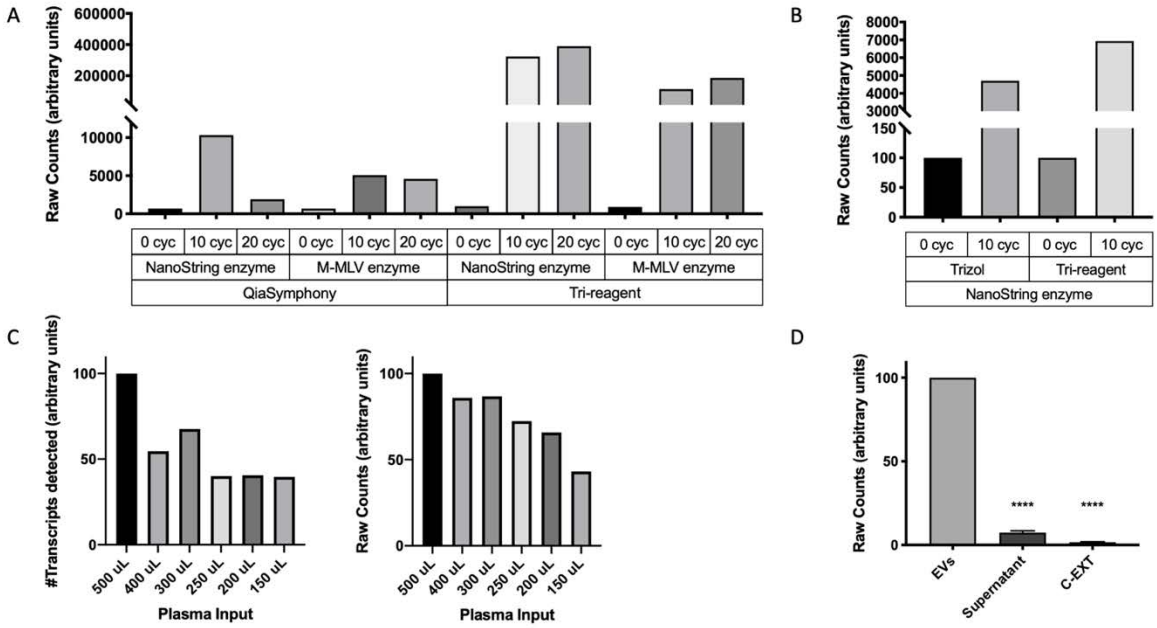


Figure 2. EV-RNA extraction, targeted pre-amplification and plasma input testing. **(A)** Total nCounter counts after automated (QiaSymphony) vs. manual (TRI reagent) RNA extraction from an EV-enriched pellet. Two different retrotranscriptases (NanoString versus M-MLV) and three pre-amplification conditions (0, 10 and 20 cycles) were tested. Results were normalized to the counts corresponding to 0 cycles. **(B)** Total nCounter counts after TRIzol LS vs. TRI reagent manual RNA isolation from an EV-enriched pellet. Results were normalized to the counts corresponding to 0 cycles. **(C)** Effect of input plasma volume (150-500 uL) on the final number of transcripts detected (left) or total nCounter counts (right) Results were normalized to the counts corresponding to 500 uL plasma. **(D)** Total nCounter counts of different fractions obtained during EV enrichment of plasma (EVs vs. supernatant $p < 0.0001$ in a one-way ANOVA with Dunnett´s multiple comparisons test; EVs versus C-EXT $p < 0.0001$). *Cyc: cycles; EVs; extracellular vesicles; C-EXT: extraction control.*

RNase A is a bovine enzyme that can be used to degrade RNA in a sample. When enriching for EVs from plasma samples, we expect two sources of RNA; extra-vesicular RNA and RNA derived from within the EVs. Our workflow for EV-RNA analysis incorporates an RNase A treatment of the EV-enriched pellets in order to remove extra-vesicular RNA not embedded into the particles. To further validate this step, we collected two EV-enriched pellets and the corresponding supernatants of healthy patients, treated them with RNase A, purified the RNA

and analyzed them with the Human Immunology V2 Panel. We found that EV-enriched pellets yielded significantly higher transcript counts compared to the supernatant (one-way ANOVA with Dunnett´s multiple comparisons, EVs vs. supernatant $p < 0.01$), indicating that the transcripts detected by nCounter are associated with EVs (**Fig. 2d**). To validate the efficacy of our RNase treatment, we also checked the expression of *GAPDH* and *CCL5* by qRT-PCR in three patient samples. Three different sample conditions were used; EVs without RNase treatment, EVs with RNase treatment and lysed EVs with RNase treatment. Results indicated that RNase treatment of intact EVs slightly reduced the *GAPDH* and *CCL5* mRNA levels, probably by removing the extra-vesicular RNA co-precipitated with the EVs (**Supplementary Fig. S3**). In contrast, once EVs were lysed, RNase treatment effectively eliminated all RNA and both *GAPDH* and *CCL5* transcripts were undetectable, confirming that transcripts purified using our workflow are indeed contained within the EVs.

EV-derived mRNA analysis on nCounter and classifier development

Based on the data presented above, we selected a workflow for subsequent experiments; starting with EV enrichment from 500 uL of plasma with the miRCURY kit, followed by manual RNA purification with TRI-reagent, retrotranscription with the Nanostring enzyme, a 10-cycle pre-amplification and final cDNA analysis by nCounter. We validated the proposed workflow by studying the EV-mRNA expression of 19 cancer patient and 10 control plasma samples using the Human Immunology V2 Panel (**Table 1**). If all samples were considered together, the average number of transcripts detected was 430 ± 79 out of the 594 transcripts in the panel (**Fig. 3a**). No significant differences were found between the number of mRNA transcripts in EVs from controls versus cancer patients (445 ± 68 and 422 ± 84 respectively, Mann-Whitney's U $p = 0.46$).

After normalization, we analyzed the differential expression (DE) of transcripts in EVs from cancer patients vs. controls (**Fig. 3b**, **Supplementary Table S1**). We found 141 mRNAs with significantly different levels; of them, 107 were upregulated and 34 downregulated in the EVs from cancer patients vs. controls. Then, we used a recursive feature elimination (RFE) method to select a gene signature predictive of the origin of the sample; a cancer patient or a control sample. The transcripts included in the final signature were *BCL10*, *CXCL11*, *CYBB* and *GBP1* and, based on their expression levels, our algorithm was able to classify plasma samples into cancer and control with a receiver operating characteristic (ROC) area under the curve (AUC) of 0.92 to 0.95 (**Fig. 3c**). The classifier also calculated signature scores for each sample, which were found to be significantly different between cancer patients and controls (Mann-Whitney U, $p < 0.01$; **Fig. 3d)**.

Improvement of classifier performance through removal of genomic DNA

The TRI-reagent-based RNA isolation incorporated in our workflow for EV-RNA purification may also co-extract EV-genomic DNA, which could bind to the nCounter probes during hybridization. In consequence, some of the counts detected can correspond to genomic DNA instead of mRNA transcripts, particularly considering that we used a 10-cycle pre-amplification step. To test if this was the case, we analyzed the effect of adding a DNase

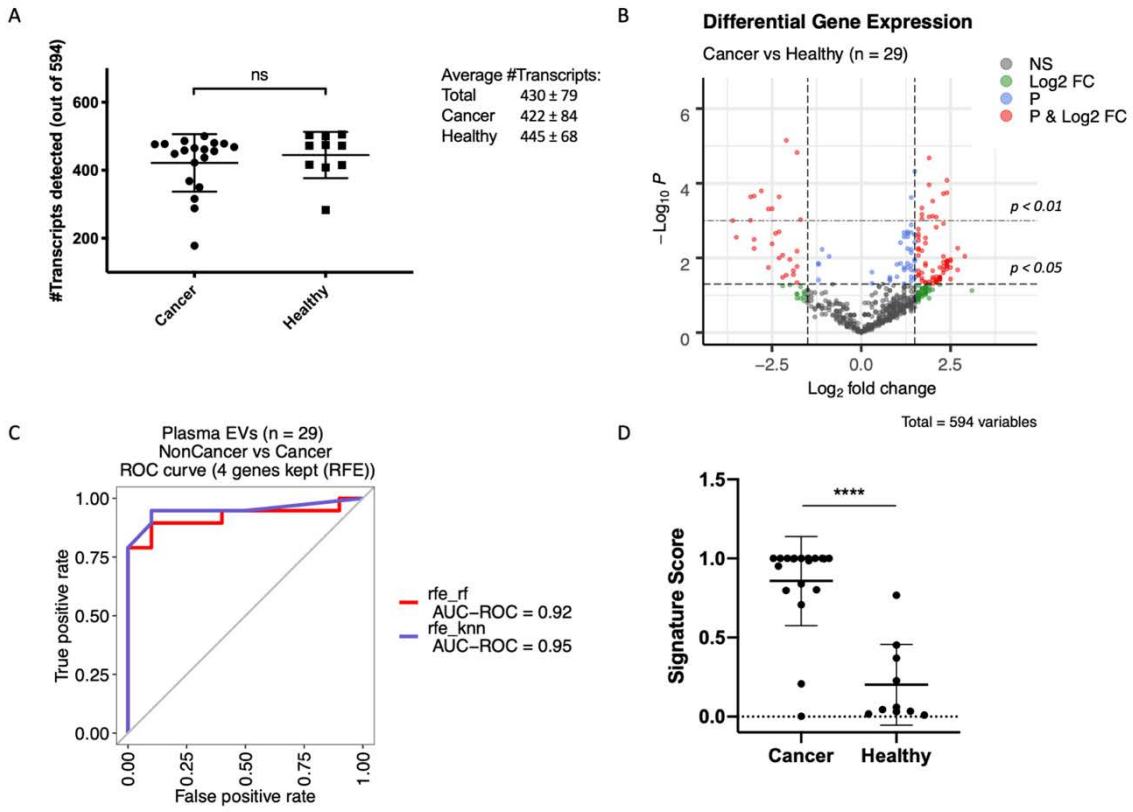


Figure 3. EV-mRNA transcript detection, differential expression analysis and development of a classifier algorithm. **(A)** Number of transcripts detected in EVs from cancer patients and healthy controls using the Human Immunology v2 nCounter panel, which targets 594 genes (two-tailed Mann-Whitney U test, $p = 0.463$). **(B)** Differential expression analysis of log2-normalized counts between cancer patients and healthy controls. The full list of genes differentially expressed is presented in Supplementary Table S1. **(C)** Area under the ROC curve of the four-gene signature, selected using recursive feature elimination (RFE), to differentiate cancer from control samples. **(D)** Scores of cancer vs. control samples based on expression of the four-gene signature ($p < 0.001$ in a two-tailed Mann-Whitney U test). *NS: not significant; EVs: extracellular vesicles; ROC: receiver operating characteristic; AUC: area under the curve; RFE: recursive feature elimination; rf: random forest; knn: k-nearest neighbors.*

treatment step after EV-RNA extraction in the 28 EV samples with remaining material. A dramatic reduction in the number of transcripts detected and the total amount of counts was observed in the EV-RNA samples after DNase treatment (**Fig. 4a**, Mann-Whitney U, $p < 0.01$ in both cases). The average number of transcripts in the 28 samples decreased from 430 ± 79 to 115 ± 66 and the total counts dropped more than 90%. As previously observed in samples without DNase treatment, the number of transcripts detected in cancer vs. control samples were not significantly different (122 ± 77 vs 103 ± 39, Mann-Whitney's U $p = 0.92$, **Fig. 4b**). In addition, there were no significant differences in the number of transcripts detected between the different tumor types; which included sarcoma, lung, rectal, anal and gastric cancer (**Supplementary Fig. S4**).

Based on these results, we decided to incorporate a DNase step in our EV-mRNA purification and analysis workflow when working with this panel (**Fig. 5**), and we confirmed that 500 µL still yielded the highest number of transcripts and total counts (**Fig. 4c**). To further validate the DNase step, we analyzed *GAPDH*, *CCL5* and *Caspase 8 (CASP8)* expression by qRT-PCR in

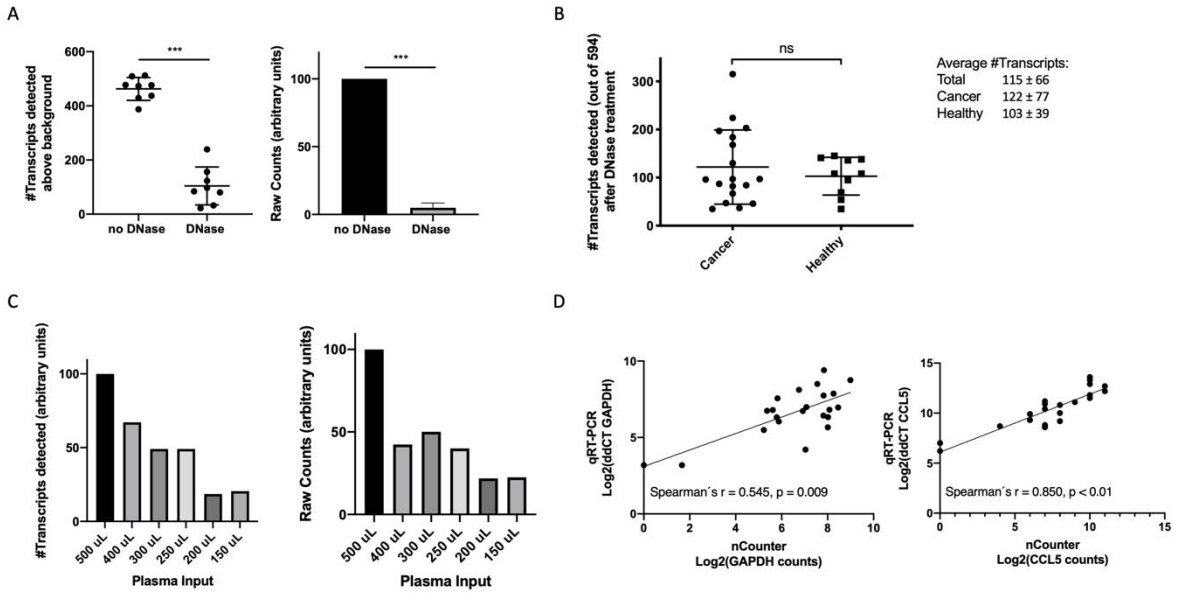


Figure 4. Effect of DNase treatment on the number of transcripts detected, total counts, and additional validation experiments. **(A)** Number of transcripts detected (left) and total nCounter counts (right) in non-DNase vs. DNase treated EV samples ($p < 0.001$ in a two-tailed Mann-Whitney U test, in both cases). **(B)** Number of transcripts detected in EVs from cancer patients and control samples using the Human Immunology v2 nCounter panel, after DNase treatment ($p = 0.916$ in a two-tailed Mann-Whitney U test). **(C)** Effect of input plasma volume (150-500 uL) on the final number of transcripts detected (left) or total nCounter counts (right), after DNase treatment. Results were normalized to the counts corresponding to 500 uL plasma. **(D)** Comparison of Log2 counts by nCounter vs. ddCT values by qRT-PCR for *GAPDH* and *CCL5* in 22 samples. Spearman´s correlation coefficient is indicated. *NS: not significant.*

the 22 EV samples with remaining RNA. We observed a statistically significant correlation between *GAPDH* and *CCL5* expression as measured by nCounter and qRT-PCR (Spearman´s r = 0.545, $p < 0.01$ and r = 0.850, $p < 0.01$, respectively; **Fig. 4d**); while *CASP8* transcripts were undetectable in DNase treated EV-RNA by both techniques.

Next, we proceeded to redesign our gene signature and classifier algorithm using the expression data derived from DNase treated EV-mRNAs. We observed that the number of differentially expressed genes in cancer vs. control samples dropped from 141 to 43 (**Fig. 6a and Supplementary Table S2**). Of them, 17 genes (40%) overlapped with those obtained using non-DNase treated samples (**Fig. 6b**). Then, we used the RFE algorithm to create a second signature, which included eight genes: *CCL5, S100A9, B2M, HLA-B, IL7R, ICAM3, ARHGDIB and PYCARD* (**Fig. 6a and Supplementary Table S3**). The algorithm based on this signature was able to classify the samples into cancer and control with a ROC AUC of 0.99 to 1.00 (**Fig. 6c**). Also, the signature scores for each sample were found to be significantly different between cancer patients and controls (Mann-Whitney U, $p < 0.01$; **Fig. 6d)**.

The significant decrease in counts after DNase treatment prompted us to further investigate the binding of nCounter probes to genomic DNA co-purified with EV-mRNA. The nCounter Human Immunology V2 Panel targets 594 gene transcripts. The probes for some genes are

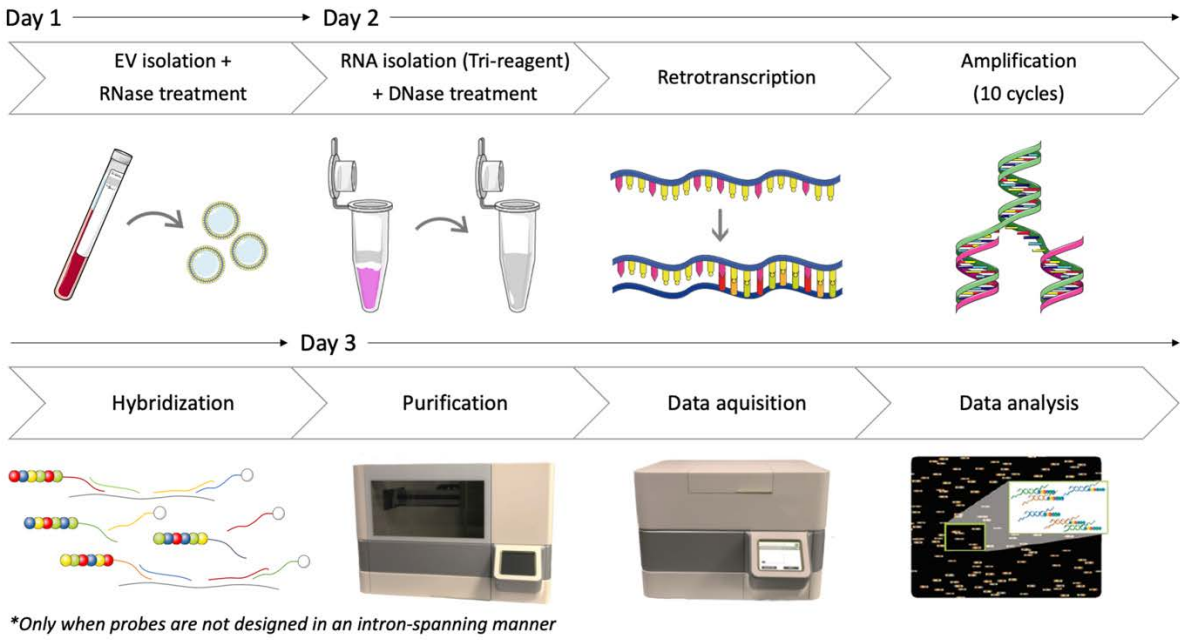


Figure 5. Final workflow for EV-RNA extraction and analysis on the nCounter platform. The miRCURY kit was used to enrich EVs from 500 uL plasma and EV-enriched preparations were treated with RNase to remove extra-vesicular RNA. Then, EV-RNA was extracted using TRI reagent and treated with DNase to remove genomic DNA. DNase treatment is only necessary when probes are not designed in an intron-spanning manner. Next, we performed retrotranscription and a 10-cycle pre-amplification, followed by hybridization, purification on the nCounter prep-station and analysis on the nCounter digital analyzer. *EVs: extracellular vesicles. Part of this figure was modified from SMART (Servier Medical Art), licensed under a Creative Common Attribution 3.0 Generic License. http://smart.servier.com/*

designed within the same exon ("non-intron spanning"); while other probes ("intron spanning") target sequences corresponding to contiguous exons of the cDNA/mRNA. Genomic DNA should only bind "non-intron spanning" probes, and we performed an additional experiment to confirm this point. We analyzed an EV sample in quadruplicate, skipping the retrotranscriptase and/or the DNase treatment (**Supplementary Table S4**). Without DNase treatment, we observed counts for "non-intron spanning" probes (such as *CASP8*, *TNFRSF8* and *B2M*) both in presence and in absence of the retrotranscriptase step. In contrast, counts for "intron spanning" probes (such as *PRKCD*, *TNFRSF10C*, *FCER1G*) were apparent only if retrotranscription was performed. Finally, DNase treatment induced a significant drop in the counts of "non-intron spanning" but, unexpectedly, also "intron spanning" probes.

Discussion

The molecules found within EVs, such as mRNAs, are often involved in intercellular communication and represent a potential source for biomarker discovery. However, lack of standardized methods and clinical validation prevents the implementation of EV-derived testing in daily practice. The nCounter platform, which allows for multiplex detection of hundreds of transcripts, has been extensively used in translational research for transcriptomic tumor characterization. In addition, the nCounter Prosigna assay, based on a 50-gene expression signature, has been fully standardized and validated at the clinical level; and received FDA approval in 2013 to predict risk of recurrence in breast cancer(21-23). However,

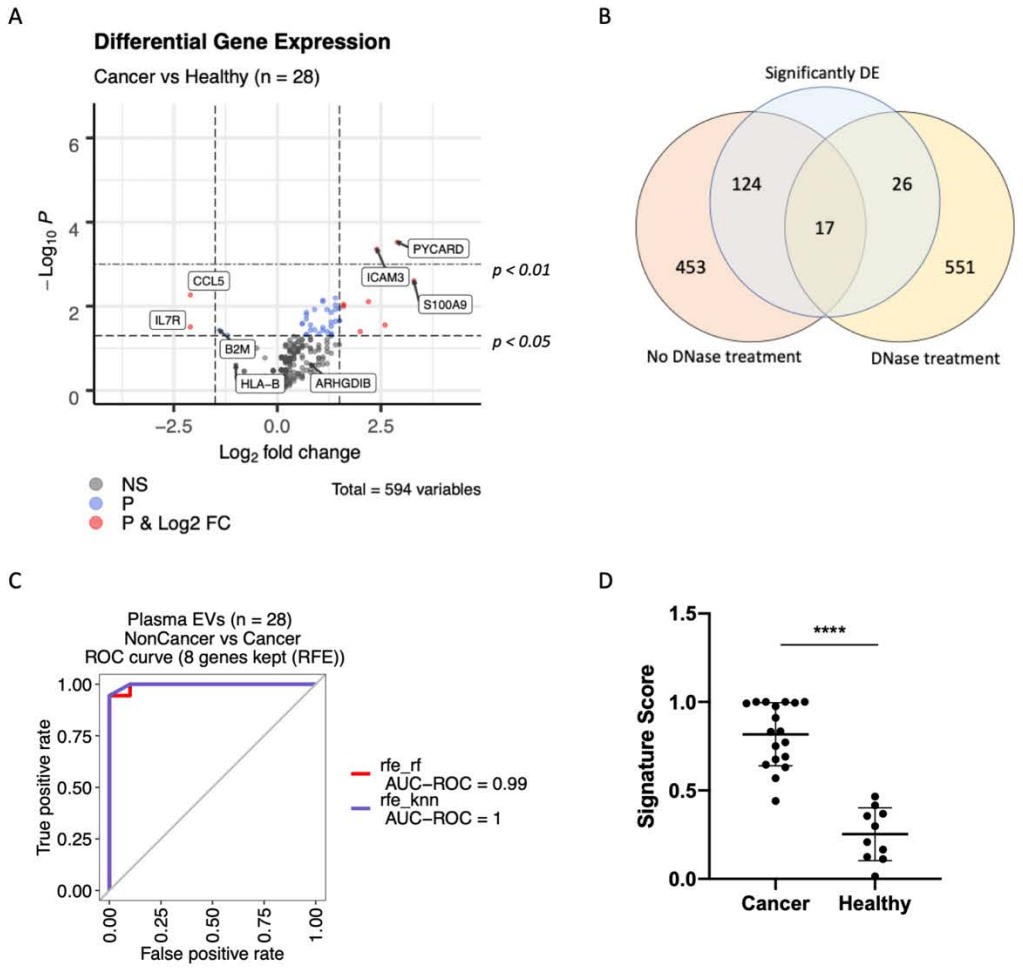


Figure 6. Effect of DNase treatment on differential expression analysis and classifier development. **(A)** Differential expression analysis of log2-normalized counts after DNase treatment between cancer patients and control samples. The full list of genes differentially expressed is presented in Supplementary Table S2. Labels indicate the eight transcripts selected for the final classification signature. **(B)** Out of 594 genes in the panel, 17 showed differential expression independently of DNase treatment. **(C)** Recursive feature elimination (RFE) was used to select an eight-gene signature that could distinguish between samples derived from cancer patients and controls. **(D)** Scores of cancer vs. control samples based on expression of the eight gene signature ($p < 0.001$ in a two-tailed Mann-Whitney U test). *DE: differentially expressed; EVs: extracellular vesicles; ROC: receiver operating characteristic; AUC: area under the curve; RFE: recursive feature elimination; rf: random forest; knn: k-nearest neighbors.*

studies investigating the performance of nCounter for mRNA analysis in liquid biopsies are scarce, particularly in the case of EVs.

Here, we present a workflow for nCounter-based analysis of plasma-derived EVs and we demonstrate that it can be used to efficiently detect transcripts in EV-enriched preparations from cancer patients and control samples. The first step of the workflow is plasma processing using the precipitation-based miRCURY kit which, as previously described(113, 114), yielded pellets enriched in EVs. Clinical laboratories do not usually have access to ultracentrifugation and precipitation kits offer several advantages, such as short turn-around time and limited technical requirements. Although 500 µL plasma was found to give the highest amount of detected transcripts and total counts by nCounter, lower plasma inputs also yielded valid

results. The second step is manual RNA isolation using TRI-reagent, which was found to outperform manual purification by TRIzol LS and automated extraction.

The extraction methods tested, including the TRI-reagent based manual extraction, did not yield sufficient RNA for direct analysis by nCounter. In consequence, after some optimization experiments, a 10-cycle pre-amplification step was added to our workflow and found to be highly reproducible. Most RNA extraction methods are known to co-purify genomic DNA (gDNA), and this was also our case. The EV-associated gDNA could be located in the interior of the vesicles or be attached to the membranous surface, and further research is needed to clarify this issue. More importantly, it has been described that simultaneous isolation of EV-derived genomic DNA during EV-RNA extraction can affect the amount of detected transcripts(115). In consequence, we tested the effect of adding a DNase treatment step to our methodology and found a significant decrease in total counts and number of transcripts detected. Interestingly, while the expression of many genes became undetectable after DNase treatment, the counts for some transcripts were maintained. These observations suggested that co-extracted EV-genomic DNA was amplified during the pre-amplification step and could hybridize to the "non-intron spanning" probes in the nCounter panel. In contrast, probes designed in an "intron-spanning" manner should not hybridize to EV-DNA. Validation experiments confirmed this point (**Supplementary Table S4**). Unexpectedly, we observed a sharp decrease also in the counts of "intron-spanning" probes after DNase treatment, strongly suggestive of EV-RNA degradation or, more likely, EV-RNA loss during the purification steps needed to remove the DNase. In consequence, appropriate design of probes should be preferred over DNase treatment when analyzing samples with low amounts of RNA that require pre-amplification. Since we were using a pre-designed panel that could not be modified, we added a DNase treatment step to our final protocol for EV-mRNA analysis (**Fig. 5)**. As an additional validation of the entire workflow, we re-analyzed 22 DNase treated samples for *GAPDH, CASP8* and *CCL5* expression by qRT-PCR. Similarly to previous reports(9, 12), we found a statistically significant correlation between the expression levels obtained by qRT-PCR and nCounter.

Next, we investigated if our workflow for EV-mRNA analysis could be used to develop gene signatures. To this end, we performed differential expression analysis of the nCounter results obtained for control and cancer samples. The signature-based algorithm obtained for EV-mRNA (DNase treated) showed an improved classifier performance in comparison with EV-nucleic acids (non-DNase treated); with ROC-AUCs of 0.99-1.00 versus 0.92-0.95 for the discrimination of control vs. cancer, respectively. This result is coincident with a previous report where a DNA removal step was shown to reduce signature noise during algorithm development and yielded better classification results(116). The genes selected for our EV-mRNA expression signature are *CCL5*, *S100A9*, *B2M*, *HLA-B*, *IL7R*, *ICAM3*, *ARHGDIB* and *PYCARD*; which are involved in several immune-related pathways such as cytokine signaling, innate immune system or lymphocyte activation (**Fig. 6a and Supplementary Table S3)**. *CCL5*, *B2M*, *HLA-B* and *IL7R* are all related to cytokine signaling and lymphocyte activation, and are thus known as immunomodulators. Interestingly, these four transcripts were downregulated in cancer samples, suggesting differences in immune system activation through

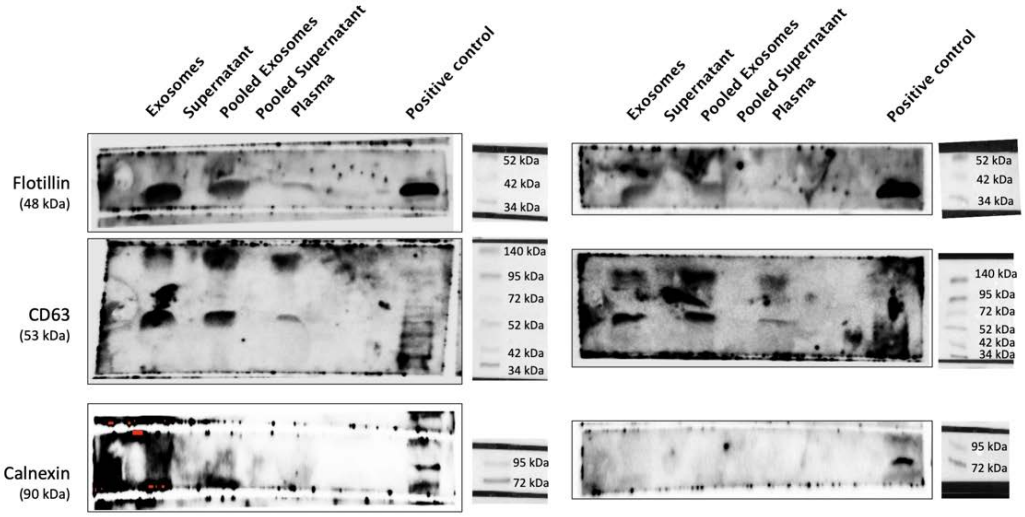
cytokine induction between cancer patients and controls. *PYCARD* is also proposed to be involved in lymphocyte activation and was found to be upregulated in cancer patients. A previous study found that *PYCARD* could suppress apoptosis of cancer cells in gastric cancer(117), potentially explaining the observed higher expression in EVs from cancer patients.

Two studies have used nCounter for analysis of mRNA isolated from blood. Kossenkov et al(75) developed a pulmonary node classifier to differentiate malignant from non-malignant nodules previously detected by low-dose CT. Since the authors made use of whole blood, the quantities of purified RNA were significantly higher than those in our EV-based study (3 μg/2.5 mL blood versus 0.01 μg/2.5 mL, respectively), avoiding the need for a pre-amplification step. Beck et al(58) used a combination of CTC-RNA and cfRNA to profile tumor-associated biomarkers and correlate them with diagnostics and survival. The quantities of RNA were similar to those obtained in our study and a pre-amplification step was also added.

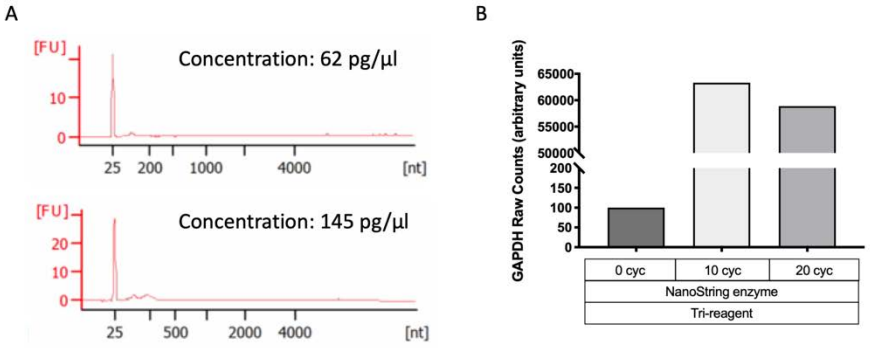
Our study shows several limitations. First, although purification of EVs from plasma yielded pellets that were found to be enriched in EVs, such precipitation techniques can also isolate a significant fraction of proteins and lipoproteins(118, 119). Second, there are no validated housekeeping genes for normalization of EV-mRNAs and we had to use the total amount of counts, as described(120-122). Third, although nCounter presents many advantages, it has also a few limitations when compared to other multiplex techniques such as RNAseq. For instance, since nCounter works with gene panels, no new transcripts can be found. Finally, our aim was to establish a workflow for the nCounter analysis of EV-mRNA and the mixed patient population we used to demonstrate the validity of our approach was not the most appropriate to develop a clinically useful signature. Much larger patient cohorts would be needed to train and validate signatures that could differentiate between cancer patients and controls.

In summary, to the best of our knowledge, this proof-of-concept study is the first to demonstrate that the nCounter platform can be used to reproducibly detect plasma EV-mRNA transcripts. Differential expression analysis can then be implemented for biomarker assay development. Our work paves the way for widespread testing of EV-mRNA expression in blood and other fluids, and subsequent selection of signatures useful in the clinical setting.

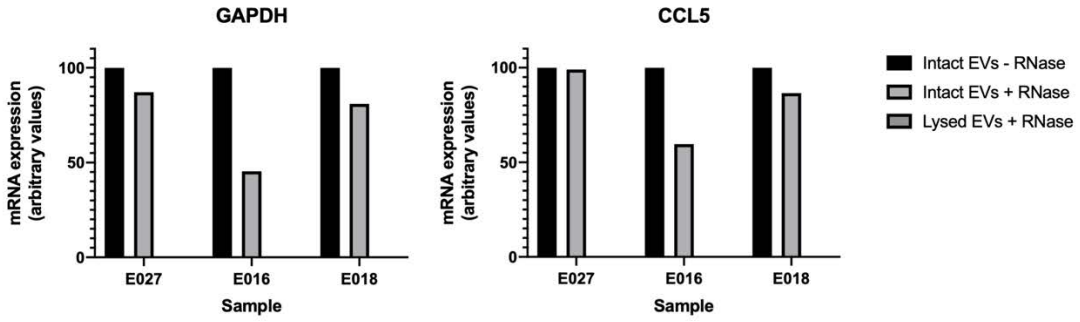
Supporting information



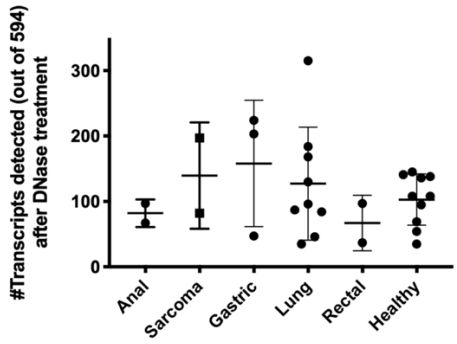
Supplementary Figure S1. Full immunoblot membranes showing expression of Flotillin, CD63 and Calnexin in EV-enriched pellets, supernatants, full plasma and a positive control sample.



Supplementary Figure S2. Bioanalyzer profiles for RNA quantity and quality and saturation of GAPDH transcripts after pre-amplification. **(A)** Bioanalyzer profiles of Tri-reagent based EV-RNA extraction from 500 µL plasma. **(B)** Raw GAPDH counts after Tri-reagent based EV-RNA extraction. Different pre-amplification conditions (0, 10 and 20 cycles) were tested, using the same plasma sample. The 0 cycle condition was used as control value for normalization. A pre-amplification of 20 cycles was found to cause saturation in the GAPDH counts. *Cyc: cycles.*



Supplementary Figure S3. qRT-PCR analysis for *GAPDH* and *CCL5* in three patient samples to determine the efficacy of RNase A treatment. For each patient three different sample conditions were used: intact EVs with- and without RNase treatment, and lysed EVs after RNase treatment. Results were normalized to the counts corresponding to EVs without RNase treatment. Undetermined values by qRT-PCR have been assigned a value of 0%.



Supplementary Figure S4. Number of transcripts detected in EVs from cancer patients, divided by tumor type, and control samples using the Human Immunology v2 nCounter panel, after DNase treatment ($p = 0.756$ in a Kruskal-Wallis test).

The supplementary tables can be found online at:
https://doi.org/10.1038/s41598-021-83132-0

Acknowledgements
The investigators wish to thank the patients for kindly accepting to donate samples to this study. We also thank all the physicians who collaborated by providing clinical information. Finally, we would like to thank colleagues and collaborators for their constructive feedback on this manuscript.

Author contributions
J.W.P.B., C.P.V. and M.A.M.V. conceptualized and designed the experiments. J.W.P.B. and A.G.C. performed the experiments. N.P. and J.W.P.B. performed data analysis. J.W.P.B. and M.A.M.V. wrote the main manuscript and prepared the figures. C.Y.H., S.W. and R.R. provided editing, comments, and experimental guidance. All authors reviewed the manuscript.

Additional information
Competing interests statement: C.Y.H. and S.W. were full-time employees of NanoString Inc. when the project was carried out. The rest of the authors declare no competing interests.

Financial support
This project has received funding from a European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement ELBA No 765492.

Chapter 4. A four-gene EV-mRNA signature to predict checkpoint inhibitor pneumonitis in lung cancer patients.

Jillian WP Bracht, Andrés Aguilar, Santi Viteri, Juan José García-Mosquera, Carlos Cabrera, Ana Drozdowskyj, Nicolas Potie, Chung-Ying Huang, Erika Aldeguer, Sonia Rodriguez, Ana Gimenez-Capitan, Silvia Calabuig, Elena Duréndez, Sarah Warren, Eloisa Jantus, Rafael Rosell, Miguel A Molina-Vila, Maria Gonzalez-Cao.

Manuscript in preparation

Abstract

Purpose: T-cell activation against self-antigens due to immune checkpoint inhibitor (ICI) treatment can induce a spectrum of toxicities, known as immune-related adverse events (irAEs). Checkpoint inhibitor pneumonitis (CIP) is one of the more lethal irAEs, but predictive factors for CIP development remain unclear. We aimed to develop a predictive gene signature to stratify patients at higher risk for CIP.

Experimental design: We retrospectively collected pre-treatment tumor tissue and plasma-EV samples from lung cancer patients with- and without CIP development. Gene expression analysis was carried out using the IO360 panel on the NanoString nCounter platform. Gene signatures with best discriminative performance were selected using a bioinformatic algorithm. In addition, a multivariate analysis was performed to explore risk factors for CIP development.

Results: Our study demonstrates that a four-gene EV-mRNA signature can discriminate between lung cancer patients with high vs. low risk of developing CIP. The EV-based CIP-signature yielded an average accuracy of 87.0% and a negative predictive value of 92.7%. Importantly, the CIP-signature was found to increase during actual CIP development in follow-up samples, while it remained low in the other time points.

Conclusion: We have developed a 4-gene CIP-signature that can be used as a predictive assay for CIP development based on plasma EV-mRNA samples from lung cancer patients under ICI treatment.

Introduction

In the last decade, several immune checkpoint inhibitors (ICIs) targeting programmed cell death protein 1 (PD-1) and programmed death-ligand 1 (PD-L1) have been approved for cancer therapy(123). Thus far, monoclonal antibodies targeting PD-1, such as nivolumab, pembrolizumab and cemiplimab, and antibodies targeting PD-L1, including atezolizumab, avelumab and durvalumab have demonstrated clinical efficacy in several malignancies, including lung cancer(124-126). T-cell activation against self-antigens due to ICI treatment can induce a spectrum of toxicities, known as immune-related adverse events (irAEs)(127). Although, in general, the development of irAEs was found to correlate with improved response rates to ICIs and longer patient survival, the type and grade of irAEs needs to be taken into account(49, 128, 129).

Checkpoint inhibitor pneumonitis (CIP) is one of the more lethal irAEs and has been reported to occur between a few days up to 24 months after treatment initiation(48, 127, 130). CIP is a non-infection induced inflammation of the lung tissue which comes with diverse clinical and radiographic manifestations and no available predictive or diagnostic test, complicating clinical management. Grade 1-2 CIP includes patients with ground-glass opacity changes and can present with symptoms such as dyspnea, cough and chest pain. When diagnosed with grade 3-4 CIP, patients experience severe symptoms including hypoxia or acute respiratory distress syndrome (ARDS) and administration of oxygen or even intubation is often required. A patient that has died due to the consequences of CIP development is classified as grade 5(131). Treatment strategies depend on the grade, but often consists of discontinuation of immunotherapy and administration of high-dose corticosteroid treatment. Multiple studies have reported shorter survival in patients that developed CIP(49, 132-134). When not detected and treated in time, CIP can lead to heart- and/or respiratory failure and death irrespective of tumor response.

While the reported incidence in literature is around 5%(135), in a retrospective analysis in our clinic and two single institution analysis by Suresh et al.(48) and Barron et al.(136) CIP was observed in 20% of patients treated with ICIs. Potential risk factors for CIP include pre-existing adverse pulmonary conditions, such as tobacco exposure, previous lung radiation, lung infections or previous interstitial lung disease(128). The pulmonary damage caused by these pathologies may induce a host immune response when exposed to anti-PD1/PD-L1 therapies(130, 137). In addition, older patients, treatment with multiple therapies and treatment with anti-PD-1 antibodies have shown higher CIP incidence(138). Despite these observations, there is still an unmet clinical need to confidently predict which patients are at higher risk for CIP and to identify the mechanisms of CIP development(131).

Liquid biopsies provide a minimally invasive and clinically relevant source of genetic information on several biological processes, including immune system activation(139). RNA molecules can be freely circulating in the blood plasma, but are much more stable when conserved in membrane-enclosed particles(69-71), such as extracellular vesicles (EVs). EVs are released from several cell types and function as intercellular communicators by transferring their contents to target cells(69, 72, 73). Consequently, RNA profiles within EVs provide an

attractive source of diagnostic and predictive biomarkers, and gene expression studies using EV-RNA are an active area of research(63, 64).

We hypothesized that a pre-treatment state of chronic lung inflammation or immune system imbalance could predict which lung cancer patients are at higher risk of developing CIP under ICI treatment, paving the way for intensive monitoring of these patients. To this end we retrospectively analyzed gene expression profiles (GEPs) of baseline tumor tissue and plasma-EV samples from lung cancer patients with- and without CIP development. We also investigated gene expression changes between pre- and on-treatment EV samples from two CIP and two non-CIP patients. Finally, a multivariate analysis was carried out to explore patterns of increased CIP risk. Our study demonstrates that a four-gene EV-mRNA signature can discriminate between lung cancer patients with high vs. low risk of developing CIP with an average accuracy of 87.0% and a negative predictive value (NPV) of 92.7%.

Materials and methods

Tumor tissue and plasma samples

This study was carried out in accordance with the principles of the Declaration of Helsinki under an approved protocol of the institutional review board of Quirón Hospitals. To explore differences in gene expression between lung cancer patients with- and without CIP development, pre-ICI-treatment formalin-fixed paraffin-embedded (FFPE) tumor tissue samples and matching biobank stored pre- and on-treatment plasma samples were retrospectively collected (**Figure 1, Table 1** and **Table 2**). Written informed consent was obtained from all participants and documented; samples were de-identified for confidentiality. Progression-free survival (PFS) was measured from the time of first anti PD-1/PD-L1 administration to the time of tumor progression by RECIST 1.1 criteria or death/loss of follow-up. Median PFS and overall survival (OS) were estimated using Kaplan-Meier analysis method. Summary statistics, frequency tables, and parametric and nonparametric statistical tests were used, as applicable. Statistical analyses were performed with the use of Prism software V8.4.3 and SAS.

Figure 1. Workflow for tissue and EV-mRNA analysis on the nCounter platform using the IO360 panel and downstream gene expression profiling analysis to differentiate between CIP risk. *FFPE: formalin-fixed paraffin-embedded; EV: extracellular vesicle; CIP: checkpoint inhibitor pneumonitis.*

Table 1. Clinical characteristics of pre-treatment tumor tissue samples from CIP and non-CIP patients.

Characteristics	CIP (n = 17)	non-CIP (n = 24)
Gender – no. (%)		
Male	11 (64.7%)	16 (66.7%)
Female	6 (35.3%)	8 (33.3%)
Age (yr) – median (range)		
Median	63 (51-82)	56 (33-76)
Smoking status – no. (%)		
Former/Current	17 (100.0%)	24 (100.0%)
Stage – no. (%)		
Stage II	0 (0.0%)	1 (4.2%)
Stage III	2 (11.8%)	0 (0.0%)
Stage IV	15 (88.2%)	23 (95.8%)
Histology – no. (%)		
Non Squamous NSCLC	12 (70.6%)	18 (75.0%)
Squamous NSCLC	4 (23.5%)	2 (8.3%)
SCLC	1 (5.9%)	4 (16.7%)
Treatment line – no. (%)		
First line	5 (29.4%)	10 (41.7%)
Pre-treated	10 (58.8%)	13 (54.2%)
Adjuvant	1 (5.9%)	1 (4.1%)
Neoadjuvant	1 (5.9%)	0 (0.0%)
ICI treatment – no. (%)		
Nivolumab	10 (58.8%)	9 (37.5%)
Pembrolizumab	3 (17.6%)	2 (8.3%)
Durvalumab	2 (11.8%)	10 (41.7%)
Avelumab	2 (11.8%)	0 (0.0%)
Atezolizumab	0 (0.0%)	3 (12.5%)
Treatment regimen – no. (%)		
Single	13 (76.5%)	21 (87.7%)
With chemo	3 (17.6%)	1 (4.1%)
With immunotherapy	0 (0.0%)	1 (4.1%)
Other combination	1 (5.9%)	1 (4.1%)
Clinical benefit* – no. (%)		
Yes	11 (64.7%)	11 (45.8%)
No	5 (29.4%)	12 (50.0%)
NA	1 (5.9%)	1 (4.2%)
CIP grade – no. (%)		
1-2	7 (41.2%)	0 (0.0%)
3-5	10 (58.8%)	0 (0.0%)
NA	0 (0.0%)	24 (100.0%)
Time to CIP (m) – median (range)		
Median	4.5 (1-14)	NA
PD-L1 IHC – no. (%)		
Positive	5 (29.4%)	5 (20.8%)
Negative	7 (41.2%)	9 (37.5%)
Unknown	5 (29.4%)	10 (41.7%)

**Clinical benefit is defined as the sum of patients with complete response (CR) or partial response (PR) plus stable disease (SD) longer than 24 weeks.*

Table 2. Clinical characteristics of pre-treatment EV-samples from CIP and non-CIP patients.

