Received: 25 August 2022 | Revised: 24 May 2023 | Accepted: 19 June 2023

DOI: 10.1002/ijc.34658

IJC INTERNATIONAL JOURNAL of CANCER

RESEARCH ARTICLE

Tumor Microenvironment and Signatures

Ex vivo expansion of lung cancer-derived disseminated cancer cells from lymph nodes identifies cells associated with metastatic progression

Steffi Treitschke[1] | Kathrin Weidele[1] | Adithi Ravikumar Varadarajan[1] | Giancarlo Feliciello[1] | Jens Warfsmann[1] | Sybille Vorbeck[1] | Bernhard Polzer[1] | Catherine Botteron[1] | Martin Hoffmann[1] | Vadim Dechand[1] | Tobias Mederer[2] | Florian Weber[1,3] | Melanie Werner-Klein[2] | Tobias Robold[4] | Hans-Stefan Hofmann[4] | Christian Werno[1] | Christoph A. Klein[1,2]

[1]Fraunhofer Institute for Toxicology and Experimental Medicine, Division of Personalized Tumor Therapy, Regensburg, Germany

[2]Experimental Medicine and Therapy Research, University of Regensburg, Regensburg, Germany

[3]Institute for Pathology, University of Regensburg, Regensburg, Germany

[4]Department of Thoracic Surgery, University Hospital Regensburg, Regensburg, Germany

Correspondence
Christoph A. Klein, Fraunhofer Institute for Toxicology and Experimental Medicine, Division of Personalized Tumor Therapy and Experimental Medicine and Therapy Research, University of Regensburg.
Email: christoph.klein@ukr.de

Funding information
Bavarian Ministry of Economic Affairs, Energy, and Technology, Grant/Award Numbers: 20-3410.1-1-2, AZ 20-3410.1-1-1; Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Grant/Award Number: SFB TRR 305-B13; Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Grant/Award Number: FOR2127-A02andFOR2127-B05

Abstract

The cellular basis of the apparent aggressiveness in lung cancer is poorly understood but likely associated with functional or molecular features of disseminated cancer cells (DCCs). DCCs from epithelial cancers are mostly detected by antibodies directed against histogenetic markers such as cytokeratin or EpCAM. It has been argued that marker-negative metastatic founder cells might escape detection. We therefore used ex vivo sphere formation for functional detection of candidate metastasis founders. We generated cell suspensions from 199 LN samples of 131 lung cancer patients and placed them into non-adherent cell culture. Sphere formation was associated with detection of DCCs using EpCAM immunocytology and with significantly poorer prognosis. The prognostic impact of sphere formation was strongly associated with high numbers of EpCAM-positive DCCs and aberrant genotypes of expanded spheres. We also noted sphere formation in patients with no evidence of lymphatic spread, however such spheres showed infrequent expression of signature genes associated with spheres from EpCAM-positive samples and displayed neither typical lung cancer mutations (*KRAS*, *TP53*, *ERBB1*) nor copy number variations, but might be linked to disease progression >5 years post curative surgery. We conclude that EpCAM identifies relevant disease-driving DCCs, that such cells can be expanded for model generation and that further research is needed to clarify the functional and prognostic role of rare EpCAM-negative sphere forming cells.

Abbreviations: ADC, adenocarcinoma; ASC, adenosquamous carcinoma; BM, bone marrow; CK, cytokeratin; CNV, copy number variation; DCC, disseminated cancer cells; DCCD, disseminated cancer cell density (number of DCC per million LN cells); EpCAM, epithelial cell adhesion molecule; GII, genome integration index; LN, lymph node; n.a., not analyzed; NEC, neuroendocrinic carcinoma; NSCLC, non-small cell lung cancer; OS, overall survival; QC, quality control; SCC, squamous cell carcinoma; SCLC, small cell lung cancer; SP, sphere; STR, short tandem repeat; TSS, tumor-specific survival; WGA, whole genome amplification; wt, wildtype; WTA, whole transcriptome amplification.

Steffi Treitschke and Kathrin Weidele are considered as co-first authors.

Christian Werno and Christoph A. Klein are considered as co-senior authors.

KEYWORDS
disseminated cancer cells, lung cancer, metastatic precursor cells, sphere culture

What's new?

Metastatic lung cancer begins with disseminated cancer cells (DCC), but some patients experience disease progression despite testing negative for DCC. To better understand the cellular basis of metastasis, these authors characterized DCCs isolated from lymph node samples of lung cancer patients and cultured isolated cells in non-adherent cell culture. Sphere formation was associated with poorer prognosis and with higher numbers of cells containing epithelial cell adhesion molecules (EpCAM). In some cases, cells formed spheres even without much EpCAM expression. Whether these cells might predict future disease progression is a subject for future study.

1 | INTRODUCTION

Lung cancer is one of the most common malignancies in Europe and the leading cause of cancer-related deaths worldwide.[1] With a prevalence of approximately 80%, non-small cell lung cancer (NSCLC) is the most frequent lung cancer subtype. The overall 5-year relative survival rate for all stages combined is 18%[2] and even 20-25% of patients with a complete resection in stage IA/B will relapse,[3,4] because metastatic dissemination has obviously occurred prior to surgery. This apparent aggressiveness of lung cancer is poorly understood since tumor volume doubling times of NSCLC are not lower (i.e., growth is not faster) than those for other epithelial cancers such as breast or colorectal cancers.[5,6]

Hence, answers might be found by detection and analysis of metastasis founder cells. Prime candidates are disseminated cancer cells (DCCs) that can be detected in regional lymph nodes (LN) or bone marrow (BM) samples long before manifestation of distant metastasis.[7,8] Several studies have linked detection of DCCs in LN or BM of NSCLC patients with poor outcome, however, data have been conflicting. Some observed an association between the presence of DCCs and poor prognosis,[9-13] whereas others did not.[14,15] We recently analyzed the prognostic impact of DCCs of NSCLC patients in more detail. First, we compared EpCAM and cytokeratin (CK)-positive cells from BM of the same patient and found that the prognostic impact of EpCAM outperformed information provided by CK-positive cells.[16] Then, we tested whether EpCAM-positive cells from LNs confer prognostic information. Here, we found detection of EpCAM-positive DCCs to be the most informative variable for outcome prediction in multivariable analysis.[17]

Although EpCAM-positive cells have thus become prime candidates for metastasis founder cells, we set out to investigate the following questions. First, we asked whether sphere formation under stem-like-cell selecting conditions is associated with disease progression; second, whether sphere formation is associated with EpCAM-positivity in immunocytologically analyzed samples and third, if spheres are formed, which molecular features were associated with expansion and possibly with disease progression. To address these issues, we collected LN samples from lung cancer patients, quantified DCCs and subjected LN suspensions to culture conditions promoting propagation of stem-like DCCs. We molecularly characterized single spheres to more specifically identify potential metastatic founder cells.

2 | METHODS

2.1 | Patients

Between 2011 and 2016, we collected lymph node (LN) samples from 131 consecutive patients diagnosed with lung cancer at the University Hospital Regensburg or at the Krankenhaus der Barmherzigen Brüder Regensburg. The baseline characteristics of the patient cohort are summarized in Table S1. Tumor stage and grading were determined according to the international system for the staging of lung cancer.[18]

2.2 | Lymph node sample processing and detection of disseminated cancer cells

Lymph nodes were bisected by the surgeon directly after removal, resulting in half LNs that were processed in our laboratory for immunocytology staining[19] and sphere culture as described below. In total, we received 199 LN samples from 131 patients. This included patients of whom we received either one bisected LN (64 patients), two bisected LNs (66 patients) and three bisected LNs (1 patient). LNs were mechanically disaggregated using a Medimax machine (Syntec International), and cells were subsequently enriched via Percoll (GE Healthcare) gradient centrifugation.[20] Cell numbers were determined with a Neubauer counting chamber. A total number of 2×10^6 cells (at a density of 10^6 cells/ml PBS) were transferred onto adhesion slides (Thermo Scientific). After 1 h sedimentation, PBS was removed and slides were air-dried overnight at room temperature. For detection and quantification of DCCs, LN samples were stained with anti-EpCAM antibody (Ber-EP4, Dako) and visualized via Alkaline Phosphatase-anti-Alkaline Phosphatase antibody (APAAP, Bio-Rad) or an AP conjugate substrate kit (Bio-Rad).[21,22] The number of EpCAM stained cells was counted and expressed as disseminated cancer cell density (DCCD = number of DCC per million LN cells). Single positive cells from selected patients were manually picked from adhesion slides using a micromanipulator and subjected to whole genome amplification and molecular analysis as described below.

2.3 | Sphere cultures

To promote non-adherent sphere formation, 10^6 pre-enriched LN cells were cultured in polyHEMA-coated 6 cm plates (12 mg/ml, Sigma Aldrich) under low-oxygen conditions (7% O_2) at 37°C in 5 ml defined sphere medium. Since no information on tumor-type specific growth conditions for DCCs from lung cancer patients was available, we compared two non-adherent, serum-free sphere culture media conditions. First, a basic sphere culture medium (medium A), which was previously shown to support sphere formation of melanoma DCCs[23] or circulating tumor cells in metastatic breast cancer.[24,25] Medium A contained DMEM/F12 culture medium supplemented with 0.5% penicillin/streptomycin, 10 µg/ml insulin (all from PAN-Biotech), 10 nM HEPES (Sigma-Aldrich), 1× B27 supplement (Gibco), 10 ng/ml epithelial growth factor, 10 ng/ml basic fibroblast growth factor, 4 µg/ml heparin (all from Sigma-Aldrich), 5 ng/ml GRO-α (R&D Systems), 20 ng/ml hyper interleukin-6 (kindly provided by S. Rose-John) and 0.2% Methylcellulose (Sigma-Aldrich). Medium B was comprised of medium A supplemented with 20% mouse embryonic fibroblast conditioned medium (R&D Systems), which was shown to support cell growth by secretion of growth factors, cytokines and nutrients.[26] Since we observed no significant difference in sphere-formation between medium A and B within DCCD = 0 ($P = .3125$) and DCCD>0 groups ($P = .5885$) the type of medium was neglected for evaluation (Figure S1). Cultures were monitored at least once a week up to 6 weeks and complemented with 500 µl fresh medium once a week. Compact cell aggregates and well-rounded 3D structures bigger than 40 µm were counted and henceforth referred to as spheres. Photographic documentation was performed for each sphere before analysis.

