nature medicine

Article

https://doi.org/10.1038/s41591-023-02232-8

Body composition and lung cancer-associated cachexia in TRACERx

Received: 20 April 2022

Accepted: 24 January 2023

Published online: 12 April 2023

Check for updates

A list of authors and their affiliations appears at the end of the paper

Cancer-associated cachexia (CAC) is a major contributor to morbidity and mortality in individuals with non-small cell lung cancer. Key features of CAC include alterations in body composition and body weight. Here, we explore the association between body composition and body weight with survival and delineate potential biological processes and mediators that contribute to the development of CAC. Computed tomography-based body composition analysis of 651 individuals in the TRACERx (TRAcking non-small cell lung Cancer Evolution through therapy (Rx)) study suggested that individuals in the bottom 20th percentile of the distribution of skeletal muscle or adipose tissue area at the time of lung cancer diagnosis, had significantly shorter lung cancer-specific survival and overall survival. This finding was validated in 420 individuals in the independent Boston Lung Cancer Study. Individuals classified as having developed CAC according to one or more features at relapse encompassing loss of adipose or muscle tissue, or body mass index-adjusted weight loss were found to have distinct tumor genomic and transcriptomic profiles compared with individuals who did not develop such features. Primary non-small cell lung cancers from individuals who developed CAC were characterized by enrichment of inflammatory signaling and epithelial–mesenchymal transitional pathways, and differentially expressed genes upregulated in these tumors included cancer-testis antigen *MAGEA6* and matrix metalloproteinases, such as *ADAMTS3*. In an exploratory proteomic analysis of circulating putative mediators of cachexia performed in a subset of 110 individuals from TRACERx, a significant association between circulating GDF15 and loss of body weight, skeletal muscle and adipose tissue was identified at relapse, supporting the potential therapeutic relevance of targeting GDF15 in the management of CAC.

Measures of body composition that distinguish skeletal muscle (SKM), visceral adipose tissue (VAT) and subcutaneous adipose tissue (SAT) are associated with clinical outcomes in various diseases, including cancer[1–4]. One extreme manifestation of altered body composition that remains poorly understood is CAC; a paraneoplastic syndrome of involuntary SKM and/or adipose tissue loss, accompanied by dysregulation of the homeostatic mechanisms that govern protein and energy balance[5,6].

Retrospective analyses have linked body composition to outcomes across multiple solid tumor types, including breast, prostate and colorectal cancers[7–9]. In non-small cell lung cancer (NSCLC), SKM wasting has been shown to be associated with cancer treatment toxicity and reduced overall survival (OS)[10]. A meta-analysis of 13 NSCLC cohort studies demonstrated an association between low SKM mass and reduced OS, but not with disease-specific survival[11]. Similarly,

e-mail: m.jamal-hanjani@ucl.ac.uk; Charles.swanton@crick.ac.uk

retrospective studies have suggested an association between low SAT and low VAT with reduced OS[12,13], again, without a significant difference in cancer-specific survival. Key areas of uncertainty include how body composition changes throughout the disease course, and what are the underlying molecular mediators of altered body composition.

Our study had three aims: (1) to systematically profile body composition using standardized methods among individuals diagnosed with early-stage lung cancer in terms of SKM, VAT and SAT, to explore associations between baseline measures and clinical outcomes; (2) to examine how these body composition measures change over time in relation to clinical outcomes and how this can be used to identify individuals who develop the CAC phenotype; and (3) to explore the tumor genomic, transcriptomic and plasma proteomic landscape for potential molecular mechanisms and mediators of CAC.

We used established computed tomography (CT) imaging analysis methods[14–16] to assess body composition at lung cancer diagnosis and relapse. To investigate the interplay between tumor biology, body composition and clinical outcomes, our analysis plan involved whole-exome and transcriptomic analysis of diagnostic bulk tumor samples, in addition to plasma proteomics, to identify tumor-intrinsic factors, as well as potential circulating mediators associated with the development of CAC.

We analyzed CT images from two large independent cohorts of individuals diagnosed with early-stage NSCLC, treated with surgical resection, plus adjuvant therapy, if indicated according to tumor stage, which included a principal cohort of 651 individuals from the TRACERx study[17,18], and a validation cohort of 420 individuals from the Boston Lung Cancer Study (BLCS)[19]. Altogether, this encompassed 1,071 diagnostic CT scans paired with detailed clinical annotation, including clinical outcomes.

SAT, VAT and SKM areas at the third lumbar vertebra level[20] were determined using a deep learning-based imaging analysis pipeline, establishing the baseline body composition database. Among the 291 individuals in TRACERx who experienced disease relapse, 274/291 individuals had available body composition and/or body weight data at diagnosis and relapse, and were evaluated to identify changes indicative of the CAC phenotype (Fig. 1). Matched bulk tumor whole-exome sequencing (WES) and RNA-sequencing (RNA-seq) data from the primary tumor were also analyzed to identify genomic and transcriptomic alterations in relation to the CAC phenotype at relapse. In a subset of individuals from this cohort with available plasma samples (total of 256 samples collected at diagnosis and relapse), plasma proteomic analysis was performed to investigate differential protein expression in relation to the CAC phenotype. By integrating longitudinal imaging, tumor and blood analyses, we provide insights into the relationship between body composition and CAC in NSCLC, characterizing the tumor genome, transcriptome and plasma proteome (Fig. 2a), establishing a platform for downstream validation and potential clinical translation.

Results

Body composition at cancer diagnosis

We assessed body composition, based on SAT, VAT and SKM cross-sectional tissue area at the level of the third lumbar vertebra (L3), in participants at the time of early-stage NSCLC diagnosis to determine the prognostic association of these parameters with lung cancer-specific survival (LCSS) as the primary outcome measure for the baseline body composition study. Body composition in 651 participants from the TRACERx study and 420 participants from the BLCS study was assessed at the time of cancer diagnosis (Fig. 1). All participants in the TRACERx and BLCS cohorts underwent primary surgical resection, and had stage I–III NSCLC. In the TRACERx cohort, 42% of participants had stage I disease (BLCS, 48%), 36% had stage II disease (BLCS, 15%) and 22% had stage III disease (BLCS, 37%). In the TRACERx cohort, 35% of participants received adjuvant therapy (55% in the BLCS cohort), 14% were current smokers and 79% were ex-smokers (37% and 51% in the BLCS cohort, respectively; Table 1).

The distribution of body SAT, VAT, SKM and body mass index (BMI) was similar in the TRACERx and BLCS cohorts (Fig. 2b,c). In the TRACERx and BLCS cohorts, men had higher VAT and SKM areas compared with women: In TRACERx, mean VAT area was 171.7 cm^2 in men versus 95.1 cm^2 in women (unpaired two-samples *t*-test, $P < 0.001$) and 198.5 cm^2 in men versus 98.3 cm^2 in women ($P < 0.001$) in the BLCS cohort. Mean SKM area was 147.4 cm^2 in men versus 99.5 cm^2 in women ($P < 0.001$) and 160.2 cm^2 in men versus 109.8 cm^2 in women ($P < 0.001$) in the TRACERx and BLCS cohorts, respectively. In both studies, women had higher SAT areas compared to men; 201.2 cm^2 versus 144.6 cm^2 ($P < 0.001$, TRACERx) and 206.5 cm^2 versus 172.6 cm^2 ($P < 0.001$, BLCS; Fig. 2b,c). There was a strong correlation between BMI and both SAT and VAT (Spearman's correlation, $r = 0.75$ and 0.73, respectively), and a weaker correlation with SKM ($r = 0.39$; Extended Data Fig. 1a); similar correlations were observed in the BLCS cohort (Extended Data Fig. 1b).

Participants in the TRACERx cohort were grouped according to the distribution of body composition into sex-adjusted bottom 20th, middle 20–80th and upper 20th percentiles, representing low, normal and high values of adipose and SKM, respectively (Fig. 2d). After adjusting for age, sex, BMI, smoking status, disease stage, histological subtype, ethnicity and adjuvant therapy use, participants in the lowest 20th percentile for SAT, VAT and SKM had worse LCSS compared to participants in the middle 20–80th percentile. The adjusted hazard ratios (HRs) were 2.09 (95% confidence interval (CI) 1.33–3.30, $P = 0.001$), 1.73 (1.10–2.72, $P = 0.019$) and 1.44 (0.95–2.19, $P = 0.088$) for SAT, VAT and SKM, respectively (Supplementary Table 1). Participants in the bottom 20th percentile for all three body compartments also had trend towards worse OS compared to those in the middle 20–80th percentile, with adjusted HRs of 1.49 (1.02–2.16, $P = 0.037$), 1.38 (0.95–2.01, $P = 0.093$) and 1.28 (0.91–1.78, $P = 0.151$) for SAT, VAT and SKM, respectively (Supplementary Table 2).

These associations between the bottom 20th percentile for body composition and poor outcome in TRACERx were substantiated in the BLCS cohort. Adjusted HRs for LCSS of 1.97 (1.24–3.14, $P = 0.004$), 1.57 (0.98–2.53; $P = 0.06$) and 1.35 (0.86–2.12, $P = 0.19$) for SAT, VAT and SKM, respectively (Fig. 2e and Supplementary Table 3). For OS, the corresponding HRs were 1.71 (95% CI 1.18–2.49, $P = 0.005$), 1.75 (95% CI 1.20–2.54, $P = 0.003$) and 1.41 (95% CI 1.00–2.01, $P = 0.05$; Supplementary Table 4).

When participants were categorized into more extreme percentiles (bottom 10%, middle 10–90%, top 10%), the poor prognostic association between the bottom centile and LCSS and OS remained similar or were stronger than those observed using the less extreme thresholds (that is, bottom 20%, middle 20–80%, top 20%; Supplementary Table 5).

These independent cohorts revealed that participants with low SAT, VAT or SKM tissue at the time of NSCLC diagnosis tended to have a shorter LCSS and OS compared to participants in the middle 20–80th percentile, indicating the potential prognostic value of baseline body composition.

Identifying the cachexia phenotype

Among the 291 participants in TRACERx with a confirmed relapse (median time to relapse 14.1 months, 95% CI 12.9–16.1), 206 had relapse body composition data and 188 participants had available abdominal CT image sections both at diagnosis and relapse for the assessment of body composition change (Fig. 1 and Methods). We endeavored to ensure consistent and reproducible longitudinal measurements, reducing the impact of investigator-dependent variation (Methods), by using an automated, deep learning-based body composition method[16], as well as manual quality control of all L3 annotations and image segmentations.