Characteristics	CIP (n = 15)	non-CIP (n = 54)
Gender – no. (%)		
Male	10 (66.7%)	41 (75.9%)
Female	5 (33.3%)	13 (24.1%)
Age (yr) – median (range)		
Median	65 (51-77)	62.5 (33-80)
Smoking status – no. (%)		
Never	0 (0.0%)	3 (5.6%)
Former/Current	15 (100.0%)	51 (94.4%)
Stage – no. (%)		
Stage II	0 (0.0%)	1 (1.9%)
Stage III	2 (13.3%)	6 (11.1%)
Stage IV	13 (86.7%)	47 (87.0%)
Histology – no. (%)		
Non Squamous NSCLC	12 (80.0%)	35 (64.8%)
Squamous NSCLC	3 (20.0%)	18 (33.3%)
SCLC	0 (0.0%)	1 (1.9%)
Treatment line – no. (%)		
First line	4 (26.7%)	8 (14.8%)
Pre-treated	10 (66.7%)	45 (83.3%)
Adjuvant	1 (6.6%)	1 (1.9%)
ICI treatment – no. (%)		
Nivolumab	9 (60.1%)	42 (77.8%)
Pembrolizumab	2 (13.3%)	1 (1.9%)
Durvalumab	2 (13.3%)	11 (20.3%)
Avelumab	2 (13.3%)	0 (0.0%)
Treatment regimen – no. (%)		
Single	13 (86.7%)	54 (100.0%)
With chemo	2 (13.3%)	0 (0.0%)
Clinical benefit* – no. (%)		
Yes	11 (73.3%)	28 (51.8%)
No	3 (20.0%)	25 (46.3%)
NA	1 (6.7%)	1 (1.9%)
CIP grade – no. (%)		
1-2	8 (53.3%)	0 (0.0%)
3-5	7 (46.7%)	0 (0.0%)
NA	0 (0.0%)	54 (100.0%)
Time to CIP (m) – median (range)		
Median	5.5 (1-14)	NA
PD-L1 IHC – no. (%)		
Positive	3 (20.0%)	15 (27.8%)
Negative	6 (40.0%)	27 (50.0%)
Unknown	6 (40.0%)	12 (22.2%)

**Clinical benefit is defined as the sum of patients with complete response (CR) or partial response (PR) plus stable disease (SD) longer than 24 weeks.*

Extracellular Vesicle Enrichment

Whole blood samples (10 mL) were collected in sterile EDTA Vacutainer tubes (BD, Plymouth, UK) and centrifugated twice at 1000 x g for 10 min at room temperature (RT). Plasma samples were stored at -80°C until further processing. The miRCURY® Exosome Serum/Plasma Kit (Qiagen, Hilden, Germany) was used to enrich for EVs from 600 µL plasma, according to the manufacturer's instructions and as previously described(140). In short, dead cells and debris (including platelets and fibrin) were cleared with thrombin and centrifugation. Precipitation Buffer was added, samples were incubated overnight at 4°C and the EV fraction was pelleted by centrifugation. EV-enriched pellets were treated with 4 µg/mL RNase A (Sigma-Aldrich, St. Louis, MO) for 1 hour (h) at 37°C, to remove extracellular RNA not associated to EVs. EV characterization and enrichment was previously performed using identical protocols(140).

FFPE and EV RNA extraction

FFPE tissue slides were stained with hematoxylin and eosin. Tumor area was evaluated by a pathologist, and samples were macro-dissected prior to RNA isolation. Typically, RNA was isolated from 2-4 slides per patient. Total RNA was extracted with the High Pure FFPE RNA Isolation Kit (Roche, Indianapolis, IN), following the manufacturer's protocols, and eluted in a final volume of 25 µL. RNA integrity and concentration were evaluated using the 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) with the RNA 6000 Pico kit (Agilent Technologies), and with the Qubit RNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA). For RNA extraction from EVs, Tri-reagent (MRC, Cincinnati, OH) was added to the EV-enriched pellets in a final volume of 1 mL and incubated at RT for 20 min to inactivate RNase A and lyse the EVs, as previously described(140). In short, chloroform: Isoamyl Alcohol (24:1) (Panreac Química SLU, Barcelona, Spain) was added and after centrifugation the aqueous upper layer was kept to precipitate RNA after adding 2.5 µL Glycogen (Merck KGaA, Darmstadt, Germany) and 500 µL 2-propanol (Merck KGaA). The final RNA pellet was washed with 75% ethanol, air dried and dissolved in 15 µL nuclease free water. To remove co-isolated DNA from the resuspended RNA, the DNA-free™ DNA Removal Kit (Thermo Fisher Scientific), was used according to manufacturer instructions. In short, 1 µL DNase buffer and 0.5 µL enzyme were added to 7.5 µL RNA sample, followed by incubation at 37°C for 30 min and DNase removal.

Gene Expression Analysis using nCounter

The nCounter® Low RNA Input Amplification Kit (NanoString Technologies, Seattle, WA) was used to retrotranscribe and pre-amplify 4 µL EV-derived RNA using 10 cycles. Retrotranscription was carried out in 0.5 mL tubes while pre-amplification, using primers targeting the genes of the Human PanCancer IO360 Panel (NanoString Technologies), was performed in 384-well plates to prevent sample evaporation. The Human PanCancer IO360 Panel (NanoString Technologies) was used to analyze EV-derived, pre-amplified cDNA according to manufacturer instructions. The FFPE-derived RNA was analyzed without pre-amplification. The IO360 panel targets 770 genes related to tumor biology, microenvironment

and immune response. In addition, the panel contains gene signatures to measure immune cell populations and tumor and immune activities. Samples were hybridized for 18 h at 65°C.

Data Normalization and Analysis
Raw nCounter counts of expressed genes were normalized in R and R studio v3.6.3 using the R package NanoStringNorm(110). Normalization was performed following several steps: technical assay variability normalization using the geometric mean of the positive control probes, background correction using the mean plus two times standard deviation (SD) of the negative control probes, and sample content normalization using either the geometric mean of the most stable set of housekeeping genes (FFPE tissue RNA) or the total amount of counts for each sample (plasma EV-RNA). Normalized counts were log2-transformed, and used for differential expression (DE) analysis. Log2 fold change (FC) of each gene was calculated as the ratio of average log-2 transformed counts of the CIP patient cohort vs. the non-CIP cohort. Volcano plots were used to visualize log2 FC on the x-axis and nominal p-values on the y-axis.

Classifier Algorithm Development
Optimal gene selection was performed using recursive feature elimination (RFE). To this end, a leave-one-out cross validation (LOOCV) algorithm was used on the full IO360 gene Panel. The number of genes to select was set at 4, 8, 16 or 770 and the amount of genes that yielded optimal performance after cross-validation was automatically selected. Classification was performed with the selected gene signature using random forest (rf), k-nearest neighbors (knn) and stochastic gradient boosting (gbm) classifiers, where appropriate, using three iterations. The model with the highest accuracy was then selected as the final model. Signature scores for each sample were derived from the final model.

Results

Gene expression profiling of FFPE tumor tissue to predict CIP development

To determine if pre-ICI-treatment gene expression profiles GEPs differ between CIP vs. non-CIP patients, we analyzed FFPE tumor tissue samples with the NanoString nCounter IO360 panel (**Figure 1**). Panel-incorporated biological signature scores, related to tumor and immune activities, were assessed. The relative abundance of some cell types was found to be significantly lower in CIP-developing patients, including neutrophil-, natural killer- (NK) and $CD56^{dim}$ NK cell scores and the NOS2 signature score (Mann-Whitney U, $p = 0.04$, $p = 0.02$, $p = 0.01$ and $p = 0.01$, respectively; **Supplementary Figure 1**). Although these signature scores found to be lower expressed in patients that develop CIP, there was no clear cutoff that could discriminate between CIP and non-CIP patients.

We then performed a differential expression (DE) analysis of all transcripts included in the panel, and discovered 54 mRNA targets with significantly different expression levels; of which most were upregulated in the CIP cohort (**Supplementary Table 1**). Hierarchical clustering led to a non-optimal separation of CIP vs non-CIP, but indicated the potential of some genes for discriminatory purposes (**Figure 2A**). A classifier algorithm was developed based on all DE genes, and classifier performance yielded receiver operating characteristic (ROC) areas

under the curve (AUC) of 0.70-0.77 with accuracies ranging from 65.9% to 73.2% (**Figure 2B**, **Supplementary Table 2**). Of note, 7 out of 27 patients that were classified as non-CIP were false negatives, resulting in a non-optimal negative predictive value (NPV) of 74.1% and a positive predictive value of 71.4%. A high NPV is useful in the clinical setting to confidently start a patient on ICI treatment, with less intense follow-up. Therefore, we aimed to improve classifier performance by selecting a gene signature with a recursive feature elimination (RFE) algorithm. The transcripts included in the signature were CDKN1A, BID, ULBP2, HDAC11, *ARID1A, IGF2R, TWIST1* and *IER3* (**Figure 2C**, **Table 3**), and annotated pathways included tumor-intrinsic factors and myeloid cell activity. Classification yielded ROC AUCs of 0.81-0.94 with accuracies of 80.5%-92.7%, and an increased NPV and positive predictive value (PPV) of 92.0% and 93.8%, respectively (**Figure 2D, Supplementary Table 2**). The Youden Index, a measure between 0 and 1 indicative of the maximum potential effectiveness of a biomarker assay, of our 8-gene tissue CIP signature was 0.84. Based on the expression of these eight genes, the algorithm calculated tissue CIP-scores for each patient, which were found to be significantly different between CIP and non-CIP patients (Mann-Whitney U, $p < 0.01$; **Figure 2E**).

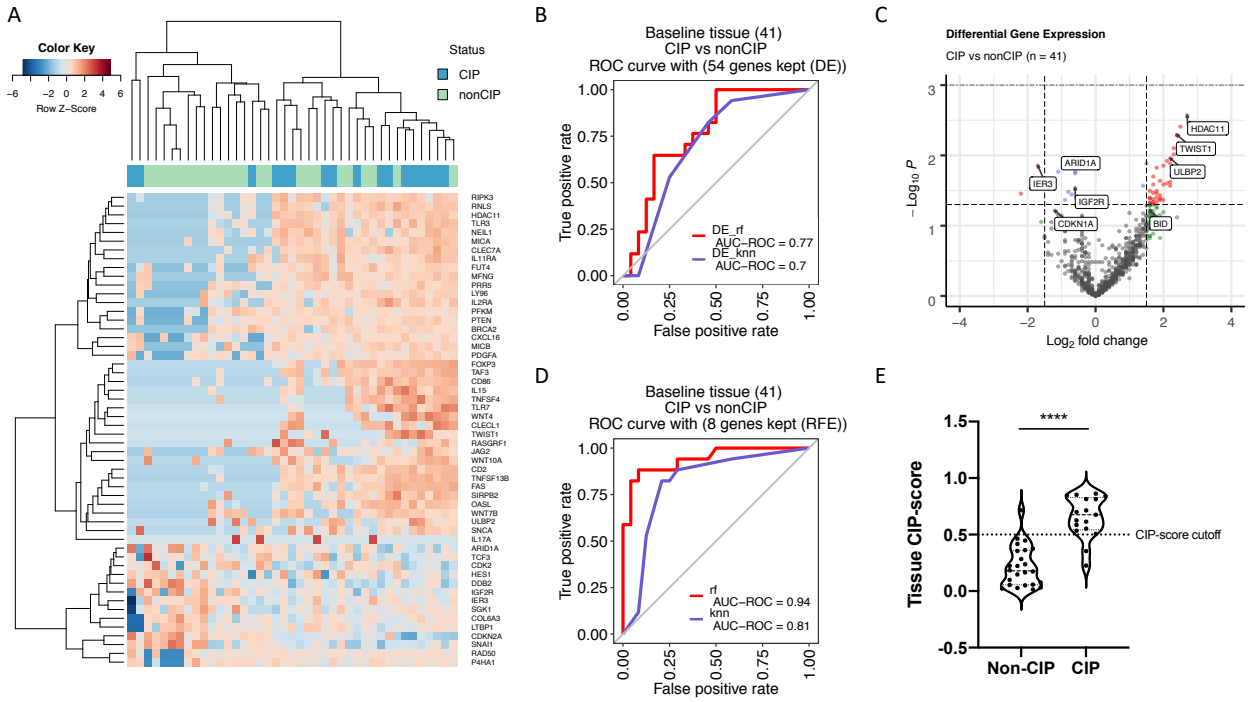


Figure 2. DE analysis of all 770 transcripts included in the panel in the tumor tissue of CIP vs non-CIP patients. **(A)** Hierarchical clustering of the patients based on the 54 mRNA transcripts that were found to be expressed at significantly different levels between CIP and non-CIP patients (**Supplementary Table 1**). **(B)** ROC curve indicative of classifier performance to distinguish between CIP and non-CIP tumor tissues using all 54 DE transcripts. Optimal classifier selection yielded accuracies of 65.9%-73.2% (**Supplementary Table 2**). **(C)** Volcano plot representing the Log2(FC) and nominal -Log10(p-values) of all transcripts included in the panel for CIP vs non-CIP tumor tissue samples. Labels indicate the eight transcripts selected for the final classification signature based on a RFE. **(D)** ROC curve indicative of classifier performance to distinguish between CIP and non-CIP tumor tissues using the RFE-selected 8-gene tissue CIP signature. Optimal classifier selection yielded accuracies of 80.5%-92.7% (**Supplementary Table 2**). **(E)** Signature scores of non-CIP vs. CIP tumor tissue samples based on expression of the eight gene signature ($p < 0.01$ in a Mann-Whitney U test). *CIP: checkpoint inhibitor pneumonitis; ROC-AUC: area under the receiver operating characteristic curve; DE: differential expression; rf: random forest; knn: k-nearest neigbors: RFE: recursive feature elimination.*

Table 3. RFE-selected 8-gene tissue CIP signature. *RFE: recursive feature elimination; CIP: checkpoint inhibitor pneumonitis; Mean Log(NC) CIP: mean of the log normalized counts in tumor tissue of CIP patients; SD: standard deviation; FC: fold change.*

Accession	Gene	Mean Log(NC) CIP	SD Log(NC) CIP	Log(FC) CIP vs nonCIP	CIP	P-value	Release of Cancer Cell Antigens	Cancer Antigen Presentation	T-cell Priming and Activation	Immune Cell Localization to Tumors	Stromal Factors	Recognition of Cancer Cells by T-cells	Killing of Cancer Cells	Myeloid Cell Activity	NK Cell Activity	Cell Cycle and Proliferation	Tumor-Intrinsic Factors	Immunometabolism	Common Signaling Pathways
NM_000389.2	CDKN1A	8,4	2,5	-1.2	Downregulated	0.061	-	-	-	-	-	-	-	+	-	+	+	-	-
NM_001196.3	BID	6,1	2,4	1.6	Upregulated	0.060	-	-	-	-	-	-	-	-	-	+	+	-	-
NM_025217.2	ULBP2	3,6	2,5	2.2	Upregulated	0.011	-	+	-	+	-	-	-	+	-	-	+	-	-
NM_001330636.1	HDAC11	5,7	2,3	2.7	Upregulated	0.003	+	-	-	-	-	-	-	-	-	-	+	-	-
NM_006015.4	ARID1A	9,4	0,8	-0.6	Downregulated	0.018	+	-	-	-	-	-	-	-	-	-	+	-	-
NM_000876.2	IGF2R	8,5	0,8	-0.6	Downregulated	0.029	-	-	-	-	-	-	+	-	-	-	-	-	-
NM_000474.3	TWIST1	2,8	2,9	2.4	Upregulated	0.005	-	-	-	-	+	-	-	-	-	-	+	-	-
NM_003897.3	IER3	8,7	2,4	-1.7	Downregulated	0.014	-	-	-	-	-	-	-	+	-	-	-	-	-

Survival- and univariate analysis related to CIP risk

To determine if CIP development correlated with progression free survival (PFS) and overall survival (OS) of the patients, Kaplan Meier analyses were carried out (**Figure 3A** and **Supplementary Figure 2**). No significant correlations were found between either CIP and non-CIP, or low-grade CIP, high-grade CIP and non-CIP (PFS, log-rank, $p = 0.15$; OS, log-rank, $p = 0.69$ and PFS, log-rank, $p = 0.22$; OS, log-rank, $p = 0.80$, respectively). Interestingly, we could detect a trend towards longer PFS for patients that developed CIP, especially in the high grade CIP cohort. In contrast, these patients tended to have a shorter OS.

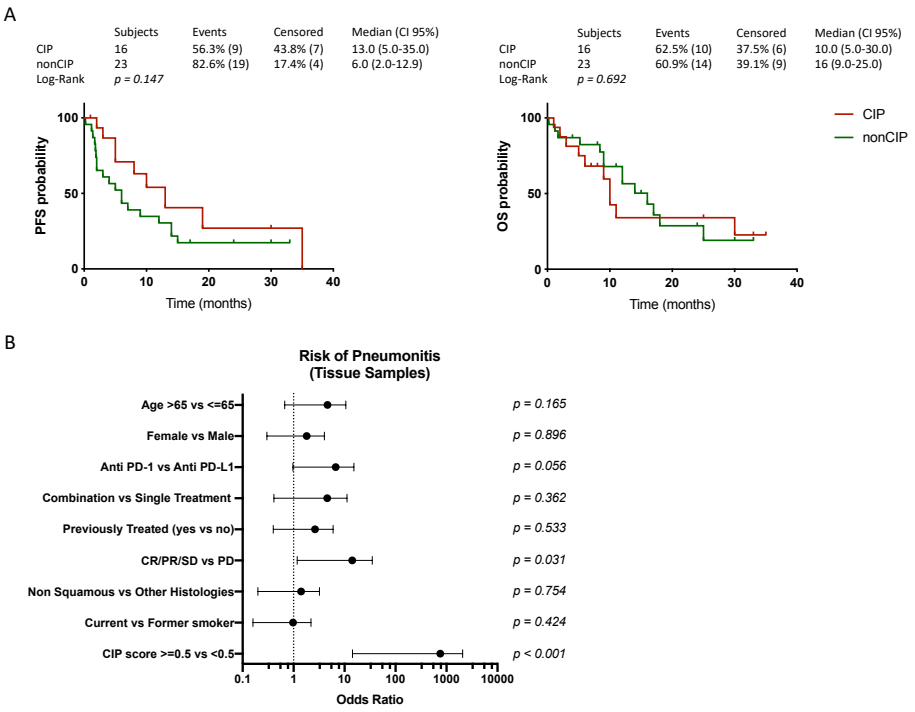


Figure 3. (A) Kaplan Meier analysis correlating the PFS and OS with development of CIP. **(B)** Univariate analysis exploring associations between patient characteristics and CIP development to determine risk factors in the tissue sample cohort. Forest plot represents the odds ratios with 95% wald confidence limits. *PFS: progression-free survival; OS: overall survival; CIP: checkpoint inhibitor pneumonitis; CR: complete response; PR: partial response; SD: stable disease; PD: progressive disease.*

We then explored if certain patient characteristics could provide risk factors for CIP development be performing a univariate analysis (**Figure 3B**). Several characteristics that have previously been associated with higher risk of CIP, such as age, gender, type of ICI treatment, combination treatment, previous treatment, treatment outcome, histology and smoking status, were evaluated. Of these characteristics, a significant association could only be found for treatment outcome. Intriguingly, our developed CIP score was found to be a significant predictive factor for CIP, with an odds ratio of 172.

Gene expression profiling of plasma-derived EV-RNA to predict CIP development

Since obtaining a liquid biopsy sample is less invasive and much faster compared to a tissue biopsy, we investigated if analyzing pre-treatment plasma-derived EV-mRNA from matched patients and an additional patient cohort would yield similar results (**Table 2**). A DE analysis yielded 57 mRNA transcripts with significantly different expression levels between CIP and non-CIP developing patients (**Figure 4A, Supplementary Table 3**). Similar to the tissue analysis, most genes were upregulated in CIP vs. non-CIP patients. Hierarchical clustering did not depict a clear distinction between CIP and non-CIP patients. We then determined if the 8-gene CIP signature that we discovered in tumor tissues (**Figure 2B**) could be applied to pre-treatment EVs. Classification based on this signature resulted in ROC AUCs of 0.45-0.59, which can be interpreted as a poor prediction capacity (**Figure 4B, Supplementary Table 4**).

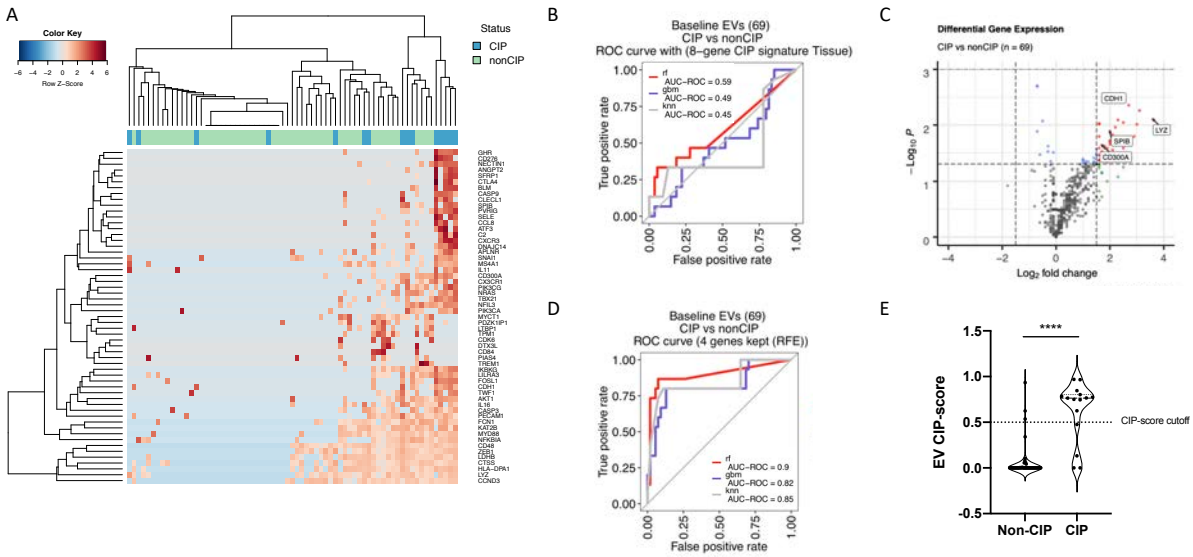


Figure 4. DE analysis of all 770 transcripts included in the panel in pre-treatment EVs of CIP vs non-CIP patients. **(A)** Hierarchical clustering of the patients based on the 57 mRNA transcripts that were found to be expressed at significantly different levels between EVs of CIP and non-CIP patients (**Supplementary Table 3**). **(B)** ROC curve indicative of low classifier performance to distinguish between CIP and non-CIP EV samples using the 8-gene CIP-signature previously developed in tumor tissue samples (**Supplementary Table 4**). **(C)** Volcano plot representing the Log2(FC) and nominal -Log10(p-values) of all transcripts included in the panel for CIP vs non-CIP EV samples. Labels indicate the four transcripts selected for the final classification signature based on a RFE. **(D)** ROC curve indicative of classifier performance to distinguish between CIP and non-CIP EVs using the RFE-selected 4-gene EV CIP signature. Optimal classifier selection yielded accuracies of 84.1% to 89.9% (**Supplementary Table 4**). **(E)** Signature scores of non-CIP vs. CIP EV samples based on expression of the four gene signature ($p < 0.01$ in a Mann-Whitney U test). *CIP: checkpoint inhibitor pneumonitis; EVs: extracellular vesicles; ROC-AUC: area under the receiver operating characteristic curve; DE: differential expression; rf: random forest; gbm: stochastic gradient boosting; knn: k-nearest neigbors: RFE: recursive feature elimination.*

Since EVs contain more systemic information compared to the tumor tissue, their cargo may not only resemble what is going on in the tumor. We then aimed to improve classifier performance by developing a new EV-specific CIP-signature. RFE was carried out on the detected EV-mRNA transcripts and yielded a 4-gene signature for CIP prediction, including *LYZ, CDH1, CD300A* and *SPIB* transcripts (**Figure 4C**, **Table 4**). Annotations of these genes include immune cell localization to tumors and myeloid cell activity. The discrimination capacity of the new classifier resulted in ROC AUCs of 0.82 to 0.90, with accuracies ranging from 84.1% to 89.9% (**Figure 4D, Supplementary Table 4**). The NPV and PPV of the EV CIP signature were 92.7% and 78.6%, respectively, with a Youden index of 0.67. Moreover, EV CIP-scores between CIP and non-CIP developing patients were found to be significantly different (Mann-Whitney U, $p < 0.01$; **Figure 4E**). Consistent with the univariate analysis in the tissue sample cohort, only our EV-CIP score was found to be associated to CIP development in the plasma sample cohort (**Supplementary Figure 3**).

Table 4. RFE-selected 4-gene EV CIP signature. *RFE: recursive feature elimination; EV: extracellular vesicle; CIP: checkpoint inhibitor pneumonitis; Mean Log(NC) CIP: mean of the log normalized counts in tumor tissue of CIP patients; SD: standard deviation; FC: fold change.*

Accession	Gene	Mean Log(NC) CIP	SD Log(NC) CIP	Log(FC) CIP vs nonCIP	CIP	P-value	Release of Cancer Cell Antigens	Cancer Antigen Presentation	T-cell Priming and Activation	Immune Cell Localization to Tumors	Stromal Factors	Recognition of Cancer Cells by T-cells	Killing of Cancer Cells	Myeloid Cell Activity	NK Cell Activity	Cell Cycle and Proliferation	Tumor-Intrinsic Factors	Immunometabolism	Common Signaling Pathways
NM_000239.2	LYZ	6.8	4.3	3.6	Upregulated	0.008	-	-	-	+	-	-	-	+	-	-	-	-	-
NM_004360.3	CDH1	3.0	3.0	2.7	Upregulated	0.004	-	-	-	+	+	-	-	+	-	-	+	-	-
NM_001256841.1	CD300A	2.2	2.6	1.7	Upregulated	0.023	-	-	-	-	-	-	-	-	-	-	-	+	-
NM_003121.3	SPIB	2.1	2.7	2.0	Upregulated	0.013	-	-	-	+	-	-	-	-	-	-	-	-	-

Follow-up EV-RNA analysis for CIP risk and diagnosis

Considering the unmet clinical need for a diagnostic CIP test and aiming to understand the mechanism of CIP development regarding changes in gene expression, we retrospectively explored follow-up plasma EV samples from 2 CIP patients and 2 non-CIP patients using the 4-gene EV CIP-signature (**Figure 5, Supplementary Table 5**). Follow-up samples were included in the sample cohort and the classifier algorithm was used to discriminate between CIP and non-CIP samples based on the CIP-score. Results indicated that the CIP signature score could not only predict CIP before starting ICI treatment, but scores also increased above the cutoff upon actual CIP development (**Figure 5A and B**). EV-mRNA was extracted from plasma samples that were obtained during first evaluation after anti-PD-1 antibody administration, and CIP-scores decreased below the cutoff in both cases. Patient 1 (**Figure 5A**) received nivolumab treatment for 9.5 months, which stabilized tumor growth, until high-grade CIP was diagnosed in June 2017. The CIP-score of the EV-mRNA at this time point increased

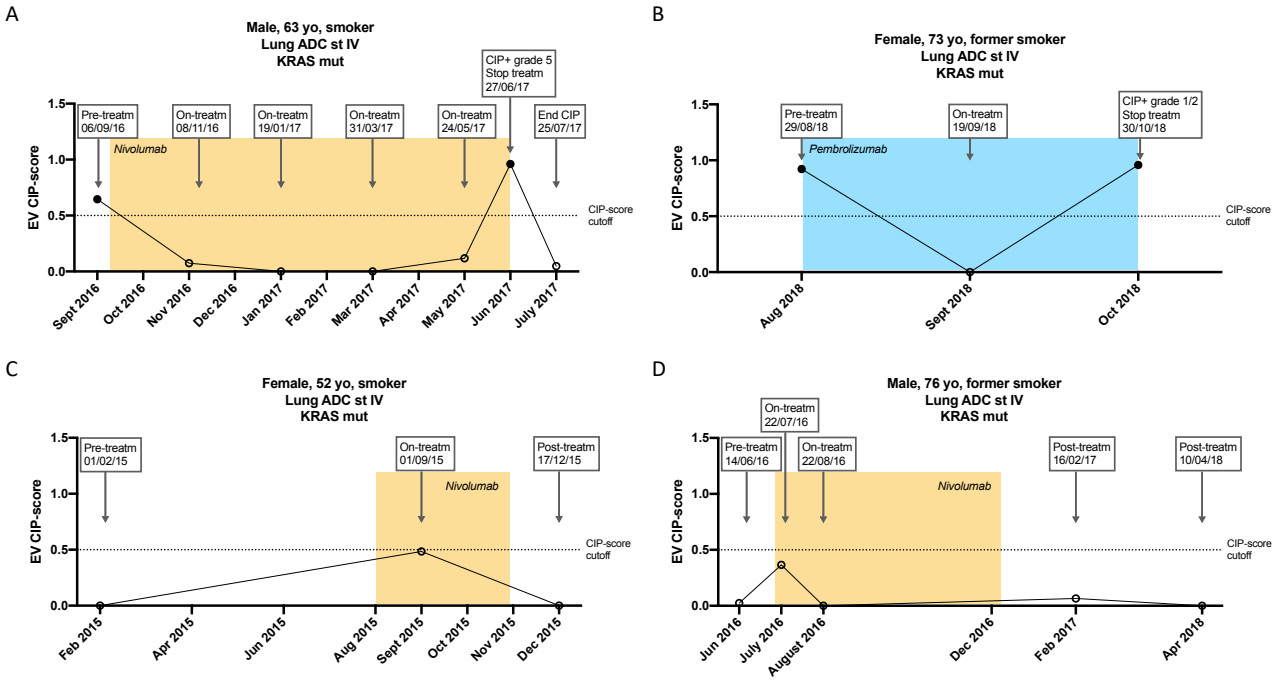


Figure 5. Results of follow-up plasma EV analysis using the 4-gene EV CIP-signature. Follow-up samples included 2 CIP patients (**A** and **B**) and 2 non-CIP patients (**C** and **D**). For CIP patients, the CIP signature scores were above cutoff levels both before starting ICI treatment, and at the time of CIP diagnosis. Signature scores remained below cutoff levels for non-CIP patients. *EV: extracellular vesicle; CIP: checkpoint inhibitor pneumonitis; ADC: adenocarcinoma; ICI: immune checkpoint inhibitor.*

above cutoff, to an even higher level than the pre-treatment CIP-score. ICI treatment was discontinued and the patient was treated with corticosteroids (methylprednisolone). The CIP-score decreased in July 2019, but unfortunately the patient died short after due to CIP-related complications. Patient 2 (**Figure 5B**) was treated with pembrolizumab for only two months when she was diagnosed with low-grade CIP. Also in this patient ICI treatment was discontinued, and corticosteroids were administered. In November 2018 the patient had a progressive disease (PD). She recovered from CIP, and continued on chemotherapy treatment.

In contrast, in patients that did not develop CIP, scores remained below the cutoff of 0.5 (**Figure 5C and D**). For Patient 3 (**Figure 5C**) we only had an available plasma sample that was collected six months before nivolumab administration, where CIP score was below cutoff levels. Less than one month after treatment initiation, the CIP score increased but remained below cut-off. Interestingly, the clinical reports mention that this patient experienced increased cough in September 2015. After two months of treatment the patient had a PD, and nivolumab was discontinued. Another two months later the CIP-score lowered to baseline levels, but the patient died from PD a few months later. Patient 4 (**Figure 5D**) started treatment at the end of June 2016. CIP-score increased one month after nivolumab administration, but lowered to baseline values in August 2016. The patient was on treatment for a total of six months until PD, and CIP-scores remained low until the beginning of 2018. The patient died one month later from a pulmonary thromboembolism.

Discussion

CIP is a serious irAE, that cannot be ignored in the current clinical setting where ICIs are increasingly used, both in previously treated advanced stage lung cancers and as a first line treatment in PD-L1 expressing tumors. Distinct clinical and radiographic manifestations of CIP create a challenging diagnostic process, where multi-disciplinary teams are often necessary. Therefore, being able to predict which patients have a higher risk to develop CIP provides an important tool for extensive monitoring of this subgroup of patients under treatment, and confidently administer ICIs to the patient group with low CIP risk. Risk factors for irAEs, especially for CIP, are currently still in exploratory phases. In the current study most patient characteristics were not found to be associated to CIP development, revealing that they cannot be used for confident treatment decision making. Progressive disease as the best response to treatment was the only characteristic that was associated to lower CIP risk in the tissue sample cohort, likely due to a shorter treatment time compared to patients with treatment response. These results are not in agreement with previously published studies, where a higher incidence of CIP was reported in patients treated with anti-PD-1 antibodies(141) and in patients treated with immunotherapy combinations(142). Also, in one study, CIP incidence was found to be higher in females(48). However, most of these results, including the results presented herein, are derived from retrospective single-center studies. Larger patient cohorts are needed to conclude if any of these characteristics provide a risk factor for CIP development.

Our retrospective gene expression analysis of pre-treatment tumor tissue and plasma-EV mRNA from lung cancer patients with- and without CIP development, indicated that predictive CIP signatures can be derived from mRNA using the nCounter platform. Results of the IO360 panel-incorporated signatures indicated a lower pre-treatment immune cell abundance in tumor tissue of CIP patients, while in general most transcripts included in the panel were upregulated in the tumor tissue of these patients. To the best of our knowledge, no previous studies have reported on the analysis of gene expression or immune cell abundance in the pre-treatment tumor tissues of CIP developing patients. However, Suresh et al.(143) performed immune cell population analysis in bronchoalveolar lavage (BAL) samples of patients during CIP. An increased fraction of activated T cells and loss of anti-inflammatory regulatory T cells (Tregs) was detected in these patients. In addition, lower expression of CTLA-4 and PD-1 in CIP patients was reported. NOS2 is an inhibitory immune signaling signature, like PD-L1 and CTLA-4, that has been associated with intratumor accumulation of Tregs(144). In agreement with their findings, we found lower NOS2 signature scores, indicative of a loss of Tregs, in pre-treatment tumor tissue of CIP patients. Further studies are needed to explore the mechanisms of CIP development related to immune population abundance in both BAL- and tumor tissue samples, before and during CIP development.

Development and optimization of a classifier algorithm yielded a predictive 8-gene CIP signature in tumor tissue, including transcripts for *CDKN1A, BID, ULBP2, HDAC11, ARID1A, IGF2R, TWIST1* and *IER3*. The average accuracy and the NPV of the signature were 86.6% and 92.0%, respectively. In the clinical setting, this allows for a more confident treatment decision making by assessing risk-benefit ratios for each patient. Reinforcing our finding, a

decrease in IGF2R protein expression in lung tumors was previously found to predispose patients to radiation-induced lung injury(145). The other transcripts have not previously been associated to CIP. Results also indicated that pre-treatment tumor tissue- and plasma EV-mRNA expression are divergent, and led to the development of an additional, EV-derived, 4-gene CIP signature based on a total of 69 plasma samples. This signature included *LYZ, CDH1, CD300A* and *SPIB* transcripts, and compared to the tissue CIP-signature we obtained an improved accuracy and NPV of 87.0% and 92.7%, respectively. However, the PPV decreased from 93.8% for the tissue CIP-score, to 78.6% in the EVs. This is likely due to the fact that the prevalence of CIP is lower in our plasma cohort (21.7%), but resembles the real-world situation better, compared to the tissue cohort (41.5%)(146).

Three out of four genes in our signature (*LYZ*, *CDH1* and *SPIB*) have been related to immune cell localization to tumors. Lysozyme (encoded by the *LYZ* gene), is an antimicrobial enzyme with important roles in the innate immune system and serum levels have been used as a marker for sarcoidosis(147). Since CIP and sarcoidosis are both characterized by an overactivation of the immune system, lysozyme could be an important player in both pathologies. Furthermore, CD300A is expressed on several immune cell types and generally acts as an inhibitory receptor(148). However, CD8+ T cells that express CD300A have a more cytotoxic phenotype and were found to infiltrate the placenta in maternal anti-fetal rejection. This processes causes a destructive inflammation of the placenta due to a maternal immune response against the fetus(149). Although none of these transcripts have been previously associated to CIP, most of them have been reported to play a role in other abnormal immune responses and further studies are needed to reveal the relationship of these genes with CIP and determine if they are CIP-specific or immune overactivation specific.

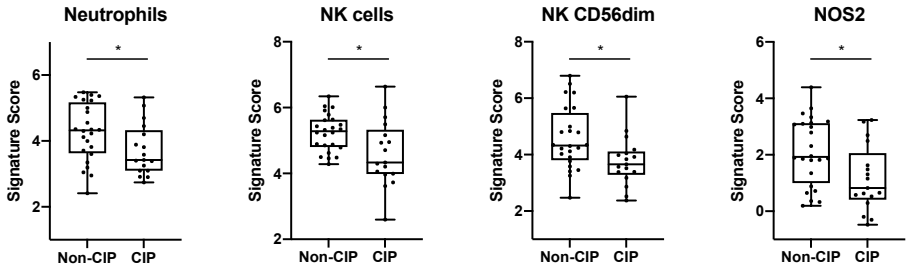
Two CIP patients were misclassified using the 8-gene tissue CIP-signature with scores of 0.226 and 0.354, respectively. Although borderline, the first patient was also misclassified with the 4-gene EV-signature (0.476). This patient had a grade 2 CIP, three months after initiation of nivolumab treatment. In contrast, the second patient was correctly classified using the EV-signature (0.802), and developed a grade 1 CIP two months after starting nivolumab treatment. Finally, one non-CIP patient was misclassified with the tissue CIP-signature (0.716), while the EV CIP-signature correctly classified this patient as non-CIP (0.000). These results emphasize the use of EV-mRNA in predictive biomarker assay development for a complete readout of the ongoing biological processes in a cancer patient. Besides the wrongly classified CIP patient mentioned above, three other CIP patients were misclassified with the 4-gene EV CIP-signature. For one of them we had a matching tissue sample, which correctly predicted CIP with a tissue CIP-score of 0.714. We did not have matched tissue for the other two samples. However, clinical data indicated that one of those samples was actually collected after CIP development upon administration of a previous treatment, and the patient did not develop CIP after collecting the plasma sample.

Differential diagnosis of CIP based on radiographic evidence is complex due to similarities with other complications. In addition, no formal diagnostic criteria or test, except for guidelines, are available(128, 143). This underscores the importance of the finding that our EV

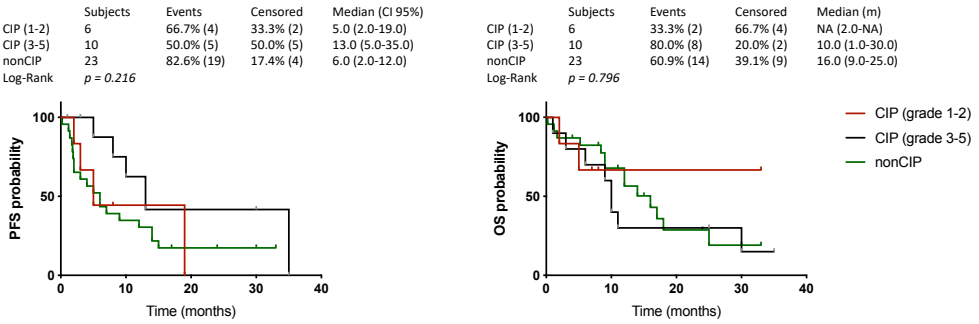
CIP-signature could not only predict higher CIP-risk, but increased upon actual CIP development in follow-up samples. Importantly, the nCounter workflow for EV-mRNA analysis can be carried out in less than three days(140), allowing for the required rapid clinical decision making.

Our study shows several limitations. First, to implement a predictive and/or diagnostic biomarker assay, extensive retrospective and prospective validation studies need to be carried out in addition to the assessed patient cohorts. Second, patients with other lung disorders (e.g. pneumonia, asthma, COPD, bronchitis and COVID-19) should be included in follow-up studies, to determine if the signature is CIP-specific, or related to a general hyper immune activation. And finally, as previously indicated, the collected tissue samples did not reflect the real-world prevalence of CIP, and therefore PPV values may reflect an over-estimation. Prospective testing of the predictive and/or diagnostic use of the EV CIP signature is ongoing, and may be extended to testing in upcoming clinical trials investigating anti PD-1/PD-L1 antibodies.

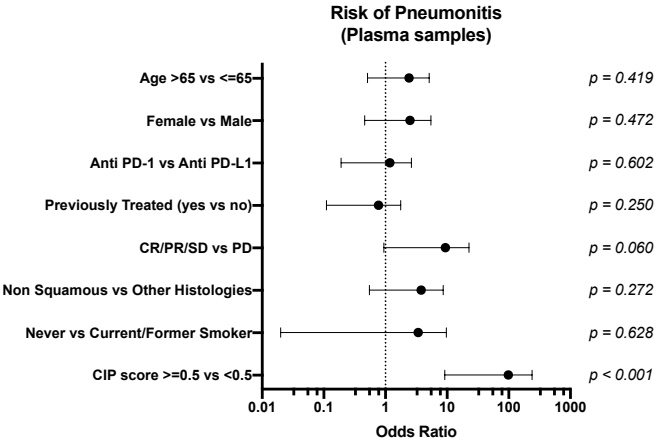
Supporting information



Supplementary Figure 1. Panel-incorporated biological signature scores, indicative of relative cell type abundance in tumor tissue from CIP vs non-CIP lung cancer patients. Differences were assessed using a Mann-Whitney U test; p = 0.04, p = 0.02, p = 0.01 and p = 0.01, respectively. *CIP: checkpoint inhibitor pneumonitis; NK: natural killer.*



Supplementary Figure 2. Kaplan Meier analysis correlating the PFS and OS with development of low- or high grade CIPs. *PFS: progression-free survival; OS: overall survival; CIP: checkpoint inhibitor pneumonitis.*



Supplementary Figure 3. Univariate analysis exploring associations between patient characteristics and CIP development to determine risk factors in the plasma sample cohort. Forest plot represents the odds ratios with 95% wald confidence limits. *CIP: checkpoint inhibitor pneumonitis; CR: complete response; PR: partial response; SD: stable disease; PD: progressive disease.*

Supplementary Table 1. Significantly DE transcripts ($p < 0.05$) in tumor tissue samples between CIP and non-CIP patients. A Log2(FC) > 0 corresponds to enrichment in tumor tissue of CIP-developing patients. *DE: differentially expressed; CIP: checkpoint inhibitor pneumonitis; P: nominal p-value; FC: fold change.*

	Transcript	P_CIP	Log2(FC)_CIP
1	HDAC11	0,0027	2,7
2	MICB	0,0039	2,5
3	TWIST1	0,0051	2,4
4	RIPK3	0,0052	2,4
5	RNLS	0,0079	2,3
6	RASGRF1	0,0097	2,3
7	ULBP2	0,011	2,2
8	PTEN	0,012	2,1
9	LY96	0,013	2,2
10	TLR7	0,013	1,5
11	IER3	0,014	-1,7
12	TNFSF13B	0,014	2
13	TAF3	0,015	1,8
14	SNCA	0,015	2,3
15	WNT10A	0,016	2,5
16	TCF3	0,017	-0,6
17	SNAI1	0,017	-1,1
18	ARID1A	0,018	-0,6
19	MICA	0,019	2,5
20	P4HA1	0,02	1,7
21	NEIL1	0,02	2,3
22	FOXP3	0,023	1,9
23	CLEC7A	0,024	2,2
24	CD2	0,025	2,1
25	RAD50	0,026	1,7
26	BRCA2	0,026	2
27	CDKN2A	0,026	-1,4
28	WNT4	0,027	1,4
29	OASL	0,027	2,2
30	IGF2R	0,029	-0,6
31	SGK1	0,029	-1,6
32	LTBP1	0,03	-1,9
33	CLECL1	0,032	1,6
34	CXCL16	0,032	1,8
35	DDB2	0,033	-0,8
36	PFKM	0,034	1,7
37	COL6A3	0,035	-2,2
38	FAS	0,035	1,8
39	HES1	0,036	-0,7
40	WNT7B	0,037	1,9
41	TNFSF4	0,04	1,6
42	FUT4	0,04	1,8
43	TLR3	0,041	1,9
44	PDGFA	0,041	1,6
45	CD86	0,042	1,7
46	IL2RA	0,043	2
47	IL17A	0,043	-0,9
48	MFNG	0,043	1,9
49	IL15	0,046	1,7
50	IL11RA	0,046	1,8
51	JAG2	0,047	1,6
52	PRR5	0,048	1,6
53	SIRPB2	0,049	1,7
54	CDK2	0,049	-0,3

Supplementary Table 2. Confusion matrices based on optimal classifiers using all DE genes between tissue of CIP and non-CIP patients (top), or the RFE selected 8-gene tissue signature (bottom). *DE: differentially expressed; CIP: checkpoint inhibitor pneumonitis; RFE: recursive feature elimination; PPV: positive predictive value; NPV: negative predictive value.*

Tissue DE genes

		Predicted		
		CIP	nonCIP	Total
Actual	CIP	10	7	17
	nonCIP	4	20	24
	Total	14	27	41

	Value	%
Sensitivity	0,59	58,82
Specificity	0,83	83,33
PPV	0,71	71,43
NPV	0,74	74,07
Accuracy	0,73	73,17
Youden´s index		0,42

Tissue RFE-selected 8-gene CIP signature

		Predicted		
		CIP	nonCIP	Total
Actual	CIP	15	2	17
	nonCIP	1	23	24
	Total	16	25	41

	Value	%
Sensitivity	0,88	88,24
Specificity	0,96	95,83
PPV	0,94	93,75
NPV	0,92	92,00
Accuracy	0,93	92,68
Youden´s index		0,84

Supplementary Table 3. Significantly DE transcripts ($p < 0.05$) in EV samples between CIP and non-CIP patients. A Log2(FC) > 0 corresponds to enrichment in EVs of CIP-developing patients. *DE: differentially expressed; CIP: checkpoint inhibitor pneumonitis; P: nominal p-value; FC: fold change.*

	Transcript	P_CIP	Log2(FC)_CIP
1	LTBP1	0,002	-0,7
2	TPM1	0,002	-0,7
3	CDH1	0,0044	2,7
4	ZEB1	0,0055	3,1
5	LYZ	0,0079	3,6
6	SNAI1	0,0081	2,3
7	PIAS4	0,0085	-0,5
8	FOSL1	0,009	2,5
9	PVRIG	0,0095	1,6
10	DNAJC14	0,0095	1,6
11	HLA-DPA1	0,0097	3
12	TBX21	0,011	2,2
13	PDZK1IP1	0,013	-0,7
14	SPIB	0,013	2
15	LILRA3	0,015	2,1
16	CD48	0,016	2,9
17	CASP9	0,016	1,6
18	AKT1	0,019	2
19	FCN1	0,02	2,2
20	KAT2B	0,02	2
21	CTSS	0,021	2,5
22	MYD88	0,021	2,2
23	NFIL3	0,022	2
24	CD300A	0,023	1,7
25	BLM	0,024	1,8
26	CDK6	0,024	-0,3
27	ANGPT2	0,026	1,6
28	NFKBIA	0,026	2,5
29	MS4A1	0,026	1,9
30	CASP3	0,026	1,8
31	CLECL1	0,027	1,5
32	PECAM1	0,028	2,1
33	GHR	0,03	1,7
34	TREM1	0,03	-0,3
35	DTX3L	0,031	-0,1
36	IL16	0,033	1,5
37	IKBKG	0,033	2,2
38	CX3CR1	0,034	1,5
39	MYCT1	0,034	-0,6
40	PIK3CG	0,035	1,6
41	CCND3	0,038	2,1
42	APLNR	0,039	1,4
43	NECTIN1	0,039	1,7
44	LDHB	0,04	2,1
45	C2	0,041	1
46	PIK3CA	0,042	1,5
47	CD276	0,043	1,6
48	TWF1	0,043	1,6
49	CXCR3	0,043	1
50	CD84	0,045	-0,2
51	CCL8	0,045	1,1
52	IL11	0,046	1,9
53	SFRP1	0,046	1,4
54	NRAS	0,046	1,2
55	SELE	0,047	1
56	CTLA4	0,047	1,4
57	ATF3	0,049	1,1

Supplementary Table 4. Confusion matrices based on optimal classifiers using the 8-gene tissue CIP signature (top), or the RFE selected 4-gene EV signature (bottom) to distinguish between EVs from CIP and nonCIP patients. *CIP: checkpoint inhibitor pneumonitis; RFE: recursive feature elimination; EVs: extracellular vesicles; PPV: positive predictive value; NPV: negative predictive value.*

8-gene tissue CIP signature

		Predicted		
		CIP	nonCIP	Total
Actual	CIP	2	13	15
	nonCIP	2	52	54
	Total	4	65	69

	Value	%
Sensitivity	0,13	13,33
Specificity	0,96	96,30
PPV	0,50	50,00
NPV	0,80	80,00
Accuracy	0,78	78,26
Youden´s index		0,10

RFE-selected 4-gene EV CIP signature

		Predicted		
		CIP	nonCIP	Total
Actual	CIP	11	4	15
	nonCIP	3	51	54
	Total	14	55	69

	Value	%
Sensitivity	0,73	73,33
Specificity	0,94	94,44
PPV	0,79	78,57
NPV	0,93	92,73
Accuracy	0,90	89,86
Youden´s index		0,68

Supplementary Table 5. Clinical characteristics of pre-treatment tumor tissue samples from CIP and non-CIP patients. *CIP: checkpoint inhibitor pneumonitis; yr: years; NSCLC: non-small cell lung cáncer; ICI: immune checkpoint inhibitor; m: months.*

Characteristics	Patient 1	Patient 2	Patient 3	Patient 4
Status	CIP	CIP	nonCIP	nonCIP
Samples				
Pre-treatment	1	1	1	1
On-treatment	5	2	1	2
Post-treatment	1	0	1	2
Gender	M	F	F	M
Age (yr)	63	73	52	76
Smoking status	Current	Former	Current	Former
Histology	Non Squamous NSCLC	Non Squamous NSCLC	Non Squamous NSCLC	Non Squamous NSCLC
Stage	IV	IV	IV	IV
Treatment line	Pre-treated	Pre-treated	Pre-treated	First line
ICI treatment	Nivolumab	Pembrolizumab	Nivolumab	Nivolumab
Clinical benefit*	Yes	No	No	No
CIP grade	Grade 5	Grade 1/2	NA	NA
Time to CIP (m)	9	2	NA	NA
PD-L1 IHC	Unknown	Positive	Unknown	Negative

**Clinical benefit is defined as the sum of patients with complete response (CR) or partial response (PR) plus stable disease (SD) longer than 24 weeks.*

Acknowledgements
The investigators wish to thank the patients for kindly accepting to donate samples to this study. We also thank all the physicians who collaborated by providing clinical information. Finally, we would like to thank colleagues and collaborators for their constructive feedback on this manuscript.