2.4 | Combined whole transcriptome- and genome amplification (WTA and WGA)

Reverse transcription of mRNA and amplification of cDNA was performed as described previously.[27,26] The quality of amplified cDNA was evaluated with a set of 3 control primers using end-point PCR (quality control; QC).[28] Genomic DNA of spheres was collected during WTA procedure, precipitated and subjected to WGA.[29,30] DNA from isolated DCCs or supernatants from spheres were amplified using *Ampli1* WGA Kit (Menarini Silicon Biosystems). The quality of amplified DNA was evaluated with a set of 4 control primers using end-point PCR and assigned to genomic integrity index 0 to 4 (GII0-GII4).[31]

2.5 | Expression of transcripts using end-point PCR

Good quality WTA products (2-3 bands in QC) were diluted 1:10 with H_2O and analyzed for the expression of tumor and leukocyte associated transcripts using end-point PCR.[28] The following primer pairs were used (5′-3′): CD45 (*PTPRC*): forward TTAGGGACACGGCT GACTTC and reverse GCTTTGCCCTGTCACAAATA; EpCAM (*EPCAM*): forward AAAGTTTGCGGACTGCACTT and reverse AGCCA CATCAGCTATGTCCA; CK19 (*KRT19*): forward AGCCGGACTGAA GAATTGAA and reverse TTCTGCCAGTGTGTCTTCCA; CD133 (*PROM1*): forward ATGACAAGCCCATCACAACA and reverse TTGCTCCTGGATTTGGAAAG. PCR products were visualized on agarose gels. Samples expressing the analyzed transcript were used as controls.

2.6 | Mutational analysis

WGA products of good quality (3-4 bands in QC, GII3 or GII4,[31]) were diluted 1:2 in H_2O and used for PCR-based amplification of KRAS, EGFR and TP53. Following primer pairs were used (5′-3′): *KRAS* exon2 (covering mutations in codon12-13): forward GTGACATGTTCTAATA TAGTCAC and reverse CTCTATTGTTGGATCATATTCGT; *EGFR* exon 18: forward TTGTCCTTCCAAATGAGCTG and reverse TGCCTTTGGTCTGTGAATTG, *EGFR* exon 20 (covering mutations in codon 790): forward AAACGTCCCTGTGCTAGGTC and reverse CATGGCAAACTCTTGCTATCC; *EGFR* exon 21 (covering mutations around codon 858): forward CAGCGGGTTACATCTTCTTTC and reverse AAACAATACAGCTAGTGGGAAGG; *TP53* exon 5 (covering mutations around codon 157 and codon 175): forward ACGCATG TTTGTTTCTTTGC and reverse AGCAATCAGTGAGGAATCAGAG; *TP53* exon 5 and 6: forward ACGCATGTTTGTTTCTTTGC and reverse AGCAATCAGTGAGGAATCAGAG, *TP53* exon 7 (covering mutations around codon 248): forward GAGGCTGAGGAAGGAGAATG and reverse AGTATGGAAGAAATCGGTAAGAGG; *TP53* exon 8 (covering mutations around codon 273): forward AGGTAGGACCTGATTTCCT TACTG and reverse AGGCATAACTGCACCCTTG. PCR products were purified with PCR purification kit (Qiagen) and sent for Sanger sequencing (Eurofins).

2.7 | Detection of copy number variations (CNVs) using array comparative genomic hybridization (aCGH) and low-pass whole genome sequencing (WGS)

Genome-wide copy number analysis of DCCs and spheres was conducted using a previously published array comparative genomic hybridization (aCGH) protocol.[32] Low-pass is a sequencing approach targeting the whole genome with average coverage <1×. Libraries were prepared using Ampli1 LowPass kit for Illumina platforms according to the manufacturer's instructions (Menarini Silicon Biosystems). Libraries were quantified using Qubit dsDNA HS Assay kit and Qubit 2.0 Fluorometer (Thermo Fisher Scientific) and average fragment sizes were assessed using Agilent High Sensitivity DNA Kit on the Agilent 2100 Bioanalyzer System (Agilent Technologies). Libraries were mixed in equimolar concentrations to obtain a 4 nM final pool ready for direct sequencing. Sequencing was performed in single read (SR) mode on a MiSeq System with MiSeq Reagent Kit v3 (150-cycle) (Illumina). For CNV profile analysis the human genome was divided into non-overlapping bins, each with a size of 500 kb. After removing adapter sequences and poor-quality bases (*BBDuk 38.84*[33]) as well as decontamination of non-human reads originating from microbial and/or fungal fresh water and/or reagents

(*BioBloom Tools 2.0.13*[34]), for each sample, reads aligned with *bwa mem 0.7.17*[35] to the reference genome hg19 were counted per bin, corrected, filtered, normalized and segmented with the bioconductor package *QDNAseq 1.26.0*.[36] The same package was used to create the *log2* (*ratio*) CNV profiles. Samples with <100.000 mapped reads were rejected from CNV-analysis. The sequencing coverage and quality statistics for each low-pass WGS sample are summarized in Table S2.

2.8 | RNA-sequencing (RNA-Seq)

Good quality WTA products from 25 spheres were processed for RNA sequencing as previously described.[37] Sequencing libraries were prepared according to the TruSeq DNA PCR-Free Library Prep Kit (Illumina). Libraries were quantified with KAPA Library Quantification Kit for Illumina Platforms (Roche), pooled in equal molar ratios and sequenced using Illumina NextSeq 2000. Fastq reads were processed using inhouse RNA-Seq workflow. The sequencing coverage and quality statistics for each RNA-Seq sample are summarized in Table S3. Initial read processing including quality control, trimming and decontamination were performed as described above in the low-pass WGS section. Following this, reads were aligned to the hg38 reference genome with STAR 2.7.10b.[38] FeatureCounts from Subread 2.0.0[39] was used to extract counts per gene per sample. Only uniquely mapping reads to exonic regions were used for counting. Library complexity (using Preseq 2.0.3[40]), genomic origin of the reads and the 5′-3′-bias (both using QualiMap 2.2.2d[41]) were evaluated from the mapping results. Final counts table were utilized to uncover the cellular heterogeneity of the spheres. Following steps were performed in R programming language 4.2.1. As first steps, gene counts were log normalized and scaled. Log (counts) as well as log (cpm) and top 500 differential expressed genes were visualized using Principle Component Analysis (PCA) and t-distributed stochastic neighbor embedding (t-SNE) clustering techniques. RunPCA and RunTSNE from Seurat 4.2.0[42] were used to construct PCA and t-SNE, respectively. Finally, using the non-parametric ssgsea method in gene set variation analysis (GSVA) package 1.44.5[43] the spheres were deconvoluted to identify different cell types. Subset of cell types from lung from CellMarker database,[44] a manually curated resource, was used as gene set in gsva against which the count table was evaluated. All the plots were made using ggplot2 3.4.2.[45]

2.9 | DNA fingerprinting

Patient origin was confirmed using short tandem repeat (STR) analysis (Geneprint10, Promega). Due to WGA of DCC and sphere samples, which includes restriction digestion by Mse I prior to STR analysis, in some cases only STR loci TH01, D21S11, D5S818, D13S317, D16S538 and vWA were accessible for analysis. Amplified fragments were detected using 3130-Avant Genetic Analyzer (Applied Biosystems). Fragment sizes were determined and evaluated using GeneMapperSoftware Version 4.1. STR-profile of cultured spheres were compared with the profile of the corresponding DCC. All analyzed patient samples matched each other. The obtained STR profiles were additionally analyzed using a public database (https://web.expasy.org/cellosaurus-str-search/). This comparison showed no match with other cell lines. Patient samples are therefore revealed as unique and not cross-contaminated or misidentified with cells of other origin.

2.10 | Statistical analysis, follow up and survival endpoints

Statistical analysis was performed with GraphPad Prism Software, version 9, using unpaired, non-parametric Mann-Whitney test, 1-way ANOVA or fisher's exact test to determine associations between two categorical variables. For survival analysis Log-rank (Mantel-Cox) test for comparing survival curves. Asterisks represent the significance as follows: ***P < .001, **P < .01, *P < .05 and ns: $P \geq .05$.

Overall survival (OS) was calculated as a period of time from the date of surgery to the date of death from any cause or the date of last follow-up. Tumor-specific survival (TSS) was defined as the duration from the date of surgery until death that was documented to be caused by lung cancer.

3 | RESULTS

3.1 | Detection of EpCAM-positive cells in lymph nodes and ex vivo cultivation

We investigated the presence of EpCAM-positive DCCs by immunocytological quantification and concomitantly cultivated cell suspensions of 199 lymph node samples from 131 lung cancer patients (Figure 1A, patient demographics listed in Table S1). DCCs were detected in 49 of 199 analyzed LNs (24.6%), with a calculated disseminated cancer cell density (DCCD = EpCAM-positive DCCs per one million LN cells) ranging from 0.5 to 750,000 (median = 3; Figure S2).