A strong correlation was observed between body weight loss and VAT loss, and between VAT and SAT loss (Spearman's $r = 0.70$

https://doi.org/10.1038/s41591-023-02232-8

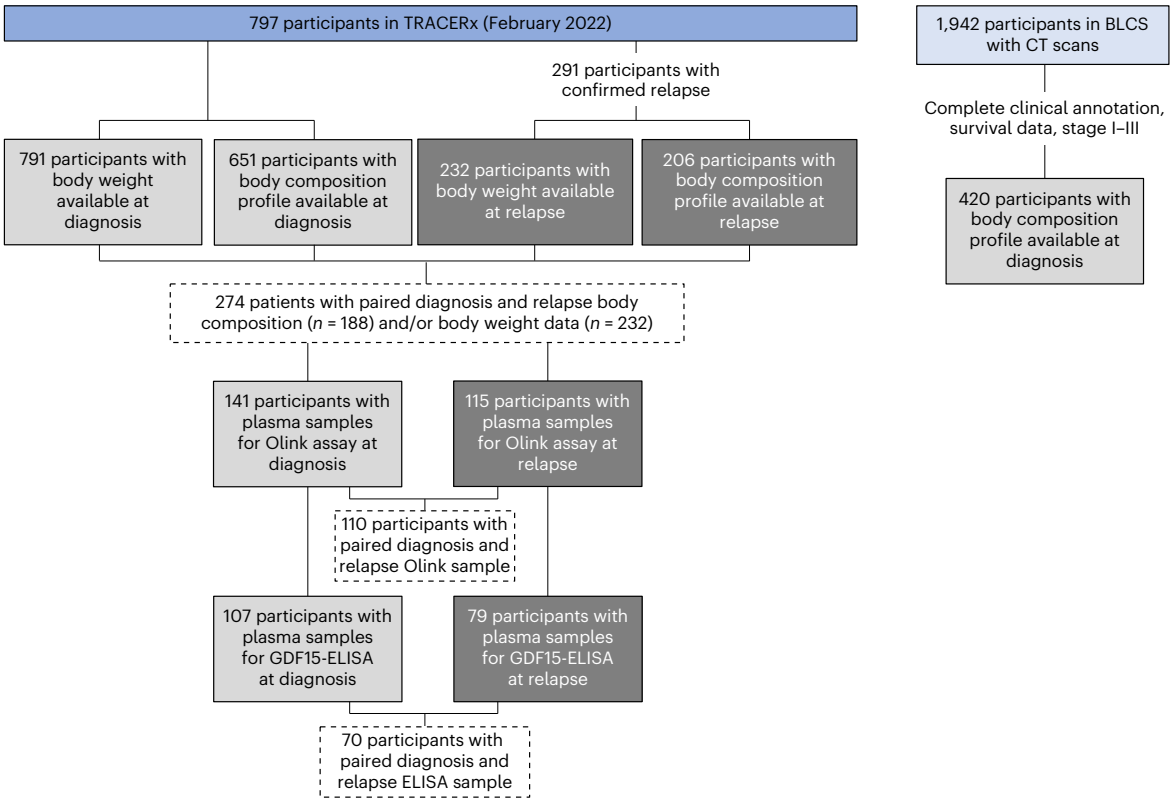


Fig. 1 | Participant and sample selection. Selection of participants and samples from the TRACERx and Boston Lung Cancer Study (BLCS) studies. All available computer tomography (CT) scans and samples at diagnosis and relapse were considered for this study. Limited availability was due to sample exhaustion or missing collection, shipment or transfer.

and 0.62; Extended Data Fig. 1c). As expected, grade 4 BMI-adjusted weight loss was also associated with loss of SAT, VAT and SKM compared to grade 0 BMI-adjusted weight loss (Methods and Extended Data Fig. 1d; $P < 0.0001$). The distribution of absolute changes (that is, gain or loss between diagnosis and relapse) in SAT, VAT, SKM (cm^2) and weight (kg) did not differ between females and males (Fig. 3a,b). Regression analyses were then used to examine the association between LCSS and varying cutoffs from 5% to 40% used to define SAT, VAT and SKM loss (Methods and Supplementary Fig. 1). A loss of ≥20% tissue area between diagnosis and relapse in SAT or VAT was associated with worse LCSS (SAT: HR 1.56 (95% CI 1.02–2.38), $P = 0.042$; VAT: HR 1.67 (95% CI 1.12–2.49), $P = 0.0111$) and worse OS (SAT: HR 1.59 (95% CI 1.08–2.36), $P = 0.019$; VAT: HR 1.72 (95% CI 1.19–2.47), $P = 0.004$) compared to participants with <20% SAT or VAT tissue area loss (Supplementary Fig. 2). A loss of ≥10% SKM tissue area in participants was associated with worse LCSS and OS compared to participants with <10% SKM tissue area loss (HR 1.80 (95% CI 1.20–2.70), $P = 0.0047$ and HR 1.85 (95% CI 1.27–2.69), $P = 0.0013$). Grade 4 BMI-adjusted weight loss in participants was associated with worse LCSS and OS compared to participants with grade 0 BMI-adjusted weight loss (HR 5.05 (95% CI 2.87–8.90), $P < 0.0001$ and HR 4.22 (95% CI 2.48–7.20), $P < 0.0001$).

To explore the occurrence of distinct patterns of loss affecting specific body compartments, participants were grouped according to isolated or co-occurring loss of SKM, VAT and SAT for the above prognostic thresholds (Fig. 3c). Among 96 participants in whom a loss occurred in any given compartment, 23 (24%) experienced losses across all three compartments, 25 (26%) experienced losses in two compartments and 48 (50%) experienced an isolated loss in one compartment (SKM ($n = 21$), VAT ($n = 22$) or SAT ($n = 5$)), suggesting distinct clinical subtypes of CAC involving different anatomical patterns of adipose and/or muscle tissue loss. Participants with co-occurring tissue loss across all three compartments had a shorter LCSS compared with participants without any tissue loss (adjusted HR: 2.6, 95% CI 1.39–4.87, $P = 0.003$; Fig. 3c and Supplementary Table 6).

Based on the above thresholds for tissue loss between diagnosis and first relapse strongly associated with LCSS and OS, we classified participants into two groups: those with a cachexia phenotype called 'CAC' ($n = 108$), defined as experiencing ≥20% SAT and/or VAT loss, and/or ≥10% SKM loss, and/or grade 4 BMI-adjusted weight loss in the interval between diagnosis and relapse, and those without called 'non-CAC' ($n = 166$; Supplementary Fig. 3). A higher proportion of men and participants with lung squamous cell carcinoma were identified in the CAC group compared to the non-CAC group (62.0% versus 51.8%, and 34.3% versus 25.9%). Smoking status and use of adjuvant therapy was similar between the two groups (Supplementary Table 7).

This categorization of participants into CAC and non-CAC groups was then used to conduct further downstream analyses to generate biological hypotheses relating to potential mediators of CAC.

Primary tumor genomic and transcriptomic features

To examine the presence of genomic and transcriptomic alterations potentially relevant to the development of CAC, we analyzed tumor sequencing data in relation to the CAC and non-CAC groups. Primary NSCLCs from participants recruited into TRACERx were subjected to multiregional WES and RNA-seq (Methods).

Distinct differential gene expression profiles, adjusted for histology and sex (Methods), were observed in primary tumors from participants in the CAC compared with the non-CAC group (Fig. 4). Gene expression of melanoma-associated antigen 6 *MAGEA6*, metalloproteinases such as *ADAMTS3*, and transcriptional and cytoskeletal regulators, such as *NR2F1* and *SPTB*, were significantly increased in the CAC group (Fig. 4a). To explore the biological relevance of

https://doi.org/10.1038/s41591-023-02232-8

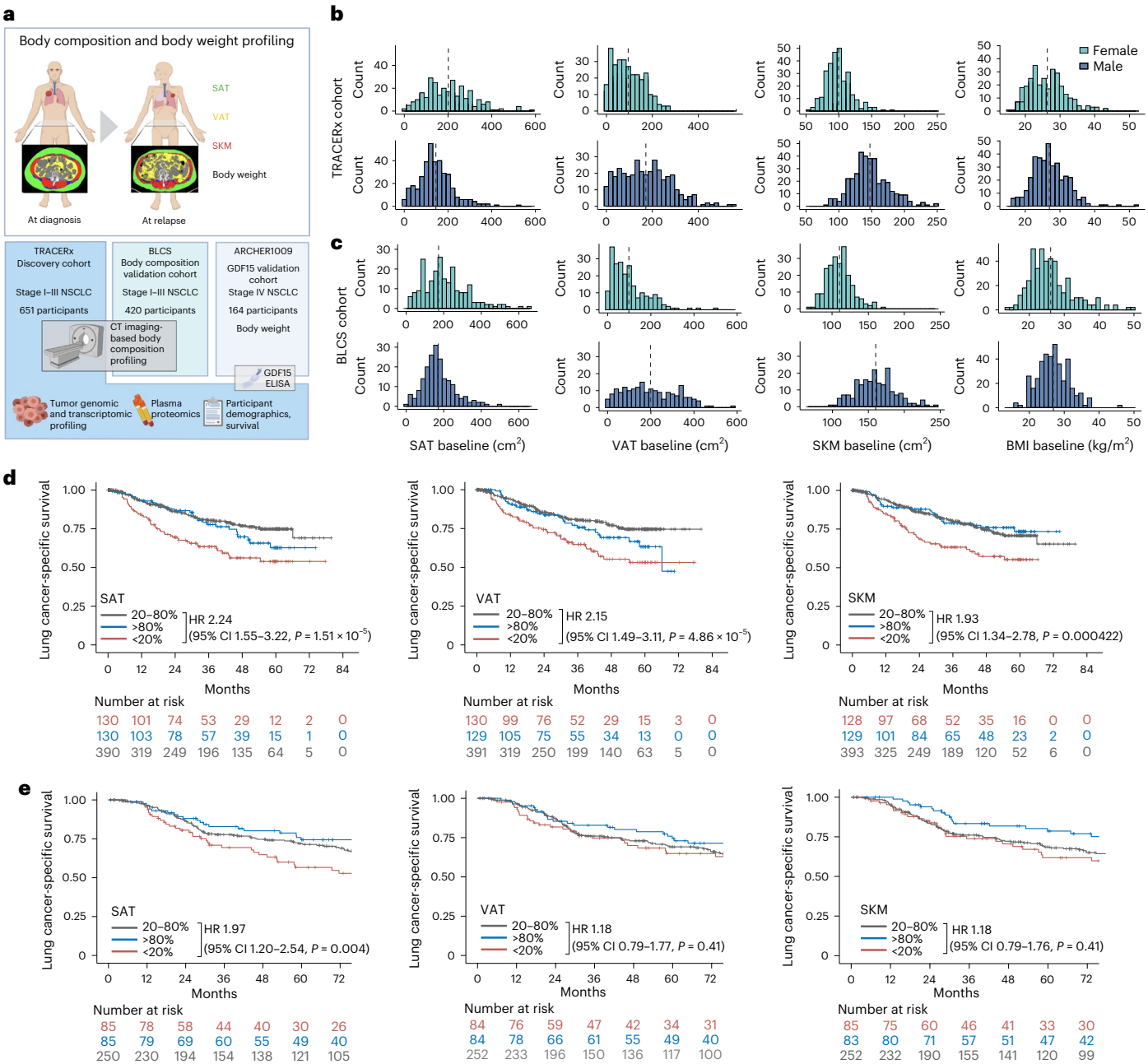


Fig. 2 | Body composition and cancer-specific survival in the TRACERx and Boston Lung Cancer Study (BLCS) studies. a, Study outline of body composition, body weight and downstream analyses. **b**,**c**, Distribution of body composition and body mass metrics according to sex at primary diagnosis in the TRACERx cohort ($n = 651$, 286 female, 265 male; **b**) and BLCS cohort ($n = 420$, 198 female, 222 male; **c**). Vertical lines indicate mean in female (green) and male (blue) participants. **d**,**e**, Lung cancers-specific survival (LCSS) in TRACERx (**d**) and BLCS (**e**) cohorts for body composition at primary diagnosis according to sex-specific bottom 20th percentile (red, <20%), mid 60th percentile (grey, 20%-80%) and upper 20th percentile (blue, >80%) of SAT, VAT and SKM. HRs, 95% CIs and unadjusted two-sided *P* values were derived from univariate Cox regression analysis of the <20% group and the 20–80% groups. SAT, subcutaneous adipose tissue; VAT, visceral adipose tissue; SKM, skeletal muscle tissue.