Financial support
This project has received funding from a European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement ELBA No 765492.

Chapter 5. NanoString-based analysis of plasma EV-mRNA for the early detection of lung cancer.

Jillian WP Bracht, Carlos Pedraz-Valdunciel, Silvia d´Ambrosi, Mafalda Antunes Ferreira, Chung-Ying Huang, Ana Gimenez-Capitan, Nicolas Potie, Danijela Koppers-Lalic, Rafael Rosell, Thomas Wurdinger, Miguel A Molina-Vila.

Manuscript in preparation

Abstract

Lung cancer remains the leading cause of cancer-related death worldwide, mainly due to the late stage diagnosis. New tools for early lung cancer detection, with low fractions of false-positive findings, are an unmet clinical need. We have previously created a pipeline for plasma extracellular vesicle (EV)-mRNA analysis on the nCounter platform and consequent gene signature development. In this study we aimed to create a specific lung cancer detection signature. EVs were enriched from 600 uL plasma of the training cohort (n = 57), including plasma from early- and late stage lung cancer patients and controls. RNA was extracted, and the pre-amplified cDNA was analyzed using the Human Immune V2 panel. A machine learning algorithm was then used to select the most optimal genes for the development of a classifier algorithm. The final model consisted of a 16-gene signature, which yielded accuracies of 81% to 84% in the training cohort. We then validated this lung cancer signature in a separately collected validation cohort (n = 43), with equally high accuracies (81%-84%). In conclusion, an nCounter-derived 16-gene EV signature can be used to accurately distinguish between early- and late stage lung cancer patients versus controls.

Introduction

With over 2 million newly diagnosed cases in 2018, lung cancer is the most commonly diagnosed cancer and the leading cause of cancer-related death worldwide(1). The often long asymptomatic phase translates into a late stage (IIIB/IV) diagnosis in 80% of the cases, with a dramatic life expectancy of only 1-2 years(150). In contrast, lung cancer patients diagnosed at early stages (I/II/IIIA) can often undergo surgery and have a significantly better clinical outcome. To this end, new tools for early lung cancer detection are urgently needed and could provide a diagnostic stage-shift, ultimately leading to prolonged patient survival.

Considering the major impact of early lung cancer detection on patient survival, trials that screen for lung cancer have been carried out since the 1960s(151). In the last decade, low-dose CT (LDCT) screening trials have gained additional interest. In 2011, the National Lung Screening Trial (NLST) reported a 20% reduction in lung cancer mortality when individuals underwent LDCT vs. chest radiography screening(152). Although other smaller studies did not observe a mortality reduction when screening with LDCT(153-155), the NELSON trial confirmed this reduction in lung cancer mortality of 24%-33%(156). Crucial characteristics of any screening test include the positive- and negative predictive values (PPV and NPV). Both describe the performance of the screening test and also rely on the prevalence of the disease. In the case of the NLST, it was shown that of all LDCT scans that had an abnormality (positive result, 24.2%), 96.4% were false positive results. The often unnecessary and potentially harmful follow-up interventions ultimately formed the basis to search for other types of- or additional screening tests(157-159).

Tumor cells, like any other cell type, continuously release nucleic acids into the bloodstream. RNA and DNA can be freely circulating within body fluids or can be loaded into membrane-bound compartments, such as tumor-educated platelets (TEPs) or extracellular vesicles (EVs). The lipid bilayer of EVs contribute to the stability of the nucleic acids, and allow for the use of biobank stored samples. EVs can be found in divergent sizes, contain a specific cargo and comprise several biological functions(69-71). They are important mediators of intercellular communication and their RNA content can provide a read-out of early oncological development(69, 72, 73). Although efforts have been made to use EV-RNA for early cancer detection(160, 161), lack of standardized protocols often yields divergent signatures and technical requirements are not suitable for implementation in the clinical laboratory setting(8, 9, 162). The NanoString nCounter platform is a multiplexed analysis platform that uses direct capturing and counting of up to 800 RNA targets(11, 92). The FDA approved Prosigna assay emphasizes the potential of the nCounter platform as a biomarker assay development tool, especially when used in diagnostic laboratories(22, 23).

We have previously developed an nCounter workflow for the detection of plasma EV-mRNA targets and the development of gene signatures(140). Using only eight gene transcripts, a machine learning algorithm was able to distinguish between EVs derived from cancer patients with different tumor types and controls with areas under the ROC curve of 0.99. In the current study we developed a specific lung cancer signature based on EV-mRNA expression analysis

with the nCounter platform. The final 16-gene signature could be used to differentiate between early- and late stage lung cancer vs. control samples. We then validated this signature in a separate cohort, yielding high accuracies.

Materials and methods

Patient samples

This study was carried out in accordance with the principles of the Declaration of Helsinki under approved protocols of the institutional review board of Quirón Hospitals and Amsterdam UMC. Plasma samples for the training and validation cohorts were retrospectively and prospectively collected from lung cancer patients and control individuals in two hospitals (**Table 1**, **Table 2**). Written informed consent was obtained from all participants and documented; samples were de-identified for confidentiality.

Table 1. Clinical characteristics of the training cohort. *no: number; yr: years.*

Characteristics	Cancer (n = 31)	Controls (n = 26)
Gender – no. (%)		
Male	16 (51.6%)	12 (46.2%)
Female	15 (48.4%)	14 (53.8%)
Age (yr) – median (range)		
Median	63 (51- 78)	60 (42-79)
Smoking status – no. (%)		
Never	7 (22.6%)	15 (57.7%)
Former	11 (35.5%)	5 (19.2%)
Current	2 (6.4%)	6 (23.1%)
Unknown	11 (35.5%)	0 (0.0%)
Stage – no. (%)		
Stage I	13 (41.9%)	NA
Stage II	3 (9.7%)	NA
Stage III	4 (12.9%)	NA
Stage IV	11 (35.5%)	NA

Extracellular Vesicle Enrichment

Whole blood samples (10 mL) were collected in sterile EDTA Vacutainer tubes (BD, Plymouth, UK) and centrifugated twice at 1000 x g for 10 min at room temperature (RT). Plasma samples were stored at -80°C until further processing. The miRCURY® Exosome Serum/Plasma Kit (Qiagen, Hilden, Germany) was used to enrich for EVs from 600 µL plasma, according to the manufacturer's instructions and as previously described(140). In short, dead cells and debris (including platelets and fibrin) were cleared with thrombin and centrifugation. Precipitation Buffer was added, samples were incubated overnight at 4°C and the EV fraction was pelleted by centrifugation. EV-enriched pellets were treated with 4 µg/mL RNase A (Sigma-Aldrich, St. Louis, MO) for 1 hour (h) at 37°C, to remove extracellular RNA not associated to EVs. EV characterization was previously performed using identical protocols(140).

Table 2. Clinical characteristics of the validation cohort. *no: number; yr: years.*

Characteristics	Cancer (n = 24)	Controls (n = 19)	Nodules (n = 3)
Gender – no. (%)			
Male	14 (58.3%)	7 (36.8%)	1 (33.3%)
Female	10 (41.7%)	12 (63.2%)	2 (66.7%)
Age (yr) – median (range)			
Median	65 (37-91)	40 (22-66)	54 (52 - 67)
Smoking status – no. (%)			
Never	5 (20.8%)	7 (36.8%)	0 (0.0%)
Former	15 (62.5%)	4 (21.1%)	1 (33.3%)
Current	4 (16.7%)	8 (42.1%)	1 (33.3%)
Unknown	0 (0.0%)	0 (0.0%)	1 (33.3%)
Stage – no. (%)			
Stage I	4 (16.7%)	NA	NA
Stage II	1 (4.2%)	NA	NA
Stage III	8 (33.3%)	NA	NA
Stage IV	11 (45.8%)	NA	NA

EV RNA extraction

TRI-reagent (MRC, Cincinnati, OH) was added to the EV-enriched pellets in a final volume of 1 mL and incubated at RT for 20 min to inactivate RNase A and lyse the EVs, as previously described(140). In short, chloroform: Isoamyl Alcohol (24:1) (Panreac Química SLU, Barcelona, Spain) was added and after centrifugation the aqueous upper layer was kept to precipitate RNA after adding 2.5 µL Glycogen (Merck KGaA, Darmstadt, Germany) and 500 µL 2-propanol (Merck KGaA). The final RNA pellet was washed with 75% ethanol, air dried and dissolved in 20 µL nuclease free water. To remove co-isolated DNA from the resuspended RNA, the DNA-free™ DNA Removal Kit (Thermo Fisher Scientific), was used according to manufacturer instructions. In short, 1 µL DNase buffer and 0.5 µL enzyme were added to 7.5 µL RNA sample, followed by incubation at 37°C for 30 min and DNase removal.

Gene Expression Analysis using nCounter

The nCounter® Low RNA Input Amplification Kit (NanoString Technologies, Seattle, WA) was used to retrotranscribe and pre-amplify 4 µL EV-derived RNA using 10 cycles. Retrotranscription was carried out in 0.5 mL tubes while pre-amplification, using primers targeting the genes of The Human Immunology V2 Panel (NanoString Technologies), was performed in 384-well plates to prevent sample evaporation. The Human Immunology V2 Panel (NanoString Technologies) was used to analyze EV-derived, pre-amplified cDNA according to manufacturer instructions. The Human Immunology V2 panel targets 594 different transcripts involved in the immune response such as cytokines, enzymes, interferons and their receptors. Samples were hybridized for 18 h at 65°C.

Data Normalization and Analysis

Raw nCounter counts of expressed genes were normalized in R and R studio v3.6.3 using the R package *NanoStringNorm*(110). Normalization was performed following several steps: technical assay variability normalization using the geometric mean of the positive control probes, background correction using the mean plus two times standard deviation (SD) of the negative control probes, and sample content normalization using the total amount of counts for each sample. Normalized counts were log2-transformed, and used for differential expression (DE) analysis. Log2 fold change (FC) of each gene was calculated as the ratio of average log-2 transformed counts of the cancer patient cohort vs. the control cohort. Volcano plots were used to visualize log2 FC on the x-axis and nominal p-values on the y-axis. GraphPad Prism software v9.0.0 was used for statistical testing and to create figures.

Classifier Algorithm Development

Optimal gene selection was performed using recursive feature elimination (RFE). To this end, a leave-one-out cross validation (LOOCV) algorithm was used on the full Human Immunology V2 Panel. The number of genes to select was set at 4, 8, 16 or 594 and the amount of genes that yielded optimal performance after cross-validation was automatically selected. Classification was performed with the selected gene signature using random forest (rf), k-nearest neighbors (knn) and stochastic gradient boosting (gbm) classifiers, where appropriate, using three iterations. The model with the highest accuracy was then selected as the final model. Signature scores for each sample were derived from the final model.

Results

Plasma EV-mRNA signature development for lung cancer detection

EVs were enriched from the plasma samples of the training cohort (n = 57) and RNA was extracted. A multiplexed gene expression analysis was then performed using the Human Immune V2 panel on the nCounter platform, including 594 gene transcripts. The training cohort consisted of 26 control samples and 31 plasma samples from lung cancer patients, including 20 early stage (stage I, II and IIIA) and 11 late stage (stage IIIB and IV) patients (**Table 1**). After data normalization we found 48 differentially expressed (DE) genes between EVs from lung cancer patients and controls (**Figure 1A**, **Supplementary Table 1**). Most genes were overexpressed in EVs from lung cancer patients. Hierarchical clustering of the training cohort, using the DE genes, did not accurately separate the cancer patients from the controls (**Figure 1B**). However, 13 out of 31 lung cancer samples clustered together. We then explored if a machine learning algorithm could distinguish between lung cancer and control samples based on the expression of all 594 genes included in the panel. Using the full gene panel, we obtained areas under the receiver operating characteristic curve (ROC-AUCs) of 0.63-0.75 and accuracies ranging from 61.4% to 73.7%, depending on the type of classifier that was used (**Figure 1C**, **Table 3**).

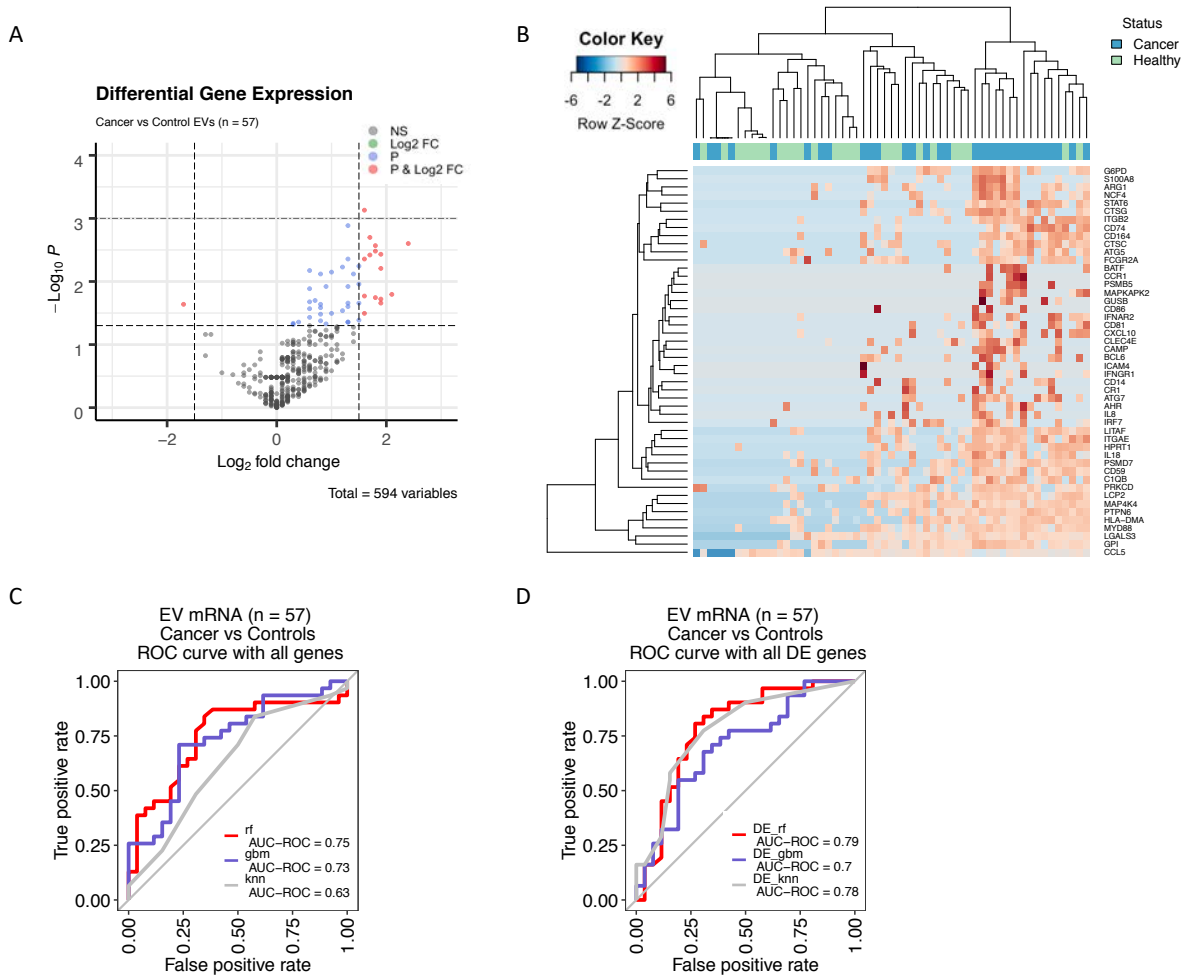


Figure 1. EV-mRNA analysis of the samples included in the training cohort, and consequent classifier algorithm development. **(A)** Volcano plot representing the Log2(FC) and nominal -Log10(p-values) of all transcripts included in the Human Immune V2 nCounter panel for lung cancer vs control samples (n = 57). **(B)** Hierarchical clustering of the patients based on the 48 mRNA transcripts that were found to be expressed at significantly different levels between EVs of lung cancer patients vs controls (**Supplementary Table 1**). **(C)** ROC curve indicative of classifier performance to distinguish between lung cancer and control EVs using all transcripts included in the panel (n=594). Optimal classifier selection yielded accuracies of 61.4% to 73.7% (**Table 3**). **(D)** ROC curve indicative of classifier performance to distinguish between lung cancer and control EVs using the previously found DE genes (**Supplementary Table 1**). Optimal classifier selection yielded accuracies of 68.4% to 77.2% (**Table 3**). *EV: extracellular vesicle; FC: fold change; ROC: receiver operating characteristic; DE: differentially expressed; NS: not significant; rf: random forest; gbm: stochastic gradient boosting; knn: k-neirest neigbors.*

Aiming to improve classifier performance, we then selected only the DE genes and developed a new classifier algorithm. ROC-AUC values increased to 0.70-0.79 with accuracies of 68.4%-77.2% in separating lung cancer patients from controls (**Figure 1D**, **Table 3**).

A final optimization step, using a recursive feature elimination (RFE) method to select the best set of genes for consequent classification of the samples, was then carried out. The obtained gene signature consisted of 16 transcripts, including: *S100A9, MYD88, LGALS3, CCL5, GPI, EEF1G, CD14, IRF7, TAGAP, B2M, FCGR2A, MAP4K4, HLA-DMA, HLA-B, ARG1, CD3D* (**Figure 2A**, **Supplementary Table 2**). Again, hierarchical clustering was not able to accurately separate lung cancer from control samples using this 16-gene signature (**Figure 2B**).

However, a machine learning algorithm that made use of these 16 transcripts could distinguish between lung cancer and control samples with accuracies ranging from 80.7% to 84.2% and ROC-AUCs of 0.86-0.91 (**Figure 2C, Table 3**). Importantly, the negative predictive value (NPV), positive predictive value (PPV) and Youden´s index value all improved using this feature selection method, with final values of 84.0%, 84.4% and 0.68, respectively. Amongst all samples included in the training cohort, 5 controls and 4 lung cancer samples were wrongly classified based on the algorithm determined cutoff value of 0.5 (**Figure 2D**).

Table 3. Performance of the machine learning algorithms in the training cohort. *DE: differentially expressed; RFE: recursive feature elimination; rf: random forest; knn: k-neirest neighbors; NPV: negative predictive value; PPV: positive predictive value.*

	Accuracy			Sensitivity	Specificity	NPV	PPV	Youden´s index
	rf	knn	gbm					
All genes (n=594)	70.2%	61.4%	73.7%	71.0%	69.2%	66.7%	73.3%	0.40
DE genes (n=48)	77.2%	68.4%	68.4%	74.2%	73.1%	70.4%	76.7%	0.47
RFE genes (n=16)	84.2%	80.7%	80.7%	87.1%	80.8%	84.0%	84.4%	0.68

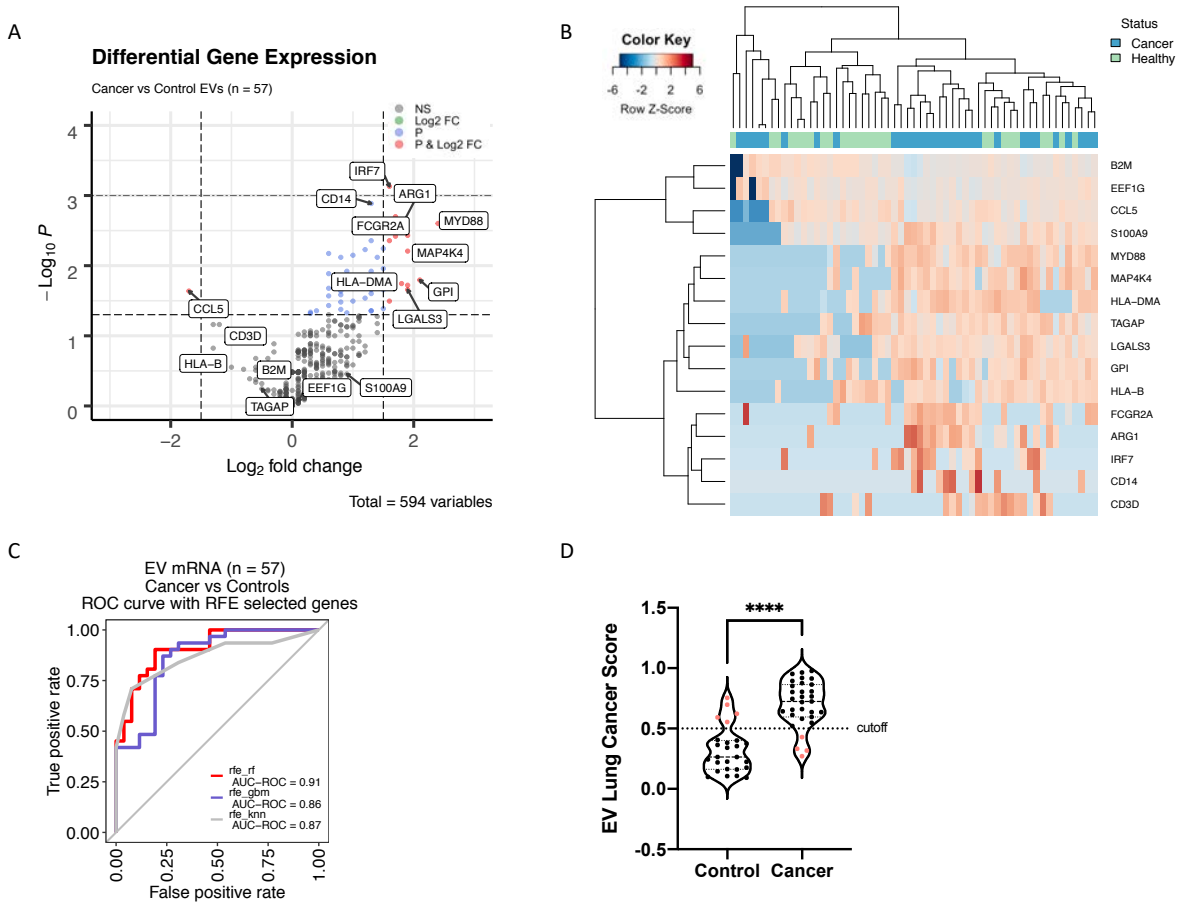


Figure 2. EV-mRNA analysis of the samples included in the training cohort, and consequent classifier algorithm development using a recursive feature elimination (RFE) method. **(A)** Volcano plot representing the Log2(FC) and nominal -Log10(p-values) of all transcripts included in panel for lung cancer vs control samples (n = 57). Labels indicate the 16 transcripts selected for the final classification signature. **(B)** Hierarchical clustering of the patients based on the 16-gene RFE-selected signature (**Supplementary Table 2**). **(C)** ROC curve indicative of

classifier performance to distinguish between lung cancer and control EVs using the RFE-selected 16-gene signature. Optimal classifier selection yielded accuracies of 80.7% to 84.2% (**Table 3**). **(D)** Signature scores of lung cancer vs control EV samples based on expression of the 16-gene signature ($p < 0.01$ in a Mann-Whitney U test). *EV: extracellular vesicle; FC: fold change; NS: not significant; ROC: receiver operating characteristic; rf: random forest; gbm: stochastic gradient boosting; knn: k-neirest neigbors.*

Validation of the plasma EV-mRNA signature for lung cancer detection

To validate our developed 16-gene lung cancer signature, we made use of an additional independently collected EV sample cohort including 19 controls, 11 late stage- and 13 early stage lung cancer samples (**Table 2**). Similar to the results of the training cohort (**Figure 2B**), hierarchical clustering could not be used to accurately separate lung cancer from control samples using the 16-gene signature that was developed based on the training cohort (**Figure 3A**). However, when applying our machine learning algorithm we obtained a high classification performance based on this signature, with ROC-AUCs of 0.9 and matching accuracies of 81.4%-83.7% (**Figure 3B**, **Table 4**). Out of the 43 samples included in the validation cohort, five lung cancer samples were wrongly classified with lung cancer scores < 0.5 (**Figure 3C**). In addition, three control samples were classified as cancer samples (scores > 0.5).

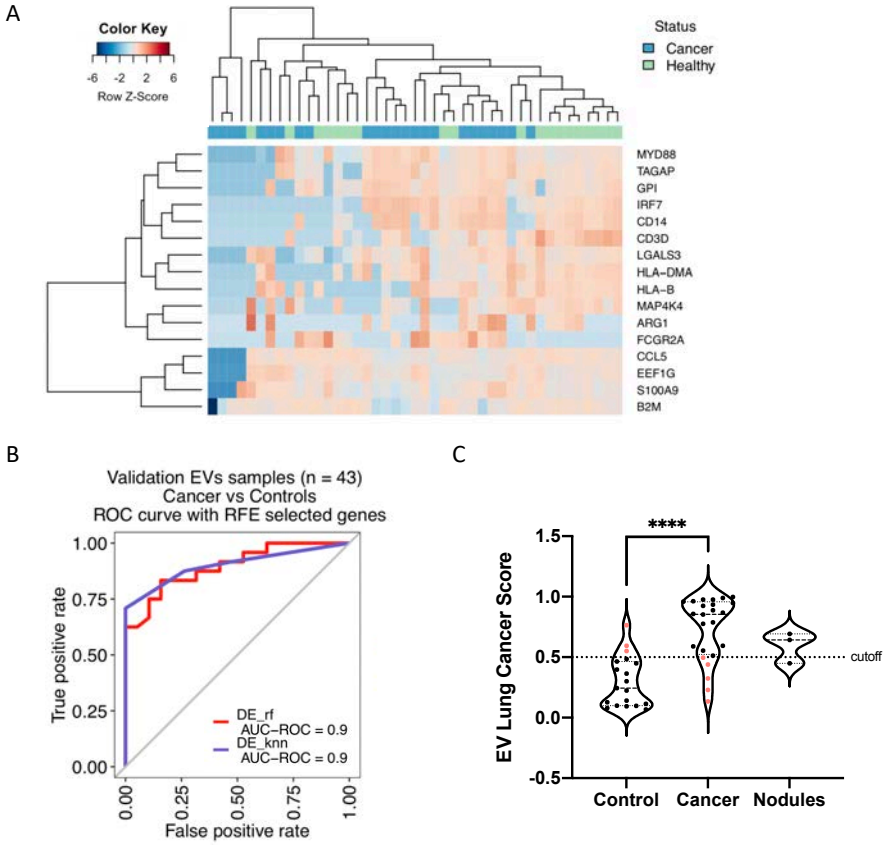


Figure 3. EV-mRNA analysis of the samples included in the validation cohort (n = 43), and consequent validation of the 16-gene lung cancer signature. **(A)** Hierarchical clustering of the patients in the validation cohort based on the 16-gene RFE-selected signature (**Supplementary Table 2**). **(B)** ROC curve indicative of classifier performance to distinguish between lung cancer and control EVs in the validation cohort, using the RFE-selected 16-gene signature. Optimal classifier selection yielded accuracies of 81.4% to 83.7% (**Table 4**). **(C)** Signature scores of lung cancer vs control and three nodule EV samples in the validation cohort, based on expression of the 16-gene signature ($p < 0.01$ in a Mann-Whitney U test for lung cancer vs control). *EV: extracellular vesicle;*

ROC: receiver operating characteristic; RFE: recursive feature elimination; rf: random forest; knn: k-neirest neigbors.

Table 4. Performance of the machine learning algorithms in the validation cohort. *RFE: recursive feature elimination; rf: random forest; knn: k-neirest neighbors; NPV: negative predictive value; PPV: positive predictive value.*

	Accuracy		Sensitivity	Specificity	NPV	PPV	Youden´s index
	rf	knn					
RFE genes (n=16)	81.4%	83.7%	79.2%	78.9%	75.0%	82.6%	0.58

We then analyzed three additional plasma EV samples from individuals with a previously detected lung nodule (**Table 2**). Our developed signature was used to let the algorithm create a lung cancer score for each sample (**Figure 3C**). Two nodule samples were classified as lung cancer with scores of 0.69 and 0.64. One nodule sample was classified as a control with a score of 0.45. Follow-up of these patients will elucidate if the nodules are indeed malignant or not.

Discussion

Lung cancer is often diagnosed at a late stage, when treatment options are not curative anymore. A blood-based biomarker assay that can detect lung cancer at an early stage allows for the diagnostic stage-shift that could reduce lung cancer mortality. In the current study, we have used the nCounter platform to develop and validate a 16-gene signature that can distinguish between plasma EV-mRNA profiles of early- or late stage lung cancer patients versus control individuals. The final models yielded accuracies ranging from 80.7% to 84.2% in the training-, and 81.4% to 83.7% in the validation cohort. To the best of our knowledge, this is the first nCounter-based EV-mRNA biomarker assay for the early detection of lung cancer.

Although low-dose computed tomography (LDCT) is currently the only recommended lung cancer screening test, the NLST has showed that 96.4% of the abnormal scans were false positives. This translates into a PPV of only 6.6%-9.7%(152, 163). The same holds true for chest radiography, where 6.9% of the tests were positive, and 94.5% were false positive results. This high fraction of false positive results can lead to unnecessary and potentially harmful follow-up interventions, in conjunction with psychological distress and anxiety for the screened individual. The unmet clinical need to detect lung cancer at an early stage with less false positive results resulted in the exploration of other types of screening tests.

In the last years, several studies have aimed to develop an EV-based lung cancer screening test. Most of these studies made use of the miRNA(164, 165) or long non-coding RNA (lncRNA)(166) content of EVs as biomarker assays. Even though some studies obtained high accuracies in separating lung cancer from controls, they often used small sample cohorts and/or did not validate their findings in an independent validation cohort. Moreover, a lack in consistency of EV enrichment and RNA extraction methodologies are likely to yield divergent signatures and thus results(167). NanoString provides a standardized platform for the analysis

of several sample types, with the option of not only analyzing RNA, but also DNA and protein panels. We have previously published a proof-of-concept study which included the establishment of a workflow to assess EV-mRNA expression on the nCounter platform(140).

Using this workflow and the Human Immune V2 gene panel, we developed the early lung cancer detection signature presented in this study with only 600 uL of human plasma input. Our 16-gene signature had a PPV of 84.0% and 82.6% in the training and validation cohorts, respectively (**Table 3** and **Table 4**). In addition, the sensitivity and specificity of our signature was around 80% in both the training and validation setting. Of the 16 transcripts included in the signature, two have been previously reported in early lung cancer detection studies. Kim et al.(168) reported that the *S100A9* protein was secreted by metaplastic bronchial cells, but not by non-metaplastic cells in culture. Moreover, Lee et al.(169) found that *CCL5* and four other serum-based transcripts could be used to differentiate non-small cell lung cancer patients from healthy controls. Confirming our findings, *CCL5* was lower expressed in the lung cancer cohort compared to the control cohort.

Even though our signature yielded high accuracies in the training and validation cohorts, some samples were misclassified. In the training cohort five controls and four lung cancer patients were wrongly classified based on the algorithm-determined cutoff value of 0.5 (**Figure 2D**). The wrongly classified lung cancer samples in the training cohort contained only early stage samples, including one stage I, one stage II and two stage IIIA lung cancer samples. In addition, five lung cancer samples and three control samples from the validation cohort were misclassified (**Figure 3B**). Of the lung cancer samples, four were derived from early stage patients, including one stage I and three stage IIIA samples. The fifth sample came from a stage IV lung cancer patient. Interestingly, cancer stage did not correlate with signature score in those wrongly classified samples. Considering the observation that most misclassified lung cancer samples were derived from patients with stage III, future addition of a larger fraction of these samples may improve classifier performance.

An EV-mRNA lung cancer signature could also provide clinical utility as a complementary tool to LDCT screening programs, particularly to lower the proportion of false positives. In this study we had plasma samples available from three individuals in which a lung nodule had been previously detected. Follow-up of the herein classified individuals (**Figure 3B**), and future cohorts of individuals with nodules, will provide preliminary information on the added value of the lung cancer signature to detect lung cancer from benign nodules.

Limitations of the current study include the high prevalence of lung cancer in the explored training and validation cohorts. The proportion of lung cancer was much higher than the actual prevalence in the population, with 54.4% lung cancer in the training and 55.8% in the validation cohort. This probably causes an overestimation of the PPV, and further studies should reveal the performance of the gene signature in a large, prospective lung cancer screening trial. In addition, patients with other lung conditions (e.g. COPD, asthma, lung embolisms) and other types of cancer should be assessed to evaluate the accuracy of our signature in these populations. Finally, precipitation-based EV enrichment can yield contaminants such as

(lipo)proteins(167). Although the utility of the signature outweighs the origin of it, deciphering where(in) the mRNA is located could improve extraction and downstream analysis.

Supporting information

Supplementary Table 1. DE genes between the lung cancer- and control samples in the training cohort. *DE: differentially expressed; P_Cancer: p-value of the DE genes; FC: fold change.*

	Transcript	P_Cancer	FC_Cancer		Transcript	P_Cancer	FC_Cancer
1	IRF7	0,00074	1,6	25	PRKCD	0,019	1,9
2	CD14	0,0013	1,3	26	CD81	0,02	0,7
3	ITGAE	0,002	1,7	27	IFNGR1	0,021	0,6
4	MYD88	0,0025	2,4	28	LGALS3	0,022	1,9
5	ARG1	0,0027	1,8	29	PSMD7	0,022	1,5
6	FCGR2A	0,0033	1,8	30	CD164	0,022	1,3
7	G6PD	0,0037	1,9	31	CAMP	0,023	0,8
8	CTSG	0,0038	1,7	32	CCL5	0,023	-1,7
9	S100A8	0,0044	1,6	33	ATG5	0,025	1,3
10	ATG7	0,0044	1,3	34	CD74	0,026	0,8
11	IL18	0,0057	1,5	35	MAPKAPK2	0,027	0,6
12	ITGB2	0,0059	1,2	36	LITAF	0,029	1,2
13	MAP4K4	0,0062	1,9	37	PTPN6	0,032	1,6
14	BATF	0,0067	0,6	38	IL8	0,032	1
15	CR1	0,0071	1	39	BCL6	0,033	0,8
16	NCF4	0,0076	1,4	40	ICAM4	0,037	0,6
17	IFNAR2	0,0076	0,8	41	LCP2	0,041	1,5
18	STAT6	0,011	1,5	42	CLEC4E	0,042	0,8
19	CTSC	0,012	1,3	43	C1QB	0,044	1,3
20	CXCL10	0,012	0,9	44	HPRT1	0,044	1,3
21	PSMB5	0,013	0,6	45	GUSB	0,044	0,4
22	GPI	0,016	2,1	46	CD86	0,046	0,3
23	CD59	0,017	1,6	47	AHR	0,047	0,9
24	HLA-DMA	0,018	1,8	48	CCR1	0,048	0,3

Supplementary Table 2. Annotations of the 16-gene signature. *Log(NC): Log normalized counts; SD: standard deviation; FC: fold change.*

	Transcript	Accession	Mean Log(NC) Cancer	SD Log(NC) Cancer	Log(FC) Cancer vs Controls	Cancer	P-value	Annotations																					
								Adaptive Immune System	Apoptosis	Cell Adhesion	Chemokine Signaling	Cytokine Signaling	Hemostasis	Host-pathogen Interaction	Immunometabolism	Innate Immune System	Lymphocyte Activation	MHC Class I Antigen Presentation	MHC Class II Antigen Presentation	NF-kB Signaling	NLR signaling	Oxidative Stress	Phagocytosis and Degradation	T Cell Receptor Signaling	TNF Family Signaling	TLR Signaling	Transcriptional Regulation	Type I Interferon Signaling	Type II Interferon Signaling
1	S100A9	NM_002965.2	8.74	4.09	0.9	Upregulated	0.350	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	+	-	-	-
2	MYD88	NM_002468.3	4.81	2.87	2.4	Upregulated	0.003	+	-	-	-	+	-	+	-	+	+	+	-	+	+	-	-	-	-	+	-	-	-
3	LGALS3	NM_001177388.1	5.90	2.79	1.9	Upregulated	0.022	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-
4	CCL5	NM_002985.2	7.62	3.32	-1.7	Downregulated	0.023	-	-	-	+	+	-	+	-	-	+	-	-	-	+	+	-	-	+	+	-	-	-
5	GPI	NM_000175.2	4.97	3.26	2.1	Upregulated	0.016	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
6	CD14	NM_000591.2	1.26	1.97	1.3	Upregulated	0.001	+	+	-	-	-	-	+	-	+	-	+	-	+	-	-	+	-	-	+	-	-	-
7	IRF7	NM_001572.3	1.89	2.30	1.6	Upregulated	0.001	-	-	-	-	+	+	+	-	+	-	-	-	-	+	-	-	-	-	+	+	+	+
8	TAGAP	NM_054114.3	4.12	2.65	-0.5	Downregulated	0.570	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
9	B2M	NM_004048.2	11.73	2.33	-0.5	Downregulated	0.480	+	-	-	-	+	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	+
10	FCGR2A	NM_021642.3	2.60	2.69	1.8	Upregulated	0.003	-	-	-	-	-	-	+	-	+	-	-	-	-	-	-	+	-	-	-	-	-	-
11	MAP4K4	NM_004834.3	3.95	2.66	1.9	Upregulated	0.006	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-
12	HLA-DMA	NM_006120.3	4.03	2.82	1.8	Upregulated	0.018	-	-	+	-	-	-	+	-	-	-	-	+	-	-	-	+	-	-	-	-	-	-
13	HLA-B	NM_005514.6	4.25	3.43	-1.0	Downregulated	0.280	+	-	+	-	+	-	+	-	+	+	+	-	-	-	-	+	-	-	-	-	+	+
14	ARG1	NM_000045.2	2.43	2.79	1.8	Upregulated	0.003	-	-	-	-	-	-	+	+	+	-	-	-	-	-	+	-	-	-	-	-	-	-
15	CD3D	NM_000732.4	1.00	1.85	-1.2	Downregulated	0.069	+	-	-	-	-	-	+	-	-	+	-	-	-	-	-	-	+	-	-	-	-	-
16	EEF1G	NM_001404.4	11.30	2.23	0.1	Upregulated	0.860	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Acknowledgements
The investigators wish to thank the patients for kindly accepting to donate samples to this study. We also thank all the physicians who collaborated by providing clinical information. Finally, we would like to thank colleagues and collaborators for their constructive feedback on this manuscript.

Financial support
This project has received funding from a European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement ELBA No 765492.

Chapter 6. Discussion, conclusion and future perspectives

Discussion

Although significant progress has been made in the biological understanding, prevention, earlier detection and treatment of malignant tumors in the last decades, the global cancer burden is still growing due to an extended life-span and continuing risk-factor exposure(1). In consequence, cancer-related deaths are expected to increase up to 17 million in 2030(170). Late stage diagnosis, ineffective treatments and acquired therapy resistance are the main causes for this high mortality rate. The concept of personalized treatment, for example based on the assessment of targetable molecular markers, has transformed clinical practice for many cancer types(3). Several technical platforms have been introduced in clinical laboratories to characterize the genomic, transcriptomic and proteomic make-up of tumors. This genotype-directed therapy was found to improve patient survival(4), and is now part of daily clinical practice. Even though tumor tissue profiling can provide significant prognostic and predictive information, advanced lung cancer and other tumors are sometimes difficult or even impossible to biopsy. This results in a tissue quantity that is insufficient for downstream profiling. Moreover, repeated sampling to explore changes in the tumor tissue upon development of resistance or adverse events is not possible(50-52). Due to these obstacles, liquid biopsies have quickly gained acceptance in the clinical setting(50). Liquid biopsies can be derived from all body fluids and can be used to complement or overcome invasive tissue biopsies(50, 52-57). The advantages of analyzing a tumor through liquid, such as repeated sampling, the minimally invasive character and heterogeneous profiling, make them a hot topic in translational research. Importantly, liquid biopsies could also provide a source for early diagnostic- or screening biomarker assays. Circulating molecules, such as DNA, RNA or proteins, can be freely present within these liquids, or can be extracted and analyzed from circulating tumor cells (CTCs)(58-60), extracellular vesicles (EVs) (61-64) and tumor-educated platelets (TEPs)(65, 66).

Their introduction to the field has seen some highly promising results, including the use of ctDNA as a complement to, or replacement for, tumor tissue biopsies. However, major efforts are still needed to bridge the gap between basic research and clinical implementation of other liquid biopsy sources, such as TEPs and EVs. Moreover, currently used technical platforms for tumor characterization are not always suitable to analyze the low quantity and quality of tumor-derived material that can be found in a liquid biopsy(6). Even when these hurdles are overcome, the cost, technical necessities, continuous need for panel updates and complex data analysis that are required for some platforms are simply not feasible in the clinical laboratory setting. In conclusion, large-scale validation and clinical implementation of liquid biopsy-based biomarker assays requires a quick, easy-to-use, relatively cheap, flexible and standardized technical platform with low input requirements. The NanoString nCounter platform meets these demands, and all types of liquid-derived circulating molecules, including RNA, DNA

and proteins can be analyzed on this platform(96). Several studies developing or validating prognostic and predictive tissue biomarker assays on nCounter have been carried out and one of those, the Prosigna™ assay, received FDA approval for clinical use. Although the NanoString platform is designed to analyze FFPE tissue samples, efforts have also been made to explore the use of liquid biopsies on this system. This provoked NanoString to develop a "Low Input Kit", a specific kit that can be used to amplify the starting material. From that moment on, studies investigating the possibility of amplified liquid biopsy-based analysis on nCounter have been performed and results show that this platform provides a valuable tool for tumor characterization, especially in the clinical setting.

This thesis focused on validating the NanoString nCounter platform for the analysis of liquid biopsy sources to develop clinically relevant biomarker assays. The nCounter platform has never been used for routine mutation detection in liquid biopsies of cancer patients. Therefore, we initially focused on validating the use of ctDNA, one of the most advanced liquid biosources in terms of clinical applications, to detect tumor-derived mutations in liquid biopsies from cancer patients. Confirming previous studies(88, 171), ctDNA could be extracted from several body fluids, including plasma, serum, cerebrospinal fluid and ascites yielding sufficient input for a pre-amplification and down-stream analysis on this platform. A retrospective analysis of 70 patient samples using the 3D Single Nucleotide Variant (SNV) Solid Tumor Panel led to a highly concordant (98.9%) detection of mutations in *EGFR*, *KRAS*, *BRAF*, *PIK3CA* and *NRAS*, compared to results obtained by NGS or PNA-Q-PCR. In addition, a prospective analysis with 91 liquid biopsy samples yielded valid results for 90 samples. Besides the previously mentioned clinically relevant mutations, we also detected mutations in *TP53*, *NFE2L2*, *CTNNB1*, *ALK*, *FBXW7* and *PTEN*. Analysis of follow-up samples revealed that the nCounter system can also be used for semi-quantitative mutation detection to follow disease evolution and detect resistance mutations. These results confirmed that nCounter can indeed be used for routine mutation testing in liquid biopsies of cancer patients. The observation that nCounter can accurately detect mutations in samples with a concentration range as low as 0.5 – 2.5 ng total DNA, highlights the clinical utility of this methodology. In contrast, NGS platforms require substantially higher input material and analyses can fail due to the high sample quality requirements(68). Other advantages of nCounter include the considerably shorter turnaround time, simplicity of the platform and the straightforward data analysis(172). The main limitation of this platform is the inability to detect mutations that are not included in the panel, although the mutations included in the panel can be adapted at any time. In addition, some mutations have a very low prevalence, making it difficult to validate their detection in average-size validation studies. Based on concordant results with other methodologies and advantages of using the nCounter platform, this platform can be accurately implemented in clinical laboratories for routine mutation testing, especially when sample input is too low to perform NGS.

After successfully validating the use of ctDNA, we explored a different liquid biopsy source; EVs. EVs are released by most cell types and, depending on the cell they derived from, contain a heterogeneous cargo. Besides proteins and lipids, EVs contain nucleic acids that are protected from degradation by a lipid bilayer(69-71). EVs are thought to function as intercellular

communication tools, by transporting their cargo to other cells(69, 72, 73). The latter has boosted the research on these vesicles, especially regarding their use as biomarkers for cancer development and progression(74). Although promising findings have been reported, the biggest hurdle in the field of EVs is the lack of standardized protocols. Vesicle enrichment can be carried out using several strategies, and downstream nucleic acid extraction and analysis techniques are even more diverse. In general, EVs have been used to study their miRNA cargo and various studies have been published where EV-miRNAs were analyzed using nCounter(79, 80). However, the nCounter platform has never been tested for the analysis of plasma-derived EV-mRNA. Thus, our second aim was to perform a proof-of-concept study where we optimized a workflow for EV enrichment, EV-mRNA purification and subsequent gene expression analysis on nCounter to develop biomarker assays. Due to the fact that clinical laboratories are usually unable to meet the technical or time requirements for ultracentrifugation, we focused on a precipitation kit for EV enrichment(114). The methodology for EV-RNA extraction and pre-amplification was optimized, and resulted in a highly reproducible final workflow that can be performed in only three days. The developed workflow was then validated in plasma samples from cancer and control individuals to measure gene expression with the Human Immune V2 594-gene panel. We were able to define a gene signature that could distinguish cancer patients from controls using a classifier algorithm. A key finding was the binding of EV-DNA to our mRNA panel, due to a non-intron spanning probe design. This required the addition of a DNase step, which significantly altered gene expression profiles but led to a higher performance of the classifier. A limitation of this study is the divergent cancer patient cohort which cannot be used to create a clinically useful gene signature. However, results of this proof-of-concept study confirm that we can use the nCounter platform for EV-mRNA analysis, and led us to implement this workflow in other projects for clinically relevant biomarker assay development.

