Genome Medicine

RESEARCH

Open Access

Genomic alterations and transcriptional phenotypes in circulating free DNA and matched metastatic tumor

Nobuyuki Takahashi[1,2†], Lorinc Pongor[1†], Shivam P. Agrawal[3], Mariya Shtumpf[3], Ankita Gurjar[3], Vinodh N. Rajapakse[1], Ahmad Shafiei[4], Christopher W. Schultz[1], Sehyun Kim[1,5], Diana Roame[6], Paula Carter[6], Rasa Vilimas[1], Samantha Nichols[1], Parth Desai[1], William Douglas Figg Sr.[6], Mohammad Bagheri[4], Vladimir B. Teif[3*] and Anish Thomas[1*]

Abstract

Background Profiling circulating cell-free DNA (cfDNA) has become a fundamental practice in cancer medicine, but the effectiveness of cfDNA at elucidating tumor-derived molecular features has not been systematically compared to standard single-lesion tumor biopsies in prospective cohorts of patients. The use of plasma instead of tissue to guide therapy is particularly attractive for patients with small cell lung cancer (SCLC), due to the aggressive clinical course of this cancer, which makes obtaining tumor biopsies exceedingly challenging.

Methods In this study, we analyzed a prospective cohort of 49 plasma samples obtained before, during, and after treatment from 20 patients with recurrent SCLC. We conducted cfDNA low-pass whole genome sequencing (0.1X coverage), comparing it with time-point matched tumor characterized using whole-exome (130X) and transcriptome sequencing.

Results A direct comparison of cfDNA and tumor biopsy revealed that cfDNA not only mirrors the mutation and copy number landscape of the corresponding tumor but also identifies clinically relevant resistance mechanisms and cancer driver alterations not detected in matched tumor biopsies. Longitudinal cfDNA analysis reliably tracks tumor response, progression, and clonal evolution. Sequencing coverage of plasma DNA fragments around transcription start sites showed distinct treatment-related changes and captured the expression of key transcription factors such as NEUROD1 and REST in the corresponding SCLC tumors. This allowed for the prediction of SCLC neuroendocrine phenotypes and treatment responses.

Conclusions cfDNA captures a comprehensive view of tumor heterogeneity and evolution. These findings have significant implications for the non-invasive stratification of SCLC, a disease currently treated as a single entity.

Keywords Circulating cell-free DNA, Circulating tumor DNA, Whole genome sequencing, Transcription factor binding site

†Nobuyuki Takahashi and Lorinc Pongor contributed equally to this work.

*Correspondence:
Vladimir B. Teif
vteif@essex.ac.uk
Anish Thomas
anish.thomas@nih.gov
Full list of author information is available at the end of the article

Background

Circulating free DNA (cfDNA) profiling is now a fundamental practice in cancer medicine, but several key questions remain. A basic question is whether the mutational profile established through cfDNA testing reliably reproduces the mutational profile derived from a tumor biopsy. Early studies, based on small numbers of patient samples, suggested low concordance between DNA alterations detected in tumor and plasma samples from the same patient [1, 2]. Subsequent studies, mostly case reports or small cohorts querying single genes, or a panel of genes, suggested high concordance between cfDNA and tumor genomic alterations [3–7]. A study used high-depth whole genome sequencing (WGS, median read depth 187X) to find concordant clonally expanded cancer driver alterations in both cfDNA and metastatic tumor, but cfDNA additionally harbored the genomes of multiple tumor subclones [8]. However, samples in this study were pre-selected based on high cfDNA fraction using targeted sequencing approaches and deep WGS of cfDNA is clinically not feasible at present. A second question is whether cfDNA testing can be applied to settings where tumor mutations are not known a priori. Next-generation sequencing-based assays afford the opportunity to broadly examine cfDNA, but most studies identify mutations in the tumor and then determine whether the same mutation is detectable in the plasma [9] or perform targeted sequencing of recurrently mutated cancer genes in cfDNA [10]. Although these approaches have substantially advanced cfDNA as a diagnostic tool, they have limited utility when tumor mutations are not known a priori, for example due to difficulties in getting tumor biopsies, and in tumors driven by dysregulated transcriptional programs.

Small cell lung cancer (SCLC) represents a paradigm to study cfDNA. SCLC represents about 15% of all lung cancers and is marked by an exceptionally high proliferative rate, strong predilection for early metastasis and poor prognosis [11]. Obtaining SCLC tumors for molecular testing is exceedingly difficult as few patients undergo surgery and the cancer is usually extensively disseminated by the time it is diagnosed [12]. At relapse, rapid disease progression generally precludes biopsies. As a result, SCLC is not included in large-scale sequencing initiatives such as the Cancer Genome Atlas (TCGA) [13]. Detection of cfDNA is well validated in patients with SCLC. Prior studies have found that cfDNA can track the disease course and identify recurrent gene alterations in *TP53* and *RB1*, although targetable recurrent genomic alterations have not been identified [14–22]. Moreover, these studies are limited by lack of corresponding tumor samples, and the approaches used do not interrogate the transcription programs which underlie SCLC heterogeneity. Indeed, aberrations in transcription regulators are the primary genetic cause of SCLC [23, 24]. Cell cycle regulators, transcription factors (TFs) and chromatin modifiers including *RB1* and *TP53*, members of *MYC* family, *SOX2, MLL1/2*, *CREBBP-EP300*, *RBL2*, and *TP73* are frequently altered in SCLC, causing aberrant expression of a broad range of genes related to neuronal and neuroendocrine differentiation and proliferation [25–28]. Importantly, SCLC transcriptional subtypes defined by differential expression of key transcription regulators [29] have therapeutic implications [30–33], but are not associated with specific mutational patterns.

Here we perform longitudinal profiling of cfDNA and time-point matched tumor from a molecularly defined prospective cohort of patients with relapsed SCLC, asking whether cfDNA reliably reproduces the genome-wide copy number aberrations and exome-wide tumor mutational profile, and going beyond mutations, whether cfDNA can be used to infer the expression of genes in the corresponding tumors (Fig. 1).

Methods

Patients

Patients with metastatic biopsy-proven SCLC enrolled on an interventional clinical trial (ClinicalTrials.gov identifier NCT02484404, $n=20$) were included in this analysis [34]. All patients provided written informed consent. Clinicopathologic data were abstracted from the medical records. The National Cancer Institute (NCI) Laboratory of Pathology confirmed the diagnoses of SCLC. All patients had received platinum-based chemotherapy for SCLC before enrolment. All patients were treated with durvalumab 1500 mg intravenously and olaparib 300 mg twice daily until disease progression or unacceptable toxicity. Through a tumor-liquid biopsy program, all patients underwent systematic tumor and plasma sampling before starting treatment, early during (2 to 4 weeks) treatment, and at the time of disease progression (median [range]: 8.1 weeks [4.1–47.7]). A matched tumor biopsy was obtained before starting treatment, during treatment, at disease progression, and while on subsequent therapies. Forty-nine plasma samples were available, including 20 obtained pre-treatment, 17 on treatment, and 12 at disease progression. Twenty-nine tumor samples were available, 18 obtained pre-treatment, 6 on treatment, 2 at progression, and 3 at a subsequent timepoint. Plasma and tumors were time-point matched in 26 cases; in most cases, the plasma samples were obtained on the day of the biopsy, or within a 30-day window. Most of the 23 plasma samples without timepoint matched tumor were obtained on-treatment or at disease progression (Additional file 1:Fig. S1). Tumor response was assessed using Response Evaluation Criteria in Solid Tumors (RECIST)

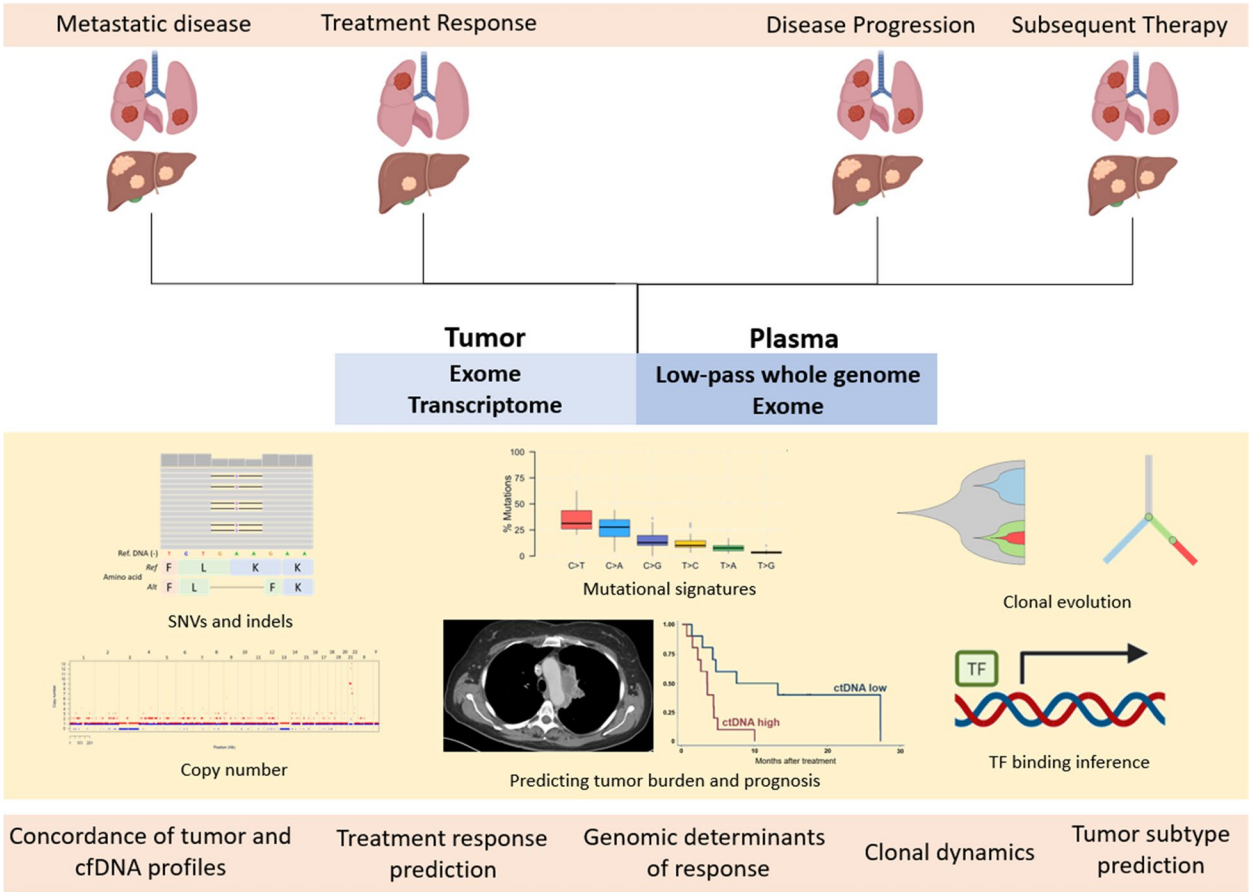

Fig. 1 Study schema. Abbreviations: cfDNA: circulating free DNA; SNV; single-nucleotide variant; indels: insertions and deletions; TF: transcription factor

v 1.1. Patient survival was followed up every 6 months by phone until death or date of cutoff. Results of clinical and exploratory studies including tumor infiltrating lymphocytes were previously reported [31, 34].