these differentially expressed genes in the development of CAC, we cross-referenced genes with significantly altered expression in the CAC group with a list of 400 genes that have been previously shown to be associated with, or functionally influence, the development of either obesity or cachexia in preclinical and clinical models, or derived from genome-wide association studies (GWAS) of obesity, BMI and cachexia (hereafter referred to as 'cachexia candidate gene list'; Supplementary Table 9)[21–35]. From this gene list, semaphorin-3A (encoded by *SEMA3A*) and insulin-like growth factor 1 (encoded by *IGF1*), which impact body weight[36,37], and potassium channel *KCNJ12*, and A-kinase anchoring protein 6 (*AKAP6*), for which genomic variants have been associated with increased BMI in GWAS[22], were found to be differentially upregulated in the CAC group (Fig. 4a,b).

Gene set enrichment analysis, adjusted for sex and histology (Methods), showed significant enrichment of several hallmark gene sets in primary tumors from participants in the CAC group compared to the non-CAC group, including epithelial–mesenchymal transition, hedgehog signaling and myogenesis, while estrogen-response pathways were significantly enriched in the non-CAC group (adjusted $P < 0.05$; Fig. 4c). Distinct differential gene expression profiles were also observed in primary tumors from participants who predominantly experienced isolated tissue loss in individual body compartments or

https://doi.org/10.1038/s41591-023-02232-8

Table 1 | Baseline participant characteristics of TRACERx and Boston Lung Cancer Study (BLCS) cohorts

	TRACERx (*N*=651)	BLCS (*N*=420)
Age, mean (s.d.), years	68.7 (9.1)	64.53 (10.1)
BMI, mean (s.d.), kg/m^2	26.7 (5.1)	26.48 (5.4)
Weight, mean (s.d.), kg	75.0 (16.9)	75.02 (18.2)
Height, mean (s.d.), m	1.67 (0.1)	1.68 (0.1)
VAT baseline, mean (s.d.), cm^2	138 (98.7)	145.5 (113.9)
SAT baseline, mean (s.d.), cm^2	169 (94.3)	190.5 (103.5)
SKM baseline, mean (s.d.), cm^2	126 (34.5)	133.6 (33.9)
Sex		
Female	286 (43.9%)	222 (52.9)
Male	365 (56.1%)	198 (47.1%)
Ethnicity		
White-European ancestry		411 (97.9%)
African American		9 (2.1%)
Non-white[a]	36 (5.5%)	
White-British-Irish	590 (90.6%)	
White-other	22 (3.4%)	
Missing	3 (0.5%)	
Smoking status		
Current smoker	89 (13.7%)	155 (36.9%)
Ex-smoker	513 (78.8%)	216 (51.4%)
Never smoked	49 (7.5%)	49 (11.7%)
NSCLC stage		
IA	131 (20.1%)	203 (48.3%)
IB	139 (21.4%)	
IIA	124 (19.0%)	63 (15.0%)
IIB	113 (17.4%)	
IIIA	138 (21.2%)	154 (36.7%)
IIIB	6 (0.9%)	
Histology		
Adenocarcinoma	362 (55.6%)	223 (53.1)
Squamous cell carcinoma	211 (32.4%)	104 (24.8)
Other	65 (10.0%)	93 (22.1)
Not available	13 (2.0%)	
Adjuvant treatment		
Adjuvant	227 (34.9%)	229 (54.5)
No adjuvant	408 (62.7%)	191 (45.5)
Missing	16 (2.5%)	

[a]The category 'non-white' includes 2 (0.3%) African, 2 (0.3%) Asian, 2 (0.3%) Black, 8 (1.2%) Caribbean, 6 (0.9%) Indian, 8 (1.2%) Middle Eastern, 1 (0.2%) Pakistani, 1 (0.2%) South American, 2 (0.3%) white and Asian, 4 (0.6%) white and Black.

body weight, for example ≥20% SAT loss versus <20% SAT loss, suggesting possibly distinct wasting mechanisms specific to adipose and muscle tissue (Extended Data Fig. 2 and Supplementary Fig. 4).

Analysis of somatic copy number alterations was performed using GISTIC2.0 to identify amplifications and deletions specific to the CAC and non-CAC groups, and the SAT/VAT/SKM and body weight loss prognostic thresholds (Fig. 4d and Supplementary Fig. 5)[38]. Multiple loci from the cachexia candidate gene list were found to be exclusively amplified in the CAC group, including chromosome 11q22.3 containing various metalloproteinases, such as *MMP1* and *MMP3*, which have been functionally implicated in tissue wasting in Drosophila models[29] and chromosome 3q27.1, containing *ADIPOQ*, which has been implicated in weight loss in humans[39].

Since inflammatory signaling pathways, such as interferon-alpha response, were observed to be upregulated in the CAC group, we further investigated whether this was concordant with increased inflammatory cell infiltration using the previously published T cell ExTRECT method, providing T cell infiltration estimates based on WES data[38,40]. No significant difference in T cell receptor-α gene (TCRA) scores, that is, T cell infiltrates, was observed between primary tumors in the CAC and non-CAC groups, and this remained so after adjusting for sex (Extended Data Fig. 3a–c). However, TCRA scores in the blood were higher in the non-CAC group, suggesting higher circulating T cell fractions in the circulation, although the differences were not significant when looking at female and male participants separately (Extended Data Fig. 3d–f). Danaher immune signature[41] and CIBERSORTx[42] RNA-seq cell type deconvolution approaches were also used to compare the abundance of inflammatory cells, but no significant difference was observed after adjusting for multiple comparisons (Extended Data Fig. 4 and Supplementary Fig. 6).

In summary, our analysis of tumor WES and RNA-seq data from participants who did and did not develop CAC revealed distinct genomic and transcriptomic profiles, including upregulated inflammatory signaling and increased expression or amplification of metalloproteinases.

Plasma proteome at cancer diagnosis and relapse

Given the potential role of circulating mediators in driving the CAC phenotype, plasma samples, where available, collected from TRACERx participants at diagnosis (141 participants: 55 in the CAC and 70 in the non-CAC group, 16 without imaging or body weight, but included for analysis with clinical demographic data) and first relapse (115 participants: 43 in the CAC and 58 in the non-CAC group, 13 without imaging or body weight, but included for analysis with clinical demographic data; Fig. 1) were subjected to unbiased proteomic profiling using the Olink Explore 3072 platform to investigate the presence of proteins associated with the development of CAC[43]. Paired diagnosis and relapse plasma samples were available for 110 participants. Differential protein expression, that is, increased or decreased plasma protein expression, in participants in the CAC group compared with the non-CAC group, at diagnosis and at relapse, was analyzed.

In the CAC group, 114 plasma proteins were found to have increased expression at the time of diagnosis compared with the non-CAC group, although these did not remain significant after adjusting for multiplicity (Benjamini–Hochberg correction; Fig. 5a). Similarly, 443 proteins were found to have increased expression at the time of relapse, of which 12 proteins, including GDF15, HSPA2, KIAA0319, TNFRSF10B and IL1RL1, remained significant after Benjamini–Hochberg correction (Fig. 5a). Concordant increased primary tumor gene expression and plasma protein expression was only identified for *HSPA2* and *SV2A* in the CAC group (Supplementary Table 10).

When considering 24 circulating proteins previously reported to play a role in CAC[44] (Supplementary Table 11), in the CAC group 14/24 proteins, including tumor necrosis factor (TNF), GDF15 and CCL3, were found to have increased plasma protein expression at the time of diagnosis compared with the non-CAC group, but the difference was not significant after Benjamini–Hochberg correction (Fig. 5b). At the time of relapse, GDF15, a known mediator of anorexia and weight loss[45], demonstrated significantly increased expression in the CAC group compared with the non-CAC group (Fig. 5b and Extended Data Fig. 5). Furthermore, a significant correlation was observed between normalized protein expression of GDF15 and loss of SAT, VAT, SKM and body weight at the time of relapse (Extended Data Fig. 6).