The use of ICIs has changed advanced cancer treatment significantly, but has also revealed new unmet clinical needs. While 20-30% of ICI treated patients have impressive and durable responses, other patients do not benefit from this type of treatment(134). PD-L1 expression is still the gold-standard for ICI response prediction and patient inclusion in clinical trials. However, this assay lacks a standardized methodology and predictive performance is suboptimal(30). More effective predictive biomarkers are needed and the complex interaction between the tumor, microenvironment and immune system may require the assessment of multiple biomarkers at the same time(16). Irrespective of response, ICI treatment can cause an overactivation of the immune system and lead to development of immune related adverse events (irAEs)(47). Although not all irAEs are detrimental, checkpoint inhibitor pneumonitis (CIP) comprises an inflammation of the lung tissue that can be lethal if not detected and treated in time. Moreover, diagnosis of CIP is still a major problem due to overlapping characteristics with other lung complications(128, 143). In agreement with our observations in the clinic, two retrospective analyses have found a 20% CIP incidence in patients treated with ICIs(48, 136). Since there are no available predictive or diagnostic tests and risk factors remain unclear, CIP development is a big concern in daily clinical practice. We hypothesized that a chronic lung inflammation or an imbalance in the immune system prior to ICI administration could be indicative of higher CIP risk, providing a window for intensive monitoring of those patients.

To this end, we aimed to develop a predictive biomarker assay for CIP development using the nCounter platform with an immune-oncology (IO360) panel comprising 770 gene transcripts. Retrospectively collected pre-treatment FFPE tumor tissues from CIP and non-CIP patients were analyzed, and a bioinformatic algorithm was used to select a gene signature that could distinguish between the two cohorts. The optimal classifier had an accuracy of 86.6% and negative predictive value (NPV) of 92% and consisted of eight genes, including *CDKN1A, BID, ULBP2, HDAC11, ARID1A, IGF2R, TWIST1* and *IER3*. Only one of those genes, *IGF2R*, had been previously linked to radiation-induced lung injury(145).

Considering our previous EV proof-of-concept results, and based on the advantages of an EV-based biomarker assay, we then analyzed pre-treatment plasma-derived EV samples for the same purpose. Interestingly and as previously reported for proteins(173), the gene expression profiles between tumor tissue and EVs differed, and the same 8-gene tissue CIP signature could not accurately distinguish CIP from non-CIP EVs. Development of a new EV-specific signature yielded a classifier with a slightly increased accuracy and NPV of 87.0% and 92.7%, respectively, based on the expression of *LYZ, CDH1, CD300A* and *SPIB*. Some of those genes were reported to be involved in other pathologies characterized by an overactivation of the immune system, such as sarcoidosis(147) or a maternal immune response against her fetus(149). Besides a more complete readout of ongoing biological processes, an EV-based assay can also be used for follow-up analysis to determine gene expression changes upon actual CIP development. We found that the 4-gene signature score could not only predict CIP development before ICI administration, but increased upon actual CIP development in follow-up samples of two patients. In contrast, in non-CIP patients the signature scores remained below the cutoff during follow-up. In conclusion, we can use both tumor tissue- and EV-mRNA to predict CIP on the nCounter platform. The possibility for follow-up analysis and their minimally invasive character emphasize the use of EVs in a predictive CIP test. In addition, the short turnaround time of nCounter allows for rapid clinical decision making in patients with high grade CIP and could potentially improve patient outcome and quality of life. Two limitations of this study should be taken into account. Firstly, an independently collected validation cohort that resembles CIP prevalence should be assessed to control for overfitting of our model. A prospective validation study is currently being carried out. Secondly, additional patient cohorts including hyper-immune activated pathologies should be evaluated to reveal if the developed signature is CIP-specific.

Predicting therapy response or adverse events could have a major clinical impact, but earlier detection of cancer is still seen as the holy grail for patient survival. Low-dose computed tomography (LDCT) is the only recommended lung cancer screening test, even though the NLST lung cancer screening trial has showed that out of all abnormal scans 96.4% were false positives(152, 163). A liquid biopsy-based test could be used to either complement, to lower the fraction of false positive results, or replace these imaging screening tests. However, a standardized platform and a test with a high accuracy are essential. We used our developed EV-mRNA workflow to explore if it could also be used for early- and late-stage lung cancer detection. From all 592 genes included in the Human Immune V2 Panel, 16 genes were selected by a machine learning algorithm due to optimal classification performance (81-84% accuracy).

These results were then validated in a separate patient- and control cohort, of which the blood samples were collected and processed in a different center. These results highlight the potential of using plasma EVs as a read-out for biological processes. Although the clinical necessity of all the developed signatures outweighs their origin, it is important to keep in mind that precipitation-based EV enrichment kits yield contaminants. Therefore, it cannot be ruled out that the signatures are (in part) derived from other fractions that can be detected in plasma, such as RNA-binding proteins or lipoproteins(167). Further research is needed to decipher where the mRNA transcripts were stored or what they were attached to in the plasma, to improve their extraction and downstream analysis.

General conclusion

In conclusion, this thesis has provided evidence that the nCounter platform can be used for the analysis of liquid biopsies in the clinical setting, specifically for ctDNA and EV-mRNA. Clinically relevant biomarker assays can be developed and validated on this standardized platform and can aid in the transition of liquid-derived biomarker assays from bench to bedside.

Future perspectives

The background of my desktop during my PhD contains the words of George Bernard Shaw: “Science never solves a problem without creating ten more”. This quote can be read in two ways; a scientific finding never ends, or – my personal interpretation - there is always more to understand. This thesis has provided a rationale for the use of nCounter to analyze liquid biopsies in the clinical setting, but also opened new questions.

Firstly, further exploration is needed to understand where the liquid biopsy derived signals are coming from. Especially in the case of EV-RNA, precipitation kits are known to capture all particles present in a plasma sample, including lipids, platelet-derived vesicles and proteins. It would be of great interest to know which type of particles are present in our samples, where those particles derived from, and if they all contain RNA.

Secondly, a biomarker assay requires large retrospective and prospective cross-validation studies before implementation in diagnostic laboratories. We are currently collecting an independent validation cohort to validate the developed CIP-signature. This cohort will include CIP and non-CIP patients, but we also aim to obtain plasma samples from patients with other pathologies where a hyper-immune activation occurs. Moreover, we have opened a new study to prospectively analyze samples from patients starting on ICI treatment, to determine CIP risk. If our 4-gene CIP signature can be validated in retrospective and prospective studies, a Prosigna-like kit could be developed for clinical use, and the panel may be expanded by adding signatures indicative of ICI treatment response.

Thirdly, although NGS platforms are not always easy to implement in clinical practice, they provide a better tool for biomarker discovery, compared to the NanoString system. Once the signatures have been detected using NGS, custom panels can be created for nCounter analyses,

allowing an easy workflow from biomarker discovery to validation and implementation in clinical studies on a standardized platform.

Finally, we have provided evidence for the use of multi-fluid derived ctDNA and plasma-derived EV-mRNA on nCounter. However, many more liquids (e.g. urine, saliva), biosources (e.g. tumor educated platelets, circulating tumor cells, white blood cells) and circulating molecules (e.g. cfRNA, circular RNA, long non-coding RNA, proteins) require similar validation studies. In addition, multi-omics approaches are currently emerging and can also be carried out on the nCounter platform, by combining DNA, RNA and protein panels.

Chapter 7. Bibliography

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.
2. Hofmarcher T, Lindgren P, Wilking N, Jonsson B. The cost of cancer in Europe 2018. Eur J Cancer. 2020;129:41-9.
3. Rosell R, Karachaliou N. Large-scale screening for somatic mutations in lung cancer. Lancet. 2016;387(10026):1354-6.
4. Kris MG, Johnson BE, Berry LD, Kwiatkowski DJ, Iafrate AJ, Wistuba, II, et al. Using multiplexed assays of oncogenic drivers in lung cancers to select targeted drugs. JAMA. 2014;311(19):1998-2006.
5. Malone ER, Oliva M, Sabatini PJB, Stockley TL, Siu LL. Molecular profiling for precision cancer therapies. Genome Med. 2020;12(1):8.
6. Teixido C, Gimenez-Capitan A, Molina-Vila MA, Peg V, Karachaliou N, Rodriguez-Capote A, et al. RNA Analysis as a Tool to Determine Clinically Relevant Gene Fusions and Splice Variants. Arch Pathol Lab Med. 2018;142(4):474-9.
7. Shendure J, Balasubramanian S, Church GM, Gilbert W, Rogers J, Schloss JA, et al. DNA sequencing at 40: past, present and future. Nature. 2017;550(7676):345-53.
8. Segundo-Val IS, Sanz-Lozano CS. Introduction to the Gene Expression Analysis. Methods Mol Biol. 2016;1434:29-43.
9. Prokopec SD, Watson JD, Waggott DM, Smith AB, Wu AH, Okey AB, et al. Systematic evaluation of medium-throughput mRNA abundance platforms. RNA. 2013;19(1):51-62.
10. Fortina P, Surrey S. Digital mRNA profiling. Nat Biotechnol. 2008;26(3):293-4.
11. Geiss GK, Bumgarner RE, Birditt B, Dahl T, Dowidar N, Dunaway DL, et al. Direct multiplexed measurement of gene expression with color-coded probe pairs. Nat Biotechnol. 2008;26(3):317-25.
12. Richard AC, Lyons PA, Peters JE, Biasci D, Flint SM, Lee JC, et al. Comparison of gene expression microarray data with count-based RNA measurements informs microarray interpretation. BMC Genomics. 2014;15:649.
13. Speranza E, Altamura LA, Kulcsar K, Bixler SL, Rossi CA, Schoepp RJ, et al. Comparison of Transcriptomic Platforms for Analysis of Whole Blood from Ebola-Infected Cynomolgus Macaques. Sci Rep. 2017;7(1):14756.
14. Goossens N, Nakagawa S, Sun X, Hoshida Y. Cancer biomarker discovery and validation. Transl Cancer Res. 2015;4(3):256-69.
15. In: Micheel CM, Nass SJ, Omenn GS, editors. Evolution of Translational Omics: Lessons Learned and the Path Forward. Washington (DC)2012.
16. Cesano A, Warren S. Bringing the next Generation of Immuno-Oncology Biomarkers to the Clinic. Biomedicines. 2018;6(1).
17. Veldman-Jones MH, Lai Z, Wappett M, Harbron CG, Barrett JC, Harrington EA, et al. Reproducible, Quantitative, and Flexible Molecular Subtyping of Clinical DLBCL Samples Using the NanoString nCounter System. Clin Cancer Res. 2015;21(10):2367-78.
18. Nielsen T, Wallden B, Schaper C, Ferree S, Liu S, Gao D, et al. Analytical validation of the PAM50-based Prosigna Breast Cancer Prognostic Gene Signature Assay and nCounter Analysis System using formalin-fixed paraffin-embedded breast tumor specimens. BMC Cancer. 2014;14:177.
19. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol. 2009;27(8):1160-7.
20. Wallden B, Storhoff J, Nielsen T, Dowidar N, Schaper C, Ferree S, et al. Development and verification of the PAM50-based Prosigna breast cancer gene signature assay. BMC Med Genomics. 2015;8:54.
21. Dowsett M, Sestak I, Lopez-Knowles E, Sidhu K, Dunbier AK, Cowens JW, et al. Comparison of PAM50 risk of recurrence score with oncotype DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. J Clin Oncol. 2013;31(22):2783-90.
22. Gnant M, Filipits M, Greil R, Stoeger H, Rudas M, Bago-Horvath Z, et al. Predicting distant recurrence in receptor-positive breast cancer patients with limited clinicopathological risk: using the PAM50 Risk of Recurrence score in 1478 postmenopausal patients of the ABCSG-8 trial treated with adjuvant endocrine therapy alone. Ann Oncol. 2014;25(2):339-45.
23. Hannouf MB, Zaric GS, Blanchette P, Brezden-Masley C, Paulden M, McCabe C, et al. Cost-effectiveness analysis of multigene expression profiling assays to guide adjuvant therapy decisions in women with invasive early-stage breast cancer. Pharmacogenomics J. 2020;20(1):27-46.

24. Reguart N, Teixido C, Gimenez-Capitan A, Pare L, Galvan P, Viteri S, et al. Identification of ALK, ROS1, and RET Fusions by a Multiplexed mRNA-Based Assay in Formalin-Fixed, Paraffin-Embedded Samples from Advanced Non-Small-Cell Lung Cancer Patients. Clin Chem. 2017;63(3):751-60.
25. Ross PM, Lee J, McKay-Fleisch J, Vellano CP, Garber J, Eterovic AK, et al., editors. Multiplex profiling of cancer driver mutations: Detection of single nucleotide variants and small InDels from small amounts of FFPE samples on the nCounter® analysis system. 66th Annual Meeting of The American Society of Human Genetics; 2016; Vancouver, Canada.
26. Meredith G, Kargl J, McKay-Fleisch J, Ross M, Garber J, Ong G, et al., editors. Multiplex detection of driver mutations in lung cancer: Simultaneous assay of single nucleotide variants (SNV) and fusion-transcripts from small amounts of FFPE samples on the nCounter® analysis system. The 66th Annual Meeting of The American Society of Human Genetics; 2016; Vancouver, Canada.
27. Topalian SL, Taube JM, Pardoll DM. Neoadjuvant checkpoint blockade for cancer immunotherapy. Science. 2020;367(6477).
28. Haslam A, Prasad V. Estimation of the Percentage of US Patients With Cancer Who Are Eligible for and Respond to Checkpoint Inhibitor Immunotherapy Drugs. JAMA Netw Open. 2019;2(5):e192535.
29. Duffy MJ, Crown J. Biomarkers for Predicting Response to Immunotherapy with Immune Checkpoint Inhibitors in Cancer Patients. Clin Chem. 2019;65(10):1228-38.
30. Gibney GT, Weiner LM, Atkins MB. Predictive biomarkers for checkpoint inhibitor-based immunotherapy. Lancet Oncol. 2016;17(12):e542-e51.
31. Bonaventura P, Shekarian T, Alcazer V, Valladeau-Guilemond J, Valsesia-Wittmann S, Amigorena S, et al. Cold Tumors: A Therapeutic Challenge for Immunotherapy. Front Immunol. 2019;10:168.
32. Danaher P, Warren S, Dennis L, D'Amico L, White A, Disis ML, et al. Gene expression markers of Tumor Infiltrating Leukocytes. J Immunother Cancer. 2017;5:18.
33. Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR, et al. IFN-gamma-related mRNA profile predicts clinical response to PD-1 blockade. J Clin Invest. 2017;127(8):2930-40.
34. Danaher P, Warren S, Lu R, Samayoa J, Sullivan A, Pekker I, et al. Pan-cancer adaptive immune resistance as defined by the Tumor Inflammation Signature (TIS): results from The Cancer Genome Atlas (TCGA). J Immunother Cancer. 2018;6(1):63.
35. Blank CU, Rozeman EA, Fanchi LF, Sikorska K, van de Wiel B, Kvistborg P, et al. Neoadjuvant versus adjuvant ipilimumab plus nivolumab in macroscopic stage III melanoma. Nat Med. 2018;24(11):1655-61.
36. Ott PA, Bang YJ, Piha-Paul SA, Razak ARA, Bennouna J, Soria JC, et al. T-Cell-Inflamed Gene-Expression Profile, Programmed Death Ligand 1 Expression, and Tumor Mutational Burden Predict Efficacy in Patients Treated With Pembrolizumab Across 20 Cancers: KEYNOTE-028. J Clin Oncol. 2019;37(4):318-27.
37. Pare L, Pascual T, Segui E, Teixido C, Gonzalez-Cao M, Galvan P, et al. Association between PD1 mRNA and response to anti-PD1 monotherapy across multiple cancer types. Ann Oncol. 2018;29(10):2121-8.
38. Cristescu R, Mogg R, Ayers M, Albright A, Murphy E, Yearley J, et al. Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. Science. 2018;362(6411).
39. Sivendran S, Chang R, Pham L, Phelps RG, Harcharik ST, Hall LD, et al. Dissection of immune gene networks in primary melanoma tumors critical for antitumor surveillance of patients with stage II-III resectable disease. J Invest Dermatol. 2014;134(8):2202-11.
40. Chen PL, Roh W, Reuben A, Cooper ZA, Spencer CN, Prieto PA, et al. Analysis of Immune Signatures in Longitudinal Tumor Samples Yields Insight into Biomarkers of Response and Mechanisms of Resistance to Immune Checkpoint Blockade. Cancer Discov. 2016;6(8):827-37.
41. Voorwerk L, Slagter M, Horlings HM, Sikorska K, van de Vijver KK, de Maaker M, et al. Immune induction strategies in metastatic triple-negative breast cancer to enhance the sensitivity to PD-1 blockade: the TONIC trial. Nat Med. 2019;25(6):920-8.
42. Sundar R, Huang KK, Qamra A, Kim KM, Kim ST, Kang WK, et al. Epigenomic promoter alterations predict for benefit from immune checkpoint inhibition in metastatic gastric cancer. Ann Oncol. 2019;30(3):424-30.
43. Schalper KA, Rodriguez-Ruiz ME, Diez-Valle R, Lopez-Janeiro A, Porciuncula A, Idoate MA, et al. Neoadjuvant nivolumab modifies the tumor immune microenvironment in resectable glioblastoma. Nat Med. 2019;25(3):470-6.
44. Cloughesy TF, Mochizuki AY, Orpilla JR, Hugo W, Lee AH, Davidson TB, et al. Neoadjuvant anti-PD-1 immunotherapy promotes a survival benefit with intratumoral and systemic immune responses in recurrent glioblastoma. Nat Med. 2019;25(3):477-86.
45. Prat A, Navarro A, Pare L, Reguart N, Galvan P, Pascual T, et al. Immune-Related Gene Expression Profiling After PD-1 Blockade in Non-Small Cell Lung Carcinoma, Head and Neck Squamous Cell Carcinoma, and Melanoma. Cancer Res. 2017;77(13):3540-50.

46. Gonzalez-Cao M, Moran T, Dalmau J, Garcia-Corbacho J, Bracht JWP, Bernabe R, et al. Assessment of the Feasibility and Safety of Durvalumab for Treatment of Solid Tumors in Patients With HIV-1 Infection: The Phase 2 DURVAST Study. JAMA Oncol. 2020.
47. Johnson DB, Balko JM. Biomarkers for Immunotherapy Toxicity: Are Cytokines the Answer? Clin Cancer Res. 2019;25(5):1452-4.
48. Suresh K, Voong KR, Shankar B, Forde PM, Ettinger DS, Marrone KA, et al. Pneumonitis in Non-Small Cell Lung Cancer Patients Receiving Immune Checkpoint Immunotherapy: Incidence and Risk Factors. J Thorac Oncol. 2018;13(12):1930-9.
49. Suresh K, Psoter KJ, Voong KR, Shankar B, Forde PM, Ettinger DS, et al. Impact of Checkpoint Inhibitor Pneumonitis on Survival in NSCLC Patients Receiving Immune Checkpoint Immunotherapy. J Thorac Oncol. 2019;14(3):494-502.
50. Bracht JWP, Mayo-de-Las-Casas C, Berenguer J, Karachaliou N, Rosell R. The Present and Future of Liquid Biopsies in Non-Small Cell Lung Cancer: Combining Four Biosources for Diagnosis, Prognosis, Prediction, and Disease Monitoring. Curr Oncol Rep. 2018;20(9):70.
51. Crowley E, Di Nicolantonio F, Loupakis F, Bardelli A. Liquid biopsy: monitoring cancer-genetics in the blood. Nat Rev Clin Oncol. 2013;10(8):472-84.
52. Brock G, Castellanos-Rizaldos E, Hu L, Coticchia C, Skog J. Liquid biopsy for cancer screening, patient stratification and monitoring. Translational Cancer Research. 2015;4(3):280-90.
53. Cai X, Janku F, Zhan Q, Fan JB. Accessing Genetic Information with Liquid Biopsies. Trends Genet. 2015;31(10):564-75.
54. Molina-Vila MA, Mayo-de-Las-Casas C, Gimenez-Capitan A, Jordana-Ariza N, Garzon M, Balada A, et al. Liquid Biopsy in Non-Small Cell Lung Cancer. Front Med (Lausanne). 2016;3:69.
55. Karachaliou N, Mayo-de-Las-Casas C, Molina-Vila MA, Rosell R. Real-time liquid biopsies become a reality in cancer treatment. Ann Transl Med. 2015;3(3):36.
56. Rolfo C, Castiglia M, Hong D, Alessandro R, Mertens I, Baggerman G, et al. Liquid biopsies in lung cancer: the new ambrosia of researchers. Biochim Biophys Acta. 2014;1846(2):539-46.
57. Parikh AR, Leshchiner I, Elagina L, Goyal L, Levovitz C, Siravegna G, et al. Liquid versus tissue biopsy for detecting acquired resistance and tumor heterogeneity in gastrointestinal cancers. Nat Med. 2019;25(9):1415-21.
58. Beck TN, Boumber YA, Aggarwal C, Pei J, Thrash-Bingham C, Fittipaldi P, et al. Circulating tumor cell and cell-free RNA capture and expression analysis identify platelet-associated genes in metastatic lung cancer. BMC Cancer. 2019;19(1):603.
59. Mondelo-Macia P, Rodriguez-Lopez C, Valina L, Aguin S, Leon-Mateos L, Garcia-Gonzalez J, et al. Detection of MET Alterations Using Cell Free DNA and Circulating Tumor Cells from Cancer Patients. Cells. 2020;9(2).
60. Kelley SO, Pantel K. A New Era in Liquid Biopsy: From Genotype to Phenotype. Clin Chem. 2019.
61. Kamyabi N, Abbasgholizadeh R, Maitra A, Ardekani A, Biswal SL, Grande-Allen KJ. Isolation and mutational assessment of pancreatic cancer extracellular vesicles using a microfluidic platform. Biomed Microdevices. 2020;22(2):23.
62. San Lucas FA, Allenson K, Bernard V, Castillo J, Kim DU, Ellis K, et al. Minimally invasive genomic and transcriptomic profiling of visceral cancers by next-generation sequencing of circulating exosomes. Ann Oncol. 2016;27(4):635-41.
63. Stevic I, Buescher G, Ricklefs FL. Monitoring Therapy Efficiency in Cancer through Extracellular Vesicles. Cells. 2020;9(1).
64. Reclusa P, Taverna S, Pucci M, Durendez E, Calabuig S, Manca P, et al. Exosomes as diagnostic and predictive biomarkers in lung cancer. J Thorac Dis. 2017;9(Suppl 13):S1373-S82.
65. Best MG, Sol N, In 't Veld S, Vancura A, Muller M, Niemeijer AN, et al. Swarm Intelligence-Enhanced Detection of Non-Small-Cell Lung Cancer Using Tumor-Educated Platelets. Cancer Cell. 2017;32(2):238-52 e9.
66. Best MG, Sol N, Kooi I, Tannous J, Westerman BA, Rustenburg F, et al. RNA-Seq of Tumor-Educated Platelets Enables Blood-Based Pan-Cancer, Multiclass, and Molecular Pathway Cancer Diagnostics. Cancer Cell. 2015;28(5):666-76.
67. Leighl NB, Page RD, Raymond VM, Daniel DB, Divers SG, Reckamp KL, et al. Clinical Utility of Comprehensive Cell-free DNA Analysis to Identify Genomic Biomarkers in Patients with Newly Diagnosed Metastatic Non-small Cell Lung Cancer. Clin Cancer Res. 2019;25(15):4691-700.
68. Alborelli I, Generali D, Jermann P, Cappelletti MR, Ferrero G, Scaggiante B, et al. Cell-free DNA analysis in healthy individuals by next-generation sequencing: a proof of concept and technical validation study. Cell Death Dis. 2019;10(7):534.

69. Margolis L, Sadovsky Y. The biology of extracellular vesicles: The known unknowns. PLoS Biol. 2019;17(7):e3000363.
70. Reclusa P, Sirera R, Araujo A, Giallombardo M, Valentino A, Sorber L, et al. Exosomes genetic cargo in lung cancer: a truly Pandora's box. Transl Lung Cancer Res. 2016;5(5):483-91.
71. Jeppesen DK, Fenix AM, Franklin JL, Higginbotham JN, Zhang Q, Zimmerman LJ, et al. Reassessment of Exosome Composition. Cell. 2019;177(2):428-45 e18.
72. Fujita Y, Yoshioka Y, Ochiya T. Extracellular vesicle transfer of cancer pathogenic components. Cancer Sci. 2016;107(4):385-90.
73. Kalluri R. The biology and function of exosomes in cancer. J Clin Invest. 2016;126(4):1208-15.
74. Nolte-'t Hoen EN, Buermans HP, Waasdorp M, Stoorvogel W, Wauben MH, t Hoen PA. Deep sequencing of RNA from immune cell-derived vesicles uncovers the selective incorporation of small non-coding RNA biotypes with potential regulatory functions. Nucleic Acids Res. 2012;40(18):9272-85.
75. Kossenkov AV, Qureshi R, Dawany NB, Wickramasinghe J, Liu Q, Majumdar RS, et al. A Gene Expression Classifier from Whole Blood Distinguishes Benign from Malignant Lung Nodules Detected by Low-Dose CT. Cancer Res. 2019;79(1):263-73.
76. Khodadadi-Jamayran A, Akgol-Oksuz B, Afanasyeva Y, Heguy A, Thompson M, Ray K, et al. Prognostic role of elevated mir-24-3p in breast cancer and its association with the metastatic process. Oncotarget. 2018;9(16):12868-78.
77. Porras TB, Kaur P, Ring A, Schechter N, Lang JE. Challenges in using liquid biopsies for gene expression profiling. Oncotarget. 2018;9(6):7036-53.
78. Wu TC, Xu K, Martinek J, Young RR, Banchereau R, George J, et al. IL1 Receptor Antagonist Controls Transcriptional Signature of Inflammation in Patients with Metastatic Breast Cancer. Cancer Res. 2018;78(18):5243-58.
79. Garcia-Contreras M, Shah SH, Tamayo A, Robbins PD, Golberg RB, Mendez AJ, et al. Plasma-derived exosome characterization reveals a distinct microRNA signature in long duration Type 1 diabetes. Sci Rep. 2017;7(1):5998.
80. Vicentini C, Calore F, Nigita G, Fadda P, Simbolo M, Sperandio N, et al. Exosomal miRNA signatures of pancreatic lesions. BMC Gastroenterol. 2020;20(1):137.
81. Hanson D, Warren S, Meredith G, Eagan M, Dennis L, Krouse M, et al. Poster: Digital Gene Expression from Low Sample Input: Highly multiplexed and robust profiling of formalin-fixed paraffin-embedded (FFPE) and fresh frozen samples from as little as 1 ng of RNA using the nCounter® Platform. 2017 (Link: file:///Users/test/Downloads/20171003_Low_Input_ASHG_poster_draft5%20(1).pdf, accessed 15/12/2020).
82. Bedin C, Enzo MV, Del Bianco P, Pucciarelli S, Nitti D, Agostini M. Diagnostic and prognostic role of cell-free DNA testing for colorectal cancer patients. Int J Cancer. 2017;140(8):1888-98.
83. de Wit S, Rossi E, Weber S, Tamminga M, Manicone M, Swennenhuis JF, et al. Single tube liquid biopsy for advanced non-small cell lung cancer. Int J Cancer. 2019;144(12):3127-37.
84. Goldman JW, Noor ZS, Remon J, Besse B, Rosenfeld N. Are liquid biopsies a surrogate for tissue EGFR testing? Ann Oncol. 2018;29(suppl_1):i38-i46.
85. Mayo-de-Las-Casas C, Jordana-Ariza N, Garzon-Ibanez M, Balada-Bel A, Bertran-Alamillo J, Viteri-Ramirez S, et al. Large scale, prospective screening of EGFR mutations in the blood of advanced NSCLC patients to guide treatment decisions. Ann Oncol. 2017;28(9):2248-55.
86. Mayo-de-Las-Casas C, Garzon Ibanez M, Jordana-Ariza N, Garcia-Pelaez B, Balada-Bel A, Villatoro S, et al. An update on liquid biopsy analysis for diagnostic and monitoring applications in non-small cell lung cancer. Expert Rev Mol Diagn. 2018;18(1):35-45.
87. Rolfo C, Mack PC, Scagliotti GV, Baas P, Barlesi F, Bivona TG, et al. Liquid Biopsy for Advanced Non-Small Cell Lung Cancer (NSCLC): A Statement Paper from the IASLC. J Thorac Oncol. 2018;13(9):1248-68.
88. Villatoro S, Mayo-de-Las-Casas C, Jordana-Ariza N, Viteri-Ramirez S, Garzon-Ibanez M, Moya-Horno I, et al. Prospective detection of mutations in cerebrospinal fluid, pleural effusion, and ascites of advanced cancer patients to guide treatment decisions. Mol Oncol. 2019;13(12):2633-45.
89. Gonzalez-Cao M, Mayo-de-Las-Casas C, Molina-Vila MA, De Mattos-Arruda L, Munoz-Couselo E, Manzano JL, et al. BRAF mutation analysis in circulating free tumor DNA of melanoma patients treated with BRAF inhibitors. Melanoma Res. 2015;25(6):486-95.
90. Szallasi Z. Detecting mutant KRAS in liquid biopsies: a biomarker searching for a role. Ann Oncol. 2017;28(4):677-8.
91. Prat A, Galvan P, Jimenez B, Buckingham W, Jeiranian HA, Schaper C, et al. Prediction of Response to Neoadjuvant Chemotherapy Using Core Needle Biopsy Samples with the Prosigna Assay. Clin Cancer Res. 2016;22(3):560-6.

92. Warren S. Simultaneous, Multiplexed Detection of RNA and Protein on the NanoString((R)) nCounter((R)) Platform. Methods Mol Biol. 2018;1783:105-20.
93. Malapelle U, Mayo de-Las-Casas C, Rocco D, Garzon M, Pisapia P, Jordana-Ariza N, et al. Development of a gene panel for next-generation sequencing of clinically relevant mutations in cell-free DNA from cancer patients. Br J Cancer. 2017;116(6):802-10.
94. Mayo-de-Las-Casas C, Velasco A, Sanchez D, Martinez-Bueno A, Garzon-Ibanez M, Gatius S, et al. Detection of somatic mutations in peritoneal lavages and plasma of endometrial cancer patients: A proof-of-concept study. Int J Cancer. 2020.
95. Ou SI, Nagasaka M, Zhu VW. Liquid Biopsy to Identify Actionable Genomic Alterations. Am Soc Clin Oncol Educ Book. 2018;38:978-97.
96. Kulkarni MM. Digital multiplexed gene expression analysis using the NanoString nCounter system. Curr Protoc Mol Biol. 2011;Chapter 25:Unit25B 10.
97. Goytain A, Ng T. NanoString nCounter Technology: High-Throughput RNA Validation. Methods Mol Biol. 2020;2079:125-39.
98. Veldman-Jones MH, Brant R, Rooney C, Geh C, Emery H, Harbron CG, et al. Evaluating Robustness and Sensitivity of the NanoString Technologies nCounter Platform to Enable Multiplexed Gene Expression Analysis of Clinical Samples. Cancer Res. 2015;75(13):2587-93.
99. Fan K, Ritter C, Nghiem P, Blom A, Verhaegen ME, Dlugosz A, et al. Circulating Cell-Free miR-375 as Surrogate Marker of Tumor Burden in Merkel Cell Carcinoma. Clin Cancer Res. 2018;24(23):5873-82.
100. Latchana N, DiVincenzo MJ, Regan K, Abrams Z, Zhang X, Jacob NK, et al. Alterations in patient plasma microRNA expression profiles following resection of metastatic melanoma. J Surg Oncol. 2018;118(3):501-9.
101. Warren S, Danaher P, Mashadi-Hossein A, Skewis L, Wallden B, Ferree S, et al. Development of Gene Expression-Based Biomarkers on the nCounter((R)) Platform for Immuno-Oncology Applications. Methods Mol Biol. 2020;2055:273-300.
102. Jan YJ, Yoon J, Chen JF, Teng PC, Yao N, Cheng S, et al. A Circulating Tumor Cell-RNA Assay for Assessment of Androgen Receptor Signaling Inhibitor Sensitivity in Metastatic Castration-Resistant Prostate Cancer. Theranostics. 2019;9(10):2812-26.
103. Tolaney SM, Guo H, Pernas S, Barry WT, Dillon DA, Ritterhouse L, et al. Seven-Year Follow-Up Analysis of Adjuvant Paclitaxel and Trastuzumab Trial for Node-Negative, Human Epidermal Growth Factor Receptor 2-Positive Breast Cancer. J Clin Oncol. 2019;37(22):1868-75.
104. Leite de Oliveira R, Bernards R. Anti-cancer therapy: senescence is the new black. EMBO J. 2018;37(10).
105. Momtaz P, Pentsova E, Abdel-Wahab O, Diamond E, Hyman D, Merghoub T, et al. Quantification of tumor-derived cell free DNA(cfDNA) by digital PCR (DigPCR) in cerebrospinal fluid of patients with BRAFV600 mutated malignancies. Oncotarget. 2016;7(51):85430-6.
106. Narrandes S, Xu W. Gene Expression Detection Assay for Cancer Clinical Use. J Cancer. 2018;9(13):2249-65.
107. Giménez-Capitán A, Bracht J, García-Mosquera J, Jordana-Ariza N, García B, Garzón-Ibañez M, et al. Multiplex detection of clinically relevant mutations in liquid biopsies of cancer patients using the nCounter Platform. Clinical Chemistry. 2020, In press.
108. Bracht JWP, Karachaliou N, Berenguer J, Pedraz-Valdunciel C, Filipska M, Codony-Servat C, et al. Osimertinib and pterostilbene in EGFR-mutation-positive non-small cell lung cancer (NSCLC). Int J Biol Sci. 2019;15(12):2607-14.
109. Puente-Massaguer E, Lecina M, Godia F. Nanoscale characterization coupled to multi-parametric optimization of Hi5 cell transient gene expression. Appl Microbiol Biotechnol. 2018;102(24):10495-510.
110. Waggott D, Chu K, Yin S, Wouters BG, Liu FF, Boutros PC. NanoStringNorm: an extensible R package for the pre-processing of NanoString mRNA and miRNA data. Bioinformatics. 2012;28(11):1546-8.
111. Emelyanov A, Shtam T, Kamyshinsky R, Garaeva L, Verlov N, Miliukhina I, et al. Cryo-electron microscopy of extracellular vesicles from cerebrospinal fluid. PLoS One. 2020;15(1):e0227949.
112. Rasch MR, Rossinyol E, Hueso JL, Goodfellow BW, Arbiol J, Korgel BA. Hydrophobic gold nanoparticle self-assembly with phosphatidylcholine lipid: membrane-loaded and janus vesicles. Nano Lett. 2010;10(9):3733-9.
113. Pasquali L, Svedbom A, Srivastava A, Rosen E, Lindqvist U, Stahle M, et al. Circulating microRNAs in extracellular vesicles as potential biomarkers for psoriatic arthritis in patients with psoriasis. J Eur Acad Dermatol Venereol. 2020;34(6):1248-56.
114. Helwa I, Cai J, Drewry MD, Zimmerman A, Dinkins MB, Khaled ML, et al. A Comparative Study of Serum Exosome Isolation Using Differential Ultracentrifugation and Three Commercial Reagents. PLoS One. 2017;12(1):e0170628.

115. Naderi A, Ahmed AA, Barbosa-Morais NL, Aparicio S, Brenton JD, Caldas C. Expression microarray reproducibility is improved by optimising purification steps in RNA amplification and labelling. BMC Genomics. 2004;5(1):9.
116. Esmaeili M, Bazrgar M, Gourabi H, Ebrahimi B, Boroujeni PB, Fakhri M. Noninvasive sexing of human preimplantation embryos using RT-PCR in the spent culture media: A proof-of-concept study. Eur J Obstet Gynecol Reprod Biol. 2020;252:89-93.
117. Deswaerte V, Nguyen P, West A, Browning AF, Yu L, Ruwanpura SM, et al. Inflammasome Adaptor ASC Suppresses Apoptosis of Gastric Cancer Cells by an IL18-Mediated Inflammation-Independent Mechanism. Cancer Res. 2018;78(5):1293-307.
118. Johnsen KB, Gudbergsson JM, Andresen TL, Simonsen JB. What is the blood concentration of extracellular vesicles? Implications for the use of extracellular vesicles as blood-borne biomarkers of cancer. Biochim Biophys Acta Rev Cancer. 2019;1871(1):109-16.
119. Sodar BW, Kittel A, Paloczi K, Vukman KV, Osteikoetxea X, Szabo-Taylor K, et al. Low-density lipoprotein mimics blood plasma-derived exosomes and microvesicles during isolation and detection. Sci Rep. 2016;6:24316.
120. Rozowsky J, Kitchen RR, Park JJ, Galeev TR, Diao J, Warrell J, et al. exceRpt: A Comprehensive Analytic Platform for Extracellular RNA Profiling. Cell Syst. 2019;8(4):352-7 e3.
121. Morhayim J, van de Peppel J, Dudakovic A, Chiba H, van Wijnen AJ, van Leeuwen JP. Molecular characterization of human osteoblast-derived extracellular vesicle mRNA using next-generation sequencing. Biochim Biophys Acta Mol Cell Res. 2017;1864(7):1133-41.
122. Koppers-Lalic D, Hackenberg M, de Menezes R, Misovic B, Wachalska M, Geldof A, et al. Noninvasive prostate cancer detection by measuring miRNA variants (isomiRs) in urine extracellular vesicles. Oncotarget. 2016;7(16):22566-78.
123. Vaddepally RK, Kharel P, Pandey R, Garje R, Chandra AB. Review of Indications of FDA-Approved Immune Checkpoint Inhibitors per NCCN Guidelines with the Level of Evidence. Cancers (Basel). 2020;12(3).
124. Borghaei H, Paz-Ares L, Horn L, Spigel DR, Steins M, Ready NE, et al. Nivolumab versus Docetaxel in Advanced Nonsquamous Non-Small-Cell Lung Cancer. N Engl J Med. 2015;373(17):1627-39.
125. Reck M, Brahmer JR. Pembrolizumab in Non-Small-Cell Lung Cancer. N Engl J Med. 2017;376(10):997.
126. Santarpia M, Giovannetti E, Rolfo C, Karachaliou N, Gonzalez-Cao M, Altavilla G, et al. Recent developments in the use of immunotherapy in non-small cell lung cancer. Expert Rev Respir Med. 2016;10(7):781-98.
127. Suresh K, Naidoo J, Lin CT, Danoff S. Immune Checkpoint Immunotherapy for Non-Small Cell Lung Cancer: Benefits and Pulmonary Toxicities. Chest. 2018;154(6):1416-23.
128. Reuss JE, Suresh K, Naidoo J. Checkpoint Inhibitor Pneumonitis: Mechanisms, Characteristics, Management Strategies, and Beyond. Curr Oncol Rep. 2020;22(6):56.
129. Haratani K, Hayashi H, Chiba Y, Kudo K, Yonesaka K, Kato R, et al. Association of Immune-Related Adverse Events With Nivolumab Efficacy in Non-Small-Cell Lung Cancer. JAMA Oncol. 2018;4(3):374-8.
130. Wu J, Hong D, Zhang X, Lu X, Miao J. PD-1 inhibitors increase the incidence and risk of pneumonitis in cancer patients in a dose-independent manner: a meta-analysis. Sci Rep. 2017;7:44173.
131. Gomatou G, Tzilas V, Kotteas E, Syrigos K, Bouros D. Immune Checkpoint Inhibitor-Related Pneumonitis. Respiration. 2020:1-11.
132. Fukihara J, Sakamoto K, Koyama J, Ito T, Iwano S, Morise M, et al. Prognostic Impact and Risk Factors of Immune-Related Pneumonitis in Patients With Non-Small-Cell Lung Cancer Who Received Programmed Death 1 Inhibitors. Clin Lung Cancer. 2019;20(6):442-50 e4.
133. Tone M, Izumo T, Awano N, Kuse N, Inomata M, Jo T, et al. High mortality and poor treatment efficacy of immune checkpoint inhibitors in patients with severe grade checkpoint inhibitor pneumonitis in non-small cell lung cancer. Thorac Cancer. 2019;10(10):2006-12.
134. Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. N Engl J Med. 2012;366(26):2443-54.
135. Le T, Minna JD, Gerber DE. Checkpoint Inhibitor Pneumonitis: Too Clinically Serious For Benefit? J Thorac Oncol. 2019;14(3):332-5.
136. Barron F, Sanchez R, Arroyo-Hernandez M, Blanco C, Zatarain-Barron ZL, Catalan R, et al. Risk of Developing Checkpoint Immune Pneumonitis and Its Effect on Overall Survival in Non-small Cell Lung Cancer Patients Previously Treated With Radiotherapy. Front Oncol. 2020;10:570233.
137. Castanon E. Anti-PD1-Induced Pneumonitis: Capturing the Hidden Enemy. Clin Cancer Res. 2016;22(24):5956-8.
138. Zhai X, Zhang J, Tian Y, Li J, Jing W, Guo H, et al. The mechanism and risk factors for immune checkpoint inhibitor pneumonitis in non-small cell lung cancer patients. Cancer Biol Med. 2020;17(3):599-611.

139. Couto N, Caja S, Maia J, Strano Moraes MC, Costa-Silva B. Exosomes as emerging players in cancer biology. Biochimie. 2018;155:2-10.
140. Bracht JWP, Gimenez-Capitan A, Huang CY, Potie N, Pedraz-Valdunciel C, Warren S, et al. Analysis of extracellular vesicle mRNA derived from plasma using the nCounter platform. Sci Rep. 2021;11(1):3712.
141. Rahouma M, Baudo M, Yahia M, Kamel M, Gray KD, Elmously A, et al. Pneumonitis as a complication of immune system targeting drugs?-a meta-analysis of anti-PD/PD-L1 immunotherapy randomized clinical trials. J Thorac Dis. 2019;11(2):521-34.
142. Naidoo J, Wang X, Woo KM, Iyriboz T, Halpenny D, Cunningham J, et al. Pneumonitis in Patients Treated With Anti-Programmed Death-1/Programmed Death Ligand 1 Therapy. J Clin Oncol. 2017;35(7):709-17.
143. Suresh K, Naidoo J, Zhong Q, Xiong Y, Mammen J, de Flores MV, et al. The alveolar immune cell landscape is dysregulated in checkpoint inhibitor pneumonitis. J Clin Invest. 2019;129(10):4305-15.
144. Jacquelot N, Yamazaki T, Roberti MP, Duong CPM, Andrews MC, Verlingue L, et al. Sustained Type I interferon signaling as a mechanism of resistance to PD-1 blockade. Cell Res. 2019;29(10):846-61.
145. Kong FM, Anscher MS, Sporn TA, Washington MK, Clough R, Barcellos-Hoff MH, et al. Loss of heterozygosity at the mannose 6-phosphate insulin-like growth factor 2 receptor (M6P/IGF2R) locus predisposes patients to radiation-induced lung injury. Int J Radiat Oncol Biol Phys. 2001;49(1):35-41.
146. Tenny S, Hoffman MR. Prevalence. StatPearls. Treasure Island (FL)2020.
147. Tomita H, Sato S, Matsuda R, Sugiura Y, Kawaguchi H, Niimi T, et al. Serum lysozyme levels and clinical features of sarcoidosis. Lung. 1999;177(3):161-7.
148. Zenarruzabeitia O, Vitalle J, Eguizabal C, Simhadri VR, Borrego F. The Biology and Disease Relevance of CD300a, an Inhibitory Receptor for Phosphatidylserine and Phosphatidylethanolamine. J Immunol. 2015;194(11):5053-60.
149. Xu Y, Tarquini F, Romero R, Kim CJ, Tarca AL, Bhatti G, et al. Peripheral CD300a+CD8+ T lymphocytes with a distinct cytotoxic molecular signature increase in pregnant women with chronic chorioamnionitis. Am J Reprod Immunol. 2012;67(3):184-97.
150. UK CR. Lung cancer incidence by stage at diagnosis.
151. Brett GZ. The value of lung cancer detection by six-monthly chest radiographs. Thorax. 1968;23(4):414-20.
152. Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med. 2011;365(5):395-409.
153. Saghir Z, Dirksen A, Ashraf H, Bach KS, Brodersen J, Clementsen PF, et al. CT screening for lung cancer brings forward early disease. The randomised Danish Lung Cancer Screening Trial: status after five annual screening rounds with low-dose CT. Thorax. 2012;67(4):296-301.
154. Infante M, Cavuto S, Lutman FR, Passera E, Chiarenza M, Chiesa G, et al. Long-Term Follow-up Results of the DANTE Trial, a Randomized Study of Lung Cancer Screening with Spiral Computed Tomography. Am J Respir Crit Care Med. 2015;191(10):1166-75.
155. Pastorino U, Rossi M, Rosato V, Marchiano A, Sverzellati N, Morosi C, et al. Annual or biennial CT screening versus observation in heavy smokers: 5-year results of the MILD trial. Eur J Cancer Prev. 2012;21(3):308-15.
156. de Koning HJ, van der Aalst CM, de Jong PA, Scholten ET, Nackaerts K, Heuvelmans MA, et al. Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. N Engl J Med. 2020;382(6):503-13.
157. Sharma M, Surani S. Exploring Novel Technologies in Lung Cancer Diagnosis: Do We Have Room for Improvement? Cureus. 2020;12(1):e6828.
158. Benzaquen J, Boutros J, Marquette C, Delingette H, Hofman P. Lung Cancer Screening, Towards a Multidimensional Approach: Why and How? Cancers (Basel). 2019;11(2).
159. Santarpia M, Liguori A, D'Aveni A, Karachaliou N, Gonzalez-Cao M, Daffina MG, et al. Liquid biopsy for lung cancer early detection. J Thorac Dis. 2018;10(Suppl 7):S882-S97.
160. Zhong Y, Ding X, Bian Y, Wang J, Zhou W, Wang X, et al. Discovery and validation of extracellular vesicle-associated miRNAs as noninvasive detection biomarkers for early-stage non-small-cell lung cancer. Mol Oncol. 2020.
161. Cha BS, Park KS, Park JS. Signature mRNA markers in extracellular vesicles for the accurate diagnosis of colorectal cancer. J Biol Eng. 2020;14:4.
162. Arechederra M, Ávila MA, Berasain C. Liquid biopsy for cancer management: a revolutionary but still limited new tool for precision medicine. Advances in Laboratory Medicine / Avances en Medicina de Laboratorio. 2020;1(3):20200009.
163. Mortani Barbosa EJ, Jr., Yang R, Hershman M. Real World Lung Cancer CT Screening Performance, Smoking Behavior, and Adherence to Recommendations: Lung-RADS Category and Smoking Status Predict Adherence. AJR Am J Roentgenol. 2020.

164. Chen J, Xu Y, Wang X, Liu D, Yang F, Zhu X, et al. Rapid and efficient isolation and detection of extracellular vesicles from plasma for lung cancer diagnosis. Lab Chip. 2019;19(3):432-43.
165. Roman-Canal B, Moiola CP, Gatius S, Bonnin S, Ruiz-Miro M, Gonzalez E, et al. EV-associated miRNAs from pleural lavage as potential diagnostic biomarkers in lung cancer. Sci Rep. 2019;9(1):15057.
166. Bai Y, Qu Y, Wu Z, Ren Y, Cheng Z, Lu Y, et al. Absolute quantification and analysis of extracellular vesicle lncRNAs from the peripheral blood of patients with lung cancer based on multi-colour fluorescence chip-based digital PCR. Biosens Bioelectron. 2019;142:111523.
167. LeBleu VS, Kalluri R. Exosomes as a Multicomponent Biomarker Platform in Cancer. Trends Cancer. 2020;6(9):767-74.
168. Kim SW, Cheon K, Kim CH, Yoon JH, Hawke DH, Kobayashi R, et al. Proteomics-based identification of proteins secreted in apical surface fluid of squamous metaplastic human tracheobronchial epithelial cells cultured by three-dimensional organotypic air-liquid interface method. Cancer Res. 2007;67(14):6565-73.
169. Lee HJ, Kim YT, Park PJ, Shin YS, Kang KN, Kim Y, et al. A novel detection method of non-small cell lung cancer using multiplexed bead-based serum biomarker profiling. J Thorac Cardiovasc Surg. 2012;143(2):421-7.
170. Thun MJ, DeLancey JO, Center MM, Jemal A, Ward EM. The global burden of cancer: priorities for prevention. Carcinogenesis. 2010;31(1):100-10.
171. Mayo-de-Las-Casas C, Velasco A, Sanchez D, Martinez-Bueno A, Garzon-Ibanez M, Gatius S, et al. Detection of somatic mutations in peritoneal lavages and plasma of endometrial cancer patients: A proof-of-concept study. Int J Cancer. 2020;147(1):277-84.
172. Yu L, Bhayana S, Jacob NK, Fadda P. Comparative studies of two generations of NanoString nCounter System. PLoS One. 2019;14(11):e0225505.
173. Hoshino A, Kim HS, Bojmar L, Gyan KE, Cioffi M, Hernandez J, et al. Extracellular Vesicle and Particle Biomarkers Define Multiple Human Cancers. Cell. 2020;182(4):1044-61 e18.