In order to investigate the proliferative potential of DCCs from LNs, LN-derived cell suspensions were cultivated under non-adherent, serum-free sphere culture conditions using two different media. We observed morphologically heterogeneous sphere structures among different patients, but also within individual patient samples. Spheres differed in size (range 40-800 µm in diameter), color (ranging from light gray to deep black, which may indicate inclusion of lung macrophages filled with tar) and compactness (ranging from clusters to compact spheroids; representative pictures in Figure 1B). Irrespective of their morphological appearance, all spherical structures were counted and documented microscopically for each LN culture.

We detected spheres growing in cultures from LNs without (DCCD = 0) and with EpCAM-positive DCCs (DCCD > 0; Figure 1C, D). In DCCD = 0 LN samples, spheres were growing in 30% of cultures (45/150), with sphere counts ranging from 0 to 190 spheres (median = 0, mean = 3.6 SD ± 17.4). In contrast, 49% of DCCD > 0 LNs (24/49) formed spheres ranging from 0 up to 132 (median = 0, mean = 10.1 SD ± 24.3; Fisher's exact test: $P = .0239$).

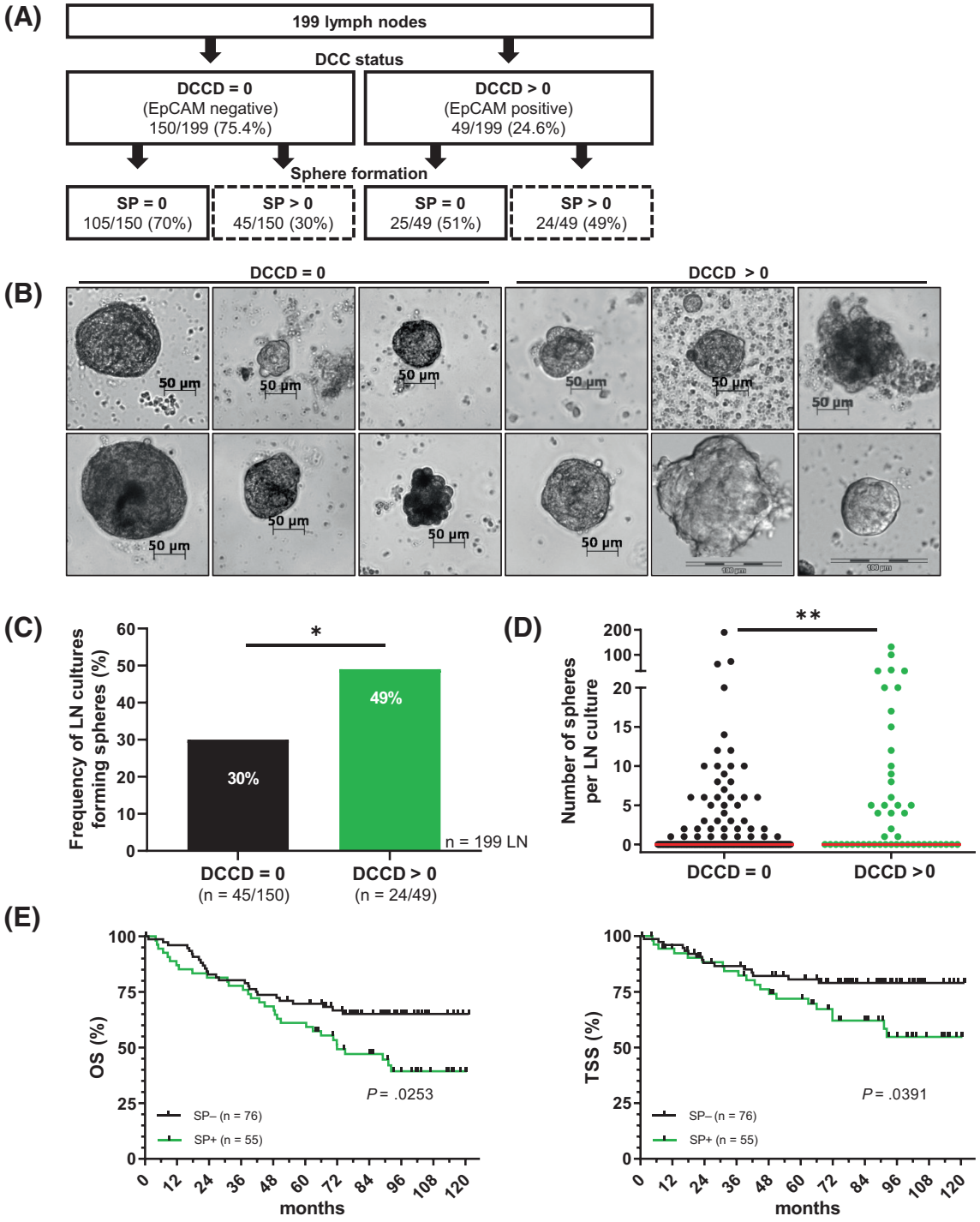
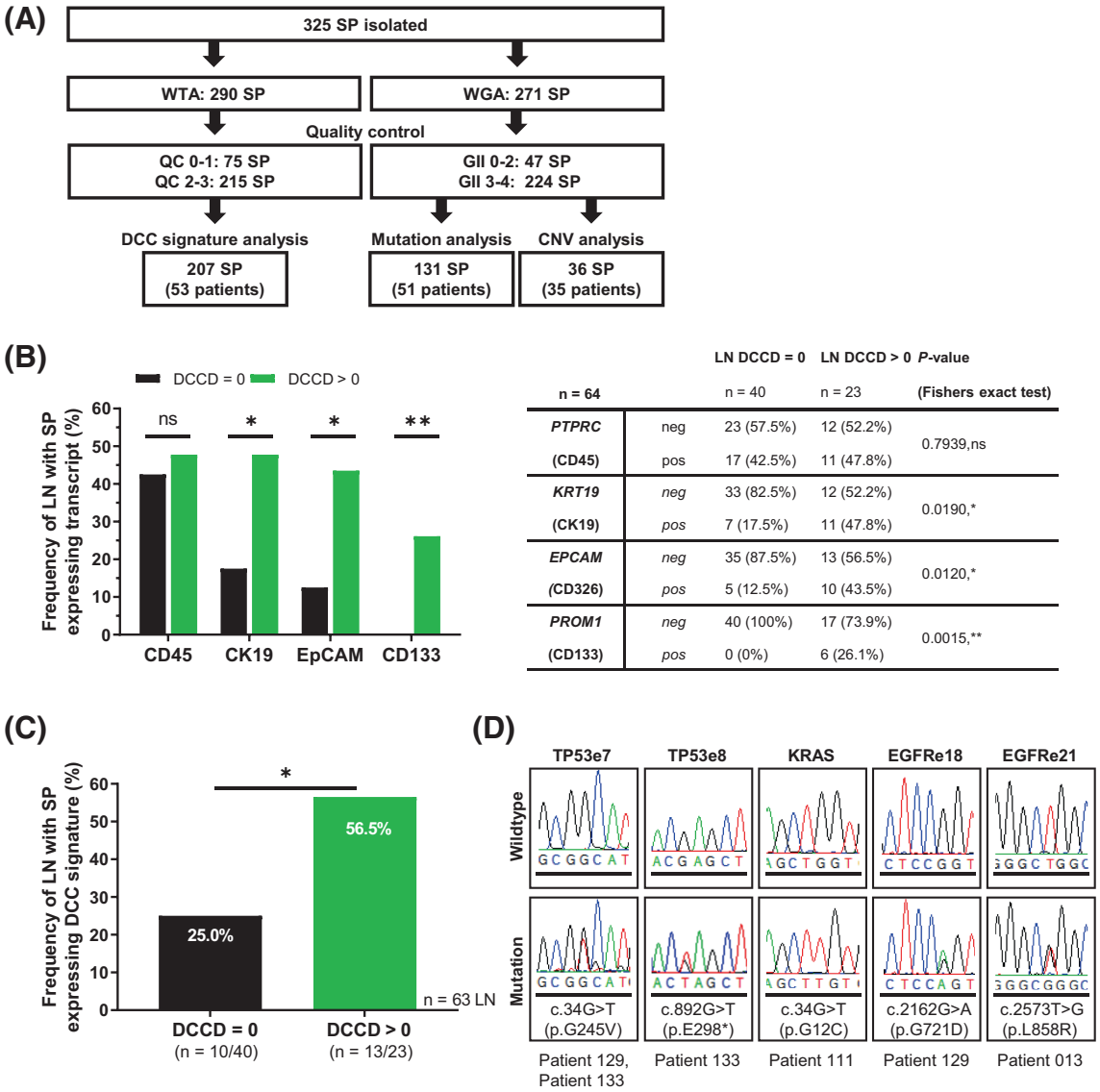


FIGURE 1 Sphere formation of lung cancer LN cell suspensions (A) Overview of lymph node samples analyzed for DCCD status and sphere formation. (B) Randomly selected images of LN-derived spheres of various patients. (C) Frequency of sphere formation in DCCD = 0 (black) compared with DCCD > 0 (green) LNs (n = 199 LN), $P = .0239$, Fisher's exact test. (D) Number of spheres per LN culture in DCCD = 0 compared with DCCD > 0 LNs (n = 199 LN), $P = .003$, Mann-Whitney test. (E) Kaplan-Meier survival analysis of patients forming spheres (SP+, n = 55, green curve) compared with patient cultures unable to develop spheres (SP−, n = 76, black curve), OS: $P = .0253$, TSS: $P = .0391$; log-rank test.