cfDNA, tumor, and germline sequencing

The DSP Circulating DNA kit from Qiagen (catalog number: 937555) was utilized to extract cfDNA from aliquots of 6.3 mL plasma which were eluted into 40–80 µL of re-suspension buffer using the Qiagen Circulating DNA kit on the QIAsymphony liquid handling system. Library preparation utilized the Kapa Hyper Prep kit with custom adapters (IDT and Broad Institute). Samples were sequenced to meet a goal of 0.1×mean coverage and Illumina NextSeq500 instruments were used for all of cfDNA sequencing with 150 bp and paired-end sequencing. Library construction was performed as previously described [35]. Hybridization and capture were performed using the relevant components of Illumina's Nextera Exome Kit and following the manufacturer's suggested protocol. For tumor sequencing, formalin-fixed, paraffin-embedded (FFPE) tumor tissue samples were prepared for whole exome sequencing (WES) and RNA sequencing. One hundred nanograms of DNA was shared to approximately 200 bp by sonication (Covaris, Woburn, MA). Exome enrichment was performed using SureSelect Clinical Research Exome Kits version 1 according to the manufacturer's instructions (Agilent, Santa Clara, CA). Paired-end sequencing (2×75 bp) was performed on an Illumina NextSeq 500 instrument. The sequences were compared to the human reference genome hg19 using internally developed somatic bioinformatic pipeline. In brief, raw sequencing data in FASTQ format were aligned against the reference human genome (hg19). RNA was extracted from FFPE tumor cores using RNeasy FFPE kits according to the manufacturer's protocol (Qiagen, Germantown, MD). For germline DNA sequencing, the patient's whole exome of peripheral blood mononuclear cells was sequenced and genotyped with the HumanOmni2.5-8v1 array (Illumina) by Personal Genome Diagnostics. Briefly, 3 µg of genomic DNA per patient sample was sequenced using the Illumina HiSeq

2000 (Illumina), generating 200 bp (2 × 100 bp reads) per fragment in the final library as described previously [32].

Variant calling, transcriptome expression, and mutational signature analysis

The Genome Analysis Toolkit (GATK) MuTect2 [36] and Strelka2 [37] were used for somatic SNV and small indel calling, respectively. ANNOVAR was used to functionally annotate genetic variants. The Sequenza software [38] was used to determine total and allele-specific DNA copy number from WES. RNA sequencing libraries were generated using TruSeq RNA Access Library Prep Kits (TruSeq RNA Exome kits; Illumina) and sequenced on NextSeq500 sequencers using 75-bp paired-end sequencing method (Illumina, San Diego, CA). Each sample was processed through an RNA sequencing data analysis pipeline where reads were mapped to the ENSEMBL human genome GRCh37 build 71 using TopHat2 [39]. Read counts for each gene between samples were transformed to Trimmed Mean of M-values-normalized Fragments per kilobase per million mapped reads (TMM-FPKM). Mutational signatures of COSMIC version 3.2 [40] were identified using the deconstructSigs [41] package (version 1.8.0), applying the same pipeline for both tissue and cfDNA samples. Mutations in maf format were processed using the *maftools* [42] package in R, which was used for statistical analysis, simple data extraction and preparing the mutation waterfall plots.

Somatic copy number alterations (SCNAs)

Somatic copy number calls were identified using *CNVkit* [43] (version 0.9.9) with default parameters. Tumor purity and ploidy were estimated by *sclust* [44] and *sequenza* [38]. The *sclust* purity values were used for adjusted copy number variation (CNV) calls using the *CNVkit* tool. Whole genome copy number representation was visualized using R (version 4.0.4) with the *rtracklayer* [45] (version 1.48.0), *ComplexHeatmap* [46] (version 2.4.3), and *ggplot2* (version 3.3.3) packages. The copy number heatmap summarizes the genome segmented into 1-Mb-sized regions, where the average CNV log2 ratio was calculated for each using the CNVkit.cns files.

Homologous recombination repair deficiency (HRD) score

Loss of heterozygosity (LOH), telomeric allelic imbalance (TAI), and large-scale state transition (LST) scores were calculated as described by Telli et al. [47]. ScarHRD was used to generate allelic imbalance profiles [48].

Phylogenetic tumor evolution

Phylogenetic trees were inferred using *PyClone* [49] (version 0.13.1) using shared variants among samples for each patient, adjusted by tumor purities calculated by *sclust*. Phylogenetic trees were prepared using the *ClonEvol* [50] (version 0.99.11) package in R (version 4.0.3). Three cfDNA and two tumor WES samples had to be removed to complete the *ClonEvol* analysis.

TF-binding site analysis

Regions with differential cfDNA occupancy were called with NucTools [51] similarly to our previous works [52–54] as detailed below. cfDNA occupancy was averaged within each 10-kb genomic window and normalized by the sequencing depth of that sample, taking into this analysis only DNA fragments with sizes 120–180 bp to account for the nucleosome protection. We determined regions which have stable nucleosome occupancy in a given condition (pre-treatment, post-treatment, progression) based on the criterion that the relative deviation of the normalized cfDNA occupancies of all individual samples with this condition determined on step 1 is < 0.5 [51]. For genomic regions which have stable nucleosome occupancy both in pre-treatment and in post-treatment conditions, we performed pairwise comparisons of the averaged normalized occupancies. We defined as "gained-nucleosome regions" regions the genomic regions where averaged normalized occupancy increased post-treatment versus pre-treatment (requiring the relative change of the average normalized occupancy > 0.4). Similarly, we defined regions where averaged normalized occupancy decreased post-treatment versus pre-treatment as "lost-nucleosome regions" (again, requiring the relative change of the average normalized occupancy > 0.4). This resulted in 267 and 342 regions where cfDNA occupancy correspondingly decreased or increased post-treatment.

These regions were then analyzed with MEME-ChIP [55] to determine transcription factors (TFs) which show enrichment of binding sites associated with differential nucleosome occupancy. The locations of binding sites of these TFs inside regions with differential cfDNA occupancy were then determined with RSAT [56]. Aggregate profiles of cfDNA occupancy around these binding sites were calculated with HOMER [57]. Analysis of binding sites of TFs experimentally profiled by chromatin immunoprecipitation sequencing (ChIP-seq) was done by finding motifs of the corresponding transcription factors inside ChIP-seq peaks with the help of gimme [58]. ChIP-seq data for *CTCF* binding in A549 SCLC cell line was obtained from GEO accession GSE175135 [59]; *CTCF* binding in healthy lung cells from GSE175135 [59]. *REST* ChIP-seq dataset in A549 cells was obtained from GSM1010749 [60]. *ASCL1* ChIP-seq in SCLC cell line DMS-53 without treatment was obtained from GSE179072 [61]. Inference of TF activity from cfDNA was also performed using the TranscriptionFactorProfiling tool [62] with default settings based on the top 50%

NEUROD1 binding sites defined in the Gene Transcription Regulation Database (GTRD) [63]. To quantify the activity of TFs, we used the calculated sequencing depth value at the center of the transcription factor binding site (TFBS).

Graph generation, statistical analysis, and code availability

All figures were generated using Origin Pro 2021 and 2024 (originlab.com), GraphPad PRISM software version 8.1.2 (GraphPad Software), R version 1.2.135 (R Foundation for Statistical Computing), and STATA software version 16.0 (Stata-Corp). All statistical tests were two-sided. P value < 0.05 was considered as statistically significant.

Results

SCLC patients have higher plasma cfDNA compared to patients with other solid tumors

cfDNA was extracted from the plasma collected at the pre-specified timepoints. The yield of cfDNA in our cohort was markedly higher than non-small cell lung cancer and breast cancer [64, 65]. We performed sparse WGS (~ 0.1 × coverage) and used a previously validated analytical approach [64] to estimate tumor fraction based on somatic copy number alterations (SCNAs) in cfDNA while accounting for subclonality and tumor ploidy. Using this approach, cfDNA tumor fraction spanned a broad range (median [interquartile range] cfDNA tumor fraction: 37.0% [11.0–49.0%]). However, all samples had detectable tumor DNA, and a high proportion of samples yielded tumor faction of > 10% (37/49, 75.5%). Most samples with lower cfDNA tumor fraction (< 10%) were collected after treatment from patients who achieved complete or partial responses ($n = 7$), or from a patient whose tumor was found to have a component of atypical carcinoid ($n = 2$). In comparison, a cohort of castration-resistant prostate cancer patients examined using a similar approach showed much lower tumor fraction (13.0% [4.0%–38.0%]) [66], and only 42% of patients with metastatic triple-negative breast cancer had tumor fraction of > 10% [67] (Fig. 2A).

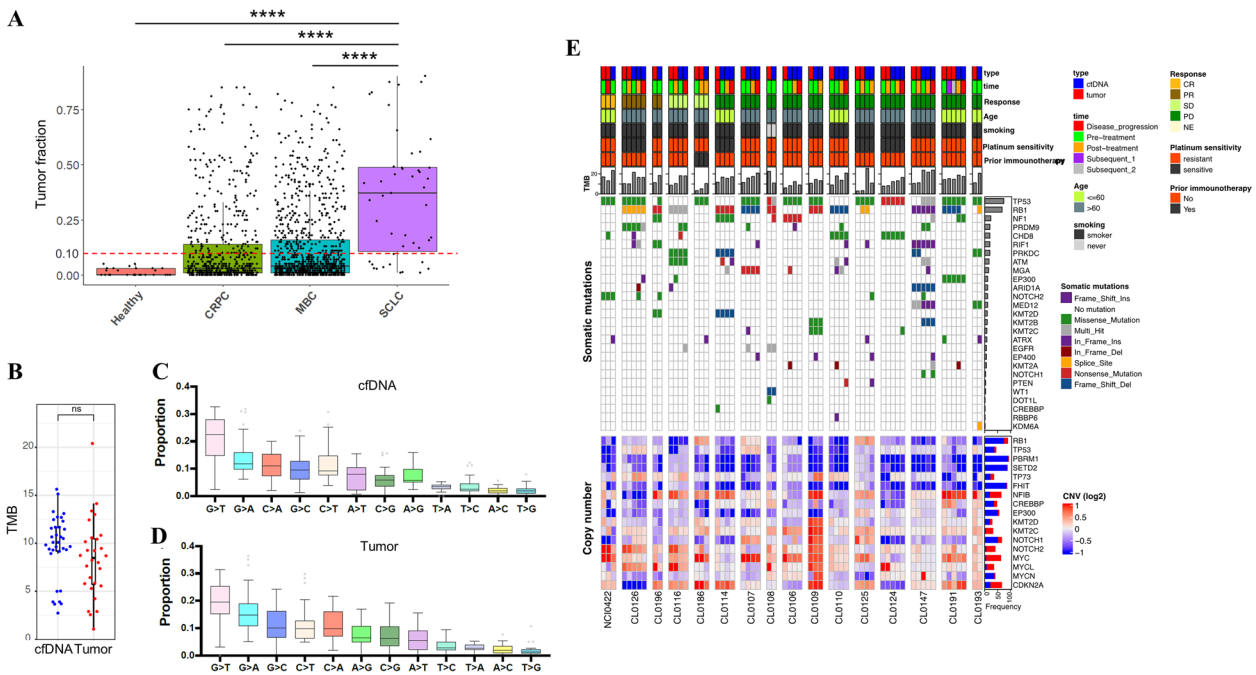

Fig. 2 Mutation profiles are highly concordant between cfDNA and tumor. **A** cfDNA tumor fraction in healthy donors and patients with CRPC, MBC, and SCLC. ****: $P < 0.0001$ by Kruskal–Wallis test followed by Dunn's multiple comparison test. **B** TMB between cfDNA and tumor samples. ns: $P > 0.05$ by Mann–Whitney U test. C, D Distributions of SNVs in cfDNA (**C**) and tumor (**D**). **E** Clinical characteristics, TMB, SNVs, and SCNAs in cfDNA and tumor. Abbreviations: cfDNA: circulating cell-free DNA; CRPC: castration resistant prostate cancer; MBC: metastatic breast cancer; TMB: tumor mutational burden; ns: not significant; SNV: single nucleotide variant; CNV: copy number variant; TIL: tumor infiltrating lymphocytes; CR: complete response; PR: partial response; SD: stable disease; PD: progressive disease; NA: not assessed; SCNA: somatic copy number alteration; SCLC: small cell lung cancer; Ins: insertion; Del: deletion. The genes in the heatmap are recurrently altered genes in SCLC [23, 68]. Platinum-sensitive is defined as disease progression ≥ 90 days after first-line platinum–based chemotherapy, and platinum-resistant as disease progression < 90 days or during first-line chemotherapy. TIL was evaluated by immunohistochemistry staining [34]