Portfolio

Curriculum Vitae

List of publications

Presentations and proceedings

Poster presentations

Curriculum vitae

Jillian Wilhelmina Paulina Bracht

Personal information

Address:	Carrer del Comte Borrell, 111-113 4-3a 08011 Barcelona, Catalonia, Spain
Date of birth:	21-01-1994
Nationality:	The Netherlands
Sex:	Female
Phone:	+31(0)653921071
E-mail:	Jill94bracht@gmail.com
LinkedIn:	https://www.linkedin.com/in/jill-bracht-872a7792/
Driving license:	B

Profile

I would describe myself as a hardworking, curious, patient and pro-active person. I am a good listener and easy to communicate with, but I also do not mind to work independently. One of my passions is public outreach, and explaining scientific projects to a broader audience in an understandable way.

Education

Feb. 2018 – Feb. 2021	**PhD** Europen Liquid Biopsy Academy (ELBA), Pangaea Oncology and Universitat Autònoma de Barcelona, PhD obtained
Aug. 2015 – Jul. 2017	**Master** Oncology, Vrije Universiteit Amsterdam, MSc obtained
Sep. 2012 – Jul. 2015	**Bachelor** Biomedical Sciences, Vrije Universiteit Amsterdam, BSc obtained
Sep. 2006 – Jun. 2012	Erfgooiers College Huizen, VWO diploma obtained in the subjects: Dutch, French, English, Mathematics, Economics, Biology, Physics and Chemistry

Work Experience and internships

Feb. 2018 – Feb. 2021 **PhD, European Liquid Biopsy Academy (ELBA) H2020-MSCA-ITN-2017, Pangaea Oncology, Medical Services Company, Barcelona, Spain**

Prof.dr. Rafael Rosell and Dr. Miguel Angel Molina-Vila
My PhD focused on the analysis of liquid biopsies (e.g. tumor educated platelets, cell-free RNA, extracellular vesicle RNA and cell-free DNA) on the nCounter NanoString platform. We have validated the detection of clinically relevant mutations and fusions, and created gene signatures indicative of early stage lung cancer, and development of adverse events during immune checkpoint inhibitor treatment. I was responsible for the whole workflow of the projects, including liquid biosource isolation, running of the nCounter system, development of a bioinformatic pipeline and managing clinical databases.

Sep. 2017 – Feb. 2018 **Associate Research Scientist, Pangaea Oncology Barcelona, Spain**

Managed several projects, designed and performed scientific experiments, wrote scientific papers and review articles and presented scientific data. Research focused mainly on lung cancer diagnostics and new *in vitro* treatment combinations to prevent therapy resistance in lung cancer patients.

Jan. 2017 – Jul. 2017 **Major internship, Hospital Universitari Dexeus – Grupo Quirónsalud, Barcelona, Spain**

Prof.dr. Rafael Rosell and Dra. Niki Karachaliou
Defined different RNA and protein biomarkers that play a role in resistance to EGFR Tyrosine Kinase Inhibitors in both EGFR mutant non-small cell lung cancer cell lines and patient-derived liquid biopsy samples. Tested the therapeutic potential of a new drug (TPX-0005) to prevent *in vitro* resistance development in cell lines.

Feb. 2016 – Aug. 2016 **Minor internship, Cancer Center Amsterdam (CCA), Amsterdam, The Netherlands**

Neuro-oncology Research Group; Prof.dr. Thomas Würdinger
RNA sequencing (Illumina) of tumor educated platelets to diagnose non-small cell lung cancer, determine the mutational profile of the tumor and to evaluate treatment response. Responsible for sample processing, from whole blood up to the library preparations for NGS. Made use of a bioinformatic pipeline and clinical data to analyze platelet signatures.

Jun. 2015 – Aug. 2015 **Research Assistant, Sanquin Blood Supply, Utrecht, The Netherlands**

Continued internship-work at Sanquin during the summer holidays as a research assistant.

Mar. 2015 – Jun. 2015 **Bachelor internship/research assistant, Sanquin Blood Supply, Amsterdam/Utrecht, The Netherlands**

Department of donor studies; Eva-Maria Merz group
Investigated if blood donations cause stress, and if so, how the donation procedure could be adjusted to make it less stressful for the donors. Assisted the donors during donation for sampling (questionnaires, saliva swabs and other relevant clinical information).

Jan. 2015 – Jan. 2017 **Volunteer, KWF (Dutch Cancer Society) as District head, Huizen, The Netherlands**

Responsible for the coordination of the fund collection in Huizen in the first week of September (recruiting volunteers, distribution of materials etc.).

Jun. 2011 – Jul. 2017 **Salesperson at Albert Heijn, Huizen, The Netherlands**

Part-time job on Saturday and during holidays, responsible for the bread counter and meat section.

Relevant Publications

- **Bracht JWP**, Gimenez-Capitan A, Huang CY et al. Analysis of extracellular vesicle mRNA derived from plasma using the nCounter platform. *Sci. Rep. 2021.* 11:3712. IF 4.0
- Gimenez-Capitan A, **Bracht JWP**, García-Mosquera JJ et al. Multiplex detection of clinically relevant mutations in liquid biopsies of cancer patients using the nCounter Platform. *Clin, Chem. 2021.* IF 7.3

- **Bracht JWP**, Karachaliou N, Bivona T et al. BRAF Mutations Classes I, II, and III in NSCLC Patients Included in the SLLIP Trial: The Need for a New Pre-Clinical Treatment Rationale. *Cancers* 2019. 11(9). pii: E1381. IF 6.1
- **Bracht JWP**, Berenguer J, Karachaliou N et al. Combining plasma-based biosources to predict treatment response in NSCLC patients. *Ann. Of Oncol* 2018. 1;29(9):2022. IF 14.2
- **Bracht JWP**, Mayo-de-las-Casas C, Berenguer J et al. The Present and Future of Liquid Biopsies in Non-Small Cell Lung Cancer: Combining Four Biosources for Diagnosis, Prognosis, Prediction, and Disease Monitoring. *Curr. Oncol. Rep* 2018. 20;20(9):70. IF 3.1
- Best M, Sol N, In 't Veld SGJG et al. Swarm Intelligence-Enhanced Detection of Non-Small-Cell Lung Cancer Using Tumor-Educated Platelets. *Cancer Cell* 2017. 32(2): 238-252. IF 22.8

Courses

Jun. 2020	BioBusiness Summerschool Online edition
Apr. 2020	Health Economics Course, University of Melbourne
Mar. 2020	EdX Soft skills Professional Certificate Program: business communication, teamwork and collaboration, critical thinking and problem solving, storytelling in the workplace, public speaking, visual presentation.
Feb. 2020	3-month Internship, CCA Amsterdam
Apr. 2019	2-month Internship, Universidad de Granada
Mar. 2019	Bioinformatics Workshop, Universidad de Granada
Jan. 2019	BCF Grant Application Course, Amsterdam
Oct. 2018	Spanish Course A2, Barcelona
Aug. 2018	PCR Workshop, Prague
Aug. 2017	Data Science: Statistical Programming with R, Utrecht
Sep. 2016	Spanish Course A1.1, Hilversum

Technical skills

Sample types:
- Experience working with FFPE tissue, blood, ascitic fluid, urine samples and cell culture medium. Including RNA, DNA, protein and EV-based analysis.

Lab techniques:
- RNA/DNA/protein/EV extraction: QiaSymphony, mirVana, miRNeasy, Trizol, miRCURY, Izon qEV, ultracentrifugation, AGC spin columns
- Quality control techniques: Nanodrop, Bioanalyzer, Qubit
- Genomic/transcriptomic analysis: (qRT-)PCR, nCounter NanoString, Illumina NGS library preparation, GeneReader
- Cell culture techniques: MTT viability assays, cell transfection, immunofluorescence, western blot, phosphoproteomics, colony formation assays

Other:
- Grant writing experience
- Project design and management

Awards and prizes

Nov. 2020 **Second price** for an abstract in the X Educational Symposium on Lung Cancer 2020: *Baseline plasma extracellular vesicle-RNA analysis can predict checkpoint inhibitor pneumonitis in lung cancer patients.* I had the opportunity to virtually present my abstract during the conference.

Nov. 2019 **Runner-up price** for an abstract in the XIII Congress on Lung Cancer 2019, Valencia: *A 770 gene panel expression analysis to predict clinical benefit of durvalumab in HIV-infected cancer patients.* I had the opportunity to present my abstract during the conference in the proffered paper session.

Sep. 2019 **Third price** for a 3-minute pitch on my PhD project: *Breaking the walls of blood vessels for cancer detection*, European Research and Innovation Days: Falling Walls Lab, Marie Sklodowska-Curie Actions, Brussels.

Fellowships

Nov. 2016: Schuurman Schimmel – van Outeren Stichting
Nov. 2016: Erasmus+ European Union grant
Nov. 2016: Stichting de Fundatie van de Vrijvrouwe van Renswoude 's Gravenhage
Nov. 2016: Stichting Bekker-la Bastide-Fonds

Computer skills

- Microsoft Office
- R and R-studio
- nSolver and Advanced analysis
- SPSS
- Graphpad/Prism
- Oligo Explorer
- Primer Express
- Image Lab

Languages

- Dutch Native tongue
- English Excellent
- Spanish Moderate

Interests

- Traveling
- Reading
- Writing
- Cooking

Other dissemination activities

Nov. 2019 **European Association for Cancer Research Ambassador**

- Helping to raise awareness of the EACR and its conferences.

Sep. 2019 **European Research and Innovation Days: Falling Walls Lab: Marie Sklodowska-Curie Actions, Brussels**

- Presented the aim of my PhD project: Breaking the walls of blood vessels for cancer detection, to the audience present at the European Research and Innovation days. The aim of this presentation was to show why my Marie Curie consortium is the most innovative and important project ongoing.

Aug. 2018 **Student Representative ELBA consortium**

- Representative of the PhD students of the Marie Curie ITN ELBA consortium.

Jun. 2016 **Lab-tour to promote blood platelet research, Amsterdam**

- We organized a lab-tour for a non-academic audience, in which we presented the project in a non-scientific and understandable manner. In addition, we organized a tour of the lab and showed the different steps of blood processing.

List of publications

Bracht JWP, Gimenez-Capitan A, Huang CY et al. Analysis of extracellular vesicle mRNA derived from plasma using the nCounter platform. *Sci. Rep.* 2021 11:3712. IF 4.0

Gimenez-Capitan A, **Bracht JWP**, García-Mosquera JJ et al. Multiplex detection of clinically relevant mutations in liquid biopsies of cancer patients using the nCounter Platform. *Clin. Chem.* 2021. IF 7.3

Attili I, Bonanno L, Karachaliou N et al. SRC and PIM1 as potential co-targets to overcome resistance in MET deregulated non-small cell lung cancer. *TLCR* 2020. IF 5.1

Santarpia M, Aguilar A, Chaib I et al. Non-Small-Cell Lung Cancer Signaling Pathways, Metabolism, and PD-1/PD-L1 Antibodies. *Cancers* 2020. 12:1475. IF 6.4

Gonzalez-Cao M, Morán T, Dalmau J et al. Assessment of the Feasibility and Safety of Durvalumab for Treatment of Solid Tumors in Patients With HIV-1 Infection: The Phase 2 DURVAST Study. *JAMA Oncol.* 2020. e200465. IF 13.9

Chaib I, Cai X, Llige D et al. Osimertinib and dihydroartemisinin: a novel drug combination targeting head and neck squamous cell carcinoma. *Ann. Transl. Med* 2019. 7(22):651. IF 3.7

Bracht JWP, Karachaliou N, Bivona T et al. BRAF Mutations Classes I, II, and III in NSCLC Patients Included in the SLLIP Trial: The Need for a New Pre-Clinical Treatment Rationale. *Cancers* 2019. 11(9). pii: E1381. IF 6.1

Codony-Servat J, Viteri S, Codony-Servat C et al. Hsp90 inhibitors enhance the antitumoral effect of osimertinib in parental and osimertinib-resistant non-small cell lung cancer cell lines. *Transl Lung Cancer Res* 2019. 8(4):340-351. IF 4.8

Bracht JWP, Karachaliou N, Berenguer J et al. Osimertinib and pterostilbene in EGFR-mutation-positive non-small cell lung cancer (NSCLC). *Int. J. Biol. Sci* 2019. 15(12): 2607-2614. IF 4.1

Karachaliou N, Codony-Servat J, **Bracht JWP** et al. Characterising acquired resistance to erlotinib in non-small cell lung cancer patients. *Expert Rev Respir Med* 2019. 19:1-10. IF 2.4

Karachaliou N, Fernandez Bruno M, **Bracht JWP** and Rosell R. Profile of alectinib for the treatment of ALK-positive Non-Small Cell Lung Cancer (NSCLC): patient selection and perspectives. *Onco Targets Ther* 2019. 12:4567-4575. IF 2.3

Bracht JWP, Karachaliou N, Berenguer J et al. PIM-1 Inhibition with AZD1208 to Prevent Osimertinib-Induced Resistance in EGFR-Mutation Positive Non-Small Cell Lung Cancer. *J Cancer Metastasis Treat* 2019. 5:22.

Codony-Servat J, Codony-Servat C, Cardona AF et al. Cancer Stem Cells Biomarkers in EGFR-Mutation-Positive Non-Small-Cell Lung Cancer. *Clin Lung Cancer* 2019. 20(3):167-177. IF 4.2

Karachaliou N, Fernandez-Bruno M, **Bracht JWP** and Rosell R. EGFR first- and second-generation TKIs - there is still place for them in EGFR-mutant NSCLC patients. *Tranls. Cancer Res* 2019. 8(Suppl 1):S23-S47. IF 1.1

Karachaliou N, **Bracht JWP**, Fernandez-Bruno M et al. Association of PALB2 Messenger RNA Expression with Platinum-Docetaxel Efficacy in Advanced Non-Small Cell Lung Cancer. *J. Thorac. Oncol* 2019. 14(2):304-310. IF 12.5

Karachaliou N, Cardona AF, **Bracht JWP** et al. Integrin-linked kinase (ILK) and src homology 2 domain-containing phosphatase 2 (SHP2): novel targets in EGFR-mutation positive non-small cell lung cancer (NSCLC). *EbioMedicine* 2019. 39:207-214. IF 6.7

Karachaliou N, **Paulina Bracht JW** and Rosell R. ARID1A Gene Driver Mutations in Lung Adenocarcinomas. *J. Thorac. Onco.* 2018. 13(12):e255-e257. IF 12.5

Karachaliou N, Fernandez-Bruno M, **Bracht JWP** and Rosell R. Challenges and unanswered questions for the next decade of immune-oncology research in NSCLC. *Transl. Lung Cancer Res* 2018. 7(6):691-702. IF 4.8

Bracht JWP, Berenguer J, Karachaliou N et al. Combining plasma-based biosources to predict treatment response in NSCLC patients. *Ann. Of Oncol* 2018. 1;29(9):2022. IF 14.2

Bracht JWP, Mayo-de-las-Casas C, Berenguer J et al. The Present and Future of Liquid Biopsies in Non-Small Cell Lung Cancer: Combining Four Biosources for Diagnosis, Prognosis, Prediction, and Disease Monitoring. *Curr. Oncol. Rep* 2018. 20;20(9):70. IF 3.1

Karachaliou N, Chaib I, Cardona AF et al. Common Co-activation of AXL and CDCP1 in EGFR-mutation-positive Non-small cell Lung Cancer Associated With Poor Prognosis. *EBioMedicine* 2018. 29:112-127. IF 6.7

Atilli I, Karachaliou N, Bonanno L et al. STAT3 as a potential immunotherapy biomarker in oncogene-addicted non-small cell lung cancer. *Ther. Adv. Med. Oncol* 2018. 10: 1-9. IF 5.7

Karachaliou N, Gonzalez-Cao M, Sosa A et al. The Combination of Checkpoint Immunotherapy and Targeted Therapy in Cancer. *Ann. Transl. Med* 2017. 5(19): 388-398.

Best M, Sol N, In ‘t Veld SGJG et al. Swarm Intelligence-Enhanced Detection of Non-Small-Cell Lung Cancer Using Tumor-Educated Platelets. *Cancer Cell* 2017. 32(2): 238-252. IF 22.8

Presentations and proceedings

Jan. 2021 **World Conference on Lung Cancer, Online**
- Selinexor can inhibit nuclear export of HMGB1, a negative predictive marker for immunotherapy response (featured poster)

Nov. 2020 **X Educational Symposium on Lung Cancer, Online**
- Baseline plasma extracellular vesicle-RNA analysis can predict checkpoint inhibitor pneumonitis in lung cancer patients

Mar. 2020 **CCA retreat, Amsterdam**
- miRNA and mRNA detection in plasma-derived extracellular vesicles (EVs) using the nCounter NanoString platform

Dec. 2019 **Primers Meeting, Badalona**
- Unexpected mild chronic stress and Cancer Depression and Cancer (DeCa)
- nCounter IO360 for immunotherapy response prediction in tumor tissue and liquid biopsy
- KRAS multi-pronged approaches and liquid biopsy: Targeting KRAS and p53 by blocking CDC7 replication

Nov. 2019 **XIII Congress on Lung Cancer, Valencia**
- A 770 gene panel expression analysis to predict clinical benefit of durvalumab in HIV-infected cancer patients

Jul. 2019 **IV Cancer Cell Signaling Pathways Conference, Badalona**
- ADRB2-BDNF tracking in cancer patients (DeCa trial)
- Differential gene expression in HIV infected cancer patients, ART-stabilized
- Driver EGFR, BRAF and KRAS mutations and co-alterations

Mar. 2019 **Reviews in Cancer Biology, Badalona**
- BRAF mutations: Classes I, II and III in NSCLC in the SLLIP trial – *Targeted treatment according to BRAF class*
- Composite of KRAS mutations in the SLLIP trial: TP53 and LKB1 co-mutations in liquid biopsy – *Restoring the LKB1 function*
- Predicting pneumonitis development in patients treated with immunotherapy

Jul. 2018 **III Cancer Cell Signaling Pathways Conference, Badalona**
- Urine and plasma exosomes for early cancer detection, monitoring targeted therapy and immunotherapy

Dec. 2017 **Mini Keynote Symposium on Translational Research Primers in Cancer, Badalona**
- Proviral integration site for Moloney murine leukemia virus-1 (PIM-1) kinase in EGFR-mutant NSCLC

Jul. 2017 **II Cancer Cell Signaling Pathways Conference, Badalona**
- Quantitative RT-PCR for gene expression profiling of EGFR mutant NSCLC cell lines before and after TKI + TPX-0005 treatment

Poster presentations

- **World Conference on Lung Cancer 2021, online:** Bracht JWP, Viteri S, Aguilar A et al. Baseline tumor immune cell infiltration and activation can predict checkpoint inhibitor pneumonitis in lung cancer patients.
- **The 2020 Early Detection of Cancer Conference, online:** Bracht JWP, Gimenez-Capitan A, Huang CY et al. NanoString-based analysis of plasma cfRNA for the early detection of non-small cell lung cancer. A pilot study.
- **IASCL Hot Topic Meeting 2020, online:** Bracht JWP, Gimenez-Capitan A, Huang CY et al. Analysis of plasma-derived extracellular vesicle (EV) cargo using the nCounter Nanostring platform.
- **AACR Virtual Annual Meeting II 2020, online:** Bracht JWP, Gimenez-Capitan A, Huang CY et al. miRNA and mRNA detection in plasma-derived extracellular vesicles (EVs) using the nCounter Nanostring platform.
- **AACR Virtual Annual Meeting II 2020, online:** Bracht JWP, Gonzalez-Cao M, Moran T et al. Transcriptomic analysis of pre-treatment tissue samples to predict clinical benefit to durvalumab in HIV-infected cancer patients.
- **AACR Virtual Annual Meeting II 2020, online:** Gimenez-Capitan A, Bracht JWP, Huang CY et al. nCounter for detection of clinically relevant alterations in exosomes of non-small cell lung cancer cells and patients.
- **AACR Advances in Liquid Biopsies 2020, Miami:** Bracht JWP, Karachaliou N, Berenguer J et al. Urine cell-free DNA (cfDNA) concentration and stability test for future clinical use.
- **XIII Congress on Lung Cancer 2019, Valencia:** Bracht JWP, Gonzalez-Cao M, Moran T et al. A 770 gene panel expression analysis to predict clinical benefit of durvalumab in HIV-infected cancer patients.
- **XIII Congress on Lung Cancer 2019, Valencia:** Gonzalez-Cao M, Viteri S, Mayo de las Casas C et al. STK11/LKB1 mutations in advanced lung cancer: gene co-expression patterns and response to immunotherapy.
- **World Conference on Lung Cancer 2019, Barcelona:** Bracht JWP, Karachaliou N, Bivona T et al. BRAF mutations: classes I, II and III in NSCLC patients included in the SLLIP trial, targeted treatment according to class.
- **World Conference on Lung Cancer 2019, Barcelona:** Aguilar-Hernandez A, Garcia Mosquerra JJ, Viteri S et al. High rate of immune-related pneumonitis in lung cancer patients treated with anti PD-1/PD-L1 antibodies.
- **American Society of Clinical Oncology 2019, Chicago:** Bracht JWP, Karachaliou N, Lanman R et al. Tracking plasma KRAS mutations (m) in lung adenocarcinoma (LUAC) patients (p).
- **AACR Annual Meeting 2019, Atlanta**: Gimenez-Capitan A, Huang CY, Bracht JWP et al. nCounter for detection of clinically relevant alterations in liquid biopsies of solid tumor patients.
- **European Lung Cancer Conference 2018, Geneva:** Bracht JWP, Karachaliou N, Berenguer J et al. Proviral Integration site for Moloney murine leukemia virus-1 (PIM-1) inhibition with AZD1208 to prevent resistance to osimertinib in EGFR mutant NSCLC.

Annexes

Annex 1. The Present and Future of Liquid Biopsies in Non-Small Cell Lung Cancer: Combining Four Biosources for Diagnosis, Prognosis, Prediction, and Disease Monitoring.

Annex 2. Multiplex Detection of Clinically Relevant Mutations in Liquid Biopsies of Cancer Patients Using a Hybridization-Based Platform.

Annex 3. Analysis of extracellular vesicle mRNA derived from plasma using the nCounter platform.

Annex 1. The present and future of liquid biopsies in non-small cell lung cancer: combining four biosources for diagnosis, prognosis, prediction, and disease monitoring (PDF)

Current Oncology Reports (2018) 20:70
https://doi.org/10.1007/s11912-018-0720-z

LUNG CANCER (H BORGHAEI, SECTION EDITOR)

The Present and Future of Liquid Biopsies in Non-Small Cell Lung Cancer: Combining Four Biosources for Diagnosis, Prognosis, Prediction, and Disease Monitoring

Jillian Wilhelmina Paulina Bracht[1] · Clara Mayo-de-las-Casas[1] · Jordi Berenguer[1] · Niki Karachaliou[1,2] · Rafael Rosell[1,3,4,5]

Abstract

Purpose of Review Liquid biopsies have potential as tools for diagnosis, prognosis, and prediction of response to therapy. Herein, we will extensively review four liquid biosources, tumor-educated platelets (TEPs), cell-free DNA (cfDNA), circulating tumor cells (CTCs), and extracellular vesicles (EVs) and we will clarify their optimal application in non-small cell lung cancer (NSCLC) diagnosis and therapy.

Recent Findings Liquid biopsies are a minimally invasive alternative to tissue biopsies—especially important in NSCLC patients—since tumor tissue is often unavailable or insufficient for complete genetic analysis. The main advantages of liquid biopsies include the possibility for repeated sampling, the lower cost, and the fact that they can reflect the complete molecular status of the patient better than a single-site biopsy. This is specifically important for lung adenocarcinoma patients since the detection of specific genetic alterations can predict response to targeted therapies.

Summary Molecular analysis is currently cardinal for therapy decision-making and disease monitoring in lung cancer patients. Liquid biopsies can make easier our daily clinical practice and if prospectively tested and validated may serve as a means for lung cancer early detection.

Keywords Liquid biopsy · Liquid biosources · NSCLC · Lung cancer · Diagnostic biomarker · Prognostic biomarker · Predictive biomarker · Resistance biomarker · TEPs · cfDNA · ctDNA · EVs · Exosomes · CTCs

Introduction

With 1.8 million newly diagnosed cases, and 1.6 million deaths in 2012, lung cancer is not only the most common type of cancer worldwide, but also the leading cause of cancer-related death [1]. This high mortality rate is mainly due to the late-stage diagnosis and acquired resistance to lung cancer therapies [2, 3]. In advanced-stage non-small cell lung cancer (NSCLC), targeted therapies that inhibit genetic alterations which drive tumor growth have demonstrated activity [4]. Specifically,

This article is part of the Topical Collection on *Lung Cancer*

✉ Niki Karachaliou
nkarachaliou@oncorosell.com

✉ Rafael Rosell
rrosell@iconcologia.net

Jillian Wilhelmina Paulina Bracht
jill94bracht@gmail.com

Clara Mayo-de-las-Casas
cmayo@panoncology.com

Jordi Berenguer
jberenguer@panoncology.com

[1] Pangaea Oncology, Laboratory of Molecular Biology, Quirón-Dexeus University Institute, Sabino Arana 5-19, 08028 Barcelona, Spain

[2] Instituto Oncológico Dr Rosell (IOR), University Hospital Sagrat Cor, QuironSalud Group, Viladomat 288, 08029 Barcelona, Spain

[3] Institut Català d'Oncologia, Hospital Germans Trias i Pujol, Carretera de Canyet, s/n, 08916 Badalona, Barcelona, Spain

[4] Institut d'Investigació en Ciències Germans Trias i Pujol, Camí de les Escoles, s/n, 08916 Badalona, Barcelona, Spain

[5] Instituto Oncológico Dr Rosell (IOR), Quirón-Dexeus University Institute, Sabino Arana 5-19, 08028 Barcelona, Spain

Published online: 20 July 2018

epidermal growth factor receptor (EGFR) mutations, anaplastic lymphoma kinase (ALK), ROS proto-oncogene 1, receptor tyrosine kinase (ROS1) fusions, as well as rearranged during transfection (RET) fusion, hepatocyte growth factor receptor (MET) mutations or amplifications, and BRAFV600E mutations can be targeted by specific inhibitors [5]. Biopsy samples from tumor tissue are currently the gold standard for both lung cancer diagnosis and for the analysis of genetic alterations. Still, obtaining a tissue biopsy is an invasive approach, limiting the possibility of repeated sampling required for monitoring treatment response and resistance to targeted therapies [6•].

NSCLC symptoms often present at a late stage, and when diagnosed, the tumor has usually already metastasized [2]. Survival increases when surgical resection—the most successful option for cure—is still a possibility [7]. Therefore, like in breast, cervical, prostate, or colon cancer [8–12], screening for NSCLC could prevent late-stage diagnosis and consequently increase survival. A screening test could be used to examine asymptomatic individuals and either diagnose NSCLC at an early stage or find individuals that are at high risk. Screening is expected to be beneficial and improve the survival of patients, considering the high incidence and mortality rate of this initially asymptomatic disease and the well-known risk factors, such as smoking.

Liquid biopsies have emerged as a crucial tool in cancer management. Tumor biomarkers can be detected in body fluids, for example, saliva, urine, and blood [13, 14]. They may be used as a minimally invasive alternative to the often difficult tissue biopsies, with the advantage that they can give a better reflection of the actual disease, whereas a tissue biopsy cannot reflect the clonal heterogeneity of a tumor [15]. Other advantages are the ease, speed, low cost, and safety of isolating body fluids compared to tumor tissue [13, 14]. Moreover, liquid biopsies can be used for repeated sampling, which can be very informative during treatment to determine changes in the mutational subtype of a tumor and thus predicting and understanding mechanisms of resistance [13, 14, 16]. Previous studies reported on the potential of liquid biopsies as sensitive and specific diagnostic biomarkers for cancer detection [17••, 18••, 19•]. In this review, we will comment on the utility of four liquid biopsy sources, tumor-educated platelets (TEPs), cell-free DNA (cfDNA), circulating tumor cells (CTCs), and extracellular vesicles (EVs) in NSCLC (Fig. 1). We will also comment on the role of liquid biopsies in early detection of lung cancer.

Cell-Free DNA as a Tool for the Management of Lung Cancer Patients

Normal and tumor cells shed cell-free fragments of DNA into the bloodstream after necrosis and apoptosis, which can be detected in plasma and serum [13, 20, 21]. The concentration of cfDNA was shown to be significantly higher in cancer patients compared to healthy individuals, correlating with cancer stage [22–24]. Part of this cfDNA in cancer patients is tumor derived and therefore known as circulating tumor DNA (ctDNA). Our group has worked with cfDNA in NSCLC since 1998 [25–28, 29•].

Since tumor DNA contains the same somatic mutations as the tumor, it is possible to discriminate between cfDNA and ctDNA and use the latter as a biomarker for NSCLC [22, 30]. However, the sensitivity of detecting ctDNA has been a major challenge because the majority of cfDNA present in the blood (> 99%) does not originate from tumor cells [22]. This makes the detection of ctDNA difficult, especially in patients with low tumor burden and therefore low levels of ctDNA [20, 22, 31]. Moreover, ctDNA cannot be used to analyze the RNA transcriptome and the proteome of the tumor [20].

Isolation of ctDNA requires only a centrifugation step after the withdrawal of 5–10 mL blood [32]. Although the fraction of ctDNA was shown to be higher in serum, plasma is a more reliable source for biomarker analysis [33, 34]. The ctDNA-containing plasma can be analyzed with real-time quantitative polymerase chain reaction (RT-qPCR), digital PCR (dPCR), beads-emulsion-amplification-and-magnetics (BEAMing), or next-generation sequencing (NGS) [35, 36] (Table 1). A limitation of using ctDNA as a biosource is the limited stability in the circulation, and therefore, blood drawn from patients should be processed quickly [20]. Importantly, ctDNA has a half-life of around 2 hours, indicating that changes in the tumor can be evaluated within hours [22].

Although the use of ctDNA for the detection of tumor-associated biomarkers in metastatic NSCLC is promising, its sensitivity depends on tumor stage [31]. Technological advances in digital genomic technologies have made it possible to detect genetic alterations in small amounts of ctDNA. It has been reported that 47% of early-stage cancer patients have detectable plasma ctDNA levels [21]. A new multi-analyte blood test, based on ctDNA and circulating proteins, was recently developed by Cohen and colleagues. The test has a sensitivity of 70% among eight different types of resectable tumors. Although the cellSEEK test confers a high specificity (> 99%), sensitivity ranged from 98% (ovarian tumors) to only 33% (breast tumors). Moreover, the median sensitivity in stage I disease was shown to be only 43% [17••], indicating that the use of ctDNA as a tool to detect early-stage (I–II) NSCLC should still be optimized.

ctDNA is valuable in detecting genetic cancer-associated alterations, monitoring tumor response to treatment, and identifying residual disease after surgery [20, 49]. The presence of specific genetic mutations—like EGFR, BRAF, or KRAS mutations—may be used for first- or second-line treatment selection. We were able to detect EGFR mutations in ctDNA of NSCLC patients with a sensitivity and specificity of 78 and 100%, respectively [29•, 50••]. With our real-time PCR assay,

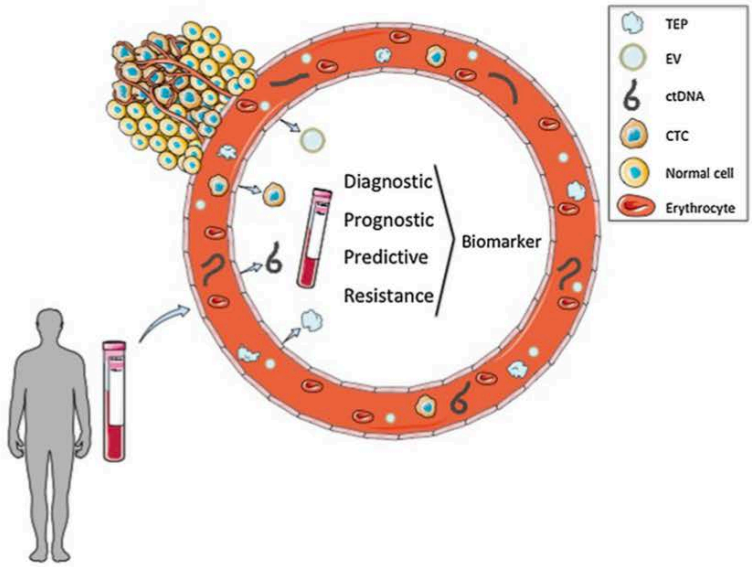


Fig. 1 Liquid biosources as biomarkers in NSCLC. Different liquid biosources—released by the tumor—can be isolated from whole blood: tumor-educated platelets (TEPs), extracellular vesicles (EVs), circulating-tumor DNA (ctDNA), and circulating tumor cells (CTCs). By combining the information that can be obtained from each biosource, it may be possible to diagnose NSCLC in asymptomatic high-risk individuals at an early stage (diagnostic biomarker), determine the location and stage of the tumor (prognostic biomarker), predict the most effective (personalized) therapy based on the mutational burden present in the tumor (predictive biomarker), and monitor therapy response by early detection of resistance (resistance biomarker), with a higher accuracy than currently used liquid biopsy assays

we performed a large-scale screening of EGFR mutations in the blood of NSCLC patients with unavailable or insufficient tissue for genetic analysis [50••]. Moreover, we designed a gene panel for the detection of 568 mutations in six different genes involved in NSCLC based on ctDNA analysis [51•]. This panel was shown to perform with 100% specificity and 90.5% sensitivity in testing relevant mutations in advanced-stage NSCLC [51•].

Table 1 Detection methods for ctDNA analysis and CTC isolation

ctDNA	Detection techniques	Detection limit	References
	qRT-PCR	0.01–10%	[36–38]
	dPCR	0.01–0.10%	[36, 39–41]
	BEAMing	0.01%	[36, 42]
	NGS	0.01–2.00%	[36, 38, 39]
CTCs	Isolation techniques	Sensitivity	References
	Filtering	49.0–100%	[20, 43••, 44]
	Microscopic scanning	95.0%	[45]
	Antibody-based capture	20.0–76.0%	[24, 46–48]

qRT-PCR, quantitative real-time polymerase chain reaction; *dPCR*, digital polymerase chain reaction; *BEAMing*, beads, emulsion, amplification, and magnetics; *NGS*, next-generation sequencing

Targeted therapies based on genomic alterations—like the EGFR tyrosine kinase inhibitors (TKIs) gefitinib, erlotinib, or afatinib—lead to acquired resistance in the majority of the patients. Since ctDNA was shown to correlate with tumor burden, it can be used to monitor response to treatment by either a ctDNA decrease when tumor burden lowers or vice versa [32]. However, since ctDNA is released upon tumor cell death, an increased fraction may also be found when a tumor is in remission and therefore timing of the blood withdrawal may be important [20].

In 2014, the Committee for Medicinal Products for Human Use of the European Medicines Agency (EMA) stated that ctDNA can be used to detect EGFR mutations when tumor samples are not available. This statement was based on data of the therascreen EGFR RGQ PCR kit [52]. The US Food and Drug Administration (FDA) approved the cobas EGFR Mutation Test v2 in June 2016 for the use of erlotinib [13, 53–56]. Shortly thereafter, the same test received approval for the detection of the T790M resistance mutation and the use of osimertinib [6•]. Moreover, the T790M and other mutations—related with tumor resistance—can be detected in ctDNA weeks to months before radiographic disease progression [22]. Thus, ctDNA could potentially be used as a prognostic, predictive, and/or diagnostic biomarker.

We conducted the Spanish Lung Liquid versus Invasive Biopsy Program (SLIPP) study. SLIPP is a multicenter observational study, set-up to validate Guardant360 as a detection method of genetic alterations in ctDNA, isolated from lung adenocarcinoma patients. Guardant360 is an NGS panel targeting ctDNA with an analytic sensitivity of almost 100% and a similar clinical sensitivity compared to tissue-derived NGS [57]. During the SLIPP study, blood samples from patients with metastatic non-squamous NSCLC were collected prior to first-line chemotherapy and upon progression. Results of the blood assay from 185 patients will be compared with the standard of care tissue testing, including, but not limited to, USA FDA-approved companion diagnostics (EGFR PCR tests and ALK FISH testing). The accrual of the SLIPP study has been completed and the results are awaited [6•].

Circulating Tumor Cells as a Biosource in NSCLC Patients

Circulating tumor cells (CTCs) are intact tumor cells, shed into the blood and lymphatic system by a tumor which may cause distant metastases [16]. Since the discovery of CTCs being present in the early course of malignant disease, they became a relevant subject for cancer research [24]. The advantage of CTCs over other blood biosources is that—besides their prognostic significance—multiple DNA-, RNA-, and protein-based assays can be applied to them to evaluate the genomic, transcriptomic, and proteomic profile of the tumor [20, 24]. However, the detection rate of CTCs in the blood of NSCLC patients is low [58, 59]. Furthermore, tumor heterogeneity cannot be reflected in a single circulating cell, and for therapy decision-making, one CTC may not be sufficient as a readout of the tumor genetic landscape [60]. The heterogeneity of surface markers and size of CTCs makes the use of this liquid biosource even more challenging [24].

The isolation of CTCs requires highly sensitive collection methods and enrichment strategies [61, 62] (Table 1). CTC isolation techniques are based on physical properties, expression of surface markers, and depletion of normal leukocytes, and consist of filtering, microscopic scanning, and antibody-based capture assays [20]. One of the methods to isolate CTCs is the CellSearch (Veridex) assay, which focuses on enrichment and immunofluorescent staining of the cells. FDA has approved CellSearch for clinical use in breast, colorectal, and prostate cancer patients [58, 63, 64]. Only in 20–40% of NSCLC patients, CTCs could be detected using this type of isolation assay [24]. On top of the difficulties in isolation, CTCs are delicate cells that are rapidly damaged or degraded after blood collection, and long-term bio-banked blood samples cannot be used for their investigation [20].

Lin et al. described a new cell-surface marker-independent method to isolate CTCs. Specifically, they developed a microfluidic labyrinth device with loops and corners to separate the cells into an analytic stream, which was able to recover 5.4 CTCs/mL blood from metastatic breast cancer patients. The device has the highest purity among label-free CTC isolation methods and is now being implemented in a clinical trial at the University of Michigan to extract CTCs from breast cancer patients [43••]. A folate receptor-based CTC detection method has been additionally investigated in NSCLC [44]. Based on this method, the amount of CTCs in early-stage NSCLC patients was significantly higher compared to individuals with benign lung diseases (9.95 CTC units compared to 6.74 units). When distinguishing between early-stage NSCLC and non-cancer individuals, a sensitivity of 67.2% was reported, while in late-stage NSCLC, the sensitivity of the method raised to 100%.

EGFR mutations were found in the CTCs of 92% of NSCLC patients with tissue-confirmed EGFR mutations [65]. Moreover, the secondary EGFR T790M mutation, the most common mechanism of resistance to EGFR TKIs, was found in CTCs of patients that had tumor progression after TKI treatment [65]. Finally, the amount of CTCs found in the blood of NSCLC patients is positively correlated with disease progression. NSCLC patients with a lower number of CTCs experience longer progression-free survival (PFS) and overall survival (OS) compared to patients with higher amounts of CTCs detected in their blood [44, 66, 67]. Several studies have investigated CTC enumeration as a biomarker for response to therapy [24, 68, 69]. The amount of CTCs increases in NSCLC patients with tumor progression, and decreases in those who respond to therapy, indicating that CTC enumeration may be a valuable tool to evaluate response and detect the acquisition of resistance much earlier than radiographic disease progression.

Although thoroughly investigated, CTCs are currently not commonly used in NSCLC, mainly due to the difficulties in their isolation and the existence of several invalidated detection techniques. Optimization of CTC detection methods and evidence for their clinical utility are necessary for the implementation of CTCs in the routine clinical practice.

The Innovative Role of Tumor-Educated Platelets in NSCLC

Blood platelets are anucleated circulating cell fragments that are derived from megakaryocytes in the bone marrow and have a lifespan of 4 to 6 days [70, 71]. Platelets adhere and aggregate on injury sites to enable wound repair, by formation of actin filaments [71]. In cancer patients, blood platelets are in an increased activated state, indicated by the elevated expression of adhesion molecules on their surface [71, 72].

Platelets contain pre-messenger RNAs (pre-mRNAs) [73]. In cancer patients, they encounter exosomes derived from CTCs, that contain tumor-associated biomolecules (like tumor-specific mRNA) [19•, 74, 75•], and platelets ingest and hereby protect these molecules from degradation [74]. When platelet surface receptors are activated by tumor-derived signals, the ingested tumor-associated biomolecules are spliced, leading to a dynamic and specific mRNA profile in the blood platelets which are subsequently transformed to tumor-educated blood platelets (TEPs) [19•, 74]. Considering that platelet isolation is a relatively simple process and that the mRNA profile of TEPs is a surrogate for tumor-related biomarkers [19•], TEPs are a promising biosource for cancer diagnostics [74, 76].

An important aspect of using TEPs in cancer diagnostics is the need for resemblance of isolated platelets to the in vivo platelet state. Biochemical or physical perturbation can lead to platelet activation, which involves shape changes and content release [77, 78]. Therefore, platelet isolation should be quick and gentle. One way to keep the platelets intact is by isolating and analyzing platelet-rich plasma [77]. However, plasma also contains components that may interfere with the platelet state. Most platelet isolation methods involve centrifugation. Normal centrifugation can activate platelets due to physical perturbation, but a 40-min density-gradient centrifugation leads to a concentrated and intact platelet pellet [77]. Moreover, it has been shown that TEPs conserve high-quality RNA for up to 48 hours at room temperature [19•]. Finally, in currently used protocols, the quantity of blood required to perform TEPs' comprehensive analyses at the RNA level is only 6 mL [19•]. In conclusion, the isolation of TEPs can be performed up to 48 hours after blood withdrawal, is minimally invasive, yields RNA of high quality, and is not time-consuming.

In 2015, Thomas Würdinger and his group showed that TEP analysis distinguishes cancer patients from healthy individuals [19•]. After isolating platelets and sequencing their RNA contents, a different platelet transcript profile was identified in cancer patients and healthy donors. A leave-one-out cross-validation (LOOCV) support vector machine (SVM) algorithm was then developed and was able to distinguish cancer among "healthy" donors with an accuracy of 96% [19•]. When exploring only NSCLC, it was found that 75 and 79% of the subjects in the training and validation cohort, respectively, were correctly classified as NSCLC patients or healthy controls [19•]. After optimization, this algorithm may be used for minimally invasive population-based screening programs, to detect NSCLC at an early stage. However, considering that in the study most of the included NSCLC samples were derived from late-stage NSCLC patients, it is still doubtful whether this algorithm is suitable for the detection of early-stage NSCLC. In 2017, we collaborated with the group of Thomas Würdinger and we optimized the accuracy of NSCLC diagnosis, by using a particle-swarm optimization (PSO) algorithm that was able to diagnose both early- and late-stage NSCLC with an accuracy of 81 and 88%, respectively [18••]. The accuracy of the PSO algorithm for the diagnosis of NSCLC will rise after adding more samples, but still we consider that it would be interesting to investigate whether different liquid biosources combined can further increase the possibility for early detection of lung cancer.

TEPs can also be used to predict treatment response to targeted therapies. Best et al. showed that their TEP algorithm predicts the presence of EGFR mutations, MET amplification, and KRAS mutations with an accuracy of 87, 91, and 90%, respectively [19•]. Moreover, we have shown that the echinoderm microtubule-associated protein-like 4-ALK (EML4-ALK) rearrangements could be found in blood platelets of NSCLC patient with an accuracy of 86% [75•]. In the same study, we monitored EML4-ALK rearrangements in the platelets of an index case for almost 3 years and we were able to detect resistance to crizotinib 2 months before radiographic disease progression [75•]. We are currently applying this method in our daily clinical practice for the molecular follow-up of EML4-ALK-rearranged NSCLC.

The Biological and Clinical Relevance of Extracellular Vesicles in NSCLC

Extracellular vesicles (EVs) can be subdivided into microvesicles, larger vesicles with a size of 50–2000 nm, and exosomes, smaller vesicles of 30–100 nm in size [79, 80]. Microvesicles are shed by cells directly through budding, while exosomes are released through exocytosis after fusion of multivesicular bodies (MVBs) with the plasma membrane of the cell [14]. Importantly, EVs were shown to be mediators in intercellular communication between the tumor and the stroma, influencing primary tumor growth, development of metastases through pre-metastatic niche formation, and drug resistance development [39, 80, 81••, 82].

Exosomes contain proteins, DNA, mRNA, and non-coding RNAs [36, 39, 80]. Moreover, they are surrounded by a lipid bilayer membrane, with which their biosources are protected from degradation in the circulation [80]. Tumor cells release a higher amount of exosomes compared to non-tumor cells [80]. Therefore, exosomes can potentially be used as a liquid biopsy source in several types of tumors, including NSCLC [39].

The isolation of exosomes requires density-gradient centrifugation, ultracentrifugation, or exosome selection based on protein markers like ALG-2 interacting protein X (ALIX), CD63, and tumor susceptibility gene 101 protein (TSG101) [39, 80, 83]. Furthermore, there are commercial kits available for exosome isolation and purification. Still, current methods have to be optimized taking into account the interference of protein and lipid complexes in the isolation process. Studies have demonstrated that exosomes can be used for the

Table 2 Comparison of the four liquid biosources

	Advantages	Limitations	Biomarker
TEPs	- Highly abundant - Easy isolation process - Long storage time - DNA-, RNA-, and protein-based information - Conservation of biosource quality	- Delicate fragments - Only RNA content was investigated - Long half-life	- Diagnostic - Prognostic - Predictive - Resistance development
ctDNA	- Highly abundant - Easy isolation process - Short half-life - Reflects changes in the tumor	- Low sensitivity due to cfDNA - No information on transcriptome or proteome - Limited stability in blood - Sensitivity depends on tumor stage - Benign disease and age can influence cfDNA	- Diagnostic - Prognostic - Predictive - Resistance development
CTCs	- DNA-, RNA-, and protein-based information - Reflects mutational burden of the tumor - Use for functional studies (cell cultures)	- Low frequency - Isolation challenging - Highly heterogeneous phenotype and genotype - Delicate cells - Limited stability in blood	- Prognostic - Predictive - Resistance development
EVs	- Conservation of biosource quality - DNA-, RNA-, and protein-based information	- Long processing time - Low sensitivity due to exosomes released by non-tumor cells	- Diagnostic - Prognostic - Predictive - Resistance development

detection of tumor biomarkers with good correlation with the tissue analysis [80]. In the case of lung cancer, it has been shown that proteins, DNA, RNA, and miRNAs, protected inside exosomes, can be exploited to diagnose the disease at an early stage [80, 84, 85]. miRNAs isolated from exosomes were able to distinguish between healthy individuals and lung cancer patients with a sensitivity and specificity of 80.3 and 92.3%, respectively [84]. Nabet et al. found that certain RNAs are significantly less shielded (and therefore active) in breast cancer patients. These RNAs promote tumor growth and therapy resistance [81••]. In another study, EGFR mutations were detected using exosomal DNA of four NSCLC cell lines [86]. Therefore, exosomal DNA, RNA, and proteins may be used to detect mutations, gene fusions, and splice variants in NSCLC patients [14]. In addition, their isolation and analysis can be used to follow response to therapy and to predict drug resistance before clinical manifestation [80]. Finally, exosomes export cell-waste products. Through this function, they can also export chemotherapeutic drugs, causing disease progression. In this case, drug resistance may be related to the quantity of exosomes, detected in blood samples [83].