Among the 141 TRACERx participants analyzed in the Olink study, in a subset of this cohort the relationship between increased GDF15 and development of the CAC phenotype at relapse had previously also been

https://doi.org/10.1038/s41591-023-02232-8

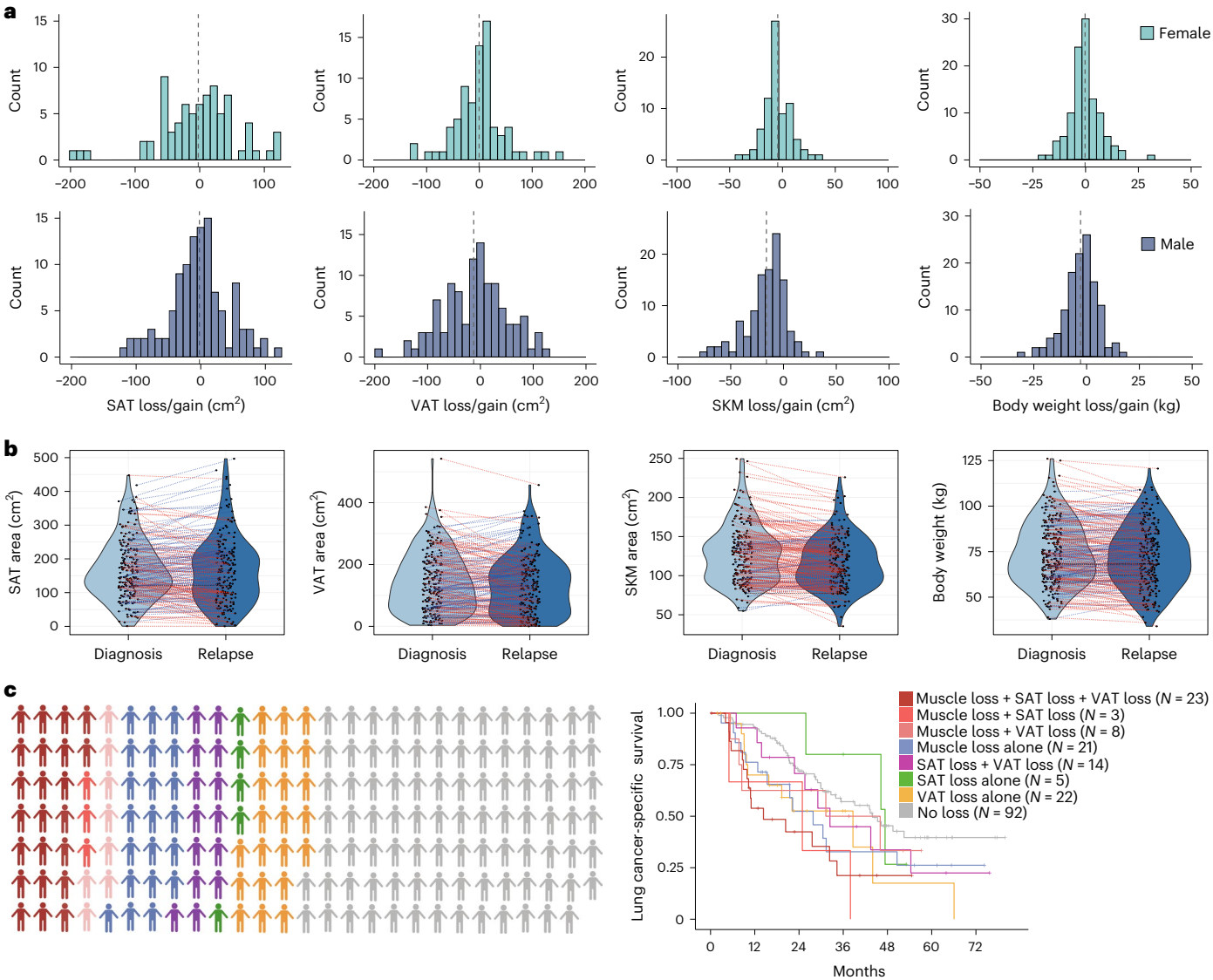


Fig. 3 | Survival outcomes according to changes in body composition between primary diagnosis and first relapse. a, Distribution of losses and gains of body composition ($n = 188$, 77 female, 111 male) and body weight ($n = 232$, 109 female, 123 male) in the TRACERx cohort between diagnosis and relapse according to sex; vertical lines indicate the mean. **b**, Dynamics of body composition ($N = 188$) and body weight ($n = 232$) between diagnosis and relapse, with red lines indicating loss, blue lines indicating gain, and gray lines indicating no changes. **c**, Subgroups of participants according to (co-)presence of SAT loss (defined as ≥20% between baseline and relapse), VAT loss (≥20%) and SKM loss (≥10%). Kaplan–Meier analysis of LCSS according to labeled subgroups.

identified using an electrochemiluminescent immunoassay (Methods), including 107 plasma samples collected at diagnosis and 79 plasma samples collected at relapse (Fig. 1). The median circulating GDF15 level in the whole cohort was 1,902 pg ml^{-1} (normal range 200–1,200 pg ml^{-1}; ref. [45]) at diagnosis, and further increased at relapse (median 2,393.5 pg ml^{-1}; Fig. 5c). Notably, serum GDF15 levels in the TRACERx cohort (median age 70 years) were higher than previously published in age-matched non-cancer volunteers (60–70 years, median plasma GDF15 levels of 866 pg ml^{-1}) and in keeping with previous reports of increased circulating GDF15 in participants with NSCLC[46–49]. Median GDF15 levels in participants who developed CAC between diagnosis and relapse were significantly higher compared with participants who did not develop CAC, both at the time of diagnosis (2,156 pg ml^{-1} versus 1,631 pg ml^{-1}; $P = 0.03$) and relapse (3,222 pg ml^{-1} versus 1,878 pg ml^{-1}, $P < 0.001$; Extended Data Fig. 7). Circulating GDF15 levels at diagnosis and at relapse were significantly associated with increased age (Extended Data Fig. 8). There was no significant association between GDF15 levels at diagnosis or at relapse and BMI at diagnosis, smoking status or number of pack years, use of adjuvant treatment and tumor stage or volume (Extended Data Fig. 8). GDF15 levels at diagnosis and relapse were not associated with time to recurrence in a Cox regression analysis (HR 1.0, 95% CI 1.0–1.0, $P = 0.269$ and $P = 0.350$, respectively).

Circulating GDF15 levels, at diagnosis and relapse, were significantly higher in participants who developed grade 4 BMI-adjusted weight loss at relapse compared to those who had stable or increased weight (grade 0; Fig. 5d,e). Increased circulating GDF15 was significantly associated with loss of body weight, as well as loss of SAT, VAT and SKM tissue (Fig. 5f–i and Supplementary Fig. 7). WES and RNA-seq data were analyzed to investigate whether increased circulating GDF15 was associated with genomic alterations and/or increased gene expression in the primary tumor. Mutations in the *GDF15* gene were found in only three participants (c.C564T, c.T312G, c.A313G). There was no significant correlation between primary tumor GDF15 gene expression and circulating GDF15 levels (Fig. 5j,k), although the power to detect any correlation was limited by the small number of relapse samples in this cohort. Furthermore, the ploidy-adjusted copy number of the *GDF15* gene on

https://doi.org/10.1038/s41591-023-02232-8

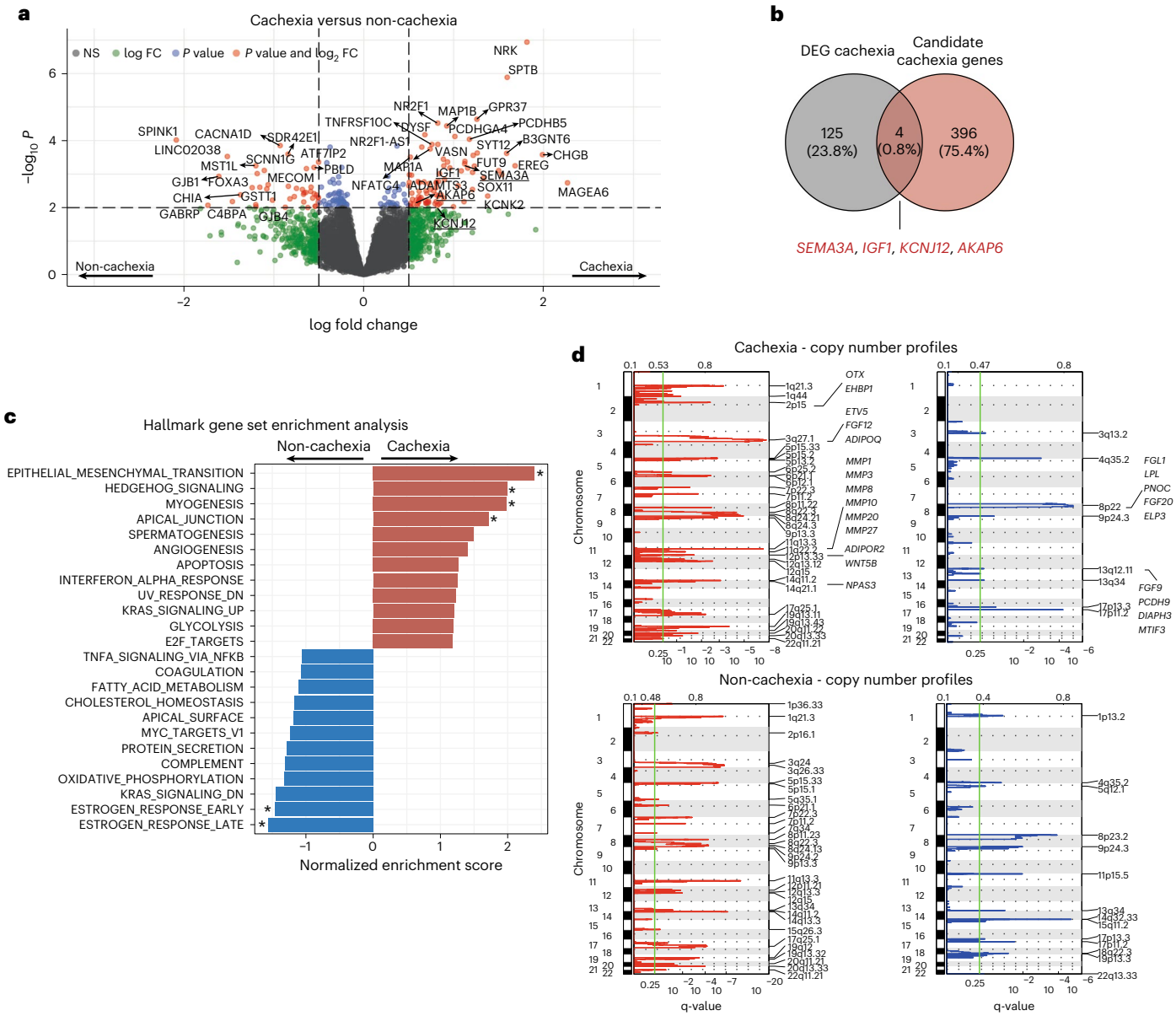


Fig. 4 | Tumor genomic and transcriptomic profiles according to cancer cachexia and non-cachexia groups. a, Differential gene expression between participants in the cachexia versus non-cachexia groups, adjusted for number of tumor regions, sex and histology. *P* values from moderated two-sided *t*-test without adjusting for multiple testing. Genes from candidate cachexia gene list are underlined. **b**, Overlap of differentially expressed genes between the cachexia group and candidate cachexia genes. **c**, Hallmark gene-set enrichment in the cachexia (red) versus non-cachexia (blue) groups, adjusted for sex and histology. The 20 gene sets with the highest enrichment scores are plotted and asterisks indicate adjusted *P* values < 0.05. Enrichment was defined as false discovery rate < 0.05 from a Fisher's exact test per gene after multiple-testing correction. **d**, GISTIC2.0 analysis of copy number alterations of cachexia (upper row) and non-cachexia (lower row) groups. *y* axes indicate chromosomal positions (1–22), red plots indicate gains, and blue plots indicate losses. *x* axes indicate *q* values. Most significant peaks are indicated on the right; regions with a false discovery rate $q \leq 0.25$ (vertical green line) are considered significant. GISTIC G-scores are plotted on the top. DEG, differentially expressed gene; GSEA, gene-set enrichment analysis; NES, normalized enrichment score.

chromosome 19p13.11q did not correlate with baseline gene expression ($R = -0.072$; Extended Data Fig. 9a) and copy number events were not associated with circulating GDF15 levels (Extended Data Fig. 9b). Participants with GDF15 copy number gain, defined as amplifications or gains of 19p13.11q, in relapsed tumor tissue, tended to have higher circulating GDF15 levels compared to participants with tumors carrying GDF15 copy number losses (4630 pg/ml versus 2671 pg/ml, $P = 0.14$; Extended Data Fig. 9c).