Mutation and copy number profiles are highly concordant between cfDNA and tumor

The observation that nearly every patient with SCLC had detectable cfDNA prompted us to examine how well the plasma cfDNA captured the tumor genomic features, a question particularly relevant in SCLCs which are difficult to biopsy. We performed WES of tissue samples with > 10% tumor fraction (average depth of 131X). Tumor fraction estimates from WES and the initial low-pass WGS were highly correlated with each other (Spearman's $r = 0.80$, $P < 0.0001$; Additional file 1:Fig. S2). WES identified a total of 40,354 (median: 1229 per sample) somatic single-nucleotide variants (SNVs) and 20,430 (median: 513 per sample) insertion or deletion (indels) among the 37 cfDNA samples with tumor faction of > 10%. This amounted to an average mutation rate of 11.85 mutations per megabase (range: 1.36–21.3) which was similar to that of tumors (Fig. 2B). The mutation burden of the cohort was higher than other cancer types (Additional file 1:Fig. S3A). G > T transversions were the most common mutations in both cfDNA and tumors, reflecting the mutagenic impact of tobacco smoking in SCLC tumorigenesis [69] (Fig. 2C, D, Additional file 1: Fig. S3B). Mutations and SCNAs were highly concordant between cfDNA and tumors (Fig. 2E, Additional file 1: Fig. S3C). Mutations and/or copy loss of *TP53* and *RB1* were found in majority of plasma and tumor (55 of 61 [90.2%] and 51 of 61, [83.6%], respectively), consistent with the frequent inactivation of these genes in SCLC [23]. Recurrent mutations and SCNAs of Myc paralogues (*MYC*, *MYCL*, *MYCN*), Notch genes (*NOTCH1*, *NOTCH2*), cell cycle regulators (*ATM*, *ATRX*, *CDH8*, *CDKN2A*), and chromatin modifiers (*ARID1A*, *KMT2B*, *KMT2C*, *KMT2D*, *EP300*) were identified in both cfDNA and tumors. Yet, there were several genes important for SCLC tumorigenesis and metastasis [68] that were recurrently mutated in cfDNA, but not detected in the corresponding tumor samples, such *TP53* and *KMT2B* (CL0147) and *NF1* (CL0191) (Fig. 2E).

Consistent with the known SCNA landscape of SCLCs [23, 68], deletions in chromosome 3p and 10q, and gains in 1p, 2p, 8q were recurrently observed in both cfDNA and tumors (Fig. 3A). SCNAs were highly concordant between cfDNA and tumors in timepoint-matched samples (median [range] Spearman's $r = 0.81$ [0.45–0.94], $P < 0.0001$ in all pairs) (Fig. 3B and C), and samples from the same patients obtained at different time points (median [range] Spearman's $r = 0.79$ [0.37–0.94], $P < 0.0001$ in all pairs, Fig. 3B and D), pointing to the limited evolution of SCNA with treatment. Finally, we assessed whether mutational signatures derived from genome-wide mutation and SCNA profiles might be comparable between plasma and tumor. These signatures hint at the causative origins of cancer, including infidelity of the DNA replication machinery, mutagen exposures, enzymatic modification of DNA, and defective DNA repair [40]. The distribution of mutational signatures in the plasma and tumor were highly concordant, with enrichment of tobacco-related single base substitutions (SBS) 4 and 5 dominant in both. The mutational signature profiles were also similar among cfDNA samples collected at different timepoints (Fig. 3E–G). HRD score estimated by tumor LOH, TAI and LST scores predicts efficacy of PARP inhibitors [47]. HRD scores evaluated by cfDNA and tumor samples were highly concordant between the tumor and plasma (Fig. 3H). Together, these observations show that plasma cfDNA recapitulates the tumor mutations and SCNAs, and genome-wide signatures that incorporate these features. In addition, cfDNA may capture key genomic alterations missed in small tumor biopsies.

cfDNA tracks the clinical course and reveal mechanisms of treatment response and resistance

Given the detectable levels of cfDNA in a high proportion of SCLC patients, with the broad dynamic range between patients, we assessed whether the cfDNA profiling could predict tumor burden. We performed volumetric segmentation, a three-dimensional assessment of computed tomography that may more accurately predict clinical outcomes than conventional evaluation by RECIST [70]. The cfDNA tumor fraction was significantly positively correlated with volumetric measurements evaluated by timepoint-matched computed tomography (Spearman's $r = 0.66$ $P < 0.0001$, Fig. 4A).

Given previous studies reporting shortening of cfDNA fragments in cancer [71–73], we analyzed the distribution of DNA fragment lengths to predict tumor burden (Additional file 1:Fig.S4). We considered cases with available plasma at all the three time points (pre-treatment, post-treatment, and disease progression), excluded samples with low cfDNA fraction, and used predictive metrics based on peaks heights of the size distribution: DNA fragments protected by the chromatosome (sizes around 165 bp), nucleosome core-particle (sizes around 150 bp), and TF binding (sizes around 50 bp) (Additional file 1:Fig. S4A). We calculated the ratio of the heights of chromatosome/nucleosome and chromatosome/TF peaks of the fragment size distribution. The ratio was shifted towards larger number of longer fragments after treatment, but this effect was not statistically significant over treatment time course for this patient cohort (Additional file 1:Fig. S4B and S4C).

Next, we assessed whether cfDNA profiles were associated with clinical characteristics and prognosis. Using the median pre-treatment cfDNA tumor fraction within

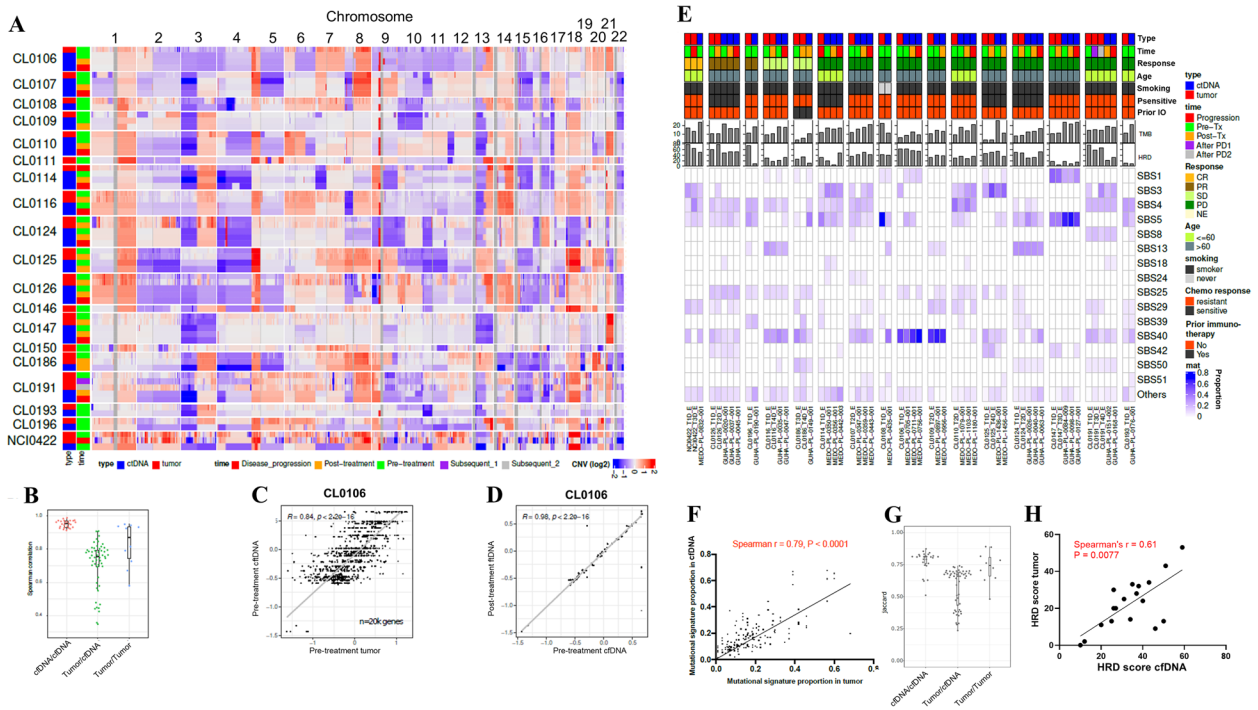

Fig. 3 Somatic copy number alterations (SCNAs) and mutational signature profiles are highly concordant between circulating cell-free DNA (cfDNA) and tumor. **A** Heatmap of SCNAs in cfDNA and tumor. In each row, samples from each patient are aligned tumor followed by cfDNA from top to bottom as indicated on the left. **B** Spearman's coefficients of correlations between cfDNA and tumor SCNA at matched or different time points. **C** Representative correlation of SCNAs between pre-treatment tumor (x-axis) and pre-treatment cfDNA (*y*-axis) in patient CL0106. Spearman's correlation coefficient (R) and *P* value are indicated. **D** Representative correlation of SCNAs between pre-treatment cfDNA (*x*-axis) and post-treatment cfDNA (*y*-axis) in patient CL0106. Spearman's correlation coefficient (R) and *P* value are indicated. **E** Clinical characteristics, TMB, HRD score, and mutational signature profiles in cfDNA and tumor COSMIC mutational signature version 3.2 [40] is computed and shown in the heatmap. Platinum-sensitive defined as disease progression ≥ 90 days after first-line platinum-based chemotherapy, and platinum-resistant disease progression < 90 days or during first-line chemotherapy. **F** Correlation of mutational signature proportions between tumor (*x*-axis) and cfDNA (*y*-axis). **G** Distribution of Jaccard index of mutational signatures between cfDNA and tumor at matched or different time points. **H** Correlation of HRD scores in pre-treatment cfDNA (*x*-axis) and tumor (*y*-axis). Abbreviations: SCNA: somatic copy number alteration; cfDNA: circulating cell-free DNA; cfDNA: circulating free DNA; SCNA: somatic copy number alteration; TMB: tumor mutational burden; HRD: homologous recombination repair deficiency; TIL: tumor infiltrating lymphocytes; CR: complete response; PR: partial response; SD: stable disease; PD: progressive disease; NE: not evaluable; NA: not assessed; COSIMIC: Catalogue of Somatic Mutations in Cancer: Psensitive: platinum sensitivity; prior IO: prior immunotherapy

the cohort, we divided the patients into low ($n=10$) or high ($n=10$) cfDNA tumor fraction groups. No significant differences were found in age, sex, smoking status, stage at diagnosis, and platinum-sensitivity between the two groups (Additional file 1:Table S1). However, progression-free survival (PFS) and overall survival (OS) durations were significantly longer in patients with low cfDNA tumor fraction than those with high cfDNA tumor fraction (median [95% confidence interval, CI] PFS and OS in low vs. high cfDNA tumor fraction groups: 2.1 months [0.9–11.0] vs. 1.0 months [0.7–1.8], $p=0.009$; 7.5 months [1.4–not estimable] vs. 3.5 months [0.7–4.4], $p=0.010$, respectively, Fig. 4B and C). After adjusting for covariates (age, sex, platinum-sensitivity) using a multivariate Cox proportional hazards model, cfDNA tumor fraction at baseline was an independent determinant of PFS and OS (Additional file 1:Table S2 and Table S3).