Liquid Biopsies and Lung Cancer Screening

In 2011, the results of the National Lung Screening Trial (NLST)—a multicenter randomized controlled trial—were published [87]. The NLST was initiated in 2002, to determine if annual screening with low-dose CT (LDCT)—compared to chest radiography—leads to a mortality reduction in individuals with increased risk for NSCLC. High-risk individuals are current or former heavy smokers, and more than 50.000 participants were randomly divided into the two screening groups. Importantly, LDCT was shown to reduce lung cancer mortality with 20% compared to chest radiography. However, other—smaller—studies did not observe a mortality reduction with LDCT screening [88–90]. Currently, the NELSON screening trial is being conducted in The Netherlands and Belgium with results expected at the end of 2018. The study investigates whether LDCT screening can reduce lung cancer mortality by more than 25% compared to no screening in high-risk participants [91, 92]. LDCT screening to find lung cancer at an early stage is now being implemented in the USA, China, and Japan [93]. Although screening could potentially reduce lung cancer mortality, of all LDCT scans and chest radiographies that indicate an abnormality, 96.4 and 94.5% of them, respectively, turn out to be false-positive results [87]. This results in unnecessary and potentially harmful follow-up interventions, on top of the psychological distress and anxiety for the screened individual [94]. Therefore, another more efficient method is needed for lung cancer screening.

If a tumor is found during screening, a tumor biopsy is pursued to confirm diagnosis, and if it is in advanced stage, tissue mutational profiling follows for therapy decision-making. However, not all lung tumors can be reached, and biopsies are invasive and have all the caveats that were previously explained. To be able to diagnose NSCLC at an early stage, we have set up a European consortium to perform a large project investigating the four blood-based biosources (cfDNA, CTCs, EVs, and TEPs). Funded by the European Union, ELBA (European Liquid Biopsy Academy) will combine the experience of small-medium enterprises (SMEs) and academic partners to develop and implement a diagnostic test for early NSCLC diagnosis [6•].

Conclusions

In recent years, research in liquid biopsies has increased significantly, and management of all types of tumors may be radically changed in the next years. The four blood-based biosources that have been reviewed here all have their advantages and disadvantages, and it is crucial to have an overview of their most effective applications (Table 2). As a diagnostic tool for NSCLC, TEPs appear to be the most promising source until now, with a diagnostic accuracy of 81% [18••]. The study by Cohen et al. [17••] indicated that a multi-analyte test leads to highly divergent results in different tumor types. While cancerSEEK had a median sensitivity of 43% in diagnosing eight types of cancer in stage 1, this sensitivity ranged from 100% in stage 1 liver cancer to only 20% in stage 1 esophageal cancer. Although the diagnostic sensitivity of CTCs shows potential in late-stage NSCLC patients, they are not the most efficient biomarker for early detection of lung cancer. Exosomes, however, contain different biosources and may turn out to be highly important for NSCLC diagnosis.

To increase survival of NSCLC patients, it is important to stratify the mutational subtype of the tumor and thus determine which targeted therapy would be most effective. TEPs were shown to be able to stratify mutational subtype of a tumor with accuracies ranging from 86% (EML4-ALK) to 91% (MET) [19•]. Moreover, ctDNA was used to detect the EGFR L858R mutation in NSCLC patients, with a sensitivity of 78% and a specificity of 100%, indicating the potential use of this assay in lung cancer management [29•]. An assay based on the information isolated from CTCs also has high accuracy; however, since CTC isolation is not standardized and still difficult, ctDNA is a more efficient and reliable source for biomarker analysis. ctDNA is protected inside exosomes, which increases its stability and also abundance compared to ctDNA in the bloodstream [80]. Therefore, isolating ctDNA from exosomes may be a more reliable approach for cancer diagnostics.

TEPs and ctDNA have potential in disease monitoring and can indicate acquired resistance much earlier than radiographic disease progression. The quantity of both ctDNA and CTCs correlates with disease stage and may also aid in detecting therapy resistance [32, 44, 66]. In metastatic breast cancer patients, ctDNA was able to indicate resistance earlier than CTCs [29•]. Finally, the quantity of exosomes is involved in drug resistance [80, 81••, 83], and an increase in exosomes is an indication of disease progression.

The biological differences among tumor types need to be reflected in the methods that are used for their diagnosis. Moreover, these differences need to be exploited to be able to predict the most effective therapy, determine prognosis, and monitor therapy resistance based on liquid biopsies. Comparing and combining different biosources in one functional assay may be the solution to fulfill the unmet needs of liquid biopsy-based diagnostics, keeping in mind that the assay may finally differ among tumor type.

Funding The work of Jillian Wilhelmina Paulina Bracht in Pangaea Oncology is supported by a European Grant (ELBA No 765492). Work in Dr. Rosell's laboratory is partially supported by a grant from La Caixa Foundation, an Instituto de Salud Carlos III grant (RESPONSE, PIE16/00011) and a European Grant (ELBA No 765492).

Compliance with Ethical Standards

Conflict of Interest Jillian Wilhelmina Paulina Bracht, Clara Mayo-de-las-Casas, Jordi Berenguer, Niki Karachaliou, and Rafael Rosell declare they have no conflict of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as:
- • Of importance
- •• Of major importance

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. CA Cancer J Clin. 2017;67(1):7–30. https://doi.org/10.3322/caac.21387.
2. Hirsch FR, Scagliotti GV, Mulshine JL, Kwon R, Curran WJ Jr, Wu YL, et al. Lung cancer: current therapies and new targeted treatments. Lancet. 2017;389(10066):299–311. https://doi.org/10.1016/S0140-6736(16)30958-8.
3. Karachaliou N, Sosa AE, Barron FB, Gonzalez Cao M, Santarpia M, Rosell R. Pharmacological management of relapsed/refractory NSCLC with chemical drugs. Expert Opin Pharmacother. 2017;18:1–10. https://doi.org/10.1080/14656566.2017.1285284.
4. Rosell R, Karachaliou N. Large-scale screening for somatic mutations in lung cancer. Lancet. 2016;387(10026):1354–6. https://doi.org/10.1016/S0140-6736(15)01125-3.
5. Jordan EJ, Kim HR, Arcila ME, Barron D, Chakravarty D, Gao J, et al. Prospective comprehensive molecular characterization of lung adenocarcinomas for efficient patient matching to approved and emerging therapies. Cancer Discov. 2017;7(6):596–609. https://doi.org/10.1158/2159-8290.CD-16-1337.
6.• Karachaliou N, Sosa AE, Molina MA, Centelles Ruiz M, Rosell R. Possible application of circulating free tumor DNA in non-small cell lung cancer patients. J Thorac Dis. 2017;9(Suppl 13):S1364–S72. https://doi.org/10.21037/jtd.2017.09.59. **The SLIPP study was used to compare the performance of the Guardant360 NGS panel in detecting genetic alterations in ctDNA versus standard of care tissue testing. Blood and tissue samples from metastatic non-squamous NSCLC patients were collected before first-line chemotherapy and upon progression.**
7. Senan S, Paul MA, Lagerwaard FJ. Treatment of early-stage lung cancer detected by screening: surgery or stereotactic ablative radiotherapy? Lancet Oncol. 2013;14(7):e270–4. https://doi.org/10.1016/S1470-2045(12)70592-2.
8. Nelson HD, Fu R, Cantor A, Pappas M, Daeges M, Humphrey L. Effectiveness of breast cancer screening: systematic review and meta-analysis to update the 2009 U.S. preventive services task force recommendation. Ann Intern Med. 2016;164(4):244–55. https://doi.org/10.7326/M15-0969.

9. Weller DP, Patnick J, McIntosh HM, Dietrich AJ. Uptake in cancer screening programmes. Lancet Oncol. 2009;10(7):693–9. https://doi.org/10.1016/S1470-2045(09)70145-7.
10. Budenholzer B. Screening for colorectal cancer. Lancet. 1997;349(9048):356–7; author reply 8. https://doi.org/10.1016/S0140-6736(05)62855-3.
11. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Zappa M, Nelen V, et al. Screening and prostate cancer mortality: results of the European randomised study of screening for prostate cancer (ERSPC) at 13 years of follow-up. Lancet. 2014;384(9959):2027–35. https://doi.org/10.1016/S0140-6736(14)60525-0.
12. Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. Br J Cancer. 2013;108(11):2205–40. https://doi.org/10.1038/bjc.2013.177.
13. Karachaliou N, Mayo-de-Las-Casas C, Molina-Vila MA, Rosell R. Real-time liquid biopsies become a reality in cancer treatment. Ann Transl Med. 2015;3(3):36. https://doi.org/10.3978/j.issn.2305-5839.2015.01.16.
14. Siravegna G, Marsoni S, Siena S, Bardelli A. Integrating liquid biopsies into the management of cancer. Nat Rev Clin Oncol. 2017;14(9):531–48. https://doi.org/10.1038/nrclinonc.2017.14.
15. Friedrich MJ. Going with the flow: the promise and challenge of liquid biopsies. JAMA. 2017;318(12):1095–7. https://doi.org/10.1001/jama.2017.10203.
16. Rolfo C, Castiglia M, Hong D, Alessandro R, Mertens I, Baggerman G, et al. Liquid biopsies in lung cancer: the new ambrosia of researchers. Biochim Biophys Acta. 2014;1846(2):539–46. https://doi.org/10.1016/j.bbcan.2014.10.001.
17.•• Cohen JD, Li L, Wang Y, Thoburn C, Afsari B, Danilova L, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science. 2018; https://doi.org/10.1126/science.aar3247. **This study describes a new multi-analyte blood test based on cfDNA and circulating proteins—called CellSEEK—that may be used to detect and localize surgically resectable cancers.**
18.•• Best MG, Sol N, In 't Veld S, Vancura A, Muller M, Niemeijer AN, et al. Swarm intelligence-enhanced detection of non-small-cell lung cancer using tumor-educated platelets. Cancer Cell. 2017;32(2):238–52 e9. https://doi.org/10.1016/j.ccell.2017.07.004. **In this paper, a Particle-Swarm Optimization algorithm—trained on TEP mRNA from healthy individuals and NSCLC patients—was shown to be able to distinguish between NSCLC patients and healthy controls with an accuracy of 81 and 88% in early- and late-stage NSCLC patients respectively.**
19.• Best MG, Sol N, Kooi I, Tannous J, Westerman BA, Rustenburg F, et al. RNA-Seq of tumor-educated platelets enables blood-based pan-cancer, multiclass, and molecular pathway cancer diagnostics. Cancer Cell. 2015;28(5):666–76. https://doi.org/10.1016/j.ccell.2015.09.018. **The study demonstrates that TEP mRNA-seq analysis can be used to train a LOOCV SVM algorithm, which can then distinguish between the blood of healthy individuals and cancer patients with an accuracy of 96%.**
20. Haber DA, Velculescu VE. Blood-based analyses of cancer: circulating tumor cells and circulating tumor DNA. Cancer Discov. 2014;4(6):650–61. https://doi.org/10.1158/2159-8290.CD-13-1014.
21. Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med. 2014;6(224):224ra24. https://doi.org/10.1126/scitranslmed.3007094.
22. Diaz LA Jr, Bardelli A. Liquid biopsies: genotyping circulating tumor DNA. J Clin Oncol. 2014;32(6):579–86. https://doi.org/10.1200/JCO.2012.45.2011.
23. Leon SA, Shapiro B, Sklaroff DM, Yaros MJ. Free DNA in the serum of cancer patients and the effect of therapy. Cancer Res. 1977;37(3):646–50.
24. Calabuig-Farinas S, Jantus-Lewintre E, Herreros-Pomares A, Camps C. Circulating tumor cells versus circulating tumor DNA in lung cancer—which one will win? Transl Lung Cancer Res. 2016;5(5):466–82. https://doi.org/10.21037/tlcr.2016.10.02.
25. Sanchez-Cespedes M, Monzo M, Rosell R, Pifarre A, Calvo R, Lopez-Cabrerizo MP, et al. Detection of chromosome 3p alterations in serum DNA of non-small-cell lung cancer patients. Ann Oncol. 1998;9(1):113–6.
26. Esteller M, Sanchez-Cespedes M, Rosell R, Sidransky D, Baylin SB, Herman JG. Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. Cancer Res. 1999;59(1):67–70.
27. Ramirez JL, Sarries C, de Castro PL, Roig B, Queralt C, Escuin D, et al. Methylation patterns and K-ras mutations in tumor and paired serum of resected non-small-cell lung cancer patients. Cancer Lett. 2003;193(2):207–16.
28. Ramirez JL, Rosell R, Taron M, Sanchez-Ronco M, Alberola V, de Las Penas R, et al. 14-3-3sigma methylation in pretreatment serum circulating DNA of cisplatin-plus-gemcitabine-treated advanced non-small-cell lung cancer patients predicts survival: the Spanish lung Cancer group. J Clin Oncol. 2005;23(36):9105–12. https://doi.org/10.1200/JCO.2005.02.2905.
29.• Karachaliou N, Mayo-de las Casas C, Queralt C, de Aguirre I, Melloni B, Cardenal F, et al. Association of EGFR L858R mutation in circulating free DNA with survival in the EURTAC trial. JAMA Oncol. 2015;1(2):149–57. https://doi.org/10.1001/jamaoncol.2014.257. **The EGFR L858R mutation was investigated in cfDNA and associated with survival in patients participating in a NSCLC trial comparing efficacy of erlotinib versus chemotherapy.**
30. Pathak AK, Bhutani M, Kumar S, Mohan A, Guleria R. Circulating cell-free DNA in plasma/serum of lung cancer patients as a potential screening and prognostic tool. Clin Chem. 2006;52(10):1833–42. https://doi.org/10.1373/clinchem.2005.062893.
31. Sacher AG, Paweletz C, Dahlberg SE, Alden RS, O'Connell A, Feeney N, et al. Prospective validation of rapid plasma genotyping for the detection of EGFR and KRAS mutations in advanced lung Cancer. JAMA Oncol. 2016;2(8):1014–22. https://doi.org/10.1001/jamaoncol.2016.0173.
32. Chen T, He R, Hu X, Luo W, Hu Z, Li J, et al. Circulating tumor DNA: a potential biomarker from solid tumors' monitor to anticancer therapies. Cancer Transl Med. 2017;3(2):64–7. https://doi.org/10.4103/ctm.ctm_6_17.
33. Thijssen MA, Swinkels DW, Ruers TJ, de Kok JB. Difference between free circulating plasma and serum DNA in patients with colorectal liver metastases. Anticancer Res. 2002;22(1A):421–5.
34. Jung M, Klotzek S, Lewandowski M, Fleischhacker M, Jung K. Changes in concentration of DNA in serum and plasma during storage of blood samples. Clin Chem. 2003;49(6 Pt 1):1028–9.
35. Heitzer E, Auer M, Ulz P, Geigl JB, Speicher MR. Circulating tumor cells and DNA as liquid biopsies. Genome Med. 2013;5(8):73. https://doi.org/10.1186/gm477.
36. Kohn L, Johansson M, Grankvist K, Nilsson J. Liquid biopsies in lung cancer-time to implement research technologies in routine care? Ann Transl Med. 2017;5(13):278. https://doi.org/10.21037/atm.2017.04.12.
37. Sorber L, Zwaenepoel K, Deschoolmeester V, Van Schil PE, Van Meerbeeck J, Lardon F, et al. Circulating cell-free nucleic acids and platelets as a liquid biopsy in the provision of personalized therapy for lung cancer patients. Lung Cancer. 2017;107:100–7. https://doi.org/10.1016/j.lungcan.2016.04.026.
38. Vendrell JA, Mau-Them FT, Beganton B, Godreuil S, Coopman P, Solassol J. Circulating cell free tumor DNA detection as a routine

tool for lung cancer patient management. Int J Mol Sci. 2017;18(2) https://doi.org/10.3390/ijms18020264.
39. Perez-Callejo D, Romero A, Provencio M, Torrente M. Liquid biopsy based biomarkers in non-small cell lung cancer for diagnosis and treatment monitoring. Transl Lung Cancer Res. 2016;5(5):455–65. https://doi.org/10.21037/tlcr.2016.10.07.
40. Zhang BO, Xu CW, Shao Y, Wang HT, Wu YF, Song YY, et al. Comparison of droplet digital PCR and conventional quantitative PCR for measuring EGFR gene mutation. Exp Ther Med. 2015;9(4):1383–8. https://doi.org/10.3892/etm.2015.2221.
41. Worrillow L, Baskaran P, Care MA, Varghese A, Munir T, Evans PA, et al. An ultra-deep sequencing strategy to detect sub-clonal TP53 mutations in presentation chronic lymphocytic leukaemia cases using multiple polymerases. Oncogene. 2016;35(40):5328–36. https://doi.org/10.1038/onc.2016.73.
42. Cai X, Janku F, Zhan Q, Fan JB. Accessing genetic information with liquid biopsies. Trends Genet. 2015;31(10):564–75. https://doi.org/10.1016/j.tig.2015.06.001.
43.•• Lin E, Rivera-Baez L, Fouladdel S, Yoon HJ, Guthrie S, Wieger J, et al. High-throughput microfluidic labyrinth for the label-free isolation of circulating tumor cells. Cell Syst. 2017;5(3):295–304 e4. https://doi.org/10.1016/j.cels.2017.08.012. **The authors describe a new label-free CTC isolation method, with highest purity among other label-free methods. The blood flows through a labyrinth device with loops and corners to separate different cell types and isolate CTCs.**
44. Yu Y, Chen Z, Dong J, Wei P, Hu R, Zhou C, et al. Folate receptor-positive circulating tumor cells as a novel diagnostic biomarker in non-small cell lung cancer. Transl Oncol. 2013;6(6):697–702.
45. Ntouroupi TG, Ashraf SQ, McGregor SB, Turney BW, Seppo A, Kim Y, et al. Detection of circulating tumour cells in peripheral blood with an automated scanning fluorescence microscope. Br J Cancer. 2008;99(5):789–95. https://doi.org/10.1038/sj.bjc.6604545.
46. Andreopoulou E, Yang LY, Rangel KM, Reuben JM, Hsu L, Krishnamurthy S, et al. Comparison of assay methods for detection of circulating tumor cells in metastatic breast cancer: AdnaGen AdnaTest BreastCancer Select/Detect versus Veridex CellSearch system. Int J Cancer. 2012;130(7):1590–7. https://doi.org/10.1002/ijc.26111.
47. Hofman V, Long E, Ilie M, Bonnetaud C, Vignaud JM, Flejou JF, et al. Morphological analysis of circulating tumour cells in patients undergoing surgery for non-small cell lung carcinoma using the isolation by size of epithelial tumour cell (ISET) method. Cytopathology. 2012;23(1):30–8. https://doi.org/10.1111/j.1365-2303.2010.00835.x.
48. Punnoose EA, Atwal S, Liu W, Raja R, Fine BM, Hughes BG, et al. Evaluation of circulating tumor cells and circulating tumor DNA in non-small cell lung cancer: association with clinical endpoints in a phase II clinical trial of pertuzumab and erlotinib. Clin Cancer Res. 2012;18(8):2391–401. https://doi.org/10.1158/1078-0432.CCR-11-3148.
49. Oxnard GR, Paweletz CP, Kuang Y, Mach SL, O'Connell A, Messineo MM, et al. Noninvasive detection of response and resistance in EGFR-mutant lung cancer using quantitative next-generation genotyping of cell-free plasma DNA. Clin Cancer Res. 2014;20(6):1698–705. https://doi.org/10.1158/1078-0432.CCR-13-2482.
50.•• Mayo-de-las-Casas C, Jordana-Ariza N, Garzon-Ibanez M, Balada-Bel A, Bertran-Alamillo J, Viteri-Ramirez S, et al. Large scale, prospective screening of EGFR mutations in the blood of advanced NSCLC patients to guide treatment decisions. Ann Oncol. 2017;0:1–8. https://doi.org/10.1093/annonc/mdx288. **Often tumor tissue if unavailable or insufficient for genetic analysis. Therefore, this study investigated the use of cfDNA as a surrogate for tumor tissue to select patients for the treatment with EGFR-TKIs.**
51.• Malapelle U, Mayo de-Las-Casas C, Rocco D, Garzon M, Pisapia P, Jordana-Ariza N, et al. Development of a gene panel for next-generation sequencing of clinically relevant mutations in cell-free DNA from cancer patients. Br J Cancer. 2017;116(6):802–10. https://doi.org/10.1038/bjc.2017.8. **A gene panel ("SiRe") based on ultra-deep sequencing of ctDNA was created to detect 568 mutations in six different genes involved in NSCLC, with sensitivity and analytical specificity of 90.5 and 100%, respectively.**
52. Douillard JY, Ostoros G, Cobo M, Ciuleanu T, Cole R, McWalter G, et al. Gefitinib treatment in EGFR mutated caucasian NSCLC: circulating-free tumor DNA as a surrogate for determination of EGFR status. J Thorac Oncol. 2014;9(9):1345–53. https://doi.org/10.1097/JTO.0000000000000263.
53. Wu YL, Zhou C, Liam CK, Wu G, Liu X, Zhong Z, et al. First-line erlotinib versus gemcitabine/cisplatin in patients with advanced EGFR mutation-positive non-small-cell lung cancer: analyses from the phase III, randomized, open-label, ENSURE study. Ann Oncol. 2015;26(9):1883–9. https://doi.org/10.1093/annonc/mdv270.
54. Park K, Yu CJ, Kim SW, Lin MC, Sriuranpong V, Tsai CM, et al. First-line erlotinib therapy until and beyond response evaluation criteria in solid tumors progression in Asian patients with epidermal growth factor receptor mutation-positive non-small-cell lung Cancer: the ASPIRATION study. JAMA Oncol. 2016;2(3):305–12. https://doi.org/10.1001/jamaoncol.2015.4921.
55. Goss G, Tsai CM, Shepherd FA, Bazhenova L, Lee JS, Chang GC, et al. Osimertinib for pretreated EGFR Thr790Met-positive advanced non-small-cell lung cancer (AURA2): a multicentre, open-label, single-arm, phase 2 study. Lancet Oncol. 2016;17(12):1643–52. https://doi.org/10.1016/S1470-2045(16)30508-3.
56. Mok T, Ladrera G, Srimuninnimit V, Sriuranpong V, Yu CJ, Thongprasert S, et al. Tumor marker analyses from the phase III, placebo-controlled, FASTACT-2 study of intercalated erlotinib with gemcitabine/platinum in the first-line treatment of advanced non-small-cell lung cancer. Lung Cancer. 2016;98:1–8. https://doi.org/10.1016/j.lungcan.2016.04.023.
57. Lanman RB, Mortimer SA, Zill OA, Sebisanovic D, Lopez R, Blau S, et al. Analytical and clinical validation of a digital sequencing panel for quantitative, highly accurate evaluation of cell-free circulating tumor DNA. PLoS One. 2015;10(10):e0140712. https://doi.org/10.1371/journal.pone.0140712.
58. Cristofanilli M, Budd GT, Ellis MJ, Stopeck A, Matera J, Miller MC, et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. N Engl J Med. 2004;351(8):781–91. https://doi.org/10.1056/NEJMoa040766.
59. Krebs MG, Hou JM, Sloane R, Lancashire L, Priest L, Nonaka D, et al. Analysis of circulating tumor cells in patients with non-small cell lung cancer using epithelial marker-dependent and -independent approaches. J Thorac Oncol. 2012;7(2):306–15. https://doi.org/10.1097/JTO.0b013e31823c5c16.
60. Alix-Panabieres C, Pantel K. Characterization of single circulating tumor cells. FEBS Lett. 2017;591(15):2241–50. https://doi.org/10.1002/1873-3468.12662.
61. Larrea E, Sole C, Manterola L, Goicoechea I, Armesto M, Arestin M, et al. New concepts in cancer biomarkers: circulating miRNAs in liquid biopsies. Int J Mol Sci. 2016;17(5) https://doi.org/10.3390/ijms17050627.
62. Maheswaran S, Haber DA. Circulating tumor cells: a window into cancer biology and metastasis. Curr Opin Genet Dev. 2010;20(1):96–9. https://doi.org/10.1016/j.gde.2009.12.002.
63. Negin BP, Cohen SJ. Circulating tumor cells in colorectal cancer: past, present, and future challenges. Curr Treat Options in Oncol. 2010;11(1–2):1–13. https://doi.org/10.1007/s11864-010-0115-3.
64. Galletti G, Portella L, Tagawa ST, Kirby BJ, Giannakakou P, Nanus DM. Circulating tumor cells in prostate cancer diagnosis and monitoring: an appraisal of clinical potential. Mol Diagn Ther. 2014;18(4):389–402. https://doi.org/10.1007/s40291-014-0101-8.

65. Maheswaran S, Sequist LV, Nagrath S, Ulkus L, Brannigan B, Collura CV, et al. Detection of mutations in EGFR in circulating lung-cancer cells. N Engl J Med. 2008;359(4):366–77. https://doi.org/10.1056/NEJMoa0800668.
66. Krebs MG, Sloane R, Priest L, Lancashire L, Hou JM, Greystoke A, et al. Evaluation and prognostic significance of circulating tumor cells in patients with non-small-cell lung cancer. J Clin Oncol. 2011;29(12):1556–63. https://doi.org/10.1200/JCO.2010.28.7045.
67. Juan O, Vidal J, Gisbert R, Munoz J, Macia S, Gomez-Codina J. Prognostic significance of circulating tumor cells in advanced non-small cell lung cancer patients treated with docetaxel and gemcitabine. Clin Transl Oncol. 2014;16(7):637–43. https://doi.org/10.1007/s12094-013-1128-8.
68. Dorsey JF, Kao GD, MacArthur KM, Ju M, Steinmetz D, Wileyto EP, et al. Tracking viable circulating tumor cells (CTCs) in the peripheral blood of non-small cell lung cancer (NSCLC) patients undergoing definitive radiation therapy: pilot study results. Cancer. 2015;121(1):139–49. https://doi.org/10.1002/cncr.28975.
69. Rolle A, Gunzel R, Pachmann U, Willen B, Hoffken K, Pachmann K. Increase in number of circulating disseminated epithelial cells after surgery for non-small cell lung cancer monitored by MAINTRAC(R) is a predictor for relapse: a preliminary report. World J Surg Oncol. 2005;3(1):18. https://doi.org/10.1186/1477-7819-3-18.
70. George JN. Platelets. Lancet. 2000;355(9214):1531–9. https://doi.org/10.1016/S0140-6736(00)02175-9.
71. Nash GF, Turner LF, Scully MF, Kakkar AK. Platelets and cancer. Lancet Oncol. 2002;3(7):425–30.
72. Menter DG, Tucker SC, Kopetz S, Sood AK, Crissman JD, Honn KV. Platelets and cancer: a casual or causal relationship: revisited. Cancer Metastasis Rev. 2014;33(1):231–69. https://doi.org/10.1007/s10555-014-9498-0.
73. McAllister SS, Weinberg RA. The tumour-induced systemic environment as a critical regulator of cancer progression and metastasis. Nat Cell Biol. 2014;16(8):717–27. https://doi.org/10.1038/ncb3015.
74. Nilsson RJ, Balaj L, Hulleman E, van Rijn S, Pegtel DM, Walraven M, et al. Blood platelets contain tumor-derived RNA biomarkers. Blood. 2011;118(13):3680–3. https://doi.org/10.1182/blood-2011-03-344408.
75.• Nilsson RJ, Karachaliou N, Berenguer J, Gimenez-Capitan A, Schellen P, Teixido C, et al. Rearranged EML4-ALK fusion transcripts sequester in circulating blood platelets and enable blood-based crizotinib response monitoring in non-small-cell lung cancer. Oncotarget. 2016;7(1):1066–75. https://doi.org/10.18632/oncotarget.6279. **The EML4-ALK rearrangement can be found in blood platelets of NSCLC patients. Moreover, it was possible to predict therapy resistance based on blood platelet analysis, 2 months before this could be seen radiographically.**
76. Leslie M. Cell biology. Beyond clotting: the powers of platelets. Science. 2010;328(5978):562–4. https://doi.org/10.1126/science.328.5978.562.
77. Hoffman M, Monroe DM, Roberts HR. A rapid method to isolate platelets from human blood by density gradient centrifugation. Am J Clin Pathol. 1992;98(5):531–3.
78. Mustard JF, Kinlough-Rathbone RL, Packham MA. Isolation of human platelets from plasma by centrifugation and washing. Methods Enzymol. 1989;169:3–11.
79. Skog J, Wurdinger T, van Rijn S, Meijer DH, Gainche L, Sena-Esteves M, et al. Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. Nat Cell Biol. 2008;10(12):1470–6. https://doi.org/10.1038/ncb1800.
80. Reclusa P, Taverna S, Pucci M, Durendez E, Calabuig S, Manca P, et al. Exosomes as diagnostic and predictive biomarkers in lung cancer. J Thorac Dis. 2017;9(Suppl 13):S1373–S82. https://doi.org/10.21037/jtd.2017.10.67.
81.•• Nabet BY, Qiu Y, Shabason JE, Wu TJ, Yoon T, Kim BC, et al. Exosome RNA unshielding couples stromal activation to pattern recognition receptor signaling in Cancer. Cell. 2017;170(2):352–66 e13. https://doi.org/10.1016/j.cell.2017.06.031. **The study describes the analysis of specific exosome-derived RNAs, which were shown to promote tumor growth and induce therapy resistance in breast cancer patients. Moreover, the RNA isolated from exosomes may be used as a diagnostic biomarker for cancer.**
82. Raposo G, Stoorvogel W. Extracellular vesicles: exosomes, microvesicles, and friends. J Cell Biol. 2013;200(4):373–83. https://doi.org/10.1083/jcb.201211138.
83. Munson P, Shukla A. Exosomes: potential in cancer diagnosis and therapy. Medicines (Basel). 2015;2(4):310–27. https://doi.org/10.3390/medicines2040310.
84. Jin X, Chen Y, Chen H, Fei S, Chen D, Cai X, et al. Evaluation of tumor-derived exosomal miRNA as potential diagnostic biomarkers for early-stage non-small cell lung cancer using next-generation sequencing. Clin Cancer Res. 2017;23(17):5311–9. https://doi.org/10.1158/1078-0432.CCR-17-0577.
85. Jakobsen KR, Paulsen BS, Baek R, Varming K, Sorensen BS, Jorgensen MM. Exosomal proteins as potential diagnostic markers in advanced non-small cell lung carcinoma. J Extracell Vesicles. 2015;4:26659. https://doi.org/10.3402/jev.v4.26659.
86. Thakur BK, Zhang H, Becker A, Matei I, Huang Y, Costa-Silva B, et al. Double-stranded DNA in exosomes: a novel biomarker in cancer detection. Cell Res. 2014;24(6):766–9. https://doi.org/10.1038/cr.2014.44.
87. National Lung Screening Trial Research T, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med. 2011;365(5):395–409. https://doi.org/10.1056/NEJMoa1102873.
88. Saghir Z, Dirksen A, Ashraf H, Bach KS, Brodersen J, Clementsen PF, et al. CT screening for lung cancer brings forward early disease. The randomised Danish Lung Cancer Screening Trial: status after five annual screening rounds with low-dose CT. Thorax. 2012;67(4):296–301. https://doi.org/10.1136/thoraxjnl-2011-200736.
89. Infante M, Cavuto S, Lutman FR, Passera E, Chiarenza M, Chiesa G, et al. Long-term follow-up results of the DANTE trial, a randomized study of lung cancer screening with spiral computed tomography. Am J Respir Crit Care Med. 2015;191(10):1166–75. https://doi.org/10.1164/rccm.201408-1475OC.
90. Pastorino U, Rossi M, Rosato V, Marchiano A, Sverzellati N, Morosi C, et al. Annual or biennial CT screening versus observation in heavy smokers: 5-year results of the MILD trial. Eur J Cancer Prev. 2012;21(3):308–15. https://doi.org/10.1097/CEJ.0b013e328351e1b6.
91. van Klaveren RJ, Oudkerk M, Prokop M, Scholten ET, Nackaerts K, Vernhout R, et al. Management of lung nodules detected by volume CT scanning. N Engl J Med. 2009;361(23):2221–9. https://doi.org/10.1056/NEJMoa0906085.
92. Ru Zhao Y, Xie X, de Koning HJ, Mali WP, Vliegenthart R, Oudkerk M. NELSON lung cancer screening study. Cancer Imaging 2011;11 Spec No A:S79–84. https://doi.org/10.1102/1470-7330.2011.9020.
93. Pedersen JH, Ashraf H. Implementation and organization of lung cancer screening. Ann Transl Med. 2016;4(8):152. https://doi.org/10.21037/atm.2016.03.59.
94. Wu GX, Raz DJ, Brown L, Sun V. Psychological burden associated with lung cancer screening: a systematic review. Clin Lung Cancer. 2016;17(5):315–24. https://doi.org/10.1016/j.cllc.2016.03.007.

Current Oncology Reports (2020) 22:52
https://doi.org/10.1007/s11912-020-00914-x

CORRECTION

Correction to: The Present and Future of Liquid Biopsies in Non-Small Cell Lung Cancer: Combining Four Biosources for Diagnosis, Prognosis, Prediction, and Disease Monitoring

Jillian Wilhelmina Paulina Bracht[1,2] · Clara Mayo-de-las-Casas[1] · Jordi Berenguer[1] · Niki Karachaliou[1,3] · Rafael Rosell[1,4,5,6]

Correction to: Current Oncology Reports (2018) 20: 70
https://doi.org/10.1007/s11912-018-0720-z

The original version of this review article unfortunately contained mistakes in the Affiliation and Funding sections. The details are given below:

1. "Universitat Autónoma de Barcelona (UAB), Barcelona, Spain" should be added under Jillian Wilhelmina Paulina Bracht.
2. The correct Funding statement is presented below:

Funding This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 765492. Work in Dr. Rosell's laboratory is partially supported by a grant from La Caixa Foundation and an Instituto de Salud Carlos III grant (RESPONSE, PIE16/ 00011).

The online version of the original article can be found at https://doi.org/10.1007/s11912-018-0720-z

✉ Niki Karachaliou
nkarachaliou@oncorosell.com

✉ Rafael Rosell
rrosell@iconcologia.net

Jillian Wilhelmina Paulina Bracht
jill94bracht@gmail.com

Clara Mayo-de-las-Casas
cmayo@panoncology.com

Jordi Berenguer
jberenguer@panoncology.com

[1] Pangaea Oncology, Laboratory of Molecular Biology, Quirón-Dexeus University Institute, Sabino Arana 5-19, 08028 Barcelona, Spain

[2] Universitat Autónoma de Barcelona (UAB), Barcelona, Spain

[3] Instituto Oncológico Dr Rosell (IOR), University Hospital Sagrat Cor, QuironSalud Group, Viladomat 288, 08029 Barcelona, Spain

[4] Institut Català d'Oncologia, Hospital Germans Trias i Pujol, Carretera de Canyet, s/n, 08916 Badalona, Barcelona, Spain

[5] Institut d'Investigació en Ciències Germans Trias i Pujol, Camí de les Escoles, s/n, 08916 Badalona, Barcelona, Spain

[6] Instituto Oncológico Dr Rosell (IOR), Quirón-Dexeus University Institute, Sabino Arana 5-19, 08028 Barcelona, Spain

Published online: 03 May 2020

Annex 2. Multiplex detection of clinically relevant mutations in liquid biopsies of cancer patients using a hybridization-based platform (PDF)

Clinical Chemistry 00:0
1–10 (2020)

Cancer Diagnostics

Multiplex Detection of Clinically Relevant Mutations in Liquid Biopsies of Cancer Patients Using a Hybridization-Based Platform

Ana Giménez-Capitán,[a] Jillian Bracht,[a,b] Juan José García,[c] Núria Jordana-Ariza,[a] Beatriz García,[a] Mónica Garzón,[a] Clara Mayo-de-las-Casas,[a] Santiago Viteri-Ramirez,[c] Alejandro Martinez-Bueno,[c] Andrés Aguilar,[c] Ivana-Gabriela Sullivan,[d] Eric Johnson,[e] Chung-Ying Huang,[e] Jay L. Gerlach,[e] Sarah Warren,[e] Joseph M. Beechem,[e] Cristina Teixidó,[f,g] Rafael Rosell,[c,h] Noemí Reguart,[g,i] and Miguel A. Molina-Vila[a,*]

BACKGROUND: With the advent of precision oncology, liquid biopsies are quickly gaining acceptance in the clinical setting. However, in some cases, the amount of DNA isolated is insufficient for Next-Generation Sequencing (NGS) analysis. The nCounter platform could be an alternative, but it has never been explored for detection of clinically relevant alterations in fluids.

METHODS: Circulating-free DNA (cfDNA) was purified from blood, cerebrospinal fluid, and ascites of patients with cancer and analyzed with the nCounter 3 D Single Nucleotide Variant (SNV) Solid Tumor Panel, which allows for detection of 97 driver mutations in 24 genes.

RESULTS: Validation experiments revealed that the nCounter SNV panel could detect mutations at allelic fractions of 0.02–2% in samples with ≥5 pg mutant DNA/μL. In a retrospective analysis of 70 cfDNAs from patients with cancer, the panel successfully detected *EGFR*, *KRAS*, *BRAF*, *PIK3CA*, and *NRAS* mutations when compared with previous genotyping in the same liquid biopsies and paired tumor tissues [Cohen kappa of 0.96 (CI = 0.92–1.00) and 0.90 (CI = 0.74–1.00), respectively]. In a prospective study including 91 liquid biopsies from patients with different malignancies, 90 yielded valid results with the SNV panel and mutations in *EGFR*, *KRAS*, *BRAF*, *PIK3CA*, *TP53*, *NFE2L2*, *CTNNB1*, *ALK*, *FBXW7*, and *PTEN* were found. Finally, serial liquid biopsies from a patient with NSCLC revealed that the semiquantitative results of the mutation analysis by the SNV panel correlated with the evolution of the disease.

CONCLUSIONS: The nCounter platform requires less DNA than NGS and can be employed for routine mutation testing in liquid biopsies of patients with cancer.

Introduction

Although genetic analysis of tumor tissue provides useful information for prognosis and treatment decision making, a significant percentage of patients with advanced-stage cancer cannot be biopsied or the amount of tumor tissue is insufficient for genetic analyses. In addition, repeated sampling for monitoring the course of the disease and detecting the emergence of mechanisms of resistance is frequently not feasible. Liquid biopsies constitute a minimally invasive, safe, and sensitive alternative in these cases; and are quickly gaining acceptance in the clinical setting (1–7).

Circulating-free DNA (cfDNA) purified from blood or other body fluids (8) is the most commonly used type of liquid biopsy. In patients with cancer, cfDNA contains a variable fraction of DNA originating in the tumor cells (circulating tumor DNA, or ctDNA) and can be used to identify clinically relevant mutations, amplifications, and gene fusions (4, 8–10), with polymerase chain reaction (PCR) based methods and targeted Next-Generation Sequencing (NGS) being the most frequently used techniques. The nCounter platform (NanoString Technologies) is a relatively novel technology initially developed for multiplex analysis of RNA molecules, and has been successfully applied for

[a]Pangaea Oncology, Laboratory of Oncology, Quirón Dexeus University Hospital, Barcelona, Spain; [b]Universitat Autònoma de Barcelona, Barcelona, Spain; [c]Dr Rosell Oncology Institute, Quirón Dexeus University Hospital, Barcelona, Spain; [d]Medical Oncology Service, Hospital de Sant Pau, Barcelona, Spain; [e]NanoString Technologies, Seattle, WA; [f]Department of Pathology, Thoracic Oncology Unit, Hospital Clínic, Barcelona, Spain; [g]Translational Genomics and Targeted Therapeutics in Solid Tumors, Institut d'Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain; [h]Cancer Biology and Precision Medicine Program, Catalán Institute of Oncology, Germans Trias i Pujol Health Sciences Institute and Hospital, Badalona, Barcelona, Spain; [i]Medical Oncology, Thoracic Oncology Unit, Hospital Clínic, Barcelona, Spain.
*Address correspondence to this author at: Laboratory of Oncology/Pangaea Oncology, Quirón Dexeus University Hospital, Sabino Arana 5-19, 08028 Barcelona, Spain. Fax +34 935460142; e-mail mamolina@panoncology.com.
Received May 12, 2020; accepted September 30, 2020.
DOI: 10.1093/clinchem/hvaa248

the detection of clinically relevant fusion transcripts and gene signatures in tumor tissues (11–13). In addition, a new hybridization probe chemistry has been developed for the detection of hotspot somatic variants in tumor tissue samples (14). However, despite the growing number of laboratories using nCounter, the platform has never been tested for the routine analysis of liquid biopsies.

In this study, we performed a retrospective validation on collection-stored cfDNA samples that revealed an excellent correlation of nCounter with other methodologies for mutation detection. Then, we prospectively analyzed fluids derived from patients with cancer and were able to detect a substantial number of relevant mutations, demonstrating that nCounter can be implemented in the clinical setting for the routine testing of liquid biopsies.

Material and Methods

PATIENTS AND CELL LINES

Fifteen liquid biopsy samples from healthy donors and 70 from patients with cancer were used for the retrospective validation of the nCounter Vantage 3D DNA Single Nucleotide Variant (SNV) Solid Tumor Panel (NanoString Technologies) (Table 1). All of them had been stored in a sample collection approved by the Spanish Ministry of Health (reference number C.0005039). Then, 91 liquid biopsies (Table 2) from 83 patients collected in 6 Spanish hospitals were analyzed with the same panel (see Tables 1 and 2 in the online Data Supplement). The study was carried out in accordance with the principles of the Declaration of Helsinki under an approved protocol of the institutional review board of Quirón Hospitals. Written informed consent was obtained from all patients and documented; samples were deidentified for patient confidentiality. Clinical information collected from each patient was limited to stage, gender, smoking status, and tumor histology. Cell lines with *EGFR*, *KRAS*, *PIK3CA*, *BRAF*, and *NRAS* mutations were used for analytical validation purposes and also as positive and negative controls (Supplemental Table 3).

DNA ISOLATION

Plasma samples (10 mL) were collected in sterile Vacutainer tubes (BD) and cerebrospinal-, pleural-, and ascitic-fluid samples (3–500 mL) in sterile containers. After 2 consecutive centrifugation steps (500*g*, 10 min), cfDNA was purified using the QIAsymphony® DSP Virus/Pathogen Midi Kit and a QIAsymphony robot (Qiagen), following the manufacturer's instructions. Initial volume was 1.2 mL, final elution volume was 30 µL. DNA concentration was estimated using Qubit®. Finally, DNA from the cell lines was purified using the DNA Easy® extraction kit (Qiagen), according to the manufacturer's instructions.

Table 1. Characteristics of the liquid biopsies included in the retrospective cohort.

Characteristics	*N* = 70	(%)
Type of tumor		
Lung cancer	49	70
Colorectal cancer	11	15.7
Breast cancer	1	1.4
Melanoma	5	7.3
Leukemia	1	1.4
Pancreatic	1	1.4
Endometrial	1	1.4
Ovarian	1	1.4
Type of fluid		
Plasma	62	88.6
Ascites	3	4.3
Serum	3	4.3
Pleural fluid	1	1.4
Cerebrospinal fluid	1	1.4
Mutations previously detected by NGS or Q-PNA-PCR		
EGFR	19	27.1
KRAS	9	12.
BRAF	11	15.7
NRAS	1	1.4
PIK3CA	3	4.3
EGFR and *PIK3CA*	2	2.8
KRAS and *PIK3CA*	1	1.4
BRAF and *PIK3CA*	1	1.4
NRAS and *KRAS*	2	2.8
No mutations	21	30

MUTATION DETECTION BY NCOUNTER

The nCounter Vantage 3D DNA SNV Solid Tumor Panel enables detection of 97 driver mutations in 24 clinically relevant genes (online Supplemental Table 4). For mutation detection using nCounter, 5 µL of purified cfDNA and a reference DNA (NanoString Technologies, provided with the panel) were subjected to a 21-cycle preamplification step in a Verity thermal cycler (Applied Biosystems), according to the manufacturer's instructions. Amplified DNA was denatured at 95 °C for 10 minutes and hybridized at 65 °C for 18–24 hours with the SNV pool, which contains mutation- and exon-specific probes that bind to DNA independently of the presence of mutations. Capture, cleanup,

Table 2. Characteristics of the samples prospectively evaluated in the study.

Characteristics	*N* = 90	(%)
Tumor type and histology		
Lung cancer	51	56.6
Adenocarcinoma	42	46.6
Squamous	2	2.2
Others	7	7.7
Colorectal cancer	20	22.2
Adenocarcinoma	19	21.1
Others	1	1.1
Breast cancer	4	4.4
Melanoma	4	4.4
Others	11	12.2
Collection time		
Basal	62	68.8
Progression	8	8.8
Follow up	18	20
Unknown	2	2.2
Type of fluid		
Plasma	87	96.7
Pleural fluid	1	1.1
Cerebrospinal fluid	2	2.2

and digital data acquisition were performed using the nCounter Prep Station™ and Digital Analyzer™ (NanoString Technologies) (Supplemental Fig. 1).

DATA ANALYSIS

Count values were exported to Excel 2016 (Microsoft) using nSolver software v.4.0 (NanoString Technologies). For each mutation, samples with count values lower than the reference DNA were automatically considered negative and excluded from further analysis. The reference consists of a wild-type DNA that does not harbor any mutation and allows estimating the "background noise" counts for every mutation targeted by the kit. The remaining mutation-specific counts were normalized using the geometric mean of the exon counts of the corresponding gene in the same sample. The same procedure was applied to the count values derived from the reference DNA. Finally, the mutation-specific normalized counts of the samples were divided by the corresponding normalized counts of the reference DNA and the result was subjected to a base-2 logarithmic transformation to obtain the log MUT values. For every mutation in the SNV panel, the average log MUT of all the negative samples for this particular mutation included in the retrospective study plus 3 SD was established as the cut-off value for positivity. The only exceptions were *KRAS* mutations, where the mean plus 2 SDs was used.

MUTATION TESTING BY PNA-Q-PCR AND NGS

Samples used in the retrospective validation had been previously genotyped by quantitative PCR in presence of a peptide-nucleic acid (PNA-Q-PCR) or NGS (4, 8, 9, 15, 16). For details about these 2 techniques, see the Supplemental Methods.

Results

ANALYTICAL VALIDATION

First, we analyzed 15 cfDNA samples purified from the blood of healthy donors. All of them tested negative for the 97 mutations targeted by the Vantage 3D DNA SNV panel. Next, using DNA from 2 mutant cell lines, we found that 5 pg of mutated genomes per μL were sufficient to detect *EGFR* 15-bp deletions and *KRAS* G12C mutations (Supplemental Table 5). Serial dilutions of a mixture of DNAs from 13 mutant cell lines spiked into a pan-negative cfDNA were employed to determine the limit of detection of the panel. Mutations in *EGFR*, *KRAS*, *NRAS*, and *PIK3CA* were detected at allelic fractions between 0.02 and 2% (Supplemental Table 6). Finally, spiked samples of 2 mutant cell lines and a cfDNA purified from the blood of a patient who tested *KRAS* G12D positive were tested on different days by different operators, and the reproducibility of the SNV panel was found to be 100% (Supplemental Table 7).