Also, the number of spheres per culture was significantly increased in LNs containing EpCAM-positive cells (DCCD > 0) compared with EpCAM-negative (DCCD = 0) LNs (Mann-Whitney test, $P = .003$). Finally, we tested whether sphere formation in LN cultures correlates with clinical outcome of the patients. We found that the presence of spheres in LN-cultures was associated with a reduced overall survival (OS) as well as tumor-specific survival (TSS) within the lung cancer patient cohort (Log-rank test, OS: $P = .0253$, TSS: 0.0391; Figure 1E).



n = 64		LN DCCD = 0 n = 40	LN DCCD > 0 n = 23	*P*-value (Fishers exact test)
PTPRC	neg	23 (57.5%)	12 (52.2%)	0.7939,ns
(CD45)	pos	17 (42.5%)	11 (47.8%)	
KRT19	*neg*	33 (82.5%)	12 (52.2%)	0.0190,*
(CK19)	*pos*	7 (17.5%)	11 (47.8%)	
EPCAM	*neg*	35 (87.5%)	13 (56.5%)	0.0120,*
(CD326)	*pos*	5 (12.5%)	10 (43.5%)	
PROM1	*neg*	40 (100%)	17 (73.9%)	0.0015,**
(CD133)	*pos*	0 (0%)	6 (26.1%)	

FIGURE 2 Molecular characterization on RNA and DNA level. (A) Overview and summary of information of isolated, amplified and tested spheres analyzed for the diverse molecular assays. (B) Frequency of LNs with spheres expressing indicated transcripts derived from DCCD = 0 (black) or DCCD > 0 (green) samples (n = 63 LN of 53 patients). (C) Frequency of LN with spheres expressing DCC-signature of DCCD = 0 (black) compared with DCCD > 0 (green) samples (n = 63 LN of 53 patients), $P = .0161$, Fisher's exact test. (D) Mutational analysis of spheres. Exemplary electropherograms show wild type and mutated regions detected in TP53 exon7 and exon8, KRAS and EGFR exon18 and exon21.

3.2 | Molecular characterization of single spheres

Therefore, we further characterized single spheres on a molecular level. We manually isolated 325 single spheres derived from DCCD = 0 and DCCD > 0 LN samples from 55 patients (median = 4, range of 1-22 per LN). In addition to phenotypical categorization of size, shape and color, we isolated RNA (for WTA) and DNA (for WGA) from each sphere and assessed WTA and WGA quality. We identified 215 LN spheres with sufficient WTA quality (QC2-3) and 224 spheres with good DNA quality (GII3-4) (Figure 2A). We observed that poor WTA and WGA quality was visually strongly associated with cellular particles smaller than 40 µm in size or very dark colored aggregates, possibly representing necrotic spheres or clusters, which were therefore excluded for additional downstream analysis.

3.3 | Expression profiling of selected transcripts in single spheres

To test if LN spheres with sufficient RNA quality comprise DCCs, we screened WTA products of 207 isolated single spheres from 63 LNs of 53 patients for a DCC-like signature by expression analysis using a panel of informative genes. In particular, we assessed expression of cell adhesion molecule *EPCAM* (CD326) and cytokeratin *KRT19* (CK19),

which are commonly described and investigated for many cancer entities including lung cancer cells.[46,47] In addition, we tested *PROM1* (CD133), a stem cell marker described to be expressed by a rare tumor cell population within lung cancer patients with the potential to form spheres,[48,49] and the leukocyte common antigen *PTPRC* (CD45) as common leukocyte marker to uncover the cellular composition of the spheres.

Comparing spheres derived from DCCD = 0 and DCCD > 0 LNs, we found no significant differences for *PTPRC* expression (Figure 2B, Fishers exact test, *PTPRC*: $P = .7939$). However, *KRT19* and *EpCAM* were more frequently expressed in spheres derived of LNs with a DCCD > 0 (Fishers exact test, *KRT19*: $P = .0190$, *EpCAM*: $P = .0120$). Moreover, the stem cell marker *PROM1* was exclusively expressed in

TABLE 1 Data of patients with spheres and their analyses

ID	Type	T	N	M	LNs	DCCD	SP	DCC signature	Mutational status	CNV status	Death[a]
(A) DCCD = 0 patients											
001	ADC	1	0	1	LN1	0	8	No	wt	Balanced	Yes
003	ADC	1	0	0	LN1	0	1	No	wt	n.a.	No
004	ASC	1	2	0	LN1	0	1	Yes	wt	Balanced	Yes
006	ADC	1	1	0	LN1	0	190	No	wt	n.a.	No
008	SCC	2	1	0	LN1	0	73	Yes	wt	Balanced	No
					LN2	0	3	Yes	wt	n.a.	
015	ADC	1	0	0	LN1	0	63	No	wt	n.a.	No
016	ADC	1	0	0	LN1	0	10	No	wt	n.a.	Yes
018	SCC	2	0	0	LN1	0	5	No	wt	Balanced	Yes
020	SCC	2	0	0	LN1	0	10	No	wt	n.a.	Yes
					LN2	0	14	No	wt	n.a.	
022	SCC	0	0	0	LN1	0	20	n.a.	n.a.	n.a.	Yes
030	SCC	1	0	0	LN1	0	2	Yes	wt	Balanced	Yes
					LN2	0	12	No	wt	n.a.	
031	SCC	1	0	0	LN1	0	1	n.a.	n.a.	n.a.	Yes
033	SCC	3	2	0	LN1	0	10	No	wt	n.a.	No
					LN2	0	1	n.a.	n.a.	n.a.	
039	ADC	3	0	0	LN1	0	12	No	wt	Balanced	No
042	NEC	1	0	0	LN1	0	3	No	wt	n.a.	No
043	ADC	2	0	0	LN1	0	1	No	wt	n.a.	No
044	SCC	2	0	0	LN1	0	10	No	wt	n.a.	Yes
047	ADC	1	0	0	LN1	0	9	Yes	wt	Balanced	Yes
056	ADC	2	0	0	LN1	0	6	No	wt	n.a.	No
060	ADC	3	0	1	LN1	0	1	No	wt	Balanced	Yes
					LN2	0	5	No	wt	n.a.	
064	ADC	1	1	0	LN1	0	6	Yes	wt	Balanced	Yes
065	SCC	1	0	0	LN1	0	4	Yes	wt	Balanced	No
068	ADC	2	1	0	LN1	0	2	No	wt	n.a.	No
074	ADC	1	0	0	LN1	0	2	No	wt	Balanced	Yes
					LN2	0	2	No	wt	n.a.	
077	SCC	2	0	0	LN1	0	7	No	n.a.	n.a.	No
					LN2	0	8	No	wt	n.a.	
084	ADC	2	0	0	LN1	0	1	No	wt	n.a.	Yes
089	ADC	1	0	0	LN1	0	2	No	wt	n.a.	Yes
104	ADC	3	2	0	LN1	0	2	No	wt	n.a.	Yes
					LN2	0	1	n.a.	n.a.	n.a.	
115	SCLC	3	0	0	LN1	0	3	No	wt	Balanced	Yes
125	ADC	4	2	0	LN1	0	6	No	n.a.	n.a.	Yes
128	ADC	3	0	0	LN1	0	6	Yes	n.a.	n.a.	No

(Continues)

TABLE 1 (Continued)

ID	Type	T	N	M	LNs	DCCD	SP	DCC signature	Mutational status	CNV status	Death[a]
(B) DCCD > 0 patients											
010	ADC	1	1	0	LN1	14	12	n.a.	wt	Aberrant	No
					LN2	0	3	No	wt	n.a.	
012	SCC	1	0	0	LN1	14	8	Yes	wt	Balanced	No
					LN2	0	6	Yes	wt	n.a.	
013	ADC	3	2	0	LN1	183	132	Yes	EGFR (p.L858R)	Aberrant	Yes
014	ADC	3	1	0	LN1	3	20	No	wt	Balanced	Yes
					LN2	0	1	n.a.	n.a.	n.a.	
019	SCC	3	1	0	LN1	1	17	No	wt	Aberrant	No
023	ADC	2	0	0	LN1	1	5	No	wt	n.a.	No
029	SCLC	1	1	0	LN1	538	1	No	wt	Balanced	Yes
					LN2	0	2	No	wt	n.a.	
035	ADC	3	1	0	LN1	2	4	Yes	wt	Balanced	Yes
040	SCC	2	1	0	LN1	1	10	Yes	wt	Balanced	No
046	SCC	2	0	0	LN1	3	5	No	wt	Balanced	Yes
051	ADC	2	0	0	LN1	3	1	Yes	wt	Balanced	Yes
059	SCC	2	0	0	LN1	17	4	No	wt	Balanced	Yes
078	ADC	2	0	0	LN1	2	2	No	n.a.	Balanced	No
					LN2	0	5	No	wt	n.a.	
079	ADC	2	0	0	LN1	2	9	Yes	wt	Balanced	No
098	ADC	1	0	0	LN1	2	5	No	wt	Balanced	Yes
106	SCC	2	2	0	LN1	18	15	No	wt	Balanced	No
111	ADC	3	1	0	LN1	250 000	35	Yes	KRAS (p.G12C)	Aberrant	Yes
113	NEC	3	2	1	LN1	750 000	100	Yes	wt	Aberrant	Yes
114	ADC	2	0	0	LN1	0.5	4	No	wt	Balanced	No
117	SCC	1	2	0	LN1	100 000	35	Yes	wt	Aberrant	No
119	SCC	3	0	1	LN1	10 000	6	Yes	wt	Aberrant	Yes
120	SCC	n.a.	2	0	LN1	1000,0	20	Yes	wt	Aberrant	Yes
129	ADC	1	0	0	LN1	423	5	Yes	TP53 (p.G245V), EGFR (p. G721D)	Aberrant	No
					LN2	0	6	Yes	n.a.	n.a.	
133	ADC	2	1	0	LN1	10 000	39	Yes	TP53 (p.G245V), (p.E298*)	Aberrant	Yes

[a]Death from any cause documented during follow-up.
*Substitution - Nonsense.