We then examined whether cfDNA tumor fraction can track tumor response or progression. As indicated in Fig. 4D, cfDNA tumor fraction declined or stabilized at low levels in patients who achieved complete or partial response to treatment. Marked reduction of cfDNA tumor fraction was observed in patients who achieved complete (NCI0422, Fig. 4E) or partial (CL0126, Additional file 1:Fig S5A) tumor responses. Both patients developed brain-only disease progression, but notably showed no detectable signals in the plasma (Fig. 4E, Additional file 1: Fig. S5A), suggesting that this approach may not be sensitive enough to detect intracranial tumor shedding. In contrast, cfDNA tumor fraction significantly

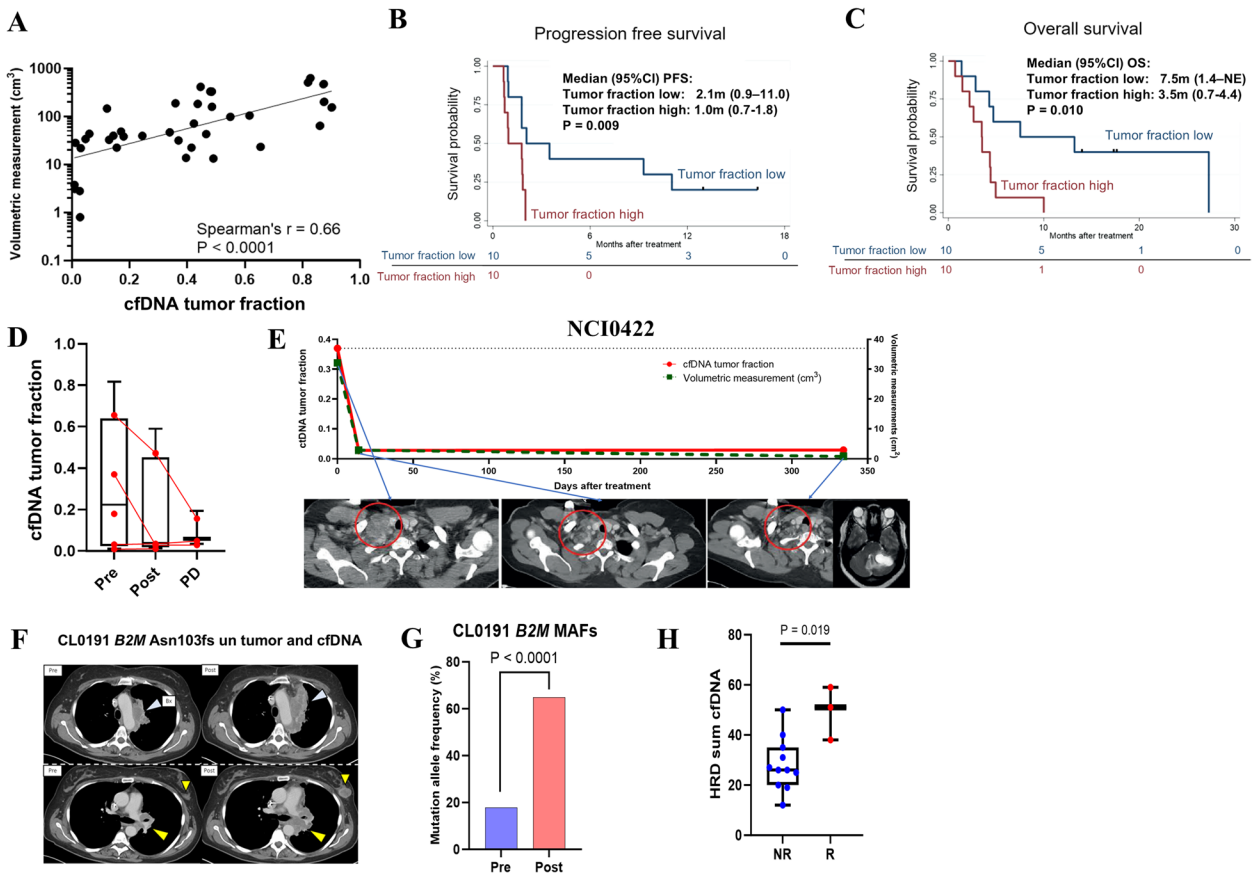

Fig. 4 cfDNA tracks the clinical course and reveal mechanisms of treatment response and resistance. **A** Correlation between cfDNA tumor fraction and volumetric measurements in time point-matched computed tomography. The *y*-axis (volumetric measurement) is logarithm transformed. **B**, **C** Kaplan–Meier curves of PFS (**B**) and OS (**C**) in patients with high vs. low cfDNA tumor fraction. High or low cfDNA tumor fraction is defined as patients whose cfDNA tumor fraction is higher or lower than the median of the cfDNA tumor fraction among all 20 samples pre-treatment. *P* values are evaluated by log-rank test. **D** Changes of cfDNA tumor fractions in patients who had CR or PR as best response. **E** Changes of cfDNA tumor fraction (red solid line, left *y*-axis) and radiological volumetric tumor measurement (green dash line, right *y*-axis) through treatment time course in a patient who had CR followed by brain only progression (NCI0422). Red circles in CT images indicate right supraclavicular lymph node metastases. **F**, **G** CT images of para-aortic lymph node metastasis (top, light blue arrowheads), left breast metastasis (bottom, small yellow arrowheads), and left mediastinal lymph node metastasis (bottom, large yellow arrowheads) in a patient who had PD as best response and with B2M Asn103fs variant in cfDNA and tumor (CL0191). "Bx" in panel **F** indicates the biopsy site for tumor sequencing. MAF changes of the B2M Asn103fs variant from pre-treatment to post-treatment cfDNA is indicated in panel **G**. *P* value is evaluated by Fisher's exact test. **H** Comparison of HRD scores derived from pre-treatment cfDNA between patients with SD or PD (= Non-responder, NR) vs. those with CR or PR (= Responder, R). *P* value is evaluated by Mann–Whitney *U* test. Abbreviations: cfDNA: circulating cell-free DNA; Tfx: tumor fraction; PFS: progression-free survival; OS: overall survival; CI: confidence interval; m: months; SCLC: small cell lung cancer; CR: complete response; PR: partial response; SD; stable disease; PD: progressive disease; HRD: homologous recombination repair deficiency; CT: computed tomography; MAF; mutation allele frequency; Bx: biopsy; cfDNA; circulating free DNA

increased over time in patients with non-responding tumors (Additional file 1:Fig. S5B). In a patient who had a minor radiographic response of a pleural lesion followed by rapid progression in liver (total volumetric measurements: from 47.2 cm^3 to 44.2 cm^3 to 205.1 cm^3 pre-treatment, post-treatment, and disease progression, respectively), the cfDNA tumor fraction tracked these radiographic changes (CL0116, Additional file 1:Fig. S5C). The cfDNA tumor fraction in a patient who had disease progression as best response steadily increased through treatment time course (CL0124, Additional file 1:Fig. S5D).

Given that SCLC responds poorly to immunotherapies despite its highly mutated genome [74], we sought to identify potential mechanisms of immunotherapy resistance from cfDNA. Truncating mutation of *B2M* (*B2M* Asn103fs), a known resistance mechanism to immune checkpoint blockade [75, 76], was identified in the cfDNA of a patient who did not respond to treatment (CL0191, Fig. 4F). Notably, the cfDNA mutation allele frequency

(MAF) of the *B2M* variant significantly increased at disease progression compared with baseline (from 17.6 to 64.9%, $p < 0.0001$ by Fisher exact test, Fig. 4G). Although anecdotal, this observation suggests clonal expansion of *B2M* as a potential resistant mechanism. Consistent with the synthetic lethality of PARP inhibition in HRD tumors [77], a cfDNA-derived HRD score [47] was significantly higher in patients who achieved tumor responses compared with those who did not (median [range] HRD score in responder vs. non-responder: 58 [38–59] vs. 26 [12–50], respectively, $P = 0.019$, Fig. 4H). Together, profiling of cfDNA can predict SCLC tumor burden, non-invasively track the disease course, and discover mechanisms of treatment response and resistance.

Longitudinal profiling of cfDNA reveal SCLC clonal architecture and track treatment responses

Since longitudinal tumor biopsies are not practical in the setting of the generally rapid progression of SCLC, the evolutionary patterns of SCLC under treatment pressure remain poorly understood. By sampling cfDNA and tumors at multiple timepoints, we evaluated the clonal evolution of SCLC through the treatment course. Phylogenic analyses revealed linear evolution in 18 of 19 patients whose longitudinal samples were successfully processed, indicating a genetic landscape that does not change markedly over treatment time course and is dominated by truncal clones harboring alterations of *TP53* and *RB1*. Prior studies have also reported low subclonal diversity in SCLC [23, 78]. However, there were few notable exceptions. In a patient who achieved complete response followed by recurrence in the brain 1 year after the treatment initiation (NCI0422), a relatively higher proportion of unique mutations were identified in the relapsed tumor compared with either pre-treatment cfDNA or tumor, whereas the pre-treatment samples harbored clonally similar cell populations (Fig. 5A and B). The median MAFs among mutations in the subclone increased at disease progression, indicating the expansion of the clone through acquired resistant mechanisms with the treatment (Fig. 5C). On the other hand, the majority of mutations were shared between cfDNA and tumor obtained at multiple time points in a patient who did not respond to treatment (CL0116, Fig. 5D and E). Three genomic clones were identified and their MAFs did not change over treatment time course (Fig. 5F).

Similar findings have been described in the context of chemotherapy wherein patients with clinical response to first-line platinum-based chemotherapy exhibited a significant increase in subclonal mutations when comparing tumors before treatment and at relapse [79]. In contrast, the number of subclonal mutations in specimens before and after chemotherapy from patients with refractory SCLC did not differ significantly. Together, the subclonal architecture of SCLC profiled non-invasively using cfDNA demonstrates a generally linear evolution with treatment, pointing to non-genetic mechanisms such as transcriptional plasticity [29, 80] in most cases.

cfDNA occupancy at TFBS predicts SCLC phenotypes and treatment response

DNA is protected from nuclease digestion through its association with a nucleosome core particle and other chromatin proteins (Additional file 1:Fig.S4A) and therefore it may be possible to infer differences in TF binding between different medical conditions from cfDNA [62, 71, 81]. We analyzed differential occupancy of cfDNA at TFBS using our recently described approach [53]. First, we identified loci with largest changes of cfDNA occupancy by scanning the genome with 10,000 base-pair (bp) sliding window and identifying regions which have similar cfDNA occupancy across all samples with the same condition (pre- or post-treatment), but significant change in occupancy post-treatment versus pre-treatment. Following these criteria (detailed in "Methods"), we identified 267 and 342 regions where cfDNA occupancy decreased or increased respectively post-treatment.