RETROSPECTIVE VALIDATION IN CLINICAL SAMPLES

A total of 70 liquid biopsies from patients with cancer were used in the retrospective validation of the nCounter SNV panel. Most of them corresponded to plasma samples ($n = 62$), but sera ($n = 3$), ascites ($n = 3$), PE and CSF ($n = 1$ each) were also represented (Table 1). Regarding tumor types, the majority of samples were from patients with lung cancer ($n = 49$), followed by colorectal cancer ($n = 11$), melanoma ($n = 5$), and other malignancies ($n = 5$). All liquid biopsies in the retrospective cohort had been previously genotyped for *EGFR*, *KRAS*, *BRAF*, *NRAS*, and *PIK3CA* hotspot mutations by NGS or PNA-Q-PCR. The cfDNAs were reanalyzed using the nCounter SNV panel, the counts for each mutation were normalized, and positive and negative samples identified as explained in Methods (see also Supplemental Fig. 2). The results obtained for 3 representative hotspot mutations are shown in Fig. 1. In all cases, the distribution of the normalized counts was

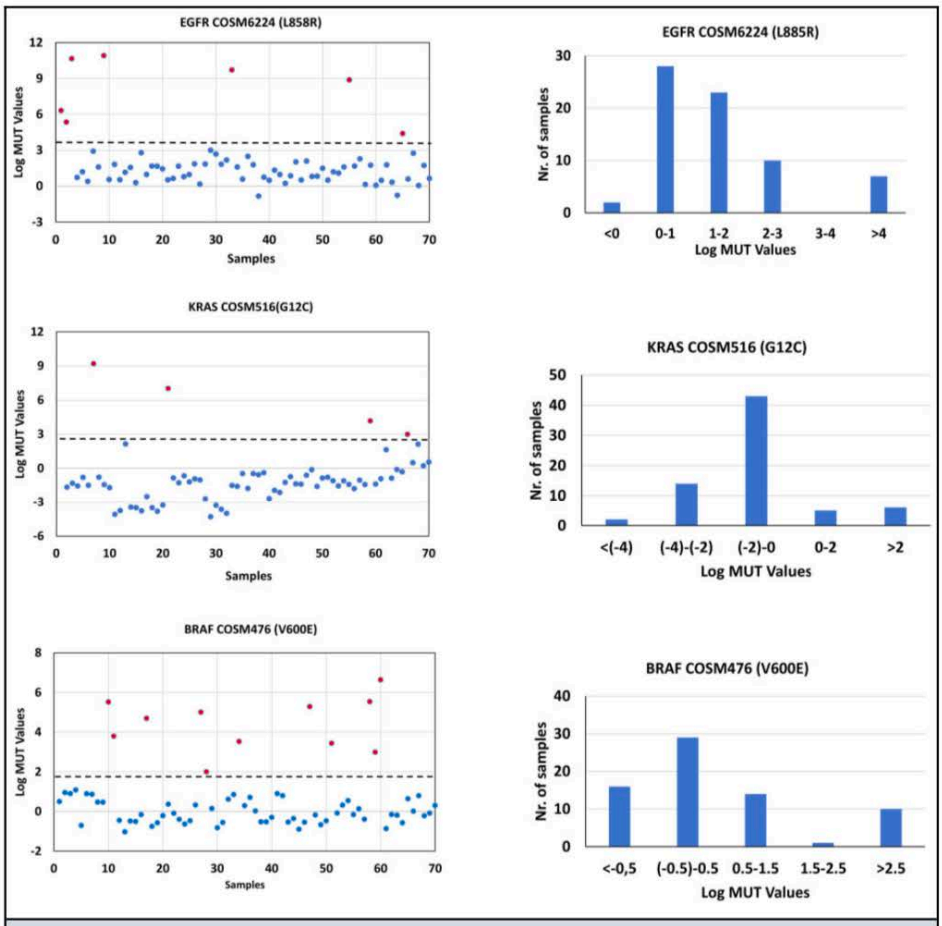


Fig. 1. Results of the mutation detection by nCounter in the retrospective cohort ($n = 70$) for 3 representative mutations, L858R in a) *EGFR*, b) G12C in *KRAS*, and c) V600E in *BRAF*. Left, individual normalized counts in the liquid biopsy samples, expressed as log MUT values. The dotted lines indicate the threshold. Right, distribution of the log MUT values.

bimodal, with the mutant samples representing a different subpopulation.

The results of the previous cfDNA genotyping were used for comparison purposes (Fig. 2). For *EGFR*, *KRAS*, *BRAF*, *NRAS*, and *PIK3CA* mutation detection, nCounter and NGS/PNA-Q-PCR showed concordance rates ranging from 97.1 to 100% and Cohen kappas from 0.91 to 1.00 (Table 3). If the 5 genes were considered together, mutation status by nCounter showed an almost perfect agreement with previous genotyping, with only 4 discordant cases, 0.96 Cohen kappa (CI = 0.92–1.00), and 98.9% concordance (CI = 97.1–99.7%).

The 4 discordant samples were further investigated (Fig. 2). One corresponded to a plasma sample with a T790M in *EGFR* detected by PNA-Q-PCR at an extremely low allelic fraction (0.004%), well below the limit of detection of the nCounter SNV panel. The only discordant sample for *BRAF* had been positive for a V600K mutation by PNA-Q-PCR, with a 2.1% allelic fraction. The plasma sample had been stored for 4 years

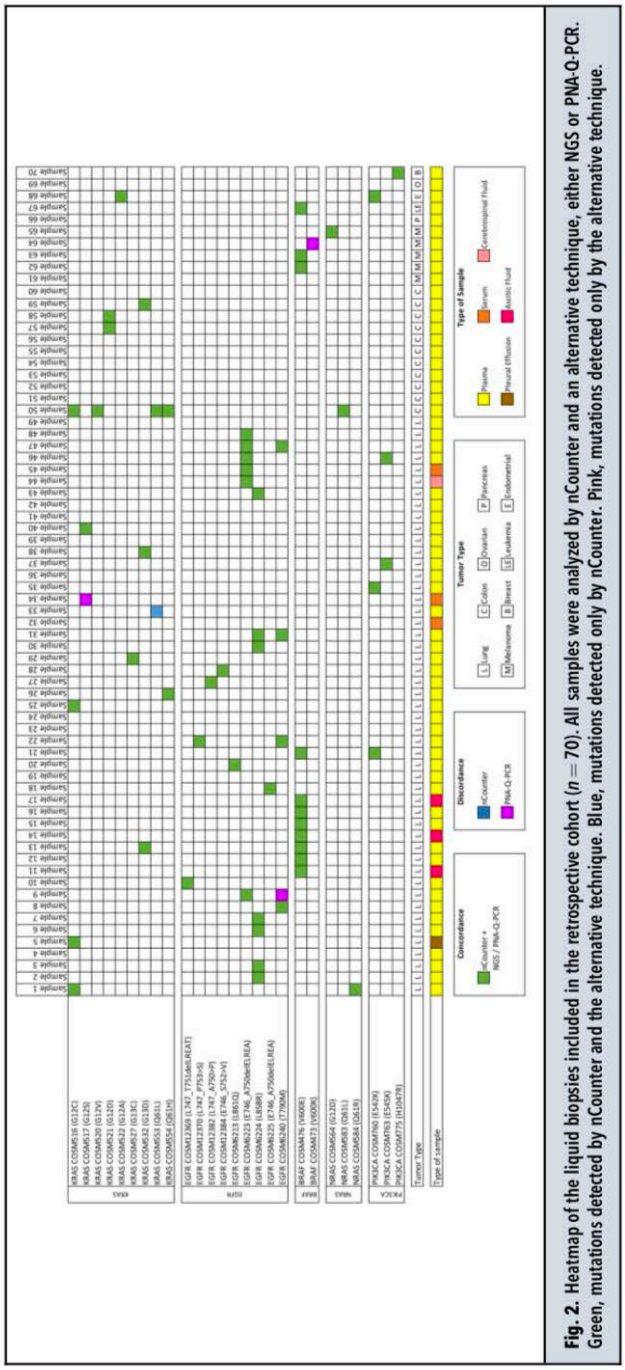


Fig. 2. Heatmap of the liquid biopsies included in the retrospective cohort ($n = 70$). All samples were analyzed by nCounter and an alternative technique, either NGS or PNA-Q-PCR. Green, mutations detected by nCounter and the alternative technique. Blue, mutations detected only by nCounter. Pink, mutations detected only by the alternative technique.

Table 3. Concordance of mutation detection by nCounter with NGS and PNA-Q-PCR in liquid biopsy samples. The 95% confidence intervals are indicated for the overall results.

	nCounter vs NGS/PNA-Q-PCR					
Genes[a]	*EGFR*	*KRAS*	*BRAF*	*NRAS*	*PIK3CA*	Overall
No. of concordant results	69	68	69	70	70	346
No. of discordant results	1	2	1	0	0	4
Sensitivity	100%	93.3%	100%	100%	100.0%	94.7% (CI = 85.4-98.9%)
Specificity	98.0%	98.2%	98.3%	100%	100.0%	99.7% (CI = 98.1-100%)
Concordance	98.6%	97.1%	98.6%	100%	100.0%	98.9% (CI = 97.1-99.7%)
Cohen kappa	0.96	0.91	0.95	1.00	1.00	0.96 (CI = 0.92-1.00)

[a]Samples not carrying a mutation in a particular gene were used as negatives for this gene, independently of the mutational status of other genes.
Abbreviations: NGS–Next-Generation Sequencing; PNA-Q-PCR-quantitative PCR in presence of a quencher-labeled peptide nucleic acid.

at the moment of the nCounter analysis and showed very low exonic counts, indicating cfDNA degradation. Regarding the 2 samples discordant for *KRAS*, one corresponded to a plasma positive for a Q61L mutation by nCounter but negative by PNA-Q-PCR and one to a serum sample harboring a G12S mutation by PNA-Q-PCR at 0.12% allelic fraction, not detected by nCounter.

Paired tissue samples with complete genotyping results were available for 30 liquid biopsies included in the retrospective study. For *EGFR*, *KRAS*, *BRAF*, *NRAS*, and *PIK3CA* mutation detection, nCounter in liquid biopsy showed 71–100% sensitivity and 100% specificity vs paired tissue (Supplemental Table 8). When the 5 genes were considered together, mutation status in liquid biopsy by nCounter showed an almost perfect agreement with previous genotyping in tissue biopsies, with a Cohen kappa of 0.90 (CI = 0.74–1.00) and a 97.3% concordance (CI = 93.1–99.2%).

Finally, we compared the log MUT values obtained by nCounter with the allelic fractions previously found by NGS or PNA-Q-PCR in the same samples. For this analysis, we selected the *KRAS*-positive liquid biopsies and we found a linear correlation between the log MUT *KRAS* values and the $\log_2$ of the *KRAS* mutant allelic fractions derived from NGS or PNA-Q-PCR ($R^2 = 0.63$; Pearson $r = 0.80$; Supplemental Fig. 3).

PROSPECTIVE ANALYSIS OF LIQUID BIOPSIES

During a 6-month period (December 2018 to June 2019), liquid biopsy samples from 83 patients with cancer were collected, submitted to DNA extraction, and prospectively analyzed using the SNV nCounter panel. Six patients had 2 or more fluid samples available (Supplemental Table 1) bringing the total number of liquid biopsies tested to 91. In all cases, the concentration of purified cfDNA was less than 1 ng/mL, as measured by Qubit. Despite these low concentrations, only one of the 91 liquid biopsies showed very low exonic counts and was considered as not evaluable.

The characteristics of the 90 liquid biopsy samples finally included in the prospective study are presented in Table 2. The majority of them corresponded to plasma samples ($n = 87$), although 2 CSF and a PE were also included. Regarding the type of malignancy, most liquid biopsies were obtained from patients with lung cancer ($n = 51$); followed by colorectal (CRC) ($n = 20$), breast ($n = 4$), melanoma ($n = 4$), prostate ($n = 3$), and other tumors ($n = 8$), including thyroid, ovarian, pancreatic, sarcoma, and kidney cancer.

The results of the mutation analysis by nCounter are presented in Fig. 3A. Among the 51 fluid samples from patients with lung cancer, mutations in *EGFR* were found in 7 samples; 3 harbored exon 19 deletions, 3 exon 21 point mutations and one a G719A mutation in exon 18. Regarding *KRAS* mutations, 7 samples were positive for the G12 ($n = 4$), Q61 ($n = 2$), or G13 ($n = 1$) positions. Hotspot mutations in *PIK3CA* were found in 4 lung cancer liquid biopsies, 3 of them coming from the same patient (see next). Finally, 11 samples harbored mutations in a variety of genes, including *PTEN*, *CTNNB1*, *GNAQ*, *NFE2L2*, *FGFR2*, *FBXW7*, *TP53*, and *ALK*. The 2 liquid biopsies positive for *ALK* mutations corresponded to patients in progression to ALK targeted therapies. In the case of the patients with CRC, 2 out of 20 liquid biopsies were positive for *KRAS* mutations and one each for *BRAF*, *NRAS*, and *CTNNB1*. Finally, among samples collected from patients with other malignancies, a G12D mutation in *NRAS* was found in 2 serial samples from a melanoma patient in progression to BRAF/MEK inhibitors, while 2 consecutive liquid biopsies from a pancreatic cancer

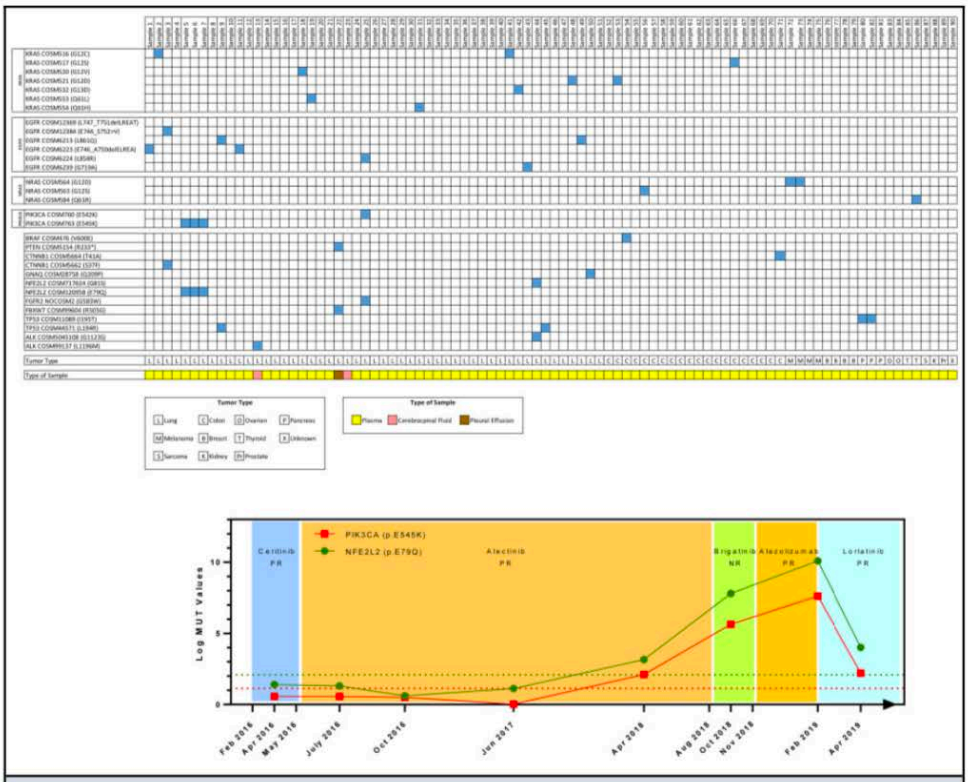


Fig. 3. a) Heatmap representing the results of the mutation detection by nCounter in the liquid biopsies included in the prospective cohort ($n = 90$). b) Case report of an *EML4-ALK* advanced NSCLC patient. Results of the mutation analysis of serial blood samples using the nCounter SNV panel are presented, together with the clinical evolution of the patient.

patient harbored the I195T mutation in *TP53* and a thyroid sample a Q61R mutation in *NRAS*.

Eighty-four of the 90 liquid biopsy samples in the prospective cohort had not been previously submitted to any kind of testing. The remaining 6 samples (4 blood and 2 CSFs) had been analyzed using liquid biopsy NGS panels, yielding invalid results. Interestingly, all of them were evaluable by nCounter, which detected drivers in 2, a *KRAS* G12D mutation in a blood sample, and a L1196M *ALK* resistance mutation in a CSF.

VALIDATION OF RESULTS OF THE PROSPECTIVE TESTING

A subset of 16 cfDNA samples from the prospective cohort with remaining material was subsequently submitted to NGS for validation purposes. The subset included samples from the 11 patients carrying mutations in genes not validated in the retrospective part of the study; such as *TP53* or *ALK* (Fig. 3, A). The NGS panel employed did not cover *CTNNB1*, *FBXW7*, *FGFR2*, or *GNAQ* and mutations in these genes could not be confirmed. For the rest of genes, NGS showed concordant results with nCounter for *EGFR* ($n = 4$), *TP53* ($n = 3$), *PIK3CA* ($n = 2$), *NRAS* ($n = 2$), *ALK* ($n = 1$), and *NFE2L2* ($n = 1$) mutations. Sequencing only failed to detect a *PTEN* mutation, while 2 samples were not evaluable due to insufficient material (Fig. 3, A and Supplemental Table 9).

MUTATION ANALYSIS OF SERIAL SAMPLES

A clinical case where serial liquid biopsies were collected will be described in further detail (Fig. 3, B). It corresponded to a patient with lung cancer, diagnosed in February 2016, harboring an *EML4-ALK* fusion and wild-type for *EGFR*, *KRAS*, *BRAF*, and *PIK3CA* in

tumor tissue at presentation. The patient started ceritinib on February 2016, which was replaced by alectinib 3 months later due to hepatic toxicity. The patient was in remission for more than 2 years, and the 4 serial blood samples obtained from April 2016 to June 2017 were pan-negative by the nCounter SNV panel. In contrast, 2 mutations were detected in a fifth sample collected in April 2018, E545K in *PIK3CA* and E79Q in *NFE2L2*, at log MUT values of 2.1 and 3.2, respectively. The patient showed radiological progression in multiple sites 4 months later, in August 2018. Alectinib was then replaced by brigatinib, a blood sample obtained in October revealed a substantial increase in the log MUT values of both mutations, which rose to 5.6 and 7.8. A subsequent radiological evaluation demonstrated lack of response to brigatinib. Atezolizumab was then administered, but the log MUT values further increased in blood, to 7.6 and 10.1; an evaluation of response on February 2019 revealed progression of the disease. Lorlatinib was finally started and the patient underwent a partial response that was accompanied by a substantial decrease in the log MUT values of the 2 mutations in plasma, which dropped to 2.2 and 4.0.

Five serial samples from the patient had remaining material after nCounter, and *PIK3CA* mutations were tested by PNA-Q-PCR for validation purposes. The results showed a good agreement with those previously obtained by nCounter (Supplemental Table 10).

Discussion

Precision oncology, based on the assessment of molecular markers predictive of treatment outcome, has transformed clinical practice for many types of cancer. Since tumor tissue is not always available or sufficient for genetic testing, liquid biopsies have quickly gained acceptance in the clinical setting (1–7). Initially, blood and other fluids from patients with cancer were mainly employed to detect clinically relevant mutations in *EGFR*, *KRAS*, *NRAS*, or *BRAF* using PCR-derived techniques targeting a limited number of exons (17). However, in the last few years, several NGS platforms have been adapted to the requirements of liquid biopsies and are being used by a growing number of laboratories.

The nCounter technology is a multiplex hybridization-based assay (18) that differs from NGS techniques, being based on direct counting of the RNA or DNA molecules (19). The technology has been adapted for the detection of mutations in DNA purified from tumor tissue (14) by the design of 3 types of probes (S, M, and T). S probes have 2 binding regions, one detecting the presence of the mutation and the other binding to a nearby wild-type sequence; while M probes act like signal attenuators of the wild-type sequences, and T probes facilitate detection (Supplemental Fig. 1). The nCounter technology has been widely used in research studies to simultaneously determine mRNA expression levels of hundreds of genes in biological samples (20, 21), including liquid biopsies of patients with cancer (22, 23). Some of these exploratory studies have led to the identification of expression-based signatures to discriminate malignant lung nodules (24) and predict outcome to immunotherapy in solid tumors (25) or drug sensitivity in prostate cancer (26). However, nCounter has never been used for routine testing of liquid biopsies of patients with cancer and the only signature in clinical use is the U.S. Food and Drug Administration-approved, tumor tissue-based Prosigna, which determines the risk of recurrence in breast cancer (11, 12, 27).

Here we have described the validation of the nCounter SNV panel, which can detect mutations and small indels in 27 genes, for the genotyping of liquid biopsy samples; followed by the implementation of the assay for the prospective testing of blood and other fluids of patients with cancer. During the validation study, we found that 5 pg of mutant DNA, purified from 1.2 mL of blood or other body fluids, was sufficient for successful analysis. Regarding limits of detection, using spiking experiments with cell lines we found values of 0.02–2% allelic fraction. The concentration of cfDNA in the liquid biopsy samples used in our study ranged between 0.1 and 0.5 ng/μL, and 5 μL were loaded in the nCounter assay; meaning that the total cfDNA input was 500–2500 pg. Since 5 pg of mutant DNA are required, the minimum allelic fractions needed for mutation detection in liquid biopsy samples by nCounter would be 0.1–0.02%, coinciding with the values found in cell line experiments. In contrast, using cfDNA inputs lower than 500 pg could lead to higher limits of detection.

The limits of detection of the nCounter SNV panel compare favorably with the requirements of liquid biopsy NGS assays (Supplemental Table 11) and the low requirement of input material explains that, among the 91 liquid biopsies prospectively analyzed with the panel, only 1 (1.1%) sample could not be evaluated. Remarkably, valid results could be obtained for 2 CSF samples, which are usually collected at small volumes and contain particularly low amounts of cfDNA (8, 28, 29). One of them corresponded to a patient who was *EML4-ALK* positive progressing to targeted therapies, where a L1196M resistant mutation was identified and used for the selection of subsequent therapies. Of note, NGS had been previously attempted with these 2 CSFs but yielded invalid results due to insufficient DNA concentration.

Finally, comparison with results obtained in tissue biopsies revealed diagnostic sensibility and specificity of 84.6% and 100%, respectively. All these values are in

the range of those reported for liquid biopsy NGS platforms such as Guardant Health or Oncomine (Supplemental Table 11). One of the limitations of our study was that the number of paired tissue samples was limited and the confidence interval calculated for the diagnostic sensitivity had a wide range, from 66.4 to 93.8%. However, we were able to compare the results obtained by nCounter in all the liquid biopsies in the retrospective cohort with the previous genotyping of the same samples by NGS or PNA-Q-PCR, showing an almost perfect agreement, with 99% concordance and a 0.96 Cohen kappa (CI = 0.92–1.00).

During prospective testing, *EGFR* mutations were found in 7/51 liquid biopsy samples from the patients with NSCLC analyzed (13.7%) (30), a percentage in the range of the frequency described in European populations, while *KRAS* mutations were detected in another 7 patients; *EGFR* and *KRAS* mutations were mutually exclusive, as expected. In the case of CRC, *KRAS*, *NRAS* or *BRAF* mutations were found in 5 out of 21 (23.8%) of liquid biopsies. This relatively low prevalence can be explained in 2 ways, (*a*) a considerable number of the patients with CRC were stage I–IIIA, with less tumor burden than advanced patients; and (*b*) 3 out of 21 samples corresponded to samples of patients in response to therapy. Finally, results obtained in serial liquid biopsies indicate that the nCounter SNV panel could be used to follow the evolution of patients with cancer. Although allelic fractions as such cannot be estimated, the log MUT values were directly dependent on allelic fractions and could be easily calculated and monitored (Fig. 4).

The nCounter platform confers several advantages over NGS techniques for mutation detection in liquid biopsies. It requires a substantially lower amount of material, has a 24–48 h turnaround time with relatively short hands-on time, sample preparation is simple compared to NGS, and data analysis is straightforward and does not require bioinformatics expertise (Supplemental Fig. 1). The main disadvantage of the nCounter platform is that, not being a sequencing technique, it cannot detect mutations other than those contained in the SNV panel, although the panel can be customized. Detection of mutations by nCounter can be particularly useful in some settings. Examples include liquid biopsies with small volumes and/or low concentrations of cfDNA, such as CSF samples (8) or pleural and peritoneal lavages (16); liquid biopsies where NGS has failed; or patients with cancer who are in urgent need of genetic testing to determine whether they are eligible for targeted therapies. Also, the nCounter SNV panel is well suited for monitoring patients in response to therapy where repeated NGS of liquid biopsies would not be cost-effective and can be spared until progression.

In summary, we have demonstrated that the nCounter SNV panel, initially developed for tumor tissue samples, shows an analytical performance similar to NGS in liquid biopsies, requires less material, and can be implemented for multiplex detection of somatic mutations in the clinical setting. Our results also pave the way for testing the performance of nCounter for the detection of other relevant alterations in liquid biopsies from patients with cancer, such as gene fusions or expression levels of genes predictive of response to immunotherapy.

Supplemental Material

Supplemental material is available at *Clinical Chemistry* online.

Nonstandard Abbreviations NGS, Next-Generation Sequencing; SNV, single nucleotide variant; cfDNA, circulating-free DNA; NSCLC, nonsmall cell lung cancer; CSF, cerebrospinal fluid; PE, pleural effusion; ctDNA, circulating tumor DNA; FFPE, formalin-fixed, paraffin embedded; PNA-Q-PCR, quantitative PCR in presence of a peptide nucleic acid; WT, wild-type.

Human Genes: *EGFR*, epidermal growth factor receptor; *KRAS*, KRAS proto-oncogene, GTPase; *BRAF*, B-Raf proto-oncogene, serine/threonine kinase; *PIK3CA*, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha; *TP53*, tumor protein p53; *NFE2L2*, nuclear factor, erythroid 2 like 2; *CTNNB1*, catenin beta 1; *ALK*, ALK receptor tyrosine kinase; *FBXW7*, F-box and WD repeat domain containing 7; *PTEN*, phosphatase and tensin homolog.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 4 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; (c) final approval of the published article; and (d) agreement to be accountable for all aspects of the article thus ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved.*

J. Bracht, statistical analysis; J.-J. García-Mosquera, provision of study material or patients; B. García, provision of study material or patients; C. Mayo-de-las-Casas, provision of study material or patients; S. Viteri, provision of study material or patients; A. Aguilar-Hernández, provision of study material or patients; I. Gabriela-Sullivan, provision of study material or patients; S. Warren, administrative support, provision of study material or patients; R. Rosell, provision of study material or patients; M.A. Molina Vila, statistical analysis, preparation of the first draft of the manuscript.

Authors' Disclosures or Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the author disclosure form. Disclosures and/or potential conflicts of interest:*

Employment or Leadership: C.-Y. Huang, S. Warren, J.M. Beechem, and J.L. Gerlach are full time employees of NanoString. S. Viteri, Pangaea Oncology.
Consultant or Advisory Role: C. Teixidó, Pfizer, Novartis, MSD, Roche, AstraZeneca, Takeda; N. Reguart, MSD, BMS, Pfizer; I. Gabriela-Sullivan, Roche, Boehringer Ingelheim, Novartis.

Stock Ownership: J.L. Gerlach, NanoString Technologies (NSTG); S. Warren, NanoString Technologies; J.M. Beechem, NanoString Technologies.
Honoraria: None declared.
Research Funding: The project received funding from a European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement ELBA No 765492. N. Reguart received funding from the Ministry of Health, Instituto de Salud Carlos III, Spain (FIS PI16/00890). Some of the reagents used in the project were provided free-of-charge by NanoString Technologies. C. Teixidó, Pfizer, Novartis; N. Reguart, Pfizer, Novartis; and Ministry of Health, Instituto de Salud Carlos III, Spain.
Expert Testimony: None declared.
Patents: None declared.
Other Remuneration: I. Gabriela-Sullivan, Roche, Pfizer, Novartis, Boehringer Ingelheim, AstraZeneca, Bristol-Myers Squibb, Merck Sharp & Dohme.
Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, preparation of manuscript, or final approval of manuscript.
Acknowledgments: We thank Ariadna Balada-Bel, María José Catalán, María González-Cao, Margarita Magem, Xavi Gonzalez, Veronica Pereira, and Rich Boykin for their contributions. The investigators also wish to thank the patients for kindly agreeing to donate samples to this study. We thank all the physicians who collaborated by providing clinical information.

References

1. Bedin C, Enzo MV, Del Bianco P, Pucciarelli S, Nitti D, Agostini M. Diagnostic and prognostic role of cell-free DNA testing for colorectal cancer patients. Int J Cancer 2017;140:1888–98.
2. Wit S, Rossi E, Weber S, Tamminga M, Manicone M, Swennenhuis JF, et al. Single tube liquid biopsy for advanced non-small cell lung cancer. Int J Cancer 2019; 144:3127–37.
3. Goldman JW, Noor ZS, Remon J, Besse B, Rosenfeld N. Are liquid biopsies a surrogate for tissue EGFR testing? Ann Oncol 2018;29:i38–i46.
4. Mayo-de-Las-Casas C, Jordana-Ariza N, Garzon-Ibanez M, Balada-Bel A, Bertran-Alamillo J, Viteri-Ramirez S, et al. Large scale, prospective screening of EGFR mutations in the blood of advanced NSCLC patients to guide treatment decisions. Ann Oncol 2017;28:2248–55.
5. Bracht JWP, Mayo-de-Las-Casas C, Berenguer J, Karachaliou N, Rosell R. The present and future of liquid biopsies in non-small cell lung cancer: combining four biosources for diagnosis, prognosis, prediction, and disease monitoring. Curr Oncol Rep 2018;20:70.
6. Mayo-de-Las-Casas C, Garzon Ibanez M, Jordana-Ariza N, Garcia-Pelaez B, Balada-Bel A, Villatoro S, et al. An update on liquid biopsy analysis for diagnostic and monitoring applications in non-small cell lung cancer. Expert Rev Mol Diagn 2018;18:35–45.
7. Rolfo C, Mack PC, Scagliotti GV, Baas P, Barlesi F, Bivona TG, et al. Liquid biopsy for advanced non-small cell lung cancer (NSCLC): a statement paper from the IASLC. J Thorac Oncol 2018;13:1248–68.
8. Villatoro S, Mayo-de-Las-Casas C, Jordana-Ariza N, Viteri RS, Garzon IM, Moya-Horno I, et al. Prospective detection of mutations in cerebrospinal fluid, pleural effusion, and ascites of advanced cancer patients to guide treatment decisions. Mol Oncol 2019;13:2633–45.
9. Gonzalez-Cao M, Mayo-de-Las-Casas C, Molina-Vila MA, De Mattos-Arruda L, Munoz-Couselo E, Manzano JL, et al. BRAF mutation analysis in circulating free tumor DNA of melanoma patients treated with BRAF inhibitors. Melanoma Res 2015;25:486–95.
10. Szallasi Z. Detecting mutant KRAS in liquid biopsies: a biomarker searching for a role. Ann Oncol 2017;28:677–8.
11. Prat A, Galvan P, Jimenez B, Buckingham W, Jeiranian HA, Schaper C, et al. Prediction of response to neoadjuvant chemotherapy using core needle biopsy samples with the Prosigna assay. Clin Cancer Res 2016;22: 560–6.
12. Nielsen T, Wallden B, Schaper C, Ferree S, Liu S, Gao D, et al. Analytical validation of the PAM50-based Prosigna breast cancer prognostic gene signature assay and nCounter analysis system using formalin-fixed paraffin-embedded breast tumor specimens. BMC Cancer 2014; 14:177.
13. Reguart N, Teixido C, Gimenez-Capitan A, Pare L, Galvan P, Viteri S, et al. Identification of ALK, ROS1, and RET fusions by a multiplexed mRNA-based assay in formalin-fixed, paraffin-embedded samples from advanced non-small-cell lung cancer patients. Clin Chem 2017;63: 751–60.
14. Warren S. Simultaneous, multiplexed detection of RNA and protein on the NanoString((r)) nCounter((r)) platform. Methods Mol Biol 2018;1783:105–20.
15. Malapelle U, Mayo de-Las-Casas C, Rocco D, Garzon M, Pisapia P, Jordana-Ariza N, et al. Development of a gene panel for next-generation sequencing of clinically relevant mutations in cell-free DNA from cancer patients. Br J Cancer 2017;116:802–10.
16. Mayo-de-Las-Casas C, Velasco A, Sanchez D, Martinez-Bueno A, Garzon-Ibanez M, Gatius S, et al. Detection of somatic mutations in peritoneal lavages and plasma of endometrial cancer patients: a proof-of-concept study. Int J Cancer 2020;147:277–284.
17. Ou SI, Nagasaka M, Zhu VW. Liquid biopsy to identify actionable genomic alterations. Am Soc Clin Oncol Educ Book 2018;38:978–97.
18. Geiss GK, Bumgarner RE, Birditt B, Dahl T, Dowidar N, Dunaway DL, et al. Direct multiplexed measurement of gene expression with color-coded probe pairs. Nat Biotechnol 2008;26:317–25.
19. Kulkarni MM. Digital multiplexed gene expression analysis using the NanoString nCounter system. Curr Protoc Mol Biol 2011;Chapter 25: Unit25B 10.
20. Goytain A, Ng T. NanoString nCounter technology: high-throughput RNA validation. Methods Mol Biol 2020; 2079:125–39.
21. Veldman-Jones MH, Brant R, Rooney C, Geh C, Emery H, Harbron CG, et al. Evaluating robustness and sensitivity of the NanoString technologies nCounter platform to enable multiplexed gene expression analysis of clinical samples. Cancer Res 2015;75:2587–93.
22. Fan K, Ritter C, Nghiem P, Blom A, Verhaegen ME, Dlugosz A, et al. Circulating cell-free miR-375 as surrogate marker of tumor burden in Merkel cell carcinoma. Clin Cancer Res 2018;24:5873–82.
23. Latchana N, DiVincenzo MJ, Regan K, Abrams Z, Zhang X, Jacob NK, et al. Alterations in patient plasma microRNA expression profiles following resection of metastatic melanoma. J Surg Oncol 2018;118:501–9.
24. Kossenkov AV, Qureshi R, Dawany NB, Wickramasinghe J, Liu Q, Majumdar RS, et al. A gene expression classifier from whole blood distinguishes benign from malignant lung nodules detected by low-dose CT. Cancer Res 2019;79:263–73.
25. Warren S, Danaher P, Mashadi-Hossein A, Skewis L, Wallden B, Ferree S, Cesano A. Development of gene expression-based biomarkers on the nCounter((r)) platform for immuno-oncology applications. Methods Mol Biol 2020;2055:273–300.
26. Jan YJ, Yoon J, Chen JF, Teng PC, Yao N, Cheng S, et al. A circulating tumor cell-RNA assay for assessment of androgen receptor signaling inhibitor sensitivity in metastatic castration-resistant prostate cancer. Theranostics 2019;9:2812–26.
27. Tolaney SM, Guo H, Pernas S, Barry WT, Dillon DA, Ritterhouse L, et al. Seven-year follow-up analysis of adjuvant paclitaxel and trastuzumab trial for node-negative, human epidermal growth factor receptor 2-positive breast cancer. J Clin Oncol 2019;37: 1868–75.
28. Li YS, Jiang BY, Yang JJ, Zhang XC, Zhang Z, Ye JY, et al. Unique genetic profiles from cerebrospinal fluid cell-free DNA in leptomeningeal metastases of EGFR-mutant non-small-cell lung cancer: a new medium of liquid biopsy. Ann Oncol 2018;29:945–52.
29. Momtaz P, Pentsova E, Abdel-Wahab O, Diamond E, Hyman D, Merghoub T, et al. Quantification of tumor-derived cell free DNA(cfDNA) by digital PCR (digPCR) in cerebrospinal fluid of patients with BRAFVv600 mutated malignancies. Oncotarget 2016;7:85430–6.
30. Rosell R, Karachaliou N. Large-scale screening for somatic mutations in lung cancer. Lancet 2016;387: 1354–6.

Annex 3. Analysis of extracellular vesicle mRNA derived from plasma using the nCounter platform (PDF)

www.nature.com/scientificreports

scientific reports

Check for updates

OPEN

Analysis of extracellular vesicle mRNA derived from plasma using the nCounter platform

Jillian W. P. Bracht[1,2✉], Ana Gimenez-Capitan[1], Chung-Ying Huang[3], Nicolas Potie[4,5], Carlos Pedraz-Valdunciel[2,6], Sarah Warren[3], Rafael Rosell[6] & Miguel A. Molina-Vila[1✉]

Extracellular vesicles (EVs) are double-layered phospholipid membrane vesicles that are released by most cells and can mediate intercellular communication through their RNA cargo. In this study, we tested if the NanoString nCounter platform can be used for the analysis of EV-mRNA. We developed and optimized a methodology for EV enrichment, EV-RNA extraction and nCounter analysis. Then, we demonstrated the validity of our workflow by analyzing EV-RNA profiles from the plasma of 19 cancer patients and 10 controls and developing a gene signature to differentiate cancer versus control samples. TRI reagent outperformed automated RNA extraction and, although lower plasma input is feasible, 500 µL provided highest total counts and number of transcripts detected. A 10-cycle pre-amplification followed by DNase treatment yielded reproducible mRNA target detection. However, appropriate probe design to prevent genomic DNA binding is preferred. A gene signature, created using a bioinformatic algorithm, was able to distinguish between control and cancer EV-mRNA profiles with an area under the ROC curve of 0.99. Hence, the nCounter platform can be used to detect mRNA targets and develop gene signatures from plasma-derived EVs.

With the growing global cancer burden, estimated at 18 million cases in 2018[1], earlier cancer detection, enhanced disease monitoring and improved therapy selection are indispensable to enhance patient survival. Liquid biopsies provide a minimally invasive, safe and sensitive surrogate for tissue biopsies. Extracellular vesicles (EVs) are double-layered phospholipid membrane vesicles that are released by most cells, including cancer cells, immune cells and even blood platelets, and can be isolated from practically any body fluid. Cells liberate highly heterogeneous EVs in terms of size (10 nm–1 µm), cargo (nucleic acids, proteins and lipids), membrane composition, biogenesis and biological function[2–4]. Several EV enrichment strategies have been described, including ultracentrifugation, size-exclusion chromatography and precipitation. Precipitation buffers capture water molecules and thereby decrease the hydration of particles, allowing their precipitation after a low-speed centrifugation.

Importantly, the active molecules that are found within EVs can be transported to local or distant target cells and execute biological functions, making EVs important mediators of intercellular communication[2,5,6]. Since the quantity of released EVs and their specific cargo is regulated by the producing cells, the RNA profiles contained within EVs could potentially be used as biomarkers for development and progression of several diseases, including cancer[7]. EVs also have the advantage of their lipid bilayer, which makes their cargo particularly stable and allows the use of biobank stored samples.

Gene expression studies using EV-RNA from cancer patients are an active area of research[8,9]. Promising findings have been reported but the lack of standardized protocols for RNA extraction and analysis, leading to inconsistent results, is hampering clinical implementation. Transcriptomic analysis studies are often conducted using quantitative real time PCR (qRT-PCR), microarrays or RNA sequencing (RNAseq), each with their own (dis)advantages[10]. While qRT-PCR provides high sensitivity and specificity with a short turnaround time, it only allows for low-throughput transcriptomic analysis of a limited number of genes. Microarrays represent a medium-throughput platform but have a narrower dynamic range of detection and are not suitable to detect

[1]Pangaea Oncology, Laboratory of Oncology, Quirón Dexeus University Hospital, Sabino Arana 5-19, 08028 Barcelona, Spain. [2]Department of Biochemistry, Molecular Biology and Biomedicine, Universitat Autónoma de Barcelona (UAB), 08193 Cerdanyola, Spain. [3]NanoString Technologies, Seattle, WA, USA. [4]Department of Genetics, Faculty of Science, University of Granada, 18071 Granada, Spain. [5]Bioinformatics Laboratory, Biotechnology Institute, Centro de Investigacion Biomedica, PTS, Avda. del Conocimiento s/n, 18100 Granada, Spain. [6]Germans Trias i Pujol Health Sciences Institute and Hospital (IGTP), Badalona, Barcelona, Spain. ✉email: jill94bracht@gmail.com; mamolina@panoncology.com

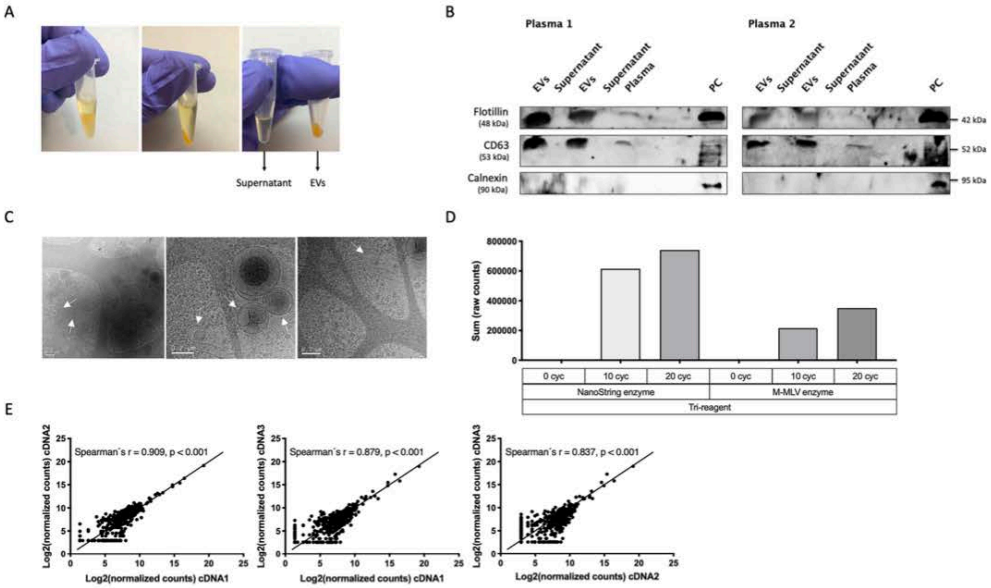


Figure 1. EV enrichment and characterization, assay reproducibility. (**A**) EV enrichment based on precipitation using the miRCURY Exosome Serum/Plasma Kit, including separation of supernatant and EV-enriched pellet. (**B**) Immunoblots showing expression of Flotillin, CD63 (both exosome markers) and Calnexin (cell specific marker) in EV-enriched pellets, supernatants, full plasma and a positive control sample. Experiments were performed in duplicates. Membranes were cut and incubated with specific antibodies for Flotillin, CD63 and Calnexin. Images were cropped for clarity purposes and full membranes can be found in Supplementary Fig. S1. (**C**) Cryogenic Electron Microscopy (cryo-EM) of EV pellets. Arrows point to extracellular vesicles with different size ranges. Scale bars are 200 nm. (**D**) Total counts by nCounter after EV-RNA extraction using TRI reagent. Two different retrotranscriptases (NanoString versus M-MLV) and three pre-amplification conditions (0, 10 and 20 cycles) were tested. (**E**) Reproducibility experiment comparing the Log2 normalized counts by nCounter from three independent cDNAs derived from a single EV-RNA sample. Spearman's correlation coefficient is indicated. EVs, extracellular vesicles; PC, positive control; Cyc, cycles.

genes expressed at either low or high levels. Finally, RNAseq is an accurate, high-throughput platform with a wide dynamic range, but limitations include longer turnaround time, high cost and complex data analysis[11,12].

In recent years, the NanoString nCounter platform has gained popularity in translational research and clinical settings. The platform provides a simple, sensitive and cost-effective solution for multiplexed analysis of up to 800 RNA targets by direct capturing and counting of individual targets. In addition, it can be used with formalin-fixed paraffin embedded (FFPE) tumor tissue and allows for low quality and quantity tissue samples[13,14]. At this respect, the nCounter-based Prosigna assay, which differentiates breast cancer subtypes and predicts the risk of recurrence based on a 50-gene signature, has been validated in the clinical practice and received FDA approval in 2013[15,16]. Another assay developed using the nCounter platform is the 18-gene Tumor Inflammation Signature (TIS), which was able to predict clinical response to PD-1 blockade in an investigational clinical trial assay[17]. These two assays, which utilize tissue samples, emphasize the potential of the nCounter platform as biomarker assay development tool, especially in diagnostic laboratories. Regarding the analysis of liquid biopsies on nCounter, several studies have investigated the potential of some materials, including cf-[18] and EV-DNA[19], CTC-RNA[20,21], leukocyte mRNA[22], cfRNA[23] and EV-miRNA[24,25], with different success rates. However, nCounter has never been tested for the analysis of EV-derived mRNA.

Here, we present a proof-of-concept study where we optimized a workflow for EV enrichment from human blood samples, EV-mRNA purification and subsequent analysis by nCounter. Then, we used the workflow to develop an EV-mRNA based gene signature to differentiate cancer versus control samples. Our work demonstrates that nCounter can be employed for biomarker discovery based on EV-mRNA.

Results

Optimization of plasma EV enrichment and EV-RNA extraction methodologies. EVs were enriched from 500 µL plasma of control samples using the miRCURY Exosome Serum/Plasma Kit (Fig. 1a). Final miRCURY sediments were submitted to western blotting, revealing enrichment in the exosome markers Flotillin and CD63, which were absent or detected at low levels in miRCURY supernatants and whole plasma samples.

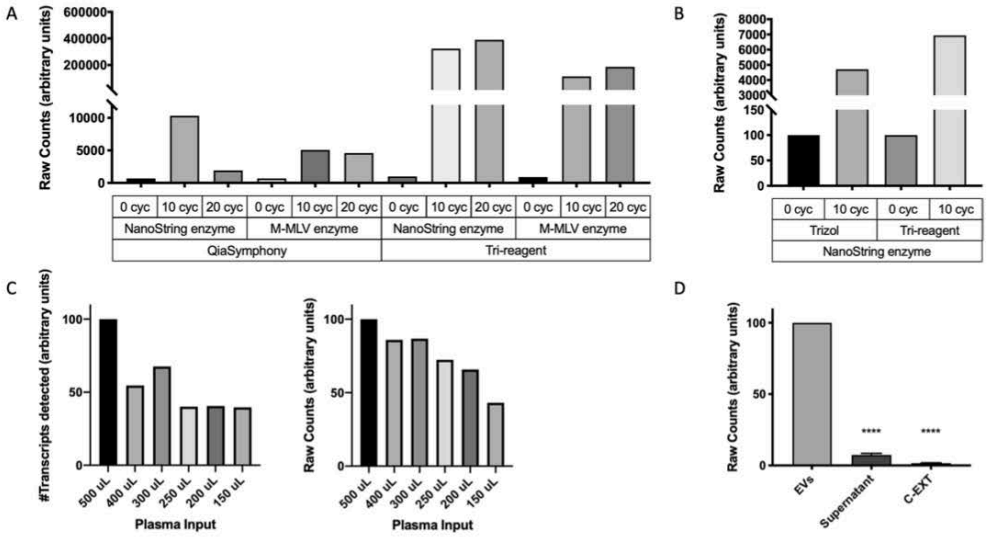


Figure 2. EV-RNA extraction, targeted pre-amplification and plasma input testing. (**A**) Total nCounter counts after automated (QiaSymphony) versus manual (TRI reagent) RNA extraction from an EV-enriched pellet. Two different retrotranscriptases (NanoString versus M-MLV) and three pre-amplification conditions (0, 10 and 20 cycles) were tested. Results were normalized to the counts corresponding to 0 cycles. (**B**) Total nCounter counts after TRIzol LS versus TRI reagent manual RNA isolation from an EV-enriched pellet. Results were normalized to the counts corresponding to 0 cycles. (**C**) Effect of input plasma volume (150–500 µL) on the final number of transcripts detected (left) or total nCounter counts (right) Results were normalized to the counts corresponding to 500 µL plasma. (**D**) Total nCounter counts of different fractions obtained during EV enrichment of plasma (EVs vs. supernatant $p < 0.0001$ in a one-way ANOVA with Dunnett's multiple comparisons test; EVs versus C-EXT $p < 0.0001$). Cyc, cycles; EVs, extracellular vesicles; C-EXT, extraction control.