LNs with DCCD > 0 (6/23; 26.1%; $P = .0015$) and all six patients expressing *PROM1* co-expressed *KRT19* and/or *EpCAM*, suggesting that the co-expression of *KRT19*, *EpCAM* and/or *PROM1* may identify DCC-derived spheres. These DCC signature genes were significantly more frequently expressed in DCCD > 0 LNs compared with DCCD = 0 LNs (DCCD > 0: 56.5%; 13/23 and DCCD = 0: 25.0%; 10/40; Fisher's exact test $P = .0161$, Figure 2C). Henceforth, patients with at least one sphere expressing either *KRT19*, *EpCAM* and/or *PROM1* will be referred to as DCC-signature positive. Of note, in most of the analyzed spheres expression of epithelial markers *KRT8* and *KRT18* was frequently observed, to a similar extent in samples derived from DCCD = 0 patients (CK8: 75%, CK18: 65%) as well as from DCCD > 0 patients (CK8: 87%, CK18: 87%; Figure S3A). Remarkably, we found heterogenous transcript expression in spheres of individual patients, including spheres expressing the DCC-signature genes as well as CD45 (exemplary data are shown for five patients in Figure S3B), indicating that in some cases non-cancer cells support sphere formation.

3.4 | Mutational analysis of single spheres

We then asked whether spheres derived from DCCD = 0 and DCCC > 0 LNs differ with regard to mutations typically found in lung cancer and analyzed 131 LN-spheres from 59 LN of 51 patients for the presence of mutations in *TP53*, *EGFR* and *KRAS*. For this, we amplified *TP53* exons 5, 6, 7 and 8, *EGFR* exons 18, 20 and 21 and *KRAS* exon 2. In total, we identified four patients with spheres harboring mutations, all derived from DCCD > 0 LNs (Figure 2D). Of note, *TP53*, *KRAS* or *EGFR* mutated spheres from these patients were also DCC-signature positive (Table 1).

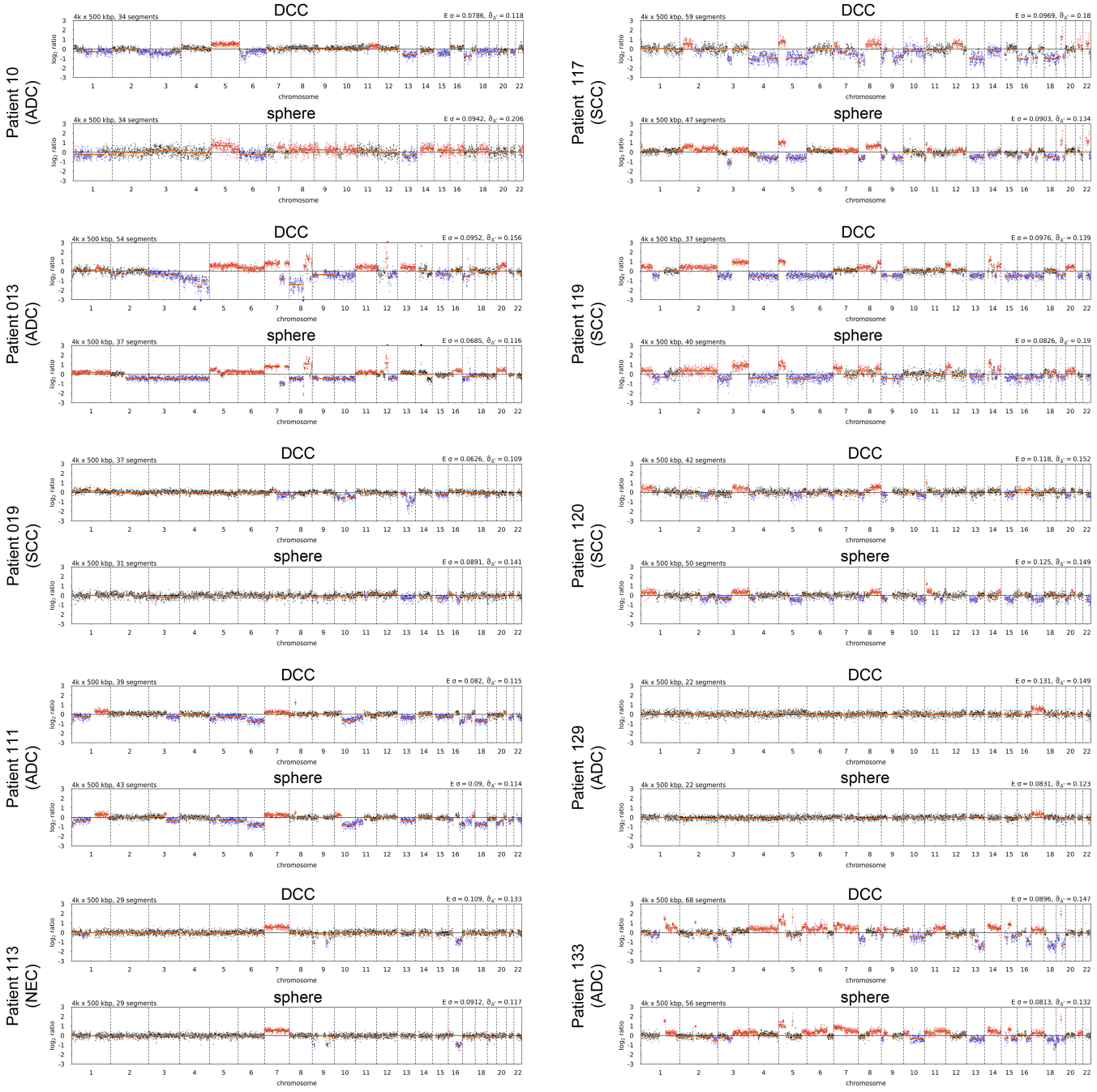


FIGURE 3 CNV analyses. Genomic profiles of matched pairs of DCC and spheres. Histogram of genomic gains (red) and losses (blue) of 10 aberrant patients are shown. ADC, adenocarcinoma; NEC, neuroendocrinic carcinoma; SCC, squamous cell carcinoma.

3.5 | Copy number variations (CNVs) of single spheres

We then tested sphere-derived DNA for copy number variations by aCGH as well as low-coverage whole genome sequencing (low-pass sequencing) of selected 35 patients. We analyzed spheres of 12 patients with DCCD = 0, comprising six patients expressing DCC-signature genes and additional six patients with a documented disease-related death and/or early disease progression, as well as spheres of 23 DCCD > 0 patients. Chromosomal aneuploidies and CNVs could not be detected in any of the patients with DCCD = 0 (Table 1A and Figure S4A). In contrast, CNV analysis of DCCD > 0 patients revealed genomic aberrations in 10 of 23 patients (43.5%, Table 1B, see Figure 3 for aberrant and Figure S4B for balanced spheres).

When comparing the genomic profiles of the analyzed spheres with the corresponding single DCCs isolated directly from the LN suspension of the same patient, we found that the profiles of ex vivo generated spheres represent matched DCC-derived profiles (Figure 3). Upon closer examination of the genomes, we recognized in several cases genomic gains on chromosome 7, containing lung cancer-related