The regions with altered cfDNA occupancy were analyzed for enrichment of TFBS (Additional file 1:Table S4) and aggregate profiles of cfDNA occupancy calculated around these TFBS (Additional file 1:Fig. S6A and S6B). Interestingly, for the class of regions where cfDNA occupancy increased post-treatment, it usually declined at the time of disease progression but remained higher than at pre-treatment levels. Among the individual TFs with marked differences in cfDNA occupancy over the treatment course, the most prominent were *NRF1* and *REST* (Fig. 6A and B), which are known to co-localize on DNA binding sites, with *REST* facilitating *NRF1* occupancy by promoting local DNA hypomethylation [82]. Plotting cfDNA occupancy profiles for patients from our cohort around sites bound by *REST* in SCLC cell line A549 showed clear distinction of the pre-treatment profile from post-treatment and disease progression (Fig. 6C). Binding sites of the chromatin organizer protein *CTCF* and its close paralog *BORIS* also showed distinct changes pre- and post-treatment (Fig. 6D and E, Additional file 1: S6C and S6D), suggesting the potential impact of treatment on three-dimensional genome organization.

We also separated our cohort into two groups based on platinum sensitivity and analyzed cfDNA occupancy at TFBS. Higher *REST* occupancy was seen at disease progression in platinum-resistant compared with platinum-sensitive cases (Fig. 6F). *NEUROD1* occupancy also changed over the treatment time-course, with numerically higher occupancy at disease progression

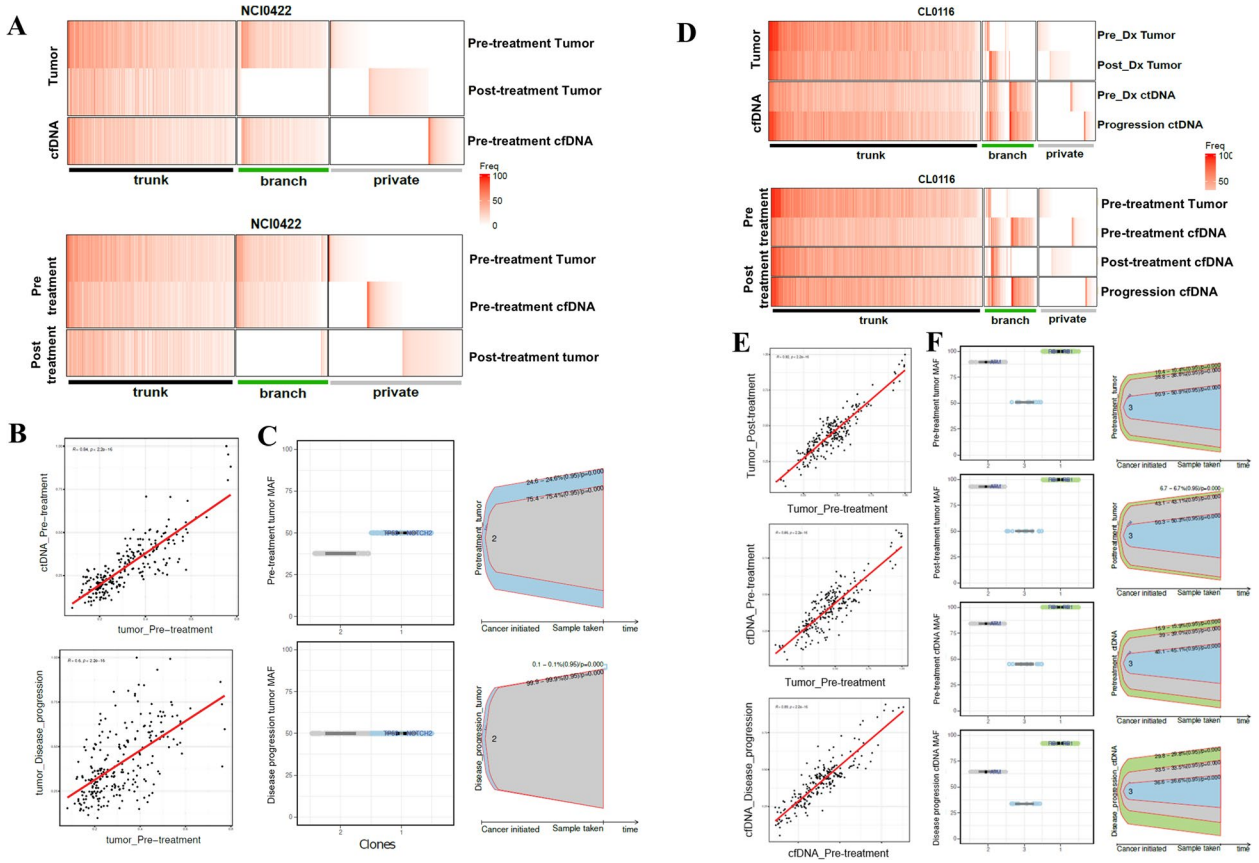

Fig. 5 Longitudinal profiling of cfDNA reveal SCLC clonal architecture and track treatment responses. **A** Mutation frequencies of variants at different time points in cfDNA and tumor from a patient who achieved complete response followed by brain only progression (NCI0422). **B** Correlations of mutation frequencies between cfDNA vs. tumor pre-treatment (top) and pre-treatment vs at disease progression tumors (bottom) in a patient who achieved complete response followed by brain only progression (NCI0422). Spearman's coefficients (R) and *P* values are indicated. **C** Visualization of genomic clones through treatment time course in a patient who achieved complete response followed by brain only progression (NCI0422). **D** Mutation frequencies of variants at different time points in cfDNA and tumor in a patient who had disease progression as the best response (CL0116). **E** Correlations of mutation frequencies between pre-treatment vs. post-treatment tumors (top), pre-treatment tumor vs. cfDNA (middle), and cfDNA pre-treatment vs at disease progression (bottom) in a patient who had disease progression as the best response (CL0116). **F** Visualization of genomic clones through treatment time course in a patient who had disease progression as the best response (CL0116). Abbreviations: cfDNA: circulating cell-free DNA; cfDNA: circulating free DNA; SCLC: small cell lung cancer; Mut. freq: mutation frequency; MAF: mutation allele frequency

in platinum-resistant compared with platinum-sensitive cases (Fig. 6G, Additional file 1: Fig. S6E). Of note, average GC content was not different genome-wide and in regions with differential cfDNA occupancy, suggesting that GC content did not confound the analysis of cfDNA occupancy (Additional file 1:Fig.S6F). The regions with differential cfDNA occupancy defined above also had very limited overlap with CNVs (only ~ 11% overlap with regions undergoing amplifications with $\log_2$fold change > 1). Thus, CNV-based and cfDNA occupancy-based analyses are complementary. The average cfDNA occupancy profiles were also not different between samples with high or low cfDNA tumor fraction (Additional file 1:Fig. S6G and S6H). Together, cfDNA TFBS occupancy analysis showed distinct treatment-related changes, which may reflect the impact of treatment on the transcriptional landscape.

The analysis performed above gave equal weights to binding sites of a given TF even if some sites contained weaker DNA binding motifs. To check whether TFBS strength influences this analysis, we also profiled SCLC-specific TFs using a recently proposed nucleosome footprint analysis method [62] which takes into account top 50% strongest TFBS from GTRD [63]. Using this approach, binding site accessibility of *NEUROD1*, an SCLC lineage defining TF [29], was significantly correlated with *NEUROD1* gene expression in the corresponding time point matched tumor (Spearman's $r =$