To explore whether the relationship between increased circulating GDF15 and the development of CAC at relapse was restricted to early-stage NSCLC, in an independent cohort of 164 participants with stage IV metastatic lung adenocarcinoma (ARCHER trial, NCT01360554), plasma GDF15 levels were measured using a GDF15 ELISA (Methods)[50]. Circulating GDF15 levels were significantly higher in participants with grade 3 or 4 BMI-adjusted weight loss over approximately 6 months after study enrollment compared to participants with grade 0 BMI-adjusted weight loss, that is, stable body weight (Extended Data Fig. 10), suggesting a similar relationship between circulating GDF15 and weight loss in the advanced disease setting. Body composition could not be assessed in this cohort since imaging was not available.

Overall, these data suggest that participants who develop features of CAC at first relapse have distinct tumor genomic and transcriptomic, as well as plasma proteomic, profiles. Circulating GDF15 demonstrated

the strongest correlation with loss of adipose tissue, SKM and body weight, emphasizing the potential for GDF15-targeted therapy in the management of CAC[51].

Discussion

The first aim of this study was to generate body composition profiles of participants with early-stage lung cancer and to elucidate their association with survival outcomes. At the time of cancer diagnosis, low SAT, VAT and SKM were each associated with poor LCSS and OS in both the TRACERx and BLCS cohorts, demonstrating the prognostic relevance and clinical utility of body composition in identifying participants with high-risk disease[52,53].

Two distinct features of CAC are loss of body weight and SKM[6]. The second aim of this study was to examine how body composition and body weight change in the context of lung cancer recurrence. We observed that, between diagnosis and first relapse, 24% of participants experienced a loss across all three body compartments, but subgroups of participants with pronounced or isolated loss of SAT, VAT or SKM were also identified, suggesting the possibility that distinct CAC clinical phenotypes may exist whereby individual, or a combination of, body composition compartments may be preferentially affected[54].

Using thresholds for a measure of loss in each individual body composition and body weight that were associated with poor LCSS and OS, we identified participants who developed CAC in the interval between diagnosis and relapse, and used this classification to address the third aim of this study, which was to explore potential molecular mechanisms and mediators driving the CAC phenotype. The genomic and transcriptomic profiles of primary tumors in participants with and without features of CAC at the time of first relapse were found to be distinct. Participants who developed CAC at relapse had increased differential expression of semaphorin-3A (*SEMA3A*), insulin-like growth factor 1 (*IGF1*), the potassium channel *KCNJ12* and A-kinase anchor protein 6 (*AKAP6*)[36]. Furthermore, multiple inflammatory pathways were upregulated and there was increased expression of several genes of interest, including lipopolysaccharide binding protein (*LBP*) and metallopeptidases, such as the ADAM metallopeptidase encoded by *ADAMTS3* (refs. [55,56]). This is in keeping with previous observations that have suggested tumor-driven systemic inflammation is a contributor to CAC[57,58]. While preclinical studies have reported increased tumor expression of interleukin-17 in the murine Lewis lung carcinoma model and parathyroid hormone-related protein[59,60] as purported mediators of cachexia, we did not see this in the TRACERx cohort.

Analysis of somatic copy number alterations in primary tumors revealed distinct copy number profiles in participants who developed CAC, including amplification of a cluster of matrix metalloproteinases on chromosome 11q22.2, including *MMP1* and *MMP3*, which have previously been functionally associated with tumor-induced muscle loss in *Drosophila* models[29], as well as amplifications involving *ADIPOQ* (adiponectin) on chromosome 3q27.1. High levels of adiponectin, a regulator of insulin sensitivity and lipid metabolism, have previously been associated with low body weight and weight loss, while low levels have been associated with obesity[39]. Overall, our transcriptomic and genomic analyses suggest specific tumor-derived factors may play a role in the development of CAC.

Finally, plasma proteomic analysis at diagnosis and relapse demonstrated significant differential plasma protein expression between participants who did and did not develop CAC, including proteins such as TNF receptor TNFRSF10B and IL1RL1 (interleukin 1 receptor-like 1). Among participants who developed CAC, the most differentially, highly expressed, candidate plasma protein at the time of relapse was GDF15; a highly conserved member of the transforming growth factor-β superfamily, which circulates at physiologically low levels in healthy states. GDF15 expression and secretion is upregulated in response to cellular stress[61,62], and elevated circulating levels have been identified in a broad range of human diseases, including cardiac, renal and respiratory failure, and notably anorexia and weight loss[63–66]. There is mounting preclinical evidence that GDF15 may be a putative druggable target, with transgenic mice overexpressing GDF15 developing a cachexia-like syndrome that can be readily reversed with neutralizing GDF15 monoclonal antibodies[64,67,51].

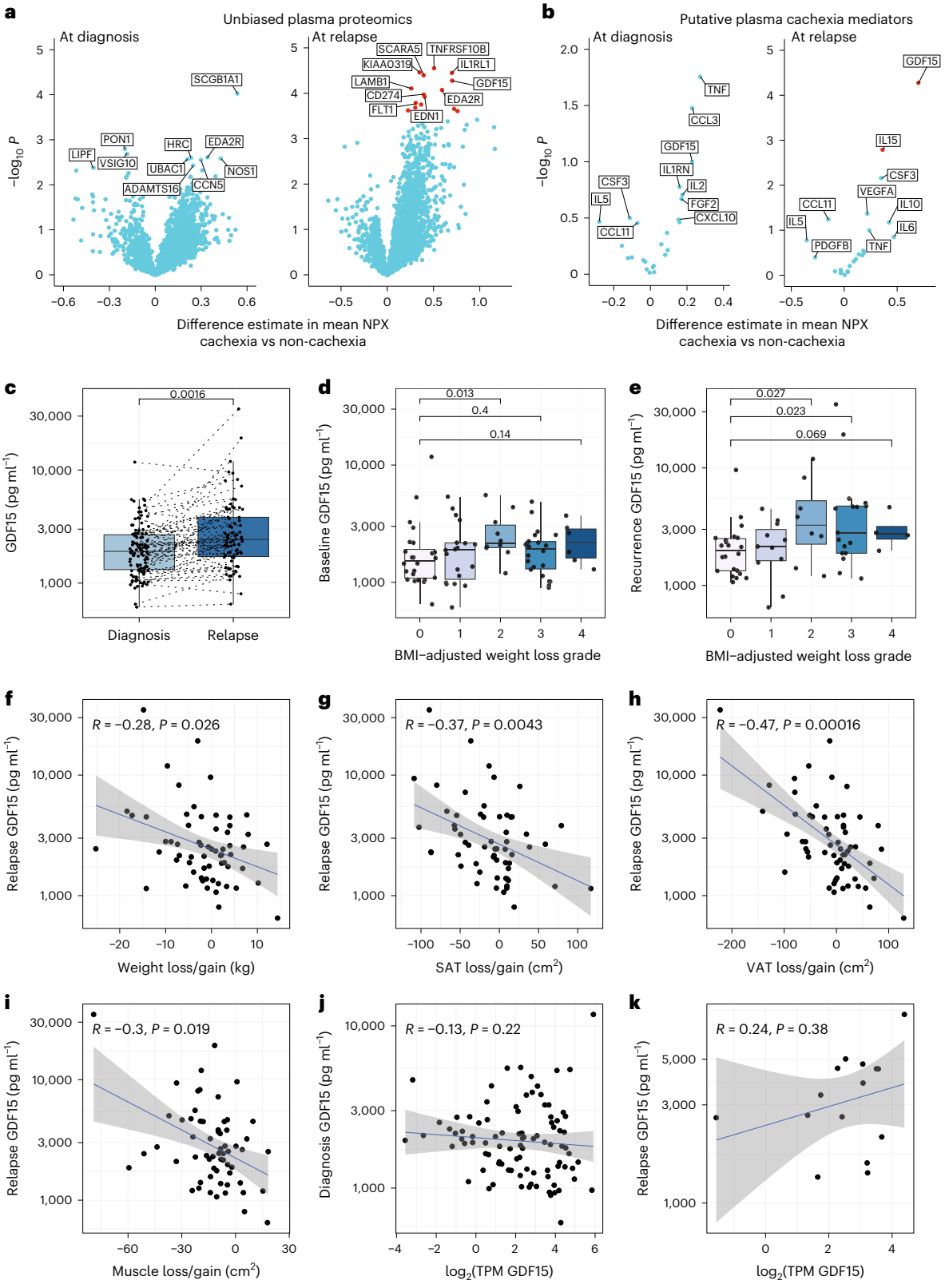
The same association between circulating GDF15 levels and body weight loss was observed in an independent cohort of participants with advanced metastatic NSCLC, suggesting that this association is agnostic of disease stage, and further corroborating the key role of GDF15 as a mediator of CAC in NSCLC. No clear evidence of increased tumor GDF15 expression was observed and plasma GDF15 levels were not clearly associated with genomic amplification of the *GDF15* locus. However, GDF15 is known to be produced in diverse tissue sites, including liver and kidney, which may act as alternative potential sources of pathological secretion[62,68]. Furthermore, GDF15 is subject to differential rates of production and clearance, mediated by other factors, such as hepatic stabilin-1 and stabilin-2 (ref. [69]).