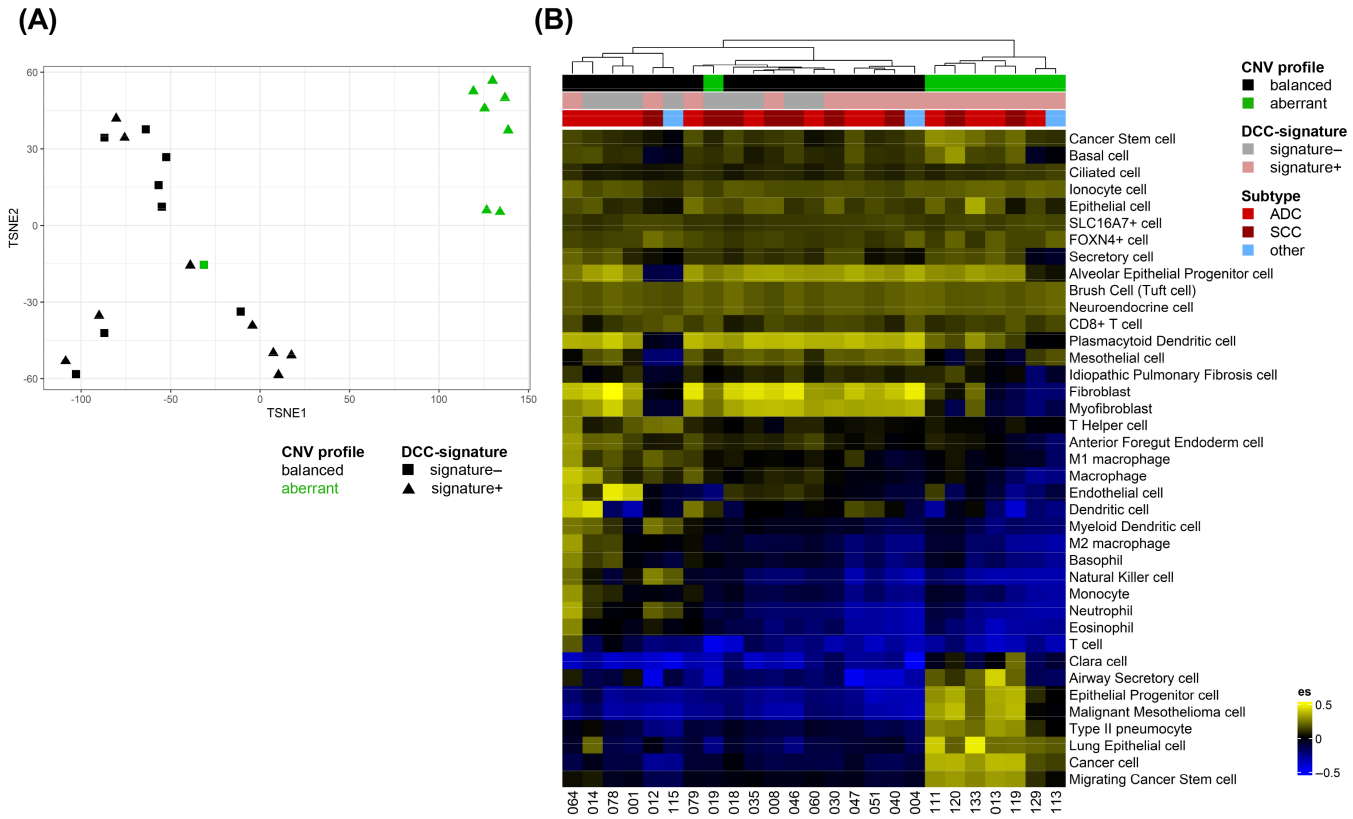


FIGURE 4 Cell composition of LN-derived spheres. (A) tSNE plot of 25 spheres of different categories (DCC-signature and CNV profile). Balanced spheres are indicated in black, aberrant in green. Signature- spheres are shown as squares, signature+ as triangles. (B) Heatmap from deconvolution analysis of 25 spheres using a subset of cell types from lung from CellMarker database. Patient ID is on x-axis, cell types are in y-axis. Phenotype categories (CNV profile, DCC-signature and subtypes) of the patient samples are highlighted on top. Heatmap colors correspond to the enrichment scores (es) of the cell type markers and range from −0.5 to 0.5 (blue to yellow).

oncogenes like *EGFR* and *BRAF*, and chromosome 8, containing *MYC*. On chromosomes 13 and 17 we frequently observed common deletions of genomic regions containing well-described tumor suppressor genes like *BRCA1* and *2*, *RB1* and *NF1* and oncogenes like *ERBB2* (Figure S5).

3.6 | Gene expression analysis of spheres

To further decipher the cell type composition within single spheres, we performed RNA-Seq of 25 selected spheres of four different categories (including 8 signature−/balanced, 9 signature+/balanced, 1 signature−/aberrant, 7 signature+/aberrant). Dimensionality reduction plots based on the top 500 most variable genes revealed that signature+/aberrant spheres were clearly separated from all other groups (Figure 4A). Furthermore, we observed a tendency that balanced spheres grouped together in three sub-clusters without a clear separation based on signature gene expression. Overall, the clustering identified clear differences in gene expression between balanced vs aberrant spheres.

To track the individual cell types within single spheres we additionally performed deconvolution analysis (Figure 4B). Aberrant spheres (with exception of the one signature−/aberrant sphere, 019) most prominently classified into specific cancer cell related categories such as "cancer cell, lung epithelial cell, malignant mesothelioma or migrating cancer stem cell." However, expression of non-cancer gene categories such as "plasmacytoid dendritic cell or fibroblast" reveal that aberrant spheres also contain additional stromal cell types. In contrast, signature- and signature+ balanced spheres scored highest in "fibroblast, myofibroblast and plasmacytoid dendritic cell" categories as well as in categories related to innate immune cells (e.g., "M1/M2 macrophage, natural killer cell, dendritic cell"). Remarkably, all spheres possibly contained (hidden) epithelial DCCs as indicated by the identified expression of "alveolar epithelial progenitor cell or epithelial cell" gene sets. Epithelial cell content of spheres did not correlate with histological subtypes.

3.7 | Identification of high-risk patients with sphere formation

All our analyses revealed that characteristic traits (e.g., signature genes, marker mutations, CNVs) were more frequently or exclusively found in spheres from patients with DCCD > 0 (Figure 5A). As DCCD > 0 patients with aberrant spheres also displayed a significantly

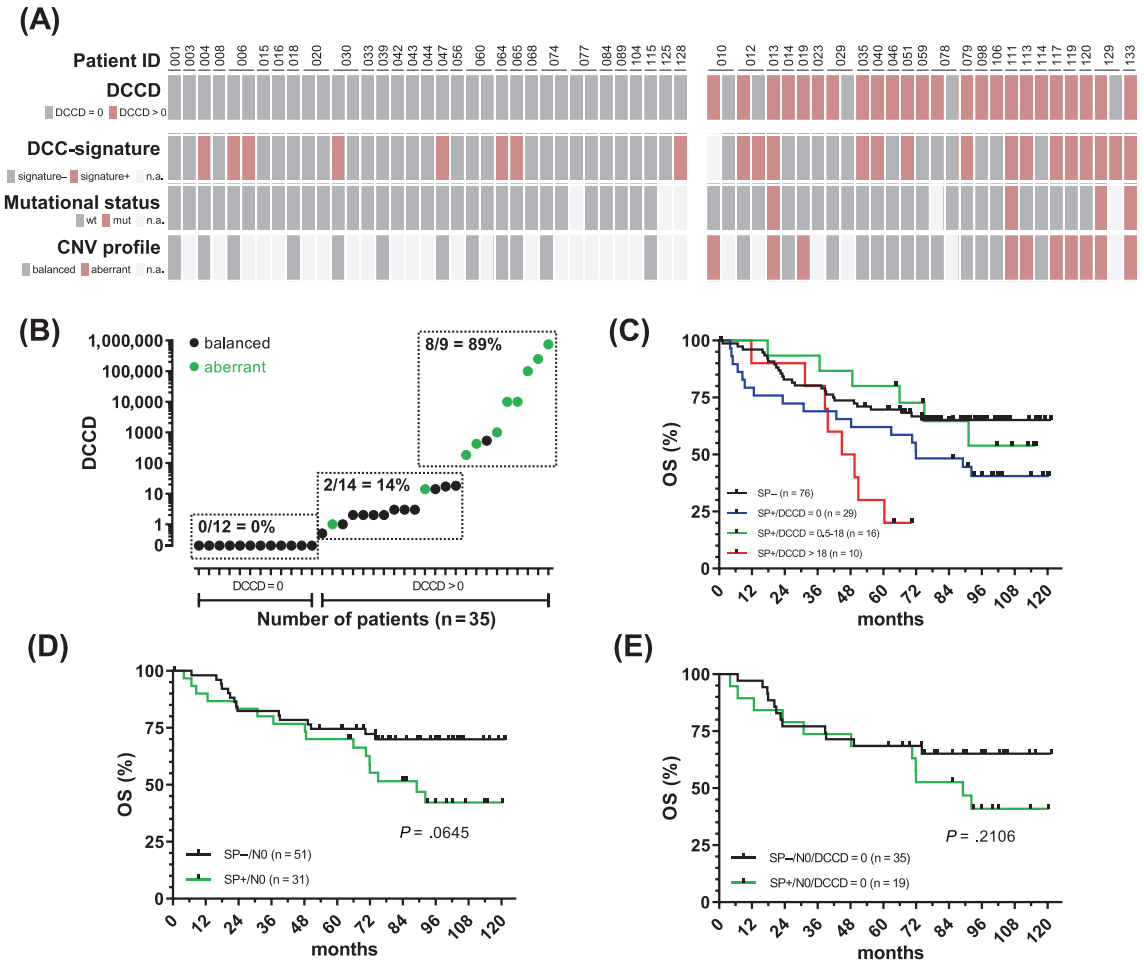


FIGURE 5 Identification of high-risk patients with sphere formation. (A) Summary of SP+ samples from 53 patients analyzed for DCCD, DCC signature, mutational status and CNV profile. DCCD: DCCD = 0 = gray, DCCD > 0 = red; DCC signature: no = gray, yes = red; mutational analysis: wt = gray, mut = red; CNV profile: balanced = gray, aberrant = red, n.a. = light gray (B) Association of LN-DCCD and CNV profile of spheres (n = 35 patients). Patients with balanced spheres are indicated in black (n = 26), patients with aberrant spheres are highlighted in green (n = 10). (C) Kaplan-Meier survival analysis of patients with cultures without sphere formation (SP−, n = 76), patients with spheres from LN cultures without DCC (SP+/DCCD = 0, n = 29), patients with spheres from LN cultures with a DCCD up to 18 (SP+/DCCD = 0.5-18, n = 16) and patients with spheres from LN cultures with a DCCD higher than 18 (SP+/DCCD > 18, n = 10). OS: SP+/DCCD > 18 vs SP−: $P = .0063$; SP+/DCCD > 18 vs SP+/DCCD = 0.5-18: $P = .0059$, log-rank test. (D) Kaplan-Meier survival analysis of lymph-node negative patients (pN0, including only M0 patients) with (green curve: SP+/N0, n = 31) and without sphere formation (black curve: SP−/N0, n = 51). OS: $P = .0645$, log-rank test. (E) Kaplan-Meier survival analysis of sphere formation including only M0 patients with no evidence of lymphatic spread (DCCD = 0 and N0). Green curve indicates patients forming spheres (SP+/N0/DCCD = 0, n = 19) compared with patients without spheres (SP−/N0/DCCD0 = 0, n = 35, black curve), OS: $P = .2106$, log-rank test.

higher DCCD than patients with genomically balanced spheres (Mann-Whitney Test, $P = .0012$, Figure S6) we asked whether aberrant sphere formation is associated with a specific DCCD range. Indeed, patients with a DCCD > 18 generated aberrant spheres in 89% of the cases compared with 14% and 0% for patients with a DCCD = 0.5-18 or DCCD = 0, respectively (Figure 5B).