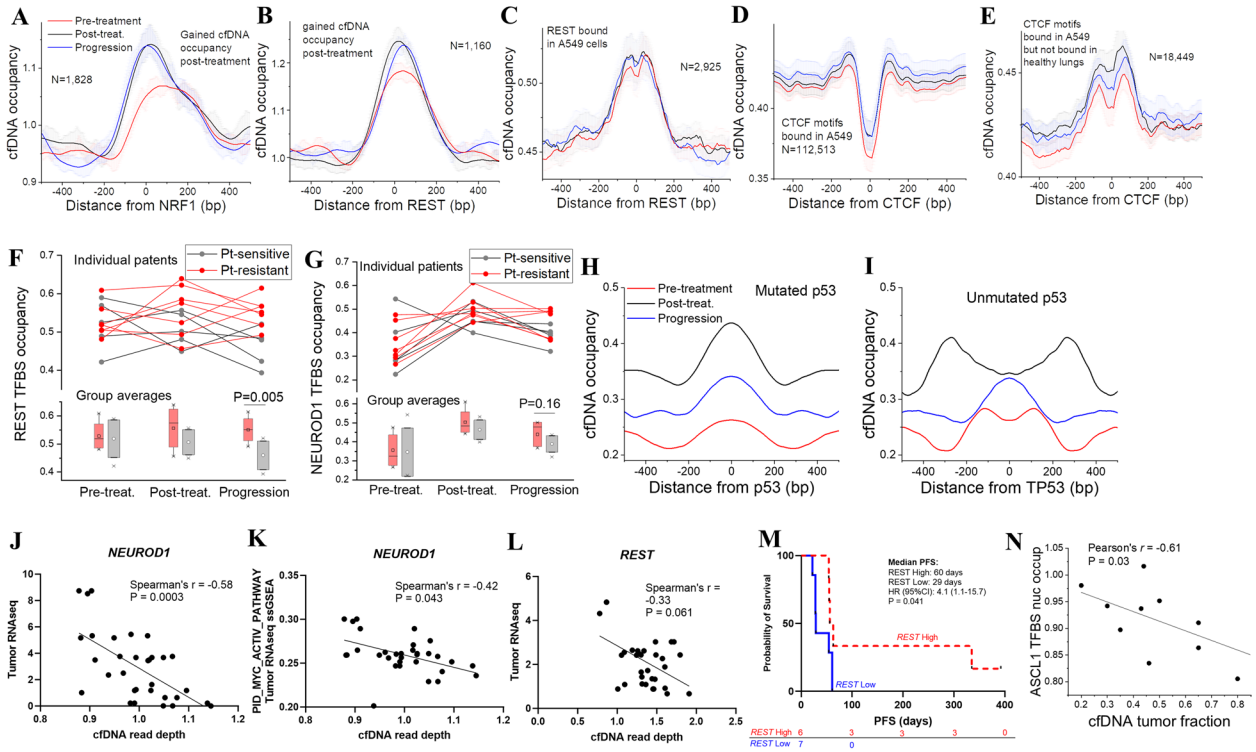

Fig. 6 cfDNA transcription factor occupancy predicts SCLC phenotypes and treatment response. **A–E** Aggregate cfDNA occupancy profiles around binding sites of *NRF1*,*REST*, and *CTCF* pre-treatment (red), post-treatment (black) and at disease progression (blue). **A** Computationally predicted *NRF1* binding sites inside regions which where cfDNA occupancy increases post-treatment. **B** cfDNA occupancy profiles around computationally predicted *REST* binding sites inside regions where cfDNA occupancy increases post-treatment. **C** cfDNA occupancy profiles around experimentally determined *REST* binding sites in A549 SCLC cells. **D** cfDNA occupancy profiles around *CTCF* binding motifs inside experimentally determined CTCF binding sites in A549 SCLC cells. **E** cfDNA occupancy profiles around a subset of *CTCF* binding sites from **D**, which do not overlap with *CTCF* sites bound in healthy lungs. **F**, **G** cfDNA occupancy at TFBSs of *REST* and *NEUROD1*. TFBSs occupancies are shown for the three time points of individual patients, along with group averages. Patients were split into two groups which were sensitive (grey circles and connecting lines) and resistant (red circles and lines) to platinum-based chemotherapy. Top: individual TFBS activity in *REST* (**F**) or *NEUROD1* (**G**) are shown. Bottom: averages within the groups of platinum-sensitive (Pt-sensitive) and resistant (Pt-resistant) in *REST* (**F**) or *NEUROD1* (**G**) are shown. *REST* sites are defined based on chromatin immunoprecipitation sequencing in A549 cells. *NEUROD1* motifs are defined computationally inside regions with increased cfDNA occupancy post-treatment vs pre-treatment. Platinum-sensitive defined as disease progression ≥ 90 days after first-line platinum-based chemotherapy, and platinum-resistant disease progression < 90 days or during first-line chemotherapy. *P* values are evaluated by Mann-Whitney *U* test. **H, I** cfDNA occupancy profiles at different timepoints around computationally predicted *TP53* binding sites inside regions which have increased cfDNA occupancy post-treatment, in samples with (**H**) vs. without (**I**) mutations in *TP53*. **J** Correlation between cfDNA read depth of *NEUROD1* binding sites (*x*-axis) and *NEUROD1* gene expression (TMM-FPKM) in timepoint-matched tumors (*y*-axis). Higher read depth indicates less TF binding, predicting less gene expression. **K** Correlation between *NEUROD1* cfDNA read depth at TFBS (*x*-axis) and the PID_MYC_ACTIV_PATHWAY scores by ssGSEA in timepoint-matched tumors (*y*-axis). Higher read depth indicates less TF binding, predicting less gene expression. **L** Correlation between cfDNA read depth at TFBS (*x*-axis) and gene expression of tumor RNA sequencing (TMM-FPKM, *y*-axis) in the gene *REST*. Higher read depth indicates less TF binding, predicting less gene expression. **M** A Kaplan-Meier curve of PFS in patients with high vs. low predicted *REST* expression. **N** Correlation between cfDNA tumor fraction and nucleosome occupancy at ASCL1 binding sites, the pre-treatment samples. Pearson's $r = -0.61$, $P = 0.03$. High vs. low predicted *REST* expression is defined as higher or lower than median predicted *REST* expression by cfDNA read depth among 13 patients whose pre-treatment cfDNA was successfully processed for the TFBS analysis. Higher predicted *REST* expression was defined as lower read depth and vice versa, given that higher read depth indicates less TF binding, predicting less gene expression. *P* value is evaluated by Log-rank test. Abbreviations: cfDNA: circulating cell-free DNA; bp: base pair; TFBS: transcriptional factor binding site; Pt: platinum-based chemotherapy; TMM-FPKM: Trimmed Mean of M-values-normalized Fragments per kilo base per million mapped reads; ssGSEA; single sample gene set enrichment analysis; PFS; progression-free survival; HR: hazard ratio; CI: confidence interval

−0.58, $P = 0.0003$; higher cfDNA read depth indicates less binding of TFs predicting less gene expression) (Fig. 6 J). Consistent with MYC driving the NEUROD1-high SCLC subtype [29, 83, 84], increased accessibility of *NEUROD1* in cfDNA was associated with upregulation of MYC transcriptional targets in the corresponding tumors (Spearman's $r = -0.51$, $P = 0.0030$) (Fig. 6 K). Binding site accessibility of *REST*, a transcriptional

repressor of neuroendocrine differentiation [29, 84] was higher in cfDNA corresponding to tumor samples with increased *REST* expression (Spearman's $r = -0.33$, $P = 0.061$, Fig. 6L). Importantly, patients with higher predicted *REST* expression based on lower *REST* read depth in cfDNA had prolonged PFS following immunotherapy than patients with lower predicted expression (Fig. 6 M). *REST* binding site accessibility was not associated with OS (Additional file 1:Fig. S7). These results are consistent with recent observations of low neuroendocrine SCLC differentiation driven by Notch signaling which targets REST being predictive of benefit from immunotherapy-based approaches, but not a prognostic marker [31, 33]. Another important TF that determines SCLC subtypes is ASCL1 [85]. We made use of a recent ASCL1 ChIP-seq dataset reported for SCLC cells [61]. Similar to many other TFs that we investigated, the cohort-average nucleosome profiles around ASCL1 binding sites were consistently similar between different time points. Interestingly, ASCL1 appeared to be a good predictor of cfDNA tumor fraction of individual patients. Nucleosome occupancy at ASCL1 binding sites was anticorrelated with cfDNA fraction, which was particularly pronounced pre-treatment (Pearson's $r = -0.61$, $P = 0.03$) (Fig. 6N). Post-treatment and in disease progression, the anticorrelation of nucleosome occupancy at ASCL1 binding sites remained but was not statistically significant (Additional file 1:Fig.S8). This suggests that ASCL1-related transcription program has strongly changed after treatment. Thus, TFBS occupancy/accessibility estimation derived from cfDNA can inform prediction of SCLC neuroendocrine phenotypes and treatment response.

Discussion

There are few systematic comparisons of the effectiveness of cfDNA at elucidating tumor-derived molecular features relative to standard single-lesion tumor biopsies in prospective cohorts of patients. The use of plasma instead of tissue to guide therapy is a particularly attractive alternative for patients with SCLC [86–89], driven by transcription addiction, and whose clinical course makes it exceedingly challenging to obtain tumor biopsies. Here, in a prospective cohort of molecularly defined patients with recurrent SCLC, treated uniformly with an immunotherapy-based combination, we find that cfDNA not only mirrored the mutation and copy number landscape of the tumor, but also its genomic signatures, while also detecting clinically relevant resistance mechanisms, and cancer driver alterations not found in matched tumor biopsies. Our study also confirmed previous observations of high cfDNA tumor fraction in patients with SCLC compared with other solid tumors [15], as well as the utility of longitudinal cfDNA analysis to reliably track tumor response and progression, and reveal mechanisms of treatment response and resistance [90]. cfDNA tumor fraction in patients with CR or PR was overall decreased post-treatment (Fig. 4D), whereas increased in patients with SD or PD (Additional file 1:Fig.S5B). Although further studies with large sample size are required, it suggests that a potential of prediction of radiological response using cfDNA.

Most applications of cfDNA to date are gene-centric focusing on somatic variants and are of limited utility when tumor mutations are not known a priori and in tumors driven by dysregulated transcriptional programs. We find that cfDNA TF binding profiles are reflective of the altered transcriptional landscape in response to treatment (before vs. after treatment vs. tumor progression) and chemo-sensitivity (platinum sensitive vs. resistant). We find a striking association between cfDNA accessibility of *NEUROD1* inferred from nucleosome footprint analysis and expression of *NEUROD1* in the corresponding tumor. A similar trend was observed between *REST* accessibility and expression, which defined tumors with low neuroendocrine differentiation and higher likelihood of response to immunotherapy [31, 33]. Another important finding is that nucleosome occupancy at binding sites of *ASCL1* shows strong anticorrelation with cfDNA fraction. Both *NEUROD1* and *ASCL1* have been proposed previously as critical markers for SCLC subtyping [85]. *NEUROD1*, *ASCL1*, and *REST* binding sites have been used recently for SCLC subtyping based on targeted high-coverage cfDNA sequencing combined with gene expression [89]. Our finding that TFBS can be used for the analysis of liquid biopsies with ultra-low sequencing coverage of cfDNA provides an important step forward to affordable tumor subtyping based on cfDNA that does not require explicit knowledge of gene expression. Our TF analyses are not limited to classical SCLC markers, as shown for a number of TFs enriched in regions with differential nucleosome occupancy, and an observation that binding sites of *CTCF* and *BORIS* are very sensitive hotspots of differential cfDNA coverage in SCLC. SCLC tumors exhibit distinct inter-tumor heterogeneity with respect to expression of neuroendocrine features, driven by expression of lineage TFs [29, 33]. Whether the subtypes engender specific therapeutic vulnerabilities is an area of active investigation [30–33]. A major barrier to clinical validation of the proposed subtypes is the limited availability of high-quality tumors for molecular analyses. The number of available sequenced cfDNA datasets from different cancer subtypes reported by different labs continues to increase exponentially [91], thus the clinical validation of this study can be expected quite fast. Our findings—the distinction between SCLC neuroendocrine phenotypes based on cfDNA accessibility of binding sites

of TFs such as *REST*, *ASCL1*, and *NEUROD1*—could allow for noninvasive characterization and treatment for patients.

Our cohort was limited by the small sample size, of whom only few patients had clinical benefit from the treatment [34]. Post-treatment plasma was collected approximately 2 weeks after treatment initiation, which might not have enough time to examine genomic dynamics through treatment time course. Future studies are needed to validate these findings in a general SCLC population and in other tumor types. Nevertheless, our cohort represents a prospective population, and the collection and processing of all samples was performed in a systematic fashion, ensuring homogeneity of pre-analytical characteristics and careful control of experimental and analytical variables. Targeted cfDNA sequencing approaches enriching cfDNA fragments covering known transcription start sites across the genome or binding sites of SCLC-specific TFs might make this approach more amenable to clinical application.

Conclusions

By direct comparisons of cfDNA versus tumor biopsy, we offer insights into non-invasive stratification and subtype-specific therapies for SCLC, now treated as a single disease, and has broad implications for mapping tumor-specific transcription factor binding on blood samples. Further studies with large cohort in a prospective manner are warranted.

Abbreviations

cfDNA	Circulating cell-free DNA
SCLC	Small cell lung cancer
TCGA	The Cancer Genome Atlas
NCI	National Cancer Institute
PCR	Polymerase chain reaction
FFPE	Formalin-fixed, paraffin-embedded
TMM-FPKM	Trimmed Mean of M-values-normalized Fragments per kilobase per million mapped reads
LOH	Loss of heterozygosity
TAI	Telomeric allelic imbalance
LST	Large-scale state transition
HMM	Hidden Markov model
ASCN	Allele specific copy number
NSR	Noise to signal ratio
PARP	Poly (ADP-ribose) polymerase
WES	Whole exome sequencing
SNV	Single-nucleotide variants
Indel	Insertion or deletion
SBS	Single base substitutions
HRD	Homologous recombination repair deficiency
SCNA	Somatic copy number alteration
RECIST	Response Evaluation Criteria in Solid Tumors
PFS	Progression-free survival
OS	Overall survival
MAF	Mutation allele frequency
TFBS	Transcription factor binding sites
GTRD	Gene Transcription Regulation Database

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13073-025-01438-4.

Additional file 1: Fig. S1. cfDNA and tumor sampling schema. Fig. S2. Copy number-predicted tumor purity correlated with cfDNA tumor fraction. Fig. S3. Mutation profiles of small cell lung cancer (SCLC) circulating free DNA (cfDNA). Fig. S4. Circulating cell free DNA (cfDNA) fragment length analysis. Fig. S5. Circulating cell free DNA (cfDNA) tracks the clinical course. Fig. S6. Dynamic changes of circulating cell-free DNA (cfDNA) occupancy around transcription factor binding sites. Fig. S7. A Kaplan-Meier curve of overall survival (OS) in patients with high vs. low predicted REST expression. Fig. S8. Correlation of cfDNA tumor fraction and nucleosome occupancy at binding sites of a transcription factor ASCL1. Table S1: Comparisons of clinical characteristics between patients with high vs. low pre-treatment circulating cell-free (cfDNA) tumor fraction. Table S2: Multivariate Cox regression analysis of progression-free survival (PFS) between patients with high vs. low pre-treatment circulating cell-free DNA (cfDNA) tumor fraction. Table S3: Multivariate Cox regression analysis of overall survival (OS) between patients with high vs. low pre-treatment circulating free DNA fraction. Table S4: Names of transcription factors (TFs) and the number of binding sites for TFs enriched in genomic regions with differential circulating cell-free DNA (cfDNA) occupancy post-treatment vs. pre-treatment.