While our study provides a NSCLC dataset integrating tumor genomics and plasma proteomics with body composition and body weight, there are limitations in the interpretation of the data. The thresholds for tissue and weight loss between diagnosis and relapse, including potential interaction between these variables, to identify individuals who develop the CAC phenotype requires validation in independent cohorts. The body composition analysis focuses on two time points in the disease course: diagnosis and first relapse. Conceivably, the cachexia phenotype may develop at subsequent time points of disease progression with increasing burden of disease, warranting further investigation. Our description of the CAC phenotype is based on changes that occur between diagnosis and first relapse, whereas the tumor genomic and transcriptomic data were mostly derived from the resected primary tumor given tissue availability. As such, our analyses are correlative and generate hypotheses; therefore, further studies to validate our findings, including functional experiments, are warranted. A selection bias cannot be excluded in the plasma proteomic analyses, since only participants with available complete baseline and relapse imaging data and sufficient amounts of banked plasma were analyzed. Future studies aiming to establish the underlying biological mechanisms of CAC would benefit from incorporating functional participant data, such as physical activity, food intake and muscle function, as well as quality-of-life measures, to reflect the complexity

Fig. 5 | Differential protein expression and associations between circulating GDF15, body composition and body weight changes, and cancer cachexia. **a**, Differential plasma proteome of participants in the cachexia versus non-cachexia groups, at diagnosis (left) and relapse (right). The ten proteins with highest differential expression are labeled; *P* values from two-sided *t*-test. Red dots indicate significant differential protein expression after Benjamini–Hochberg correction. **b**, Differential plasma protein expression of putative cachexia mediators in the cachexia versus non-cachexia group, at diagnosis (left) and relapse (right). The ten proteins with highest differential expression are labeled; *P* values from two-sided *t*-test. Red dots indicate significant differential protein expression after Benjamini–Hochberg correction. **c**, Plasma GDF15 levels in participants at diagnosis (baseline, $n = 107$) or first recurrence ($n = 79$) of NSCLC in the TRACERx cohort (two-sided Wilcoxon test). **d**,**e**, Baseline ($n = 80$; **d**) and recurrence ($n = 61$; **e**) GDF15 levels according to weight change category in the TRACERx cohort (two-sided Wilcoxon test). **f**–**i** Spearman's correlation of recurrence GDF15 levels and loss/gain of body weight ($n = 62$; **f**), SAT (**g**), VAT (**h**) and muscle ($n = 61$; **i**). **j**,**k**, Spearman's correlation of diagnosis (**j**) and recurrence (**k**) GDF15 levels and GDF15 gene expression as transcripts per million (TPM). All Wilcoxon tests are two-sided. Box plots represent lower quartile, median and upper quartile, and whiskers extend to a maximum of 1.5 times the interquartile range beyond the box. Points indicate individual data points. Gray error bands represent the 95% CI of the fitted linear model. *y* axes in **d**–**i** represent $\log_{10}$ scales.

https://doi.org/10.1038/s41591-023-02232-8

of CAC and to capture its related constitutional symptoms. Overall, this study demonstrates the significant, independent and potentially prognostic impact of altered body composition on clinical outcomes (LCSS and OS) in NSCLC. The presence of specific body composition changes, either predominant loss of adipose or muscle tissue, in subgroups of participants suggests distinct clinical subtypes of CAC, which may be driven by unique biological mechanisms warranting further investigation. We show that automated pipeline technologies unlock the potential to leverage CT imaging embedded in medical oncological practice to identify individuals at risk of developing CAC, simultaneously providing the scientific means to study potential drivers and mechanisms of cachexia pathophysiology. Amidst the plethora of proposed pro-cachectic mediators, GDF15 emerges as a differentially expressed, and clinically measurable, protein, with a mounting evidence base, establishing its potential to translate to a biotherapeutic target.



Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-023-02232-8.

References

1. Calle, E. E., Rodriguez, C., Walker-Thurmond, K. & Thun, M. J. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of US adults. *N. Engl. J. Med.* **348**, 1625–1638 (2003).
2. Guo, Y. et al. Body mass index and mortality in chronic obstructive pulmonary disease: a dose-response meta-analysis. *Medicine* **95**, e4225 (2016).
3. Lavie, C. J., McAuley, P. A., Church, T. S., Milani, R. V. & Blair, S. N. Obesity and cardiovascular diseases: implications regarding fitness, fatness, and severity in the obesity paradox. *J. Am. Coll. Cardiol.* **63**, 1345–1354 (2014).
4. Martin, L. et al. Cancer cachexia in the age of obesity: skeletal muscle depletion is a powerful prognostic factor, independent of body mass index. *J. Clin. Oncol.* **31**, 1539–1547 (2013).
5. Baracos, V. E., Martin, L., Korc, M., Guttridge, D. C. & Fearon, K. C. H. Cancer-associated cachexia. *Nat. Rev. Dis. Primers* **4**, 17105 (2018).
6. Fearon, K. et al. Definition and classification of cancer cachexia: an international consensus. *Lancet Oncol.* **12**, 489–495 (2011).
7. Brown, J. C. et al. The deterioration of muscle mass and radiodensity is prognostic of poor survival in stage I–III colorectal cancer: a population-based cohort study (C-SCANS). *J. Cachexia Sarcopenia Muscle* **9**, 664–672 (2018).
8. Caan, B. J. et al. Association of muscle and adiposity measured by computed tomography with survival in patients with nonmetastatic breast cancer. *JAMA Oncol.* **4**, 798–804 (2018).
9. Lee, J. S. et al. Subcutaneous fat distribution is a prognostic biomarker for men with castration resistant prostate cancer. *J. Urol.* **200**, 114–120 (2018).
10. Baracos, V. E., Reiman, T., Mourtzakis, M., Gioulbasanis, I. & Antoun, S. Body composition in patients with non-small cell lung cancer: a contemporary view of cancer cachexia with the use of computed tomography image analysis. *Am. J. Clin. Nutr.* **91**, 1133s–1137s (2010).
11. Yang, M., Shen, Y., Tan, L. & Li, W. Prognostic value of sarcopenia in lung cancer: a systematic review and meta-analysis. *Chest* **156**, 101–111 (2019).
12. Popinat, G. et al. Subcutaneous fat mass measured on multislice computed tomography of pretreatment PET/CT is a prognostic factor of stage IV non-small cell lung cancer treated by nivolumab. *Oncoimmunology* **8**, e1580128 (2019).
13. Tan, H. et al. Preoperative body composition combined with tumor metabolism analysis by PET/CT is associated with disease-free survival in patients with NSCLC. *Contrast Media Mol. Imaging* **2022**, 7429319 (2022).
14. Mourtzakis, M. et al. A practical and precise approach to quantification of body composition in cancer patients using computed tomography images acquired during routine care. *Appl. Physiol. Nutr. Metab.* **33**, 997–1006 (2008).
15. Shen, W. et al. Total body skeletal muscle and adipose tissue volumes: estimation from a single abdominal cross-sectional image. *J. Appl. Physiol.* **97**, 2333–2338 (2004).
16. Dabiri, S. et al. Deep learning method for localization and segmentation of abdominal CT. *Comput. Med. Imaging Graph* **85**, 101776 (2020).
17. Frankell, A. M. et al. The natural history of NSCLC and impact of subclonal selection in TRACERx. *Nature* (In press) (2023).
18. Jamal-Hanjani, M. et al. Tracking the evolution of non-small-cell lung cancer. *N. Engl. J. Med.* **376**, 2109–2121 (2017).
19. Christiani, D. B. Lung Cancer Study. https://sites.sph.harvard.edu/blcs/ (2022).
20. Anyene, I. et al. Body composition from single versus multi-slice abdominal computed tomography: concordance and associations with colorectal cancer survival. *J. Cachexia Sarcopenia Muscle* **13**, 2974–2984 (2022).
21. Speliotes, E. K. et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.* **42**, 937–948 (2010).
22. Locke, A. E. et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518**, 197–206 (2015).
23. Wen, W. et al. Meta-analysis of genome-wide association studies in East Asian-ancestry populations identifies four new loci for body mass index. *Hum. Mol. Genet* **23**, 5492–5504 (2014).
24. Akiyama, M. et al. Genome-wide association study identifies 112 new loci for body mass index in the Japanese population. *Nat. Genet.* **49**, 1458–1467 (2017).
25. Winkler, T. W. et al. The influence of age and sex on genetic associations with adult body size and shape: a large-scale genome-wide interaction study. *PLoS Genet* **11**, e1005378 (2015).
26. Johns, N. et al. New genetic signatures associated with cancer cachexia as defined by low skeletal muscle index and weight loss. *J. Cachexia Sarcopenia Muscle* **8**, 122–130 (2017).
27. Solheim, T. S. et al. Is there a genetic cause for cancer cachexia? A clinical validation study in 1,797 patients. *Br. J. Cancer* **105**, 1244–1251 (2011).
28. Baranski, T. J. et al. A high-throughput, functional screen of human body mass index GWAS loci using tissue-specific RNAi *Drosophila melanogaster* crosses. *PLOS Genet.* **14**, e1007222 (2018).
29. Lodge, W. et al. Tumor-derived MMPs regulate cachexia in a *Drosophila* cancer model. *Dev. Cell* **56**, 2664–2680 (2021).
30. Ding, G. et al. Coordination of tumor growth and host wasting by tumor-derived Upd3. *Cell Rep.* **36**, 109553 (2021).
31. Kwon, Y. et al. Systemic organ wasting induced by localized expression of the secreted insulin/IGF antagonist ImpL2. *Dev. Cell* **33**, 36–46 (2015).
32. Song, W. et al. Tumor-derived ligands trigger tumor growth and host wasting via differential MEK activation. *Dev. Cell* **48**, 277–286 (2019).
33. Newton, H. et al. Systemic muscle wasting and coordinated tumour response drive tumourigenesis. *Nat. Commun.* **11**, 4653 (2020).
34. Figueroa-Clarevega, A. & Bilder, D. Malignant *Drosophila* tumors interrupt insulin signaling to induce cachexia-like wasting. *Dev. Cell* **33**, 47–55 (2015).
35. Kim, J. et al. Tumor-induced disruption of the blood–brain barrier promotes host death. *Dev. Cell* **56**, 2712–2721 (2021).
36. van der Klaauw, A. A. et al. Human semaphorin 3 variants link melanocortin circuit development and energy balance. *Cell* **176**, 729–742 (2019).
37. Trobec, K., von Haehling, S., Anker, S. D. & Lainscak, M. Growth hormone, insulin-like growth factor 1, and insulin signaling—a pharmacological target in body wasting and cachexia. *J. Cachexia Sarcopenia Muscle* **2**, 191–200 (2011).
38. Mermel, C.H., Schumacher, S.E., Hill, B. et al. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* **12**, R41 (2011).
39. Achari, A. E. & Jain, S. K. Adiponectin, a therapeutic target for obesity, diabetes and endothelial dysfunction. *Int J. Mol. Sci.* **18**, 1321 (2017).
40. Bentham, R. et al. Using DNA sequencing data to quantify T cell fraction and therapy response. *Nature* **597**, 555–560 (2021).