Finally, we analyzed whether formation of aberrant spheres has an impact on survival of all patients within our cohort. We performed subcategory analyses to compare patients that did not generate spheres (SP−) with SP+/DCCD = 0, SP+/DCCD = 0.5-18 and SP+/DCCD > 18 patients with sphere formation. Only the latter group showed a significantly reduced OS and a non-significant TSS trend when compared with either the SP− ($P = .0063$) or the SP+/DCCD = 0.5-18 ($P = .0059$) patients, indicating that formation of aberrant spheres is associated with accelerated progression (Figure 5C and S7A).

Interestingly, when comparing survival of patients forming spheres or not (SP+ vs SP−) without evidence of distant metastases (M0) and lymphatic spread (pN0 and pN0/DCCD = 0), we observed a late divergence of survival curves and a borderline increase in risk in patients forming spheres in all analyses (pN0: OS: $P = .065$, TSS: $P = .0890$; Figure 5D and S7B; pN0/DCCD = 0: OS: $P = .2106$, TSS: $P = .0489$; Figure 5E and S7C). The late divergence of the survival curves may suggest that undetected DCCs with sphere formation potential are associated with late disease progression.

4 | DISCUSSION

Here, we show that sphere formation under stem-like-cell selecting conditions is associated with patient outcome. Spheres were characterized by expression of selected signature genes, cancer-related gene sets, CNVs and more easily generated from samples with a DCCD > 18. We also noted that sphere forming samples of DCCD = 0 patients tended to reduced survival.

Metastasis founder cells must be comprised among DCCs, however currently used markers fail to unambiguously identify them. DCCs are detected by staining cells from mesenchymal organs via antibodies directed against CK (for bone marrow DCCs) and EpCAM (for lymph node DCCs), commonly used in lung cancer and a large number of other carcinomas to identify patients at risk of relapse.[50] However, since in all tested cancers a significant proportion of marker-negative patients progress, it has long been argued that marker-negative cancer cells exist. We therefore used sphere formation under stem-like-cell selecting conditions and then compared it to the immunocytological result.

The observed sphere formation rate of immunocytological-positive LNs was higher than for published lung cancer tissue biopsies (49% vs 40%).[51] Sphere numbers were considerably lower compared with the corresponding DCC numbers, indicating that only a subpopulation of DCCs was competent to survive and proliferate in the applied culture conditions. Interestingly, sphere formation was also documented in 30% of DCCD = 0 samples. None of the previous studies using comparable sphere culture conditions reported generation of non-tumor cell-derived spheres, however, only few validated the malignant origin of spheres generated from patient-derived material.[51]

When we compared spheres from patients with DCCD = 0 vs DCCD > 0 LNs, spheres of both groups expressed CD45 at a comparable frequency, precluding DCC-origin confirmation based on this exclusion marker. However, LN-spheres from patients with a DCCD > 0 more frequently expressed epithelial markers EpCAM, CK19 and the stem cell marker CD133, indicating that those spheres may comprise cells of epithelial origin and stem cell-like properties. Since the absence of EpCAM transcript in spheres derived from EpCAM-negative samples was somehow expected, we also searched for genetic hints of malignancy in spheres derived from DCCD = 0 samples, however, none were found. In contrast, low-pass sequencing to detect CNVs confirmed tumorigenic origin in 43% of tested DCC-positive patients, with four of them also harboring typical gene mutations (*TP53*, *KRAS* and *EGFR*).

Some DCC-negative and -positive samples expressed one or several of the signature genes but did not exhibit CNVs that would provide unequivocal evidence of malignancy. Gene expression analysis of such spheres revealed that they mainly consist of fibroblasts or innate immune cells but we also found genes characteristic like "epithelial or alveolar epithelial progenitor cells". This might indicate the existence of sphere-forming cancer cells that lack typical mutations and/or are present in low abundance within spheres, so that CNV are not detected. If present, such DCCs left behind in patients may have the potential to progress after years of clinical latency—possibly after they acquired disease-driving genetic alterations. This reasoning appears to be supported by the borderline significant association (OS: $P = .065$, Figure 5D) of sphere forming samples and disease progression in early stage (pN0) patients. In this group, Kaplan-Meier survival analysis revealed separation of the survival curves rather late, that is, only after 65 months, consistent with ongoing cellular evolution. However, we cannot conclusively decide whether the detection of the epithelial gene sets in samples negative by DCC immunocytology indicates the presence of hidden DCCs in the tested spheres or not, since we did not obtain thoracic lymph nodes from non-cancer patients as controls. To clarify these issues further studies such as single cell RNA-Seq and more extensive follow-up studies will be needed.

These observations support our starting hypothesis that detection of marker-negative DCC may be uncovered by growth in culture, since currently used markers fail to unambiguously identify distant-disease driving cells in all patient samples. Comparing sphere formation under stem-like-cell selecting conditions with immunocytological results side-by-side provides further support for this reasoning.

In the case of more extensive lymph node involvement (DCCD > 18), sphere formation with expression of gene sets reflecting epithelial and cancer stem cell profiles and detection of CNVs place patients at highest risk. In addition, samples with mutations (n = 4, 17% of all DCCD > 0 sphere forming patients) also displayed CNVs, and 8/10 samples with CNVs also expressed the DCC signature genes. Apparently, a high number of DCCs in LNs reflects a proliferative phenotype and genotype that also enables sphere formation. In addition, specific genomic aberrations are acquired during proliferation within the LNs,[17] which may drive sphere formation. This is consistent with observations in melanoma, where we previously found that xenograft formation and acquisition of specific alterations within the LN are intimately linked.[23] On the other hand, not all samples with high DCCD formed spheres and therefore more investigations are needed to identify the cellular and (epi)genetic changes required for metastatic colony formation and ex vivo expansion.

Our conclusions are inherently linked to the methods applied, that is, non-adherent sphere culture conditions. Alternative approaches like organoid culture[52] may be tested in the future. For the time being, our sphere assay, when applied to samples with a DCCD > 18 generates cell models that (i) reflect the genotype of the DCCs detected by immunocytology and (ii) are associated with disease progression. They can therefore be considered representative models for drug testing experiments to uncover vulnerabilities in early systemic lung cancer. Whether spheres devoid of malignancy indicators comprise the earliest metastasis founders or reflect a cancer-associated epiphenomenon, is open for future studies.

AUTHOR CONTRIBUTIONS

Conception and design of the study: Steffi Treitschke, Kathrin Weidele, Christian Werno and Christoph A. Klein. Development of methodology: Steffi Treitschke, Kathrin Weidele, Christian Werno and Christoph A. Klein. Acquisition of data: Steffi Treitschke, Kathrin Weidele, Giancarlo Feliciello, Sybille Vorbeck, Catherine Botteron, Melanie Werner-Klein, Bernhard Polzer, Hans-Stefan Hofmann, Tobias Robold, Florian Weber, Tobias Mederer, Christian Werno and Christoph A. Klein. Analysis and interpretation of data: Steffi Treitschke, Kathrin Weidele, Giancarlo

Feliciello, Bernhard Polzer, Martin Hoffmann, Adithi Ravikumar Varadarajan, Jens Warfsmann, Vadim Dechand, Christian Werno and Christoph A. Klein. Writing of the manuscript: Steffi Treitschke, Kathrin Weidele, Christian Werno and Christoph A. Klein. Review, and/or revision: all authors. The study reported in the paper has been performed by the authors, unless clearly specified in the text.

ACKNOWLEDGEMENTS

We thank Judith Proske, Siegfried Rein, Irene Nebeja, Theresa Kettl and Thomas Schamberger for excellent technical assistance and Matthias Winzinger and Wolfgang Dietmaier for their support in fingerprint analysis. We would like to express gratitude to the patients, who participated in the study. The graphical abstract was created with biorender.com. Open Access funding enabled and organized by Projekt DEAL.

FUNDING INFORMATION

Bavarian ministry of economic affairs, energy and technology, AZ 20-3410.1-1-1and 20-3410.1-1-2; Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—SFB TRR 305-B13 (to CW); FOR 2127-A02 (to CK), FOR 2127-B05 (to MWK).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interests.

DATA AVAILABILITY STATEMENT

All low-pass genome sequencing profiles are provided in the manuscript or as supplemental information. The raw RNA-Seq data are available to the EGA under the accession number EGAS00001007369. Other data that support the findings of this study are available from the corresponding author upon reasonable request.

ETHICS STATEMENT

The study was approved by the University of Regensburg Ethics Committee (ethics vote numbers 07-079). All patients provided written informed consent.