Acknowledgements

The authors gratefully acknowledge the contributions of the patients who participated in the study.

Authors' contributions

NT, LP, and VNR performed data analysis and interpretation. AG, SPA, MS, and VT performed analyses of cfDNA occupancy at TFBS. AS and MB performed volumetric segmentation of CT scan. CWS, SK, and PD provided supports for interpretations and discussion about results. DR, PC, and WDF corresponded specimen storage. NT, RV, SN, PD, and AT provided clinical supports during the clinical trial. NT, LP, VNR, SK, DP, and AT contributed to manuscript design and interpretation of data. NT, LP and AT wrote the manuscript with comments and contributions from all authors. All authors read and approved the final manuscript.

Funding

Open access funding provided by the National Institutes of Health This study was supported by the Center for Cancer Research, the Intramural Program of the National Cancer Institute (ZIA BC 011793). AT reports research funding from AstraZeneca, Tarveda, EMD Serono, and Prolynx. VBT was supported by Cancer Research UK grants EDDPMA-Nov21\100044 and SEB-PCTA-2022/100001 and Pancreatic Cancer UK grant EDG2023-24_Teif.

Data availability

Sequencing data (cfDNA WGS, tumor WES, RNA sequencing) used in this study is deposited in dbGaP (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs003689.v1.p1) [92]. Codes for the GTRD analysis are available in https://github.com/PeterUlz/TranscriptionFactorProfiling/tree/master/Ref/GTRD [62, 93]. Codes for TF-binding site analysis are available in cfDNAtools https://github.com/TeifLab/cfDNAtools [52, 53] and NucTools https://homeveg.github.io/nuctools/ [51].

Declarations

Ethics approval and consent to participate

The trial was conducted under a NCI Center for Cancer Research–sponsored investigational new drug application with institutional review board approval (15-c-0145). Written informed consent was obtained from all patients as well as to use and share data and specimens collected for the study. The study was carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving humans.

Consent for publication
Informed consent for publication was obtained from the patients who participated in the study.

Competing interests
AT reports research funding from AstraZeneca, Tarveda, EMD Serono, and Prolynx. The remaining authors declare that they have no competing interests.

Author details
[1]Developmental Therapeutics Branch, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20892, USA. [2]Department of Medical Oncology, National Cancer Center Hospital East, Kashiwa, Japan. [3]School of Life Sciences, University of Essex, Colchester, UK. [4]Department of Radiology and Imaging Sciences, Center for Cancer Research, National Cancer Institute, Bethesda, USA. [5]Department of Internal Medicine, Seoul National University Bundang Hospital, Seoul National University College of Medicine, Seongnam, Korea. [6]Genitourinary Malignancies Branch, Center for Cancer Research, National Cancer Institute, Bethesda, USA.

Received: 17 September 2023 Accepted: 11 February 2025
Published online: 25 February 2025

References

1. Kuderer NM, Burton KA, Blau S, Rose AL, Parker S, Lyman GH, Blau CA. Comparison of 2 Commercially Available Next-Generation Sequencing Platforms in Oncology. Jama Oncol. 2017;3(7):996–8.
2. Chae YK, Davis AA, Carneiro BA, Chandra S, Mohindra N, Kalyan A, et al. Concordance between genomic alterations assessed by next-generation sequencing in tumor tissue or circulating cell-free DNA. Oncotarget. 2016;7(40):65364–73.
3. Thierry AR, Pastor B, Jiang ZQ, Katsiampoura AD, Parseghian C, Loree JM, et al. Circulating DNA Demonstrates Convergent Evolution and Common Resistance Mechanisms during Treatment of Colorectal Cancer. Clin Cancer Res. 2017;23(16):4578–91.
4. 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.
5. Goyal L, Saha SK, Liu LY, Siravegna G, Leshchiner I, Ahronian LG, et al. Polyclonal Secondary FGFR2 Mutations Drive Acquired Resistance to FGFR Inhibition in Patients with FGFR2 Fusion-Positive Cholangiocarcinoma. Cancer Discov. 2017;7(3):252–63.
6. Russo M, Siravegna G, Blaszkowsky LS, Corti G, Crisafulli G, Ahronian LG, et al. Tumor Heterogeneity and Lesion-Specific Response to Targeted Therapy in Colorectal Cancer. Cancer Discov. 2016;6(2):147–53.
7. 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.
8. Herberts C, Annala M, Sipola J, Ng SWS, Chen XE, Nurminen A, et al. Deep whole-genome ctDNA chronology of treatment-resistant prostate cancer. Nature. 2022;608(7921):199–208.
9. Powles T, Assaf ZJ, Davarpanah N, Banchereau R, Szabados BE, Yuen KC, et al. ctDNA guiding adjuvant immunotherapy in urothelial carcinoma. Nature. 2021;595(7867):432-+.
10. Cohen JD, Li L, Wang YX, Thoburn C, Afsari B, Danilova L, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science. 2018;359(6378):926-+.
11. Rudin CM, Brambilla E, Faivre-Finn C, Sage J. Small-cell lung cancer. Nat Rev Dis Primers. 2021;7(1):3.
12. Thomas A, Pattanayak P, Szabo E, Pinsky P. Characteristics and Outcomes of Small Cell Lung Cancer Detected by CT Screening. Chest. 2018;154(6):1284–90.
13. Thomas A, et al. Translational research: A patient-centered approach to bridge the valley of death. Cancer Cell. 2022. PMID: 35561672.
14. Almodovar K, Iams WT, Meador CB, Zhao Z, York S, Horn L, et al. Longitudinal Cell-Free DNA Analysis in Patients with Small Cell Lung Cancer Reveals Dynamic Insights into Treatment Efficacy and Disease Relapse. J Thorac Oncol. 2018;13(1):112–23.
15. Nong J, Gong Y, Guan Y, Yi X, Yi Y, Chang L, et al. Circulating tumor DNA analysis depicts subclonal architecture and genomic evolution of small cell lung cancer. Nat Commun. 2018;9(1):3114.
16. Mohan S, Foy V, Ayub M, Leong HS, Schofield P, Sahoo S, et al. Profiling of Circulating Free DNA Using Targeted and Genome-wide Sequencing in Patients with SCLC. J Thorac Oncol. 2020;15(2):216–30.
17. Mondelo-Macía P, García-González J, León-Mateos L, Castillo-García A, López-López R, Muinelo-Romay L, Díaz-Peña R. Current Status and Future Perspectives of Liquid Biopsy in Small Cell Lung Cancer. Biomedicines. 2021;9(1):48.
18. Fernandez-Cuesta L, Perdomo S, Avogbe PH, Leblay N, Delhomme TM, Gaborieau V, et al. Identification of Circulating Tumor DNA for the Early Detection of Small-cell Lung Cancer. EBioMedicine. 2016;10:117–23.
19. Devarakonda S, Sankararaman S, Herzog BH, Gold KA, Waqar SN, Ward JP, et al. Circulating Tumor DNA Profiling in Small-Cell Lung Cancer Identifies Potentially Targetable Alterations. Clin Cancer Res. 2019;25(20):6119–26.
20. Board RE, Williams VS, Knight L, Shaw J, Greystoke A, Ranson M, et al. Isolation and extraction of circulating tumor DNA from patients with small cell lung cancer. Ann N Y Acad Sci. 2008;1137:98–107.
21. Du M, Thompson J, Fisher H, Zhang P, Huang CC, Wang L. Genomic alterations of plasma cell-free DNAs in small cell lung cancer and their clinical relevance. Lung cancer (Amsterdam, Netherlands). 2018;120:113–21.
22. Iams WT, Almodovar K, Meador C, York SJ, Horn L, Raymond C, et al. Longitudinal monitoring of circulating tumor DNA and peripheral T cell repertoire in patients with small cell lung cancer. J Clin Oncol. 2017;35(15_suppl):8571-.
23. George J, Lim JS, Jang SJ, Cun Y, Ozretic L, Kong G, et al. Comprehensive genomic profiles of small cell lung cancer. Nature. 2015;524(7563):47–53.
24. Sun Y, et al. CDK7 Inhibition Synergizes with Topoisomerase I Inhibition in Small Cell Lung Cancer Cells by Inducing Ubiquitin-Mediated Proteolysis of RNA Polymerase II. Mol Cancer Ther. 2022. PMID: 35830858.
25. Kim DW, Kim KC, Kim KB, Dunn CT, Park KS. Transcriptional deregulation underlying the pathogenesis of small cell lung cancer. Transl Lung Cancer Res. 2018;7(1):4–20.
26. Pongor LS, Schultz CW, Rinaldi L, Wangsa D, Redon CE, Takahashi N, et al. Extrachromosomal DNA Amplification Contributes to Small Cell Lung Cancer Heterogeneity and Is Associated with Worse Outcomes. Cancer Discov. 2023;13(4):928–49.
27. Thomas A, et al. Small cell lung cancer: Time to revisit DNA-damaging chemotherapy. Sci Transl Med. 2016. PMID: 27384345.
28. Takahashi N, et al. Replication stress defines distinct molecular subtypes across cancers. Cancer Res Commun. 2022. PMID: 36381660.
29. Rudin CM, Poirier JT, Byers LA, Dive C, Dowlati A, George J, et al. Molecular subtypes of small cell lung cancer: a synthesis of human and mouse model data. Nat Rev Cancer. 2019;19(5):289–97.
30. Thomas A, Takahashi N, Rajapakse VN, Zhang X, Sun Y, Ceribelli M, et al. Therapeutic targeting of ATR yields durable regressions in small cell lung cancers with high replication stress. Cancer Cell. 2021;39(4):566-79.e7.
31. Roper N, Velez MJ, Chiappori A, Kim YS, Wei JS, Sindiri S, et al. Notch signaling and efficacy of PD-1/PD-L1 blockade in relapsed small cell lung cancer. Nat Commun. 2021;12(1):3880.
32. Lissa D, Takahashi N, Desai P, Manukyan I, Schultz CW, Rajapakse V, et al. Heterogeneity of neuroendocrine transcriptional states in metastatic small cell lung cancers and patient-derived models. Nat Commun. 2022;13(1):2023.
33. Gay CM, Stewart CA, Park EM, Diao L, Groves SM, Heeke S, et al. Patterns of transcription factor programs and immune pathway activation define four major subtypes of SCLC with distinct therapeutic vulnerabilities. Cancer Cell. 2021;39(3):346-60.e7.
34. Thomas A, Vilimas R, Trindade C, Erwin-Cohen R, Roper N, Xi L, et al. Durvalumab in Combination with Olaparib in Patients with Relapsed SCLC: Results from a Phase II Study. J Thorac Oncol. 2019;14(8):1447–57.
35. Fisher S, Barry A, Abreu J, Minie B, Nolan J, Delorey TM, et al. A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. Genome Biol. 2011;12(1):R1.
36. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol. 2013;31(3):213–9.