41. Danaher, P. et al. Gene expression markers of tumor infiltrating leukocytes. *J. Immunother. Cancer* **5**, 18 (2017).
42. Newman, A. M. et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* **37**, 773–782 (2019).
43. Assarsson, E. et al. Homogenous 96-Plex PEA immunoassay exhibiting high sensitivity, specificity and excellent scalability. *PLoS ONE* **9**, e95192 (2014).
44. Lerner, L. et al. MAP3K11/GDF15 axis is a critical driver of cancer cachexia. *J. Cachexia Sarcopenia Muscle* **7**, 467–482 (2016).
45. Lockhart, S. M., Saudek, V. & O'Rahilly, S. GDF15: a hormone conveying somatic distress to the brain. *Endocr. Rev.* **41**, bnaa007 (2020).
46. Wollert, K. C. et al. An automated assay for growth differentiation factor 15. *J. Appl. Lab. Med.* **1**, 510–521 (2017).
47. Cai, D. et al. Extensive serum biomarker analysis in patients with non-small-cell lung carcinoma. *Cytokine* **126**, 154868 (2020).
48. Roche. Elecsys GDF-15 datasheet. https://diagnostics.roche.com/global/en/products/params/elecsys-gdf-15.html (2020).
49. Welsh, P. et al. Reference ranges for GDF-15, and risk factors associated with GDF-15, in a large general population cohort. *Clin. Chem. Lab. Med.* **60**, 1820–1829 (2022).
50. Ramalingam, S. S. et al. Dacomitinib versus erlotinib in patients with advanced-stage, previously treated non-small-cell lung cancer (ARCHER 1009): a randomised, double-blind, phase 3 trial. *Lancet Oncol.* **15**, 1369–1378 (2014).
51. Kim-Muller, J. Y. et al. GDF15 neutralization restores muscle function and physical performance in a mouse model of cancer cachexia. *Cell Rep.* **42**, 111947 (2023).
52. He, X. et al. Age- and sex-related differences in body composition in healthy subjects aged 18 to 82 years. *Medicine* **97**, e11152 (2018).
53. Bamia, C., Trichopoulou, A., Lenas, D. & Trichopoulos, D. Tobacco smoking in relation to body fat mass and distribution in a general population sample. *Int. J. Obes.* **28**, 1091–1096 (2004).
54. Sartori, R., Romanello, V. & Sandri, M. Mechanisms of muscle atrophy and hypertrophy: implications in health and disease. *Nat. Commun.* **12**, 330 (2021).
55. Bindels, L. B. et al. Increased gut permeability in cancer cachexia: mechanisms and clinical relevance. *Oncotarget* **9**, 18224–18238 (2018).
56. Cal, S. & López-Otín, C. ADAMTS proteases and cancer. *Matrix Biol.* **44-46**, 77–85 (2015).
57. de Matos-Neto, E. M. et al. Systemic inflammation in cachexia—is tumor cytokine expression profile the culprit? *Front. Immunol.* **6**, 629 (2015).
58. Webster, J. M., Kempen, L. J. A. P., Hardy, R. S. & Langen, R. C. J. Inflammation and skeletal muscle wasting during cachexia. *Front. Physiol.* **11**, 597675 (2020).
59. Ying, L. et al. IL-17A contributes to skeletal muscle atrophy in lung cancer-induced cachexia via JAK2/STAT3 pathway. *Am. J. Physiol.* **322**, C814–C824 (2022).
60. Kir, S. et al. Tumour-derived PTH-related protein triggers adipose tissue browning and cancer cachexia. *Nature* **513**, 100–104 (2014).
61. Rochette, L., Zeller, M., Cottin, Y. & Vergely, C. Insights into mechanisms of GDF15 and receptor GFRAL: therapeutic targets. *Trends Endocrinol. Metab.* **31**, 939–951 (2020).
62. Mulderrig, L. et al. Aldehyde-driven transcriptional stress triggers an anorexic DNA damage response. *Nature* **600**, 158–163 (2021).
63. Tsai, V. W., Lin, S., Brown, D. A., Salis, A. & Breit, S. N. Anorexia-cachexia and obesity treatment may be two sides of the same coin: role of the TGF-β superfamily cytokine MIC-1/GDF15. *Int J. Obes.* **40**, 193–197 (2016).
64. Johnen, H. et al. Tumor-induced anorexia and weight loss are mediated by the TGF-β superfamily cytokine MIC-1. *Nat. Med.* **13**, 1333–1340 (2007).
65. Lerner, L. et al. Plasma growth differentiation factor 15 is associated with weight loss and mortality in cancer patients. *J. Cachexia Sarcopenia Muscle* **6**, 317–324 (2015).
66. Tsai, V. W. et al. Treatment with the TGF-β superfamily cytokine MIC-1/GDF15 reduces the adiposity and corrects the metabolic dysfunction of mice with diet-induced obesity. *Int J. Obes.* **42**, 561–571 (2018).
67. Emmerson, P. J. et al. The metabolic effects of GDF15 are mediated by the orphan receptor GFRAL. *Nat. Med.* **23**, 1215–1219 (2017).
68. Patel, S. et al. Endogenous GDF15 and FGF21 additively alleviate hepatic steatosis and insulin resistance in obese mice. Preprint at *bioRxiv* https://doi.org/10.1101/2022.06.08.495255 (2022).
69. Schledzewski, K. et al. Deficiency of liver sinusoidal scavenger receptors stabilin-1 and -2 in mice causes glomerulofibrotic nephropathy via impaired hepatic clearance of noxious blood factors. *J. Clin. Invest* **121**, 703–714 (2011).

Othman Al-Sawaf[1,2,3,111], Jakob Weiss[4,5,6,111], Marcin Skrzypski[7,111], Jie Min Lam[1,2,8], Takahiro Karasaki[1,2,3], Francisco Zambrana[9], Andrew C. Kidd[10], Alexander M. Frankell[1,3], Thomas B. K. Watkins[1,3], Carlos Martínez-Ruiz[1,11], Clare Puttick[1,3,11], James R. M. Black[1,11], Ariana Huebner[1,3,11], Maise Al Bakir[1,3], Mateo Sokač[12,13,14], Susie Collins[15], Selvaraju Veeriah[1], Neil Magno[1], Cristina Naceur-Lombardelli[1], Paulina Prymas[1], Antonia Toncheva[1], Sophia Ward[1,3,16], Nick Jayanth[17], Roberto Salgado[18,19], Christopher P. Bridge[20], David C. Christiani[21], Raymond H. Mak[4,5], Camden Bay[22], Michael Rosenthal[22], Naveed Sattar[23], Paul Welsh[23], Ying Liu[24,25], Norbert Perrimon[24,25], Karteek Popuri[26], Mirza Faisal Beg[27], Nicholas McGranahan[1,11], Allan Hackshaw[17], Danna M. Breen[28], Stephen O'Rahilly[29], Nicolai J. Birkbak[12,13,14], Hugo J. W. L. Aerts[4,5,30], TRACERx Consortium*, Mariam Jamal-Hanjani[1,2,8,112] ✉ & Charles Swanton[1,3,8,112] ✉

[1]Cancer Research UK Lung Cancer Centre of Excellence, University College London Cancer Institute, London, UK. [2]Cancer Metastasis Laboratory, University College London Cancer Institute, London, UK. [3]Cancer Evolution and Genome Instability Laboratory, The Francis Crick Institute, London, UK. [4]Artificial Intelligence in Medicine (AIM) Program, Mass General Brigham, Harvard Medical School, Boston, MA, USA. [5]Department of Radiation

Oncology, Brigham and Women's Hospital, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA. [6]Department of Diagnostic and Interventional Radiology, University Freiburg, Freiburg, Germany. [7]Department of Oncology and Radiotherapy, Medical University of Gdańsk, Gdańsk, Poland. [8]Department of Oncology, University College London Hospitals, London, UK. [9]Infanta Sofía University Hospital, Madrid, Spain. [10]Institute of Infection, Immunity & Inflammation, University of Glasgow, Glasgow, UK. [11]Cancer Genome Evolution Research Group, Cancer Research UK Lung Cancer Centre of Excellence, University College London Cancer Institute, London, UK. [12]Department of Clinical Medicine, Aarhus University, Aarhus, Denmark. [13]Department of Molecular Medicine, Aarhus University Hospital, Aarhus, Denmark. [14]Bioinformatics Research Centre, Aarhus University, Aarhus, Denmark. [15]Early Clinical Development, Pfizer UK Ltd, Cambridge, UK. [16]Advanced Sequencing Facility, The Francis Crick Institute, London, UK. [17]Cancer Research UK & UCL Cancer Trials Centre, London, UK. [18]Department of Pathology, ZAS Hospitals, Antwerp, Belgium. [19]Division of Research, Peter MacCallum Cancer Centre, Melbourne, Australia. [20]MGH & BWH Center for Clinical Data Science, Boston, MA, USA. [21]Department of Medicine, Massachusetts General Hospital/Harvard Medicine School, and Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [22]Department of Radiology, Brigham and Women's Hospital and Dana-Farber Cancer Institute, and Harvard Medical School, Boston, MA, USA. [23]School of Cardiovascular and Metabolic Health, University of Glasgow, Glasgow, UK. [24]Department of Genetics, Harvard Medical School, Boston, USA. [25]Howard Hughes Medical Institute, Harvard Medical School, Boston, USA. [26]Department of Computer Science, Memorial University of Newfoundland, St. John's, Newfoundland and Labrador, Burnaby, Canada. [27]School of Engineering Science, Simon Fraser University, Burnaby, British Colombia, Canada. [28]Internal Medicine Research Unit, Pfizer, Cambridge, MA, USA. [29]Wellcome Trust-MRC Institute of Metabolic Science and NIHR Cambridge Biomedical Research Centre, University of Cambridge, Cambridge, UK. [30]Radiology and Nuclear Medicine, CARIM & GROW, Maastricht University, Maastricht, The Netherlands. [111]These authors contributed equally: Othman Al-Sawaf, Jakob Weiss, Marcin Skrzypski. [112]These authors jointly supervised this work: Nicolai J. Birkbak, Hugo J.W.L. Aerts, Mariam Jamal-Hanjani, Charles Swanton. *A list of authors and their affiliations appears at the end of the paper. ✉e-mail: m.jamal-hanjani@ucl.ac.uk; Charles.swanton@crick.ac.uk