ORCID

Kathrin Weidele https://orcid.org/0009-0008-9313-5057
Bernhard Polzer https://orcid.org/0000-0002-3797-9975
Christoph A. Klein https://orcid.org/0000-0001-7128-1725

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*. 2020;70:7-30.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. *CA Cancer J Clin*. 2017;67:7-30.
3. Uramoto H, Tanaka F. Prediction of recurrence after complete resection in patients with NSCLC. *Anticancer Res*. 2012;32:3953-3960.
4. Brustugun OT, Grønberg BH, Fjellbirkeland L, et al. Substantial nation-wide improvement in lung cancer relative survival in Norway from 2000 to 2016. *Lung Cancer*. 2018;122:138-145.
5. Friberg S, Mattson S. On the growth rates of human malignant tumors: implications for medical decision making. *J Surg Oncol*. 1997;65:284-297.
6. Klein CA. Parallel progression of primary tumours and metastases. *Nat Rev Cancer*. 2009;9:302-312.
7. Riethdorf S, Wikman H, Pantel K. Review: biological relevance of disseminated tumor cells in cancer patients. *Int J Cancer*. 2008;123:1991-2006.
8. He Z, Xia Y, Tang S, Chen Y, Chen L. Detection of occult tumor cells in regional lymph nodes is associated with poor survival in pN0 non-small cell lung cancer: a meta-analysis. *J Thorac Dis*. 2016;8:375-385.
9. Wu J, Ohta Y, Minato H, et al. Nodal occult metastasis in patients with peripheral lung adenocarcinoma of 2.0 cm or less in diameter. *Ann Thorac Surg*. 2001;71:1772-1777. discussion 1777-8.
10. Kubuschok B, Passlick B, Izbicki JR, Thetter O, Pantel K. Disseminated tumor cells in lymph nodes as a determinant for survival in surgically resected non-small-cell lung cancer. *J Clin Oncol*. 1999;17:19-24.
11. Passlick B, Izbicki JR, Kubuschok B, Thetter O, Pantel K. Detection of disseminated lung cancer cells in lymph nodes: impact on staging and prognosis. *Ann Thorac Surg*. 1996;61:177-182; discussion 183.
12. Osaki T, Oyama T, Gu C-D, et al. Prognostic impact of micrometastatic tumor cells in the lymph nodes and bone marrow of patients with completely resected stage I non-small-cell lung cancer. *J Clin Oncol*. 2002;20:2930-2936.
13. Yasumoto K, Osaki T, Watanabe Y, Kato H, Yoshimura T. Prognostic value of cytokeratin-positive cells in the bone marrow and lymph nodes of patients with resected nonsmall cell lung cancer: a multicenter prospective study. *Ann Thorac Surg*. 2003;76:194-201. discussion 202.
14. Rud AK, Borgen E, Mælandsmo GM, et al. Clinical significance of disseminated tumour cells in non-small cell lung cancer. *Br J Cancer*. 2013;109:1264-1270.
15. Rud AK, Boye K, Fodstad Ø, et al. Detection of disseminated tumor cells in lymph nodes from patients with early stage non-small cell lung cancer. *Diagn Pathol*. 2016;11:50.
16. Mederer T, Elsner F, Robold T, et al. EpCAM-positive disseminated cancer cells in bone marrow impact on survival of early-stage NSCLC patients. *Lung Cancer (Amsterdam, Netherlands)*. 2022;167:73-77.
17. Elsner F, Hoffmann M, Fahrioglu-Yamaci R, et al. Disseminated cancer cells detected by immunocytology in lymph nodes of NSCLC patients are highly prognostic and undergo parallel molecular evolution. *J Pathol*. 2022;258:250-263.
18. Mountain CF. The international system for staging lung cancer. *Sem Surg Onc*. 2000;18:106-115.
19. Ulmer A, Dietz K, Hodak I, et al. Quantitative measurement of melanoma spread in sentinel lymph nodes and survival. *PLoS Med*. 2014;11:e1001604.
20. Ulmer A, Fischer JR, Schanz S, et al. Detection of melanoma cells displaying multiple genomic changes in histopathologically negative sentinel lymph nodes. *Clin Cancer Res*. 2005;11:5425-5432.
21. Hosseini H, Obradović MMS, Hoffmann M, et al. Early dissemination seeds metastasis in breast cancer. *Nature*. 2016;540:552-558.
22. Klein CA, Blankenstein TJF, Schmidt-Kittler O, et al. Genetic heterogeneity of single disseminated tumour cells in minimal residual cancer. *Lancet (London, England)*. 2002;360:683-689.
23. Werner-Klein M, Scheitler S, Hoffmann M, et al. Genetic alterations driving metastatic colony formation are acquired outside of the primary tumour in melanoma. *Nat Commun*. 2018;9:595.
24. Zhang L, Ridgway LD, Wetzel MD, et al. The identification and characterization of breast cancer CTCs competent for brain metastasis. *Sci Transl Med*. 2013;5:180ra48.
25. Yu M, Bardia A, Aceto N, et al. Cancer therapy. Ex vivo culture of circulating breast tumor cells for individualized testing of drug susceptibility. *Science (New York, N.Y.)*. 2014;345:216-220.
26. Fang D, Nguyen TK, Leishear K, et al. A tumorigenic subpopulation with stem cell properties in melanomas. *Cancer Res*. 2005;65:9328-9337.
27. Hartmann CH, Klein CA. Gene expression profiling of single cells on large-scale oligonucleotide arrays. *Nucleic Acids Res*. 2006;34:e143.

28. Gužvić M, Braun B, Ganzer R, et al. Combined genome and transcriptome analysis of single disseminated cancer cells from bone marrow of prostate cancer patients reveals unexpected transcriptomes. *Cancer Res*. 2014;74:7383-7394.
29. Klein CA, Schmidt-Kittler O, Schardt JA, Pantel K, Speicher MR, Riethmüller G. Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells. *Proc Natl Acad Sci U S A*. 1999;96:4494-4499.
30. Klein CA, Seidl S, Petat-Dutter K, et al. Combined transcriptome and genome analysis of single micrometastatic cells. *Nat Biotechnol*. 2002;20:387-392.
31. Polzer B, Medoro G, Pasch S, et al. Molecular profiling of single circulating tumor cells with diagnostic intention. *EMBO Mol Med*. 2014;6:1371-1386.
32. Czyż ZT, Hoffmann M, Schlimok G, Polzer B, Klein CA. Reliable single cell array CGH for clinical samples. *PloS One*. 2014;9:e85907.
33. DOE Joint Genome Institute. BBTools: DOE Joint Genome Institute. 2022 https://jgi.doe.gov/data-and-tools/software-tools/bbtools/
34. Chu J, Sadeghi S, Raymond A, et al. BioBloom tools: fast, accurate and memory-efficient host species sequence screening using bloom filters. *Bioinformatics (Oxford, England)*. 2014;30:3402-3404.
35. Li H. *Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM*; arXiv. 2013:3:1303.3997. https://doi.org/10.48550/arXiv.1303.3997
36. Scheinin I, Sie D, Bengtsson H, et al. DNA copy number analysis of fresh and formalin-fixed specimens by shallow whole-genome sequencing with identification and exclusion of problematic regions in the genome assembly. *Genome Res*. 2014;24:2022-2032.
37. Werner-Klein M, Grujovic A, Irlbeck C, et al. Interleukin-6 trans-signaling is a candidate mechanism to drive progression of human DCCs during clinical latency. *Nat Commun*. 2020;11:4977.
38. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)*. 2013;29:15-21.
39. Liao Y, Smyth GK, Shi W. The subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res*. 2013;41:e108.
40. Deng C, Daley T, Smith AD. Applications of species accumulation curves in large-scale biological data analysis. *Quant Biol (Beijing, China)*. 2015;3:135-144.
41. Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics (Oxford, England)*. 2016;32:292-294.
42. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021;184:3573-3587.e29.
43. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. 2013;14:7.
44. Zhang X, Lan Y, Xu J, et al. CellMarker: a manually curated resource of cell markers in human and mouse. *Nucleic Acids Res*. 2019;47:D721-D728.
45. ggplot2. *Elegant Graphics for Data Analysis*. Cham, Switzerland: Springer; 2016.
46. Broers JL, Ramaekers FC, Rot MK, et al. Cytokeratins in different types of human lung cancer as monitored by chain-specific monoclonal antibodies. *Cancer Res*. 1988;48:3221-3229.
47. Balzar M, Winter MJ, de Boer CJ, Litvinov SV. The biology of the 17-1A antigen (ep-CAM). *J Mol Med (Berl)*. 1999;77:699-712.
48. Eramo A, Lotti F, Sette G, et al. Identification and expansion of the tumorigenic lung cancer stem cell population. *Cell Death Differ*. 2008;15:504-514.
49. Tirino V, Desiderio V, d'Aquino R, de Francesco F, Pirozzi G, Galderisi U, Cavaliere C, de Rosa A de, Papaccio G. Detection and characterization of CD133+ cancer stem cells in human solid tumours. *PLoS One* 2008;3, 3.
50. Riethdorf S, Pantel K. Disseminated tumor cells in bone marrow and circulating tumor cells in blood of breast cancer patients: current state of detection and characterization. *Pathobiol J Immunopathol Mol Cell Biol*. 2008;75:140-148.
51. Herreros-Pomares A, de Maya-Girones JD, Calabuig-Fariñas S, et al. Lung tumorspheres reveal cancer stem cell-like properties and a score with prognostic impact in resected non-small-cell lung cancer. *Cell Death Dis*. 2019;10:660.
52. Pauli C, Hopkins BD, Prandi D, et al. Personalized in vitro and in vivo cancer models to guide precision medicine. *Cancer Discov*. 2017;7:462-477.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Treitschke S, Weidele K, Varadarajan AR, et al. Ex vivo expansion of lung cancer-derived disseminated cancer cells from lymph nodes identifies cells associated with metastatic progression. *Int J Cancer*. 2023;1-14. doi:10.1002/ijc.34658