37. Kim S, Scheffler K, Halpern AL, Bekritsky MA, Noh E, Källberg M, et al. Strelka2: fast and accurate calling of germline and somatic variants. Nat Methods. 2018;15(8):591–4.
38. Favero F, Joshi T, Marquard AM, Birkbak NJ, Krzystanek M, Li Q, et al. Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. Ann Oncol. 2015;26(1):64–70.
39. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013;14(4):R36.
40. Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, et al. The repertoire of mutational signatures in human cancer. Nature. 2020;578(7793):94–101.
41. Rosenthal R, McGranahan N, Herrero J, Taylor BS, Swanton C. DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. Genome Biol. 2016;17:31.
42. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res. 2018;28(11):1747–56.
43. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. PLoS Comput Biol. 2016;12(4):e1004873.
44. Cun Y, Yang TP, Achter V, Lang U, Peifer M. Copy-number analysis and inference of subclonal populations in cancer genomes using Sclust. Nat Protoc. 2018;13(6):1488–501.
45. Lawrence M, Gentleman R, Carey V. rtracklayer: an R package for interfacing with genome browsers. Bioinformatics (Oxford, England). 2009;25(14):1841–2.
46. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics (Oxford, England). 2016;32(18):2847–9.
47. Telli ML, Timms KM, Reid J, Hennessy B, Mills GB, Jensen KC, et al. Homologous Recombination Deficiency (HRD) Score Predicts Response to Platinum-Containing Neoadjuvant Chemotherapy in Patients with Triple-Negative Breast Cancer. Clin Cancer Res. 2016;22(15):3764–73.
48. Sztupinszki Z, Diossy M, Krzystanek M, Reiniger L, Csabai I, Favero F, et al. Migrating the SNP array-based homologous recombination deficiency measures to next generation sequencing data of breast cancer. NPJ Breast Cancer. 2018;4:16.
49. Roth A, Khattra J, Yap D, Wan A, Laks E, Biele J, et al. PyClone: statistical inference of clonal population structure in cancer. Nat Methods. 2014;11(4):396–8.
50. Dang HX, White BS, Foltz SM, Miller CA, Luo J, Fields RC, Maher CA. ClonEvol: clonal ordering and visualization in cancer sequencing. Ann Oncol. 2017;28(12):3076–82.
51. Vainshtein Y, Rippe K, Teif VB. NucTools: analysis of chromatin feature occupancy profiles from high-throughput sequencing data. BMC Genomics. 2017;18(1):158.
52. Piroeva KV, McDonald C, Xanthopoulos C, Fox C, Clarkson CT, Mallm JP, et al. Nucleosome repositioning in chronic lymphocytic leukemia. Genome Res. 2023;33(10):1649–61.
53. Jacob DR, Guiblet WM, Mamayusupova H, Shtumpf M, Ciuta I, Ruje L, et al. Nucleosome reorganisation in breast cancer tissues. Clin Epigenetics. 2024;16(1):50.
54. Shtumpf M, Jeong S, Bikova M, Mamayusupova H, Ruje L, Teif VB. Aging clock based on nucleosome reorganisation derived from cell-free DNA. Aging Cell. 2024:23:e14100.
55. Machanick P, Bailey TL. MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics. 2011;27(12):1696–7.
56. Nguyen NTT, Contreras-Moreira B, Castro-Mondragon JA, Santana-Garcia W, Ossio R, Robles-Espinoza CD, et al. RSAT 2018: regulatory sequence analysis tools 20th anniversary. Nucleic Acids Res. 2018;46(W1):W209–14.
57. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010;38(4):576–89.
58. van Heeringen SJ, Veenstra GJ. GimmeMotifs: a de novo motif prediction pipeline for ChIP-sequencing experiments. Bioinformatics. 2011;27(2):270–1.
59. ENCODE_Project_Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57–74.
60. Gertz J, Savic D, Varley KE, Partridge EC, Safi A, Jain P, et al. Distinct properties of cell-type-specific and shared transcription factor binding sites. Mol Cell. 2013;52(1):25–36.
61. Costanzo F, Martinez Diez M, Santamaria Nunez G, Diaz-Hernandez JI, Genes Robles CM, Diez Perez J, et al. Promoters of ASCL1- and NEUROD1-dependent genes are specific targets of lurbinectedin in SCLC cells. EMBO Mol Med. 2022;14(4):e14841.
62. Ulz P, Perakis S, Zhou Q, Moser T, Belic J, Lazzeri I, et al. Inference of transcription factor binding from cell-free DNA enables tumor subtype prediction and early detection. Nat Commun. 2019;10(1):4666.
63. Kolmykov S, Yevshin I, Kulyashov M, Sharipov R, Kondrakhin Y, Makeev VJ, et al. GTRD: an integrated view of transcription regulation. Nucleic Acids Res. 2021;49(D1):D104–11.
64. Adalsteinsson VA, Ha G, Freeman SS, Choudhury AD, Stover DG, Parsons HA, et al. Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. Nat Commun. 2017;8(1):1324.
65. Szpechcinski A, Chorostowska-Wynimko J, Struniawski R, Kupis W, Rudzinski P, Langfort R, et al. Cell-free DNA levels in plasma of patients with non-small-cell lung cancer and inflammatory lung disease. Br J Cancer. 2015;113(3):476–83.
66. Choudhury AD, Werner L, Francini E, Wei XX, Ha G, Freeman SS, et al. Tumor fraction in cell-free DNA as a biomarker in prostate cancer. JCI Insight. 2018;3(21):e122109.
67. Stover DG, Parsons HA, Ha G, Freeman SS, Barry WT, Guo H, et al. Association of Cell-Free DNA Tumor Fraction and Somatic Copy Number Alterations With Survival in Metastatic Triple-Negative Breast Cancer. J Clin Oncol. 2018;36(6):543–53.
68. Kim KB, Dunn CT, Park KS. Recent progress in mapping the emerging landscape of the small-cell lung cancer genome. Exp Mol Med. 2019;51(12):1–13.
69. Pfeifer GP, Hainaut P. On the origin of G –> T transversions in lung cancer. Mutat Res. 2003;526(1–2):39–43.
70. Ogunleye OA, Hsu LY, Sun CC, Nierman P, Wiestner A, Jones EC, Bagheri H. Splenic Size and Volume Measurements in Patients with Chronic Lymphocytic Leukemia. Acad Radiol. 2023;30(6):1056–65.
71. Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J. Cell-free DNA Comprises an In Vivo Nucleosome Footprint that Informs Its Tissues-Of-Origin. Cell. 2016;164(1–2):57–68.
72. Udomruk S, Orrapin S, Pruksakorn D, Chaiyawat P. Size distribution of cell-free DNA in oncology. Crit Rev Oncol Hematol. 2021;166:103455.
73. Underhill HR, Kitzman JO, Hellwig S, Welker NC, Daza R, Baker DN, et al. Fragment Length of Circulating Tumor DNA. PLoS Genet. 2016;12(7):e1006162.
74. Yarchoan M, Hopkins A, Jaffee EM. Tumor Mutational Burden and Response Rate to PD-1 Inhibition. N Engl J Med. 2017;377(25):2500–1.
75. Gettinger S, Choi J, Hastings K, Truini A, Datar I, Sowell R, et al. Impaired HLA Class I Antigen Processing and Presentation as a Mechanism of Acquired Resistance to Immune Checkpoint Inhibitors in Lung Cancer. Cancer Discov. 2017;7(12):1420–35.
76. Sade-Feldman M, Jiao YJ, Chen JH, Rooney MS, Barzily-Rokni M, Eliane JP, et al. Resistance to checkpoint blockade therapy through inactivation of antigen presentation. 2017;8(1):1136.
77. Lord CJ, Ashworth A. PARP inhibitors: Synthetic lethality in the clinic. Science. 2017;355(6330):1152–8.
78. Zhou H, Hu Y, Luo R, Zhao Y, Pan H, Ji L, et al. Multi-region exome sequencing reveals the intratumoral heterogeneity of surgically resected small cell lung cancer. Nat Commun. 2021;12(1):5431.
79. George J, Maas L, Abedpour N, Cartolano M, Kaiser L, Fischer RN, et al. Evolutionary trajectories of small cell lung cancer under therapy. Nature. 2024;627(8005):880–9.
80. Wagner AH, Devarakonda S, Skidmore ZL, Krysiak K, Ramu A, Trani L, et al. Recurrent WNT pathway alterations are frequent in relapsed small cell lung cancer. Nat Commun. 2018;9(1):3787.
81. Ulz P, Thallinger GG, Auer M, Graf R, Kashofer K, Jahn SW, et al. Inferring expressed genes by whole-genome sequencing of plasma DNA. Nat Genet. 2016;48(10):1273–8.
82. Domcke S, Bardet AF, Adrian Ginno P, Hartl D, Burger L, Schubeler D. Competition between DNA methylation and transcription factors determines binding of NRF1. Nature. 2015;528(7583):575–9.
83. Mollaoglu G, Guthrie MR, Böhm S, Brägelmann J, Can I, Ballieu PM, et al. MYC Drives Progression of Small Cell Lung Cancer to a Variant Neuroendocrine Subtype with Vulnerability to Aurora Kinase Inhibition. Cancer Cell. 2017;31(2):270–85.

84. Ireland AS, Micinski AM, Kastner DW, Guo B, Wait SJ, Spainhower KB, et al. MYC Drives Temporal Evolution of Small Cell Lung Cancer Subtypes by Reprogramming Neuroendocrine Fate. Cancer Cell. 2020;38(1):60-78.e12.
85. Borromeo MD, Savage TK, Kollipara RK, He M, Augustyn A, Osborne JK, et al. ASCL1 and NEUROD1 Reveal Heterogeneity in Pulmonary Neuroendocrine Tumors and Regulate Distinct Genetic Programs. Cell Rep. 2016;16(5):1259–72.
86. Chemi F, Pearce SP, Clipson A, Hill SM, Conway AM, Richardson SA, et al. cfDNA methylome profiling for detection and subtyping of small cell lung cancers. Nat Cancer. 2022;3(10):1260–70.
87. Fialkoff G, Takahashi N, Sharkia I, Gutin J, Pongor L, et al. bioRxiv 2022.06.24.497386. https://doi.org/10.1101/2022.06.24.497386.
88. Heeke S, Gay CM, Estecio MR, Tran H, Morris BB, Zhang B, et al. Tumor- and circulating-free DNA methylation identifies clinically relevant small cell lung cancer subtypes. Cancer Cell. 2024;42(2):225–37 e5.
89. Hiatt JB, Doebley AL, Arnold HU, Adil M, Sandborg H, Persse TW, et al. Molecular phenotyping of small cell lung cancer using targeted cfDNA profiling of transcriptional regulatory regions. Sci Adv. 2024;10(15):eadk2082.
90. Hench IB, Hench J, Tolnay M. Liquid Biopsy in Clinical Management of Breast, Lung, and Colorectal Cancer. Front Med. 2018;5:9.
91. Shtumpf M, Piroeva KV, Agrawal SP, Jacob DR, Teif VB. NucPosDB: a database of nucleosome positioning in vivo and nucleosomics of cell-free DNA. Chromosoma. 2022;131(1–2):19–28.
92. Takahashi N, Pongor L, Agrawal SP, Shtumpf M, Gurjar A et al. Genomic Alterations and Transcriptional Phenotypes in Circulating Tumor DNA and Matched Metastatic Tumor. dbGaP (dbGaP Study Accession: phs003689.v1.p1). 2025. https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs003689.v1.p1.
93. Kolmykov S, Yevshin I, Kulyashov M, Sharipov R, Kondrakhin Y, Makeev VJ, et al. TranscriptionFactorProfiling. Github. https://github.com/PeterUlz/TranscriptionFactorProfiling/tree/master/Ref/GTRD.

Publisher's Note