TRACERx Consortium

Charles Swanton[1,3,8,112], Mariam Jamal-Hanjani[1,2,8,112], Nicholas McGranahan[1,11], Othman Al-Sawaf[1,2,3,111], Jie Min Lam[1,2,8], Takahiro Karasaki[1,2,3], Alexander M. Frankell[1,3], Thomas B. K. Watkins[3], Carlos Martínez-Ruiz[1,11], Clare Puttick[1,3,11], James R. M. Black[1,11], Ariana Huebner[1,3,11], Maise Al Bakir[1,3], Selvaraju Veeriah[1], Cristina Naceur-Lombardelli[1], Paulina Prymas[1], Antonia Toncheva[1], Sophia Ward[1,3,16], Roberto Salgado[18,19], Allan Hackshaw[17], Nicolai J. Birkbak[1,3,12,13,14], Hugo Jwl Aerts[4,5,30], Jason F. Lester[31], Amrita Bajaj[32], Apostolos Nakas[32], Azmina Sodha-Ramdeen[32], Keng Ang[32], Mohamad Tufail[32], Mohammed Fiyaz Chowdhry[32], Molly Scotland[32], Rebecca Boyles[32], Sridhar Rathinam[32], Claire Wilson[33], Domenic Marrone[33], Sean Dulloo[33], Dean A. Fennell[32,33], Gurdeep Matharu[34], Jacqui A. Shaw[34], Joan Riley[34], Lindsay Primrose[34], Ekaterini Boleti[35], Heather Cheyne[36], Mohammed Khalil[36], Shirley Richardson[36], Tracey Cruickshank[36], Gillian Price[37,38], Keith M. Kerr[38,39], Sarah Benafif[8], Kayleigh Gilbert[40], Babu Naidu[41], Akshay J. Patel[42], Aya Osman[42], Christer Lacson[42], Gerald Langman[42], Helen Shackleford[42], Madava Djearaman[42], Salma Kadiri[42], Gary Middleton[42,43], Angela Leek[44], Jack Davies Hodgkinson[44], Nicola Totten[44], Angeles Montero[45], Elaine Smith[45], Eustace Fontaine[45], Felice Granato[45], Helen Doran[45], Juliette Novasio[45], Kendadai Rammohan[45], Leena Joseph[45], Paul Bishop[45], Rajesh Shah[45], Stuart Moss[45], Vijay Joshi[45], Philip Crosbie[45,46,47], Fabio Gomes[48], Kate Brown[48], Mathew Carter[48], Anshuman Chaturvedi[47,48], Lynsey Priest[47,48], Pedro Oliveira[47,48], Colin R. Lindsay[49], Fiona H. Blackhall[49], Matthew G. Krebs[49], Yvonne Summers[49], Alexandra Clipson[47,50], Jonathan Tugwood[47,50], Alastair Kerr[47,50], Dominic G. Rothwell[47,50], Elaine Kilgour[47,50], Caroline Dive[47,50], Roland F. Schwarz[51,52], Tom L. Kaufmann[52,53], Gareth A. Wilson[3], Rachel Rosenthal[3], Peter Van Loo[54,55,56], Zoltan Szallasi[57,58,59], Judit Kisistok[12,13,14], Mateo Sokac[12,13,14], Miklos Diossy[57,58,60], Jonas Demeulemeester[56,61,62], Abigail Bunkum[1,2,63], Aengus Stewart[64], Alastair Magness[64], Andrew Rowan[3], Angeliki Karamani[65], Benny Chain[65], Brittany B. Campbell[3], Carla Castignani[56,66], Chris Bailey[3], Christopher Abbosh[1], Clare E. Weeden[64], Claudia Lee[3], Corentin Richard[1], Crispin T. Hiley[1,3], David A. Moore[1,3,67], David R. Pearce[65], Despoina Karagianni[65], Dhruva Biswas[1,3,68], Dina Levi[64], Elena Hoxha[65], Elizabeth Larose Cadieux[56,66], Emilia L. Lim[1,3], Emma Colliver[3], Emma Nye[69], Eva Grönroos[64], Felip Gálvez-Cancino[65], Foteini Athanasopoulou[1,3,16], Francisco Gimeno-Valiente[1], George Kassiotis[70,71], Georgia Stavrou[65], Gerasimos Mastrokalos[65], Haoran Zhai[1,3], Helen L. Lowe[65], Ignacio Garcia Matos[65], Jacki Goldman[64], James L. Reading[65], Javier Herrero[68], Jayant K. Rane[3,65], Jerome Nicod[16], John A. Hartley[65], Karl S. Peggs[72,73], Katey S. S. Enfield[3], Kayalvizhi Selvaraju[65], Kerstin Thol[1,11], Kevin Litchfield[1,74], Kevin W. Ng[70], Kezhong Chen[65], Krijn Dijkstra[75,76], Kristiana Grigoriadis[1,3,11], Krupa Thakkar[1], Leah Ensell[65], Mansi Shah[65], Marcos Vasquez Duran[65], Maria Litovchenko[65], Mariana Werner Sunderland[1], Mark S. Hill[3], Michelle Dietzen[1,3,11], Michelle Leung[1,3,11], Mickael Escudero[64], Mihaela Angelova[3], Miljana Tanić[66,77], Monica Sivakumar[1], Nnennaya Kanu[1], Olga Chervova[65], Olivia Lucas[1,3,8,63], Oriol Pich[3], Philip Hobson[64], Piotr Pawlik[65], Richard Kevin Stone[69], Robert Bentham[1,11], Robert E. Hynds[65], Roberto Vendramin[64], Sadegh Saghafinia[1], Saioa López[65], Samuel Gamble[65], Seng Kuong Anakin Ung[65], Sergio A. Quezada[1,78], Sharon Vanloo[1], Simone Zaccaria[1,63], Sonya Hessey[1,2,63], Stefan Boeing[64], Stephan Beck[66], Supreet Kaur Bola[65], Tamara Denner[64], Teresa Marafioti[67], Thanos P. Mourikis[65], Victoria Spanswick[65], Vittorio Barbè[64], Wei-Ting Lu[64], William Hill[64], Wing Kin Liu[1,2], Yin Wu[65], Yutaka Naito[64], Zoe Ramsden[64], Catarina Veiga[79], Gary Royle[80], Charles-Antoine Collins-Fekete[81], Francesco Fraioli[82], Paul Ashford[83], Tristan Clark[84], Martin D. Forster[1,8], Siow Ming Lee[1,8], Elaine Borg[67], Mary Falzon[67], Dionysis Papadatos-Pastos[8], James Wilson[8], Tanya Ahmad[8], Alexander James Procter[85], Asia Ahmed[85], Magali N. Taylor[85], Arjun Nair[85,86], David Lawrence[87], Davide Patrini[87], Neal Navani[88,89], Ricky M. Thakrar[88,89], Sam M. Janes[88], Emilie Martinoni Hoogenboom[90], Fleur Monk[90], James W. Holding[90], Junaid Choudhary[90], Kunal Bhakhri[90], Marco Scarci[90], Martin Hayward[90], Nikolaos Panagiotopoulos[90], Pat Gorman[90], Reena Khiroya[67], Robert CM. Stephens[90], Yien Ning Sophia Wong[90], Steve Bandula[90], Abigail Sharp[17], Sean Smith[17], Nicole Gower[17], Harjot Kaur Dhanda[17], Kitty Chan[17], Camilla Pilotti[17], Rachel Leslie[17], Anca Grapa[91], Hanyun Zhang[91], Khalid AbdulJabbar[91], Xiaoxi Pan[91], Yinyin Yuan[92], David Chuter[93], Mairead MacKenzie[93], Serena Chee[94], Aiman Alzetani[94], Judith Cave[95], Lydia Scarlett[94], Jennifer Richards[94], Papawadee Ingram[94], Silvia Austin[94], Eric Lim[96,97], Paulo De Sousa[97], Simon Jordan[97], Alexandra Rice[97], Hilgardt Raubenheimer[97], Harshil Bhayani[97], Lyn Ambrose[97], Anand Devaraj[97], Hema Chavan[97], Sofina Begum[97], Silviu I. Buderi[97], Daniel Kaniu[97], Mpho Malima[97], Sarah Booth[97], Andrew G. Nicholson[98,99], Nadia Fernandes[97], Pratibha Shah[97], Chiara Proli[97], Madeleine Hewish[100,101], Sarah Danson[102], Michael J. Shackcloth[103], Lily Robinson[104],

Peter Russell[104], Kevin G. Blyth[105,106,107], Craig Dick[108], John Le Quesne[105,106,109], Alan Kirk[110], Mo Asif[110], Rocco Bilancia[110], Nikos Kostoulas[110] & Mathew Thomas[110]

[31]Singleton Hospital, Swansea Bay University Health Board, Swansea, UK. [32]University Hospitals of Leicester NHS Trust, Leicester, UK. [33]University of Leicester, Leicester, UK. [34]Cancer Research Centre, University of Leicester, Leicester, UK. [35]Royal Free Hospital, Royal Free London NHS Foundation Trust, London, UK. [36]Aberdeen Royal Infirmary NHS Grampian, Aberdeen, UK. [37]Department of Medical Oncology, Aberdeen Royal Infirmary NHS Grampian, Aberdeen, UK. [38]University of Aberdeen, Aberdeen, UK. [39]Department of Pathology, Aberdeen Royal Infirmary NHS Grampian, Aberdeen, UK. [40]The Whittington Hospital NHS Trust, London, UK. [41]Birmingham Acute Care Research Group, Institute of Inflammation and Ageing, University of Birmingham, Birmingham, UK. [42]University Hospital Birmingham NHS Foundation Trust, Birmingham, UK. [43]Institute of Immunology and Immunotherapy, University of Birmingham, Birmingham, UK. [44]Manchester Cancer Research Centre Biobank, Manchester, UK. [45]Wythenshawe Hospital, Manchester University NHS Foundation Trust, Wythenshawe, UK. [46]Division of Infection, Immunity and Respiratory Medicine, University of Manchester, Manchester, UK. [47]Cancer Research UK Lung Cancer Centre of Excellence, University of Manchester, Manchester, UK. [48]The Christie NHS Foundation Trust, Manchester, UK. [49]Division of Cancer Sciences, The University of Manchester and The Christie NHS Foundation Trust, Manchester, UK. [50]Cancer Research UK Manchester Institute Cancer Biomarker Centre, University of Manchester, Manchester, UK. [51]Institute for Computational Cancer Biology, Center for Integrated Oncology (CIO), Cancer Research Center Cologne Essen (CCCE), Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany. [52]Berlin Institute for the Foundations of Learning and Data (BIFOLD), Berlin, Germany. [53]Berlin Institute for Medical Systems Biology, Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC), Berlin, Germany. [54]Department of Genetics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. [55]Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. [56]Cancer Genomics Laboratory, The Francis Crick Institute, London, UK. [57]Danish Cancer Society Research Center, Copenhagen, Denmark. [58]Computational Health Informatics Program, Boston Children's Hospital, Boston, MA, USA. [59]Department of Bioinformatics, Semmelweis University, Budapest, Hungary. [60]Department of Physics of Complex Systems, ELTE Eötvös Loránd University, Budapest, Hungary. [61]Integrative Cancer Genomics Laboratory, Department of Oncology, KU Leuven, Leuven, Belgium. [62]VIB – KU Leuven Center for Cancer Biology, Leuven, Belgium. [63]Computational Cancer Genomics Research Group, University College London Cancer Institute, London, UK. [64]The Francis Crick Institute, London, UK. [65]University College London Cancer Institute, London, UK. [66]Medical Genomics, University College London Cancer Institute, London, UK. [67]Department of Cellular Pathology, University College London Hospitals, London, UK. [68]Bill Lyons Informatics Centre, University College London Cancer Institute, London, UK. [69]Experimental Histopathology, The Francis Crick Institute, London, UK. [70]Retroviral Immunology Group, The Francis Crick Institute, London, UK. [71]Department of Infectious Disease, Faculty of Medicine, Imperial College London, London, UK. [72]Department of Haematology, University College London Hospitals, London, UK. [73]Cancer Immunology Unit, Research Department of Haematology, University College London Cancer Institute, London