Sediments, supernatants and plasmas were negative for the cell-specific marker calnexin (Fig. 1b, Supplementary Fig. S1). Cryogenic electron microscopy (cryo-EM), a commonly used technique for EV characterization[26,27], was used to visualize the miRCURY sediments, revealing EVs with the classical morphology and a diameter of 100–300 nm, in agreement with the reported 10–1 µm size range (Fig. 1c)[2–4].

TRIzol LS and TRI reagent are mixtures of phenol, guanidine isothiocyanate and other components routinely used for nucleic acid extractions. We found that the quantity of RNA that could be isolated from EV-enriched sediments using TRI reagent was too low to be determined by the Qubit RNA High Sensitivity Assay Kit (Thermo Fisher Scientific). Bioanalyzer profiles of two representative samples revealed RNA concentrations < 150 pg/µL, insufficient for nCounter analysis using the Human Immunology V2 Panel, which we had selected for our study (Supplementary Fig. S2). Therefore, we tested retrotranscription and pre-amplification of the EV-mRNA with the nCounter Low RNA Input Amplification Kit, using primers targeting the genes of the Panel. Two reverse transcriptases were compared, the M-MLV and the enzyme provided by the kit, together with 10 versus 20-cycles for the pre-amplification step. Results indicated that the retrotranscriptase provided by the kit was more efficient in terms of final counts and that both 10 and 20 cycles yielded sufficient raw counts for successful nCounter analysis with the Human Immunology V2 Panel (Fig. 1d). However, 20-cycles of pre-amplification led to saturation for some genes with higher expression levels (Supplementary Fig. S2). In consequence, 10 cycles were selected for the final workflow.

To test the reproducibility of the steps described above, we retrotranscribed and pre-amplified the same EV-RNA sample on three independent reactions. Then, we compared the results obtained when submitting the three resulting cDNA samples to nCounter analysis. A strong correlation was found between the normalized counts for each individual gene obtained in the different cDNAs, represented by a Spearman's r of 0.84–0.91, $p < 0.01$ (Fig. 1e).

All the experiments so far described had been performed with 500 µL plasma samples and TRI reagent. Two additional RNA extraction methods were tested on EV-enriched preparations from control samples, the automated QiaSymphony and the manual TRIzol LS Reagent based isolation (Fig. 2a). When considering the total number of counts by nCounter, TRI reagent was found to outperform both QiaSymphony and TRIzol LS, independently of the retrotranscriptase or the number of cycles used for the pre-amplification step (Fig. 2b). Finally, we also tested the effect of plasma input volume on downstream analysis of EV-RNA on the nCounter platform. Both the total counts and number of transcripts detected were higher with an initial plasma volume of 500 µL (Fig. 2c).

Characteristics	Cancer patients (n = 19)	Controls (n = 10)
Gender—no. (%)		
Male	11 (57.9)	5 (50.0)
Female	8 (42.1)	5 (50.0)
Age—year		
Median	62	42
Range	45–78	24–53
Tumor type—no. (%)		
Lung (ADC)	10 (52.7)	–
Gastric (ADC)	3 (15.8)	–
Anal (SCC)	2 (10.5)	–
Rectal (ADC)	2 (10.5)	–
Sarcoma (UPS)	2 (10.5)	–
Stage—no. (%)		–
Stage III	2 (10.5)	–
Stage IV	17 (89.5)	–

Table 1. Clinical characteristics of the cancer patients and controls included in the study. ADC, adenocarcinoma; SCC, squamous cell carcinoma; UPS, undifferentiated pleomorphic sarcoma.

RNase A is a bovine enzyme that can be used to degrade RNA in a sample. When enriching for EVs from plasma samples, we expect two sources of RNA; extra-vesicular RNA and RNA derived from within the EVs. Our workflow for EV-RNA analysis incorporates an RNase A treatment of the EV-enriched pellets in order to remove extra-vesicular RNA not embedded into the particles. To further validate this step, we collected two EV-enriched pellets and the corresponding supernatants of healthy patients, treated them with RNase A, purified the RNA and analyzed them with the Human Immunology V2 Panel. We found that EV-enriched pellets yielded significantly higher transcript counts compared to the supernatant (one-way ANOVA with Dunnett's multiple comparisons, EVs versus supernatant $p < 0.01$), indicating that the transcripts detected by nCounter are associated with EVs (Fig. 2d). To validate the efficacy of our RNase treatment, we also checked the expression of *GAPDH* and *CCL5* by qRT-PCR in three patient samples. Three different sample conditions were used; EVs without RNase treatment, EVs with RNase treatment and lysed EVs with RNase treatment. Results indicated that RNase treatment of intact EVs slightly reduced the *GAPDH* and *CCL5* mRNA levels, probably by removing the extra-vesicular RNA co-precipitated with the EVs (Supplementary Fig. S3). In contrast, once EVs were lysed, RNase treatment effectively eliminated all RNA and both *GAPDH* and *CCL5* transcripts were undetectable, confirming that transcripts purified using our workflow are indeed contained within the EVs.

EV-derived mRNA analysis on nCounter and classifier development. Based on the data presented above, we selected a workflow for subsequent experiments; starting with EV enrichment from 500 µL of plasma with the miRCURY kit, followed by manual RNA purification with TRI reagent, retrotranscription with the Nanostring enzyme, a 10-cycle pre-amplification and final cDNA analysis by nCounter. We validated the proposed workflow by studying the EV-mRNA expression of 19 cancer patient and 10 control plasma samples using the Human Immunology V2 Panel (Table 1). If all samples were considered together, the average number of transcripts detected was 430 ± 79 out of the 594 transcripts in the panel (Fig. 3a). No significant differences were found between the number of mRNA transcripts in EVs from controls versus cancer patients (445 ± 68 and 422 ± 84 respectively, Mann–Whitney's U $p = 0.46$).

After normalization, we analyzed the differential expression (DE) of transcripts in EVs from cancer patients versus controls (Fig. 3b, Supplementary Table S1). We found 141 mRNAs with significantly different levels; of them, 107 were upregulated and 34 downregulated in the EVs from cancer patients versus controls. Then, we used a recursive feature elimination (RFE) method to select a gene signature predictive of the origin of the sample; a cancer patient or a control sample. The transcripts included in the final signature were *BCL10*, *CXCL11*, *CYBB* and *GBP1* and, based on their expression levels, our algorithm was able to classify plasma samples into cancer and control with a receiver operating characteristic (ROC) area under the curve (AUC) of 0.92 to 0.95 (Fig. 3c). The classifier also calculated signature scores for each sample, which were found to be significantly different between cancer patients and controls (Mann–Whitney U, $p < 0.01$; Fig. 3d).

Improvement of classifier performance through removal of genomic DNA. The TRI reagent-based RNA isolation incorporated in our workflow for EV-RNA purification may also co-extract EV-genomic DNA, which could bind to the nCounter probes during hybridization. In consequence, some of the counts detected can correspond to genomic DNA instead of mRNA transcripts, particularly considering that we used a 10-cycle pre-amplification step. To test if this was the case, we analyzed the effect of adding a DNase treatment step after EV-RNA extraction in the 28 eV samples with remaining material. A dramatic reduction in the number of transcripts detected and the total amount of counts was observed in the EV-RNA samples after DNase treatment (Fig. 4a, Mann–Whitney U, $p < 0.01$ in both cases). The average number of transcripts in the 28 sam-

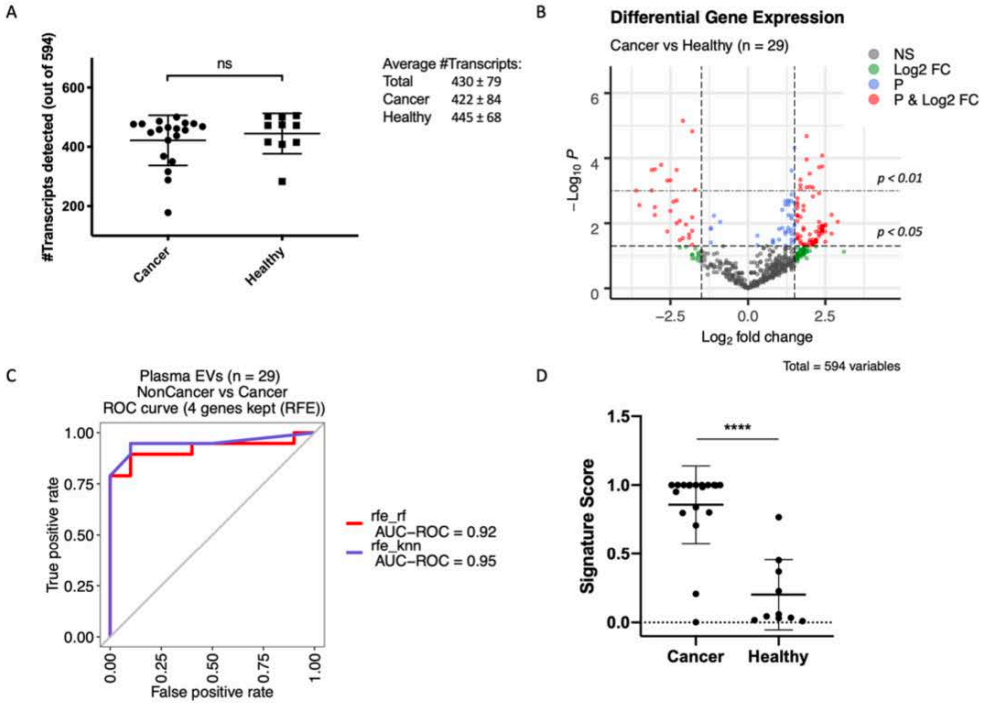


Figure 3. EV-mRNA transcript detection, differential expression analysis and development of a classifier algorithm. (**A**) Number of transcripts detected in EVs from cancer patients and healthy controls using the Human Immunology v2 nCounter panel, which targets 594 genes (two-tailed Mann–Whitney U test, $p = 0.463$). (**B**) Differential expression analysis of log2-normalized counts between cancer patients and healthy controls. The full list of genes differentially expressed is presented in Supplementary Table S1. (**C**) Area under the ROC curve of the four-gene signature, selected using recursive feature elimination (RFE), to differentiate cancer from control samples. (**D**) Scores of cancer versus control samples based on expression of the four-gene signature ($p < 0.001$ in a two-tailed Mann–Whitney U test). NS, not significant; EVs, extracellular vesicles; ROC, receiver operating characteristic; AUC, area under the curve; RFE, recursive feature elimination; rf, random forest; knn, k-nearest neighbors.

ples decreased from 430 ± 79 to 115 ± 66 and the total counts dropped more than 90%. As previously observed in samples without DNase treatment, the number of transcripts detected in cancer versus control samples were not significantly different (122 ± 77 vs 103 ± 39, Mann–Whitney's U $p = 0.92$, Fig. 4b). In addition, there were no significant differences in the number of transcripts detected between the different tumor types; which included sarcoma, lung, rectal, anal and gastric cancer (Supplementary Fig. S4).

Based on these results, we decided to incorporate a DNase step in our EV-mRNA purification and analysis workflow when working with this panel (Fig. 5), and we confirmed that 500 µL still yielded the highest number of transcripts and total counts (Fig. 4c). To further validate the DNase step, we analyzed *GAPDH*, *CCL5* and *Caspase 8 (CASP8)* expression by qRT-PCR in the 22 eV samples with remaining RNA. We observed a statistically significant correlation between *GAPDH* and *CCL5* expression as measured by nCounter and qRT-PCR (Spearman's r = 0.545, $p < 0.01$ and r = 0.850, $p < 0.01$, respectively; Fig. 4d); while *CASP8* transcripts were undetectable in DNase treated EV-RNA by both techniques.

Next, we proceeded to redesign our gene signature and classifier algorithm using the expression data derived from DNase treated EV-mRNAs. We observed that the number of differentially expressed genes in cancer versus control samples dropped from 141 to 43 (Fig. 6a and Supplementary Table S2). Of them, 17 genes (40%) overlapped with those obtained using non-DNase treated samples (Fig. 6b). Then, we used the RFE algorithm to create a second signature, which included eight genes: *CCL5, S100A9, B2M, HLA-B, IL7R, ICAM3, ARHGDIB and PYCARD* (Fig. 6a and Supplementary Table S3). The algorithm based on this signature was able to classify the samples into cancer and control with a ROC AUC of 0.99 to 1.00 (Fig. 6c). Also, the signature scores for each sample were found to be significantly different between cancer patients and controls (Mann–Whitney U, $p < 0.01$; Fig. 6d).

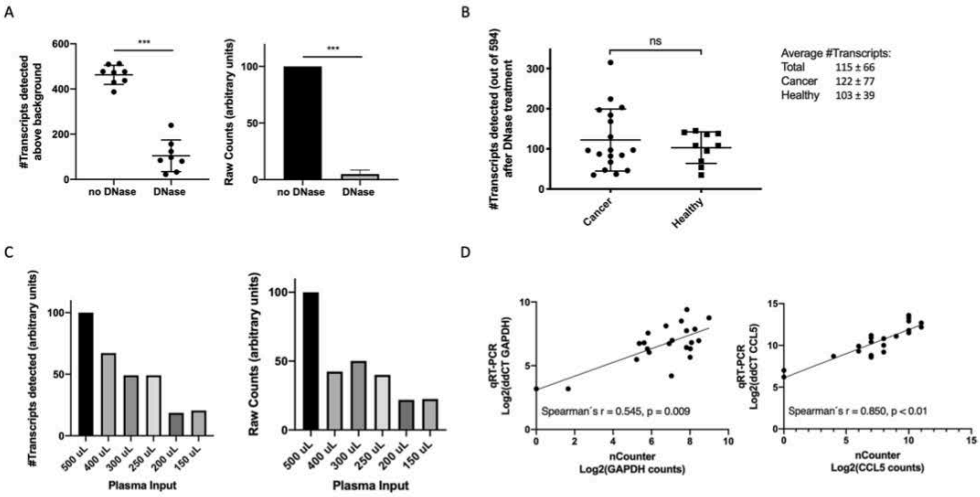


Figure 4. Effect of DNase treatment on the number of transcripts detected, total counts, and additional validation experiments. (**A**) Number of transcripts detected (left) and total nCounter counts (right) in non-DNase vs. DNase treated EV samples ($p < 0.001$ in a two-tailed Mann–Whitney U test, in both cases). (**B**) Number of transcripts detected in EVs from cancer patients and control samples using the Human Immunology v2 nCounter panel, after DNase treatment ($p = 0.916$ in a two-tailed Mann–Whitney U test). (**C**) Effect of input plasma volume (150–500 µL) on the final number of transcripts detected (left) or total nCounter counts (right), after DNase treatment. Results were normalized to the counts corresponding to 500 µL plasma. (**D**) Comparison of Log2 counts by nCounter versus ddCT values by qRT-PCR for *GAPDH* and *CCL5* in 22 samples. Spearman's correlation coefficient is indicated. NS, not significant.

The significant decrease in counts after DNase treatment prompted us to further investigate the binding of nCounter probes to genomic DNA co-purified with EV-mRNA. The nCounter Human Immunology V2 Panel targets 594 gene transcripts. The probes for some genes are designed within the same exon ("non-intron spanning"); while other probes ("intron spanning") target sequences corresponding to contiguous exons of the cDNA/mRNA. Genomic DNA should only bind "non-intron spanning" probes, and we performed an additional experiment to confirm this point. We analyzed an EV sample in quadruplicate, skipping the retrotranscriptase and/or the DNase treatment (Supplementary Table S4). Without DNase treatment, we observed counts for "non-intron spanning" probes (such as *CASP8*, *TNFRSF8* and *B2M*) both in presence and in absence of the retrotranscriptase step. In contrast, counts for "intron spanning" probes (such as *PRKCD*, *TNFRSF10C*, *FCER1G*) were apparent only if retrotranscription was performed. Finally, DNase treatment induced a significant drop in the counts of "non-intron spanning" but, unexpectedly, also "intron spanning" probes.

Discussion

The molecules found within EVs, such as mRNAs, are often involved in intercellular communication and represent a potential source for biomarker discovery. However, lack of standardized methods and clinical validation prevents the implementation of EV-derived testing in daily practice. The nCounter platform, which allows for multiplex detection of hundreds of transcripts, has been extensively used in translational research for transcriptomic tumor characterization. In addition, the nCounter Prosigna assay, based on a 50-gene expression signature, has been fully standardized and validated at the clinical level; and received FDA approval in 2013 to predict risk of recurrence in breast cancer[15,16,28]. However, studies investigating the performance of nCounter for mRNA analysis in liquid biopsies are scarce, particularly in the case of EVs.

Here, we present a workflow for nCounter-based analysis of plasma-derived EVs and we demonstrate that it can be used to efficiently detect transcripts in EV-enriched preparations from cancer patients and control samples. The first step of the workflow is plasma processing using the precipitation-based miRCURY kit which, as previously described[29,30], yielded pellets enriched in EVs. Clinical laboratories do not usually have access to ultracentrifugation and precipitation kits offer several advantages, such as short turn-around time and limited technical requirements. Although 500 µL plasma was found to give the highest amount of detected transcripts and total counts by nCounter, lower plasma inputs also yielded valid results. The second step is manual RNA isolation using TRI reagent, which was found to outperform manual purification by TRIzol LS and automated extraction.

The extraction methods tested, including the TRI reagent based manual extraction, did not yield sufficient RNA for direct analysis by nCounter. In consequence, after some optimization experiments, a 10-cycle pre-amplification step was added to our workflow and found to be highly reproducible. Most RNA extraction methods are

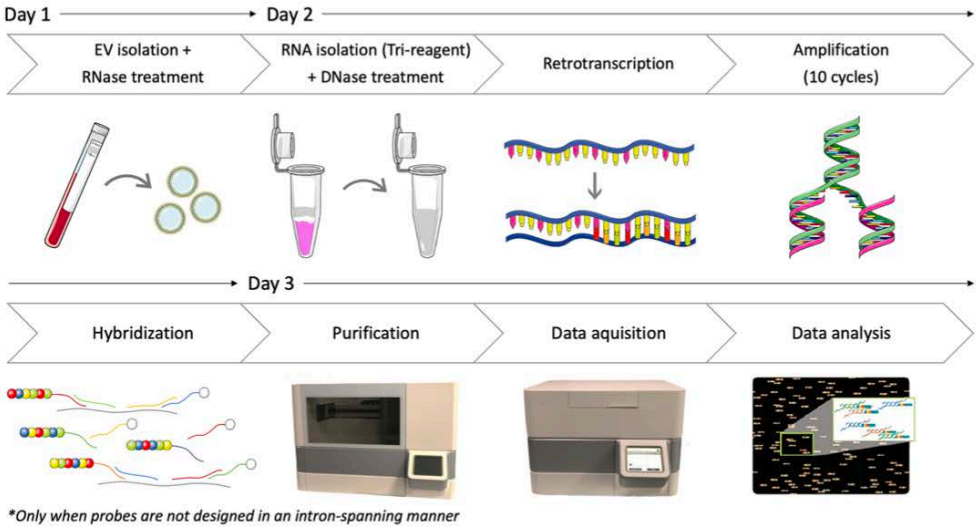


Figure 5. Final workflow for EV-RNA extraction and analysis on the nCounter platform. The miRCURY kit was used to enrich EVs from 500 µL plasma and EV-enriched preparations were treated with RNase to remove extra-vesicular RNA. Then, EV-RNA was extracted using TRI reagent and treated with DNase to remove genomic DNA. DNase treatment is only necessary when probes are not designed in an intron-spanning manner. Next, we performed retrotranscription and a 10-cycle pre-amplification, followed by hybridization, purification on the nCounter prep-station and analysis on the nCounter digital analyzer. EVs, extracellular vesicles. Part of this figure was modified from SMART (Servier Medical Art), licensed under a Creative Common Attribution 3.0 Generic License. http://smart.servier.com/.

known to co-purify genomic DNA (gDNA), and this was also our case. The EV-associated gDNA could be located in the interior of the vesicles or be attached to the membranous surface, and further research is needed to clarify this issue. More importantly, it has been described that simultaneous isolation of EV-derived genomic DNA during EV-RNA extraction can affect the amount of detected transcripts[31]. In consequence, we tested the effect of adding a DNase treatment step to our methodology and found a significant decrease in total counts and number of transcripts detected. Interestingly, while the expression of many genes became undetectable after DNase treatment, the counts for some transcripts were maintained. These observations suggested that co-extracted EV-genomic DNA was amplified during the pre-amplification step and could hybridize to the "non-intron spanning" probes in the nCounter panel. In contrast, probes designed in an "intron-spanning" manner should not hybridize to EV-DNA. Validation experiments confirmed this point (Supplementary Table S4). Unexpectedly, we observed a sharp decrease also in the counts of "intron-spanning" probes after DNase treatment, strongly suggestive of EV-RNA degradation or, more likely, EV-RNA loss during the purification steps needed to remove the DNase. In consequence, appropriate design of probes should be preferred over DNase treatment when analyzing samples with low amounts of RNA that require pre-amplification. Since we were using a pre-designed panel that could not be modified, we added a DNase treatment step to our final protocol for EV-mRNA analysis (Fig. 5). As an additional validation of the entire workflow, we re-analyzed 22 DNase treated samples for *GAPDH, CASP8* and *CCL5* expression by qRT-PCR. Similarly to previous reports[12,32], we found a statistically significant correlation between the expression levels obtained by qRT-PCR and nCounter.

Next, we investigated if our workflow for EV-mRNA analysis could be used to develop gene signatures. To this end, we performed differential expression analysis of the nCounter results obtained for control and cancer samples. The signature-based algorithm obtained for EV-mRNA (DNase treated) showed an improved classifier performance in comparison with EV-nucleic acids (non-DNase treated); with ROC-AUCs of 0.99–1.00 versus 0.92–0.95 for the discrimination of control versus cancer, respectively. This result is coincident with a previous report where a DNA removal step was shown to reduce signature noise during algorithm development and yielded better classification results[33]. The genes selected for our EV-mRNA expression signature are *CCL5*, *S100A9*, *B2M*, *HLA-B*, *IL7R*, *ICAM3*, *ARHGDIB* and *PYCARD*; which are involved in several immune-related pathways such as cytokine signaling, innate immune system or lymphocyte activation (Fig. 6a and Supplementary Table S3). *CCL5*, *B2M*, *HLA-B* and *IL7R* are all related to cytokine signaling and lymphocyte activation, and are thus known as immunomodulators. Interestingly, these four transcripts were downregulated in cancer samples, suggesting differences in immune system activation through cytokine induction between cancer patients and controls. *PYCARD* is also proposed to be involved in lymphocyte activation and was found to be upregulated in

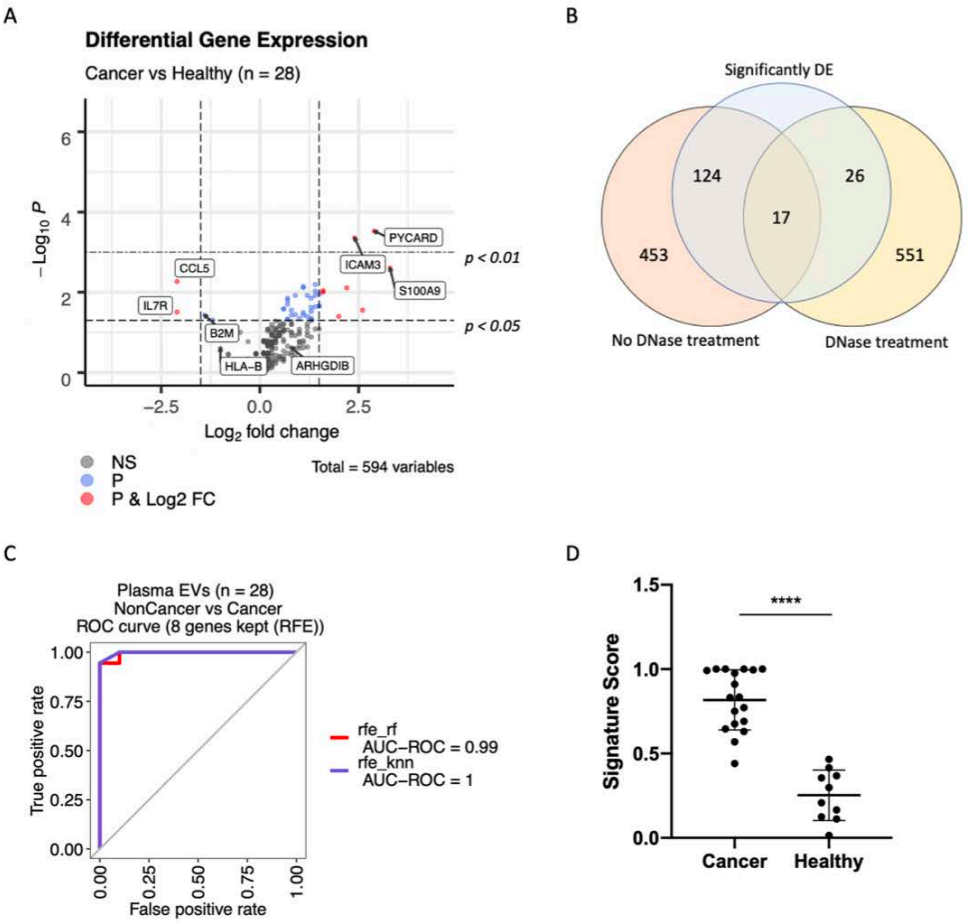


Figure 6. Effect of DNase treatment on differential expression analysis and classifier development. (**A**) Differential expression analysis of log2-normalized counts after DNase treatment between cancer patients and control samples. The full list of genes differentially expressed is presented in Supplementary Table S2. Labels indicate the eight transcripts selected for the final classification signature. (**B**) Out of 594 genes in the panel, 17 showed differential expression independently of DNase treatment. (**C**) Recursive feature elimination (RFE) was used to select an eight-gene signature that could distinguish between samples derived from cancer patients and controls. (**D**) Scores of cancer versus control samples based on expression of the eight gene signature ($p < 0.001$ in a two-tailed Mann–Whitney U test). DE, differentially expressed; EVs, extracellular vesicles; ROC, receiver operating characteristic; AUC, area under the curve; RFE, recursive feature elimination; rf, random forest; knn, k-nearest neighbors.

cancer patients. A previous study found that *PYCARD* could suppress apoptosis of cancer cells in gastric cancer[34], potentially explaining the observed higher expression in EVs from cancer patients.

Two studies have used nCounter for analysis of mRNA isolated from blood. Kossenkov et al[23] developed a pulmonary node classifier to differentiate malignant from non-malignant nodules previously detected by low-dose CT. Since the authors made use of whole blood, the quantities of purified RNA were significantly higher than those in our EV-based study (3 µg/2.5 mL blood versus 0.01 µg/2.5 mL, respectively), avoiding the need for a pre-amplification step. Beck et al[21] used a combination of CTC-RNA and cfRNA to profile tumor-associated biomarkers and correlate them with diagnostics and survival. The quantities of RNA were similar to those obtained in our study and a pre-amplification step was also added.

Our study shows several limitations. First, although purification of EVs from plasma yielded pellets that were found to be enriched in EVs, such precipitation techniques can also isolate a significant fraction of proteins and lipoproteins[35,36]. Second, there are no validated housekeeping genes for normalization of EV-mRNAs and we had to use the total amount of counts, as described[37–39]. Third, although nCounter presents many advantages, it has also a few limitations when compared to other multiplex techniques such as RNAseq. For instance, since nCounter works with gene panels, no new transcripts can be found. Finally, our aim was to establish a workflow for the nCounter analysis of EV-mRNA and the mixed patient population we used to demonstrate the validity of our approach was not the most appropriate to develop a clinically useful signature. Much larger patient cohorts would be needed to train and validate signatures that could differentiate between cancer patients and controls.

In summary, to the best of our knowledge, this proof-of-concept study is the first to demonstrate that the nCounter platform can be used to reproducibly detect plasma EV-mRNA transcripts. Differential expression analysis can then be implemented for biomarker assay development. Our work paves the way for widespread testing of EV-mRNA expression in blood and other fluids, and subsequent selection of signatures useful in the clinical setting.

Methods

Patient samples. This study was carried out in accordance with the principles of the Declaration of Helsinki under an approved protocol of the institutional review board of Quirón Hospitals. We selected for the study all blood samples from advanced stage cancer patients arriving to our institution in a two month period with sufficient volume for EV extraction after routine genetic testing (n = 19). Blood samples from 10 healthy controls were also collected (Table 1). Written informed consent was obtained from all participants and documented; samples were de-identified for confidentiality. Clinical information collected from each participant was limited to gender, age, tumor type and stage.

Extracellular vesicle enrichment. Whole blood samples (10 mL) were collected in sterile EDTA Vacutainer tubes (BD, Plymouth, UK) and centrifugated twice at 1000 × g for 10 min at room temperature (RT). Plasma samples were stored at − 80 °C until further processing. The miRCURY Exosome Serum/Plasma Kit (Qiagen, Hilden, Germany) was used to enrich for EVs from 500 µL plasma, according to the manufacturer's instructions, unless otherwise specified. In short, dead cells and debris (including platelets and fibrin) were cleared with thrombin and centrifugation. Precipitation Buffer was added, samples were incubated overnight at 4 °C and the EV fraction was pelleted by centrifugation. Supernatants were collected and stored for separate analysis. EV enriched pellets were resuspended for further processing.

EV characterization by western blot. EV enriched pellets were resuspended in 300 µL ice-cold radioimmunoprecipitation assay buffer containing protease inhibitor mixture (Roche Applied Science, Penzberg, Germany), as previously described[40]. Samples were incubated on ice for 30 min, homogenized and centrifugated 15 min, 12.000 × g at 4 °C. Supernatants (lysates) were collected and 80 µg proteins were electrophoresed on 10% SDS–polyacrylamide gels (Life Technologies, Carlsbad, CA, USA) and transferred to PVDF membranes (Bio-Rad Laboratories Inc., Hercules, CA, USA). Membranes were blocked in Odyssey Blocking Buffer (Li-Cor Biosciences, Lincoln, NE, USA). All target proteins were immunoblotted with appropriate primary and horseradish peroxidase (HRP)-conjugated secondary antibodies (Supplementary Table S5). Chemiluminescent bands were detected in a ChemiDoc MP Imaging System (Bio-Rad Laboratories Inc.).

EV characterization with cryogenic electron microscopy. Cryogenic Electron Microscopy (Cryo-EM) was performed by the Microscopy Service of the Universitat Autonoma de Barcelona (UAB), and used for direct visualization of EVs using the TEM JEOL 2011 200 kV. As previously described[41], 2 µL volume of the resuspended EV sample was added to a carbon TEM grid. The grid was transferred onto the cyro-preparation chamber of a Leica electron microscope, containing a liquid ethane bath cooled to -180 °C. Using a piece of filter paper, the EV solution was taken off the grid and plunged into the liquid ethane. The orifice trapped frozen EV solution was assembled into a plunger (Leica EM GP) and blotted with Whatman No. 1 filter paper. The grid was placed in a liquid nitrogen bath, and then loaded into a liquid nitrogen-cooled TEM grid holder. The grid holder was placed into a JEOL 2011 TEM microscope. Imaging was performed using a Gatan UltraScan US1000 CCD camera and data was analyzed with Digital Micrograph 1.8.

RNA extraction. EV-enriched pellets were treated with 4 µg/mL RNase A from bovine pancreas (Sigma-Aldrich, St. Louis, MO) for 1 h at 37 °C, to remove extra-vesicular RNA not associated to EVs. For TRIzol LS Reagent (Thermo Fisher Scientific, Waltham, MA) and TRI Reagent (MRC, Cincinnati, OH) extraction, TRIzol solutions were added to a final volume of 1 mL and incubated at RT for 20 min to inactivate RNase A and lyse the EVs. Then, 200 µL Chloroform: Isoamyl Alcohol (24:1) (Panreac Química SLU, Barcelona, Spain) was added and samples were vigorously vortexed and centrifuged at 12,000 × g for 15 min at 4 °C. The aqueous upper layer was kept and RNA was precipitated adding 2.5 µL Glycogen (Merck KGaA, Darmstadt, Germany) and 500 µL 2-propanol (Merck KGaA), incubating 10 min at RT and centrifugating for 10 min, 12,000 × g, at 4 °C. The final RNA pellet was washed with 75% ethanol, air dried and dissolved in 20 µL nuclease free water. The QIAsymphony DSP Virus/Pathogen Kit was also tested in the automated QIAsymphony SP System (Qiagen) for RNA extraction from EVs, according to the manufacturer's instructions.

Electrophoretic analysis of RNA. The approximate quantity and size distribution of the isolated EV-RNA was evaluated using the Bioanalyzer RNA 6000 Pico Assay (Agilent Technologies, Santa Clara, CA) according to manufacturer instructions.

DNase treatment. In order to remove co-isolated DNA, the EV-RNA samples were treated with the DNA-free DNA Removal Kit (Thermo Fisher Scientific), according to manufacturer instructions. In short, 1 µL DNase buffer and 0.5 µL enzyme were added to 7.5 µL RNA sample, followed by incubation at 37 °C for 30 min and DNase removal.

Gene expression analysis using nCounter. The nCounter Low RNA Input Amplification Kit (NanoString Technologies, Seattle, WA) was used to retrotranscribe and pre-amplify 4 µL EV-derived RNA using 10 cycles. Retrotranscription was carried out in 0.5 mL tubes while pre-amplification, using primers targeting the genes of the Human Immunology V2 Panel (NanoString Technologies), was performed in 384-well plates to prevent sample evaporation. In parallel, a Moloney Murine Leukema Virus (M-MLV) Reverse Transcriptase Enzyme (Thermo Fisher Scientific) was also tested for cDNA synthesis. The Human Immunology V2 Panel (NanoString Technologies) was used to analyze EV-derived, pre-amplified cDNA according to manufacturer instructions. This panel targets 594 general genes involved in the immune response such as cytokines, enzymes, interferons and their receptors. Samples were hybridized for 18 h at 65 °C.

Data normalization and analysis. Raw nCounter counts of expressed genes were normalized in R and R studio (version 3.6.3, https://cran.r-project.org/bin/macosx/) using the R package *NanoStringNorm*[42]. Normalization was performed following several steps: technical assay variability normalization using the geometric mean of the positive control probes, background correction using the mean plus two times standard deviation (SD) of the negative control probes, and sample content normalization using the total amount of counts for each sample. Normalized counts were log2-transformed, and used for differential expression (DE) analysis. Log2 fold change (FC) of each gene was calculated as the ratio of average log-2 transformed counts of the cancer patient cohort versus the control cohort. Volcano plots were used to visualize log2 FC on the x-axis and nominal p-values on the y-axis. GraphPad Prism software (version 9.0.0; https://www.graphpad.com/scientific-software/prism/) was used for other statistical testing and to create figures.

Classifier algorithm development. Optimal gene selection was performed using recursive feature elimination (RFE). To this end, a leave-one-out cross validation (LOOCV) algorithm was used on the full Human Immunology V2 gene Panel. The number of genes to select was set at 4, 8, 16 or 579 and the amount of genes that yielded optimal performance after cross-validation was automatically selected. Classification was performed with the selected gene signature using random forest (RF) and k-nearest neighbors (KNN) classifiers with three iterations. The model with the highest accuracy was then selected as the final model. Signature scores for each sample were derived from the final model.

Gene expression analysis using qRT-PCR. Complementary DNA (cDNA) was synthesized using the M-MLV Reverse Transcriptase Enzyme (Thermo Fisher Scientific). Hereafter, cDNA was added to TaqMan Universal Master Mix (Applied Biosystems) in a 12.5 µL reaction with specific primers and probe designed for each gene. The primer and probe sets were designed using Primer Express Software (version 3.0.1; https://www.thermofisher.com/order/catalog/product/4363993#/4363993) according to their Ref Seq (http://www.ncbi.nlm.nih.gov/LocusLink) Gene-specific primers were designed as follows: GAPDH, forward: 5′-TGACCTCAACTACATGGTTTACATGTT-3′ and reverse: 5′-TGACGGTGCCATGGAATTT-3′; Caspase 8, forward: 5′-CAGGGCTCAAATTTCTGCCTAC-3′ and reverse: 5′-GAAGAAGTGAGCAGATCAGAATTGAG-3′; CCL5, forward: 5′ CATCTGCCTCCCCATATTCCT 3′ and reverse: 5′ AGTGGGCGGGCAATGTAG 3′. Quantification of gene expression was performed using the QuantStudio 7 Flex Real-Time PCR System (Thermo Fisher Scientific). Expression levels of mRNA were expressed as arbitrary units based on Ct values. Commercial RNAs were used as controls (liver and lung; Stratagene, La Jolla, CA). In all quantitative experiments, a sample was considered not evaluable when the standard deviation of the Ct values was > 0.30 in 2 independent analyses.

Data availability

Supplementary information is available for this paper.

Received: 22 October 2020; Accepted: 29 January 2021
Published online: 12 February 2021

References

1. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **68**, 394–424. https://doi.org/10.3322/caac.21492 (2018).
2. Margolis, L. & Sadovsky, Y. The biology of extracellular vesicles: the known unknowns. *PLoS Biol.* **17**, e3000363. https://doi.org/10.1371/journal.pbio.3000363 (2019).
3. Reclusa, P. *et al.* Exosomes genetic cargo in lung cancer: a truly Pandora's box. *Transl. Lung Cancer Res.* **5**, 483–491. https://doi.org/10.21037/tlcr.2016.10.06 (2016).
4. Jeppesen, D. K. *et al.* Reassessment of exosome composition. *Cell* **177**, 428-445 e418. https://doi.org/10.1016/j.cell.2019.02.029 (2019).
5. Fujita, Y., Yoshioka, Y. & Ochiya, T. Extracellular vesicle transfer of cancer pathogenic components. *Cancer Sci.* **107**, 385–390. https://doi.org/10.1111/cas.12896 (2016).

6. Kalluri, R. The biology and function of exosomes in cancer. *J. Clin. Invest.* **126**, 1208–1215. https://doi.org/10.1172/JCI81135 (2016).
7. Nolte-'t Hoen, E. N. *et al.* Deep sequencing of RNA from immune cell-derived vesicles uncovers the selective incorporation of small non-coding RNA biotypes with potential regulatory functions. *Nucleic Acids Res.* **40**, 9272–9285. https://doi.org/10.1093/nar/gks658 (2012).
8. Stevic, I., Buescher, G. & Ricklefs, F. L. Monitoring therapy efficiency in cancer through extracellular vesicles. *Cells* https://doi.org/10.3390/cells9010130 (2020).
9. Reclusa, P. *et al.* Exosomes as diagnostic and predictive biomarkers in lung cancer. *J. Thorac. Dis.* **9**, S1373–S1382. https://doi.org/10.21037/jtd.2017.10.67 (2017).
10. Narrandes, S. & Xu, W. Gene expression detection assay for cancer clinical use. *J. Cancer* **9**, 2249–2265. https://doi.org/10.7150/jca.24744 (2018).
11. Segundo-Val, I. S. & Sanz-Lozano, C. S. Introduction to the gene expression analysis. *Methods Mol. Biol.* **1434**, 29–43. https://doi.org/10.1007/978-1-4939-3652-6_3 (2016).
12. Prokopec, S. D. *et al.* Systematic evaluation of medium-throughput mRNA abundance platforms. *RNA* **19**, 51–62. https://doi.org/10.1261/rna.034710.112 (2013).
13. Geiss, G. K. *et al.* Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat. Biotechnol.* **26**, 317–325. https://doi.org/10.1038/nbt1385 (2008).
14. Warren, S. Simultaneous, multiplexed detection of RNA and protein on the nanostring((R)) nCounter((R)) platform. *Methods Mol. Biol.* **1783**, 105–120. https://doi.org/10.1007/978-1-4939-7834-2_5 (2018).
15. Gnant, M. *et al.* Predicting distant recurrence in receptor-positive breast cancer patients with limited clinicopathological risk: using the PAM50 risk of recurrence score in 1478 postmenopausal patients of the ABCSG-8 trial treated with adjuvant endocrine therapy alone. *Ann. Oncol.* **25**, 339–345. https://doi.org/10.1093/annonc/mdt494 (2014).
16. Hannouf, M. B. *et al.* Cost-effectiveness analysis of multigene expression profiling assays to guide adjuvant therapy decisions in women with invasive early-stage breast cancer. *Pharmacogenomics J.* **20**, 27–46. https://doi.org/10.1038/s41397-019-0089-x (2020).
17. Ayers, M. *et al.* IFN-gamma-related mRNA profile predicts clinical response to PD-1 blockade. *J. Clin. Invest.* **127**, 2930–2940. https://doi.org/10.1172/JCI91190 (2017).
18. Giménez-Capitán, A. *et al.* Multiplex detection of clinically relevant mutations in liquid biopsies of cancer patients using the nCounter Platform. *Clin. Chem.* (2020) (**in press**).
19. Kamyabi, N. *et al.* Isolation and mutational assessment of pancreatic cancer extracellular vesicles using a microfluidic platform. *Biomed. Microdev.* **22**, 23. https://doi.org/10.1007/s10544-020-00483-7 (2020).
20. Porras, T. B., Kaur, P., Ring, A., Schechter, N. & Lang, J. E. Challenges in using liquid biopsies for gene expression profiling. *Oncotarget* **9**, 7036–7053. https://doi.org/10.18632/oncotarget.24140 (2018).
21. Beck, T. N. *et al.* Circulating tumor cell and cell-free RNA capture and expression analysis identify platelet-associated genes in metastatic lung cancer. *BMC Cancer* **19**, 603. https://doi.org/10.1186/s12885-019-5795-x (2019).
22. Wu, T. C. *et al.* IL1 receptor antagonist controls transcriptional signature of inflammation in patients with metastatic breast cancer. *Cancer Res.* **78**, 5243–5258. https://doi.org/10.1158/0008-5472.CAN-18-0413 (2018).
23. Kossenkov, A. V. *et al.* A gene expression classifier from whole blood distinguishes benign from malignant lung nodules detected by low-dose CT. *Cancer Res.* **79**, 263–273. https://doi.org/10.1158/0008-5472.CAN-18-2032 (2019).
24. Garcia-Contreras, M. *et al.* Plasma-derived exosome characterization reveals a distinct microRNA signature in long duration Type 1 diabetes. *Sci. Rep.* **7**, 5998. https://doi.org/10.1038/s41598-017-05787-y (2017).
25. Vicentini, C. *et al.* Exosomal miRNA signatures of pancreatic lesions. *BMC Gastroenterol.* **20**, 137. https://doi.org/10.1186/s12876-020-01287-y (2020).
26. Emelyanov, A. *et al.* Cryo-electron microscopy of extracellular vesicles from cerebrospinal fluid. *PLoS ONE* **15**, e0227949. https://doi.org/10.1371/journal.pone.0227949 (2020).
27. Rasch, M. R. *et al.* Hydrophobic gold nanoparticle self-assembly with phosphatidylcholine lipid: membrane-loaded and Janus vesicles. *Nano Lett.* **10**, 3733–3739. https://doi.org/10.1021/nl102387n (2010).
28. Dowsett, M. *et al.* Comparison of PAM50 risk of recurrence score with oncotype DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. *J. Clin. Oncol.* **31**, 2783–2790. https://doi.org/10.1200/JCO.2012.46.1558 (2013).
29. Pasquali, L. *et al.* Circulating microRNAs in extracellular vesicles as potential biomarkers for psoriatic arthritis in patients with psoriasis. *J. Eur. Acad. Dermatol. Venereol.* **34**, 1248–1256. https://doi.org/10.1111/jdv.16203 (2020).
30. Helwa, I. *et al.* A comparative study of serum exosome isolation using differential ultracentrifugation and three commercial reagents. *PLoS ONE* **12**, e0170628. https://doi.org/10.1371/journal.pone.0170628 (2017).
31. Naderi, A. *et al.* Expression microarray reproducibility is improved by optimising purification steps in RNA amplification and labelling. *BMC Genomics* **5**, 9. https://doi.org/10.1186/1471-2164-5-9 (2004).
32. Richard, A. C. *et al.* Comparison of gene expression microarray data with count-based RNA measurements informs microarray interpretation. *BMC Genomics* **15**, 649. https://doi.org/10.1186/1471-2164-15-649 (2014).
33. Esmaeili, M. *et al.* Noninvasive sexing of human preimplantation embryos using RT-PCR in the spent culture media: a proof-of-concept study. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **252**, 89–93. https://doi.org/10.1016/j.ejogrb.2020.06.023 (2020).
34. Deswaerte, V. *et al.* Inflammasome adaptor ASC suppresses apoptosis of gastric cancer cells by an il18-mediated inflammation-independent mechanism. *Cancer Res.* **78**, 1293–1307. https://doi.org/10.1158/0008-5472.CAN-17-1887 (2018).
35. Johnsen, K. B., Gudbergsson, J. M., Andresen, T. L. & Simonsen, J. B. What is the blood concentration of extracellular vesicles? Implications for the use of extracellular vesicles as blood-borne biomarkers of cancer. *Biochim. Biophys. Acta Rev. Cancer* **109–116**, 2019. https://doi.org/10.1016/j.bbcan.2018.11.006 (1871).
36. Sodar, B. W. *et al.* Low-density lipoprotein mimics blood plasma-derived exosomes and microvesicles during isolation and detection. *Sci. Rep.* **6**, 24316. https://doi.org/10.1038/srep24316 (2016).
37. Rozowsky, J. *et al.* exceRpt: a comprehensive analytic platform for extracellular RNA profiling. *Cell Syst.* **8**, 352-357 e353. https://doi.org/10.1016/j.cels.2019.03.004 (2019).
38. Morhayim, J. *et al.* Molecular characterization of human osteoblast-derived extracellular vesicle mRNA using next-generation sequencing. *Biochim. Biophys. Acta Mol. Cell Res.* **1133–1141**, 2017. https://doi.org/10.1016/j.bbamcr.2017.03.011 (1864).
39. Koppers-Lalic, D. *et al.* Noninvasive prostate cancer detection by measuring miRNA variants (isomiRs) in urine extracellular vesicles. *Oncotarget* **7**, 22566–22578. https://doi.org/10.18632/oncotarget.8124 (2016).
40. Bracht, J. W. P. *et al.* Osimertinib and pterostilbene in EGFR-mutation-positive non-small cell lung cancer (NSCLC). *Int. J. Biol. Sci.* **15**, 2607–2614. https://doi.org/10.7150/ijbs.32889 (2019).
41. Puente-Massaguer, E., Lecina, M. & Godia, F. Nanoscale characterization coupled to multi-parametric optimization of Hi5 cell transient gene expression. *Appl. Microbiol. Biotechnol.* **102**, 10495–10510. https://doi.org/10.1007/s00253-018-9423-5 (2018).
42. Waggott, D. *et al.* NanoStringNorm: an extensible R package for the pre-processing of NanoString mRNA and miRNA data. *Bioinformatics* **28**, 1546–1548. https://doi.org/10.1093/bioinformatics/bts188 (2012).

Acknowledgements
The investigators wish to thank the patients for kindly accepting to donate samples to this study. We also thank all the physicians who collaborated by providing clinical information. Finally, we would like to thank colleagues and collaborators for their constructive feedback on this manuscript.

Author contributions
J.W.P.B., C.P.V. and M.A.M.V. conceptualized and designed the experiments. J.W.P.B. and A.G.C. performed the experiments. N.P. and J.W.P.B. performed data analysis. J.W.P.B. and M.A.M.V. wrote the main manuscript and prepared the figures. C.Y.H., S.W. and R.R. provided editing, comments, and experimental guidance. All authors reviewed the manuscript.

Funding
This project has received funding from a European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement ELBA No. 765492.

Competing interests
C.Y.H. and S.W. were full-time employees of NanoString Inc. when the project was carried out. The rest of the authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-83132-0.

Correspondence and requests for materials should be addressed to J.W.P.B. or M.A.M.-V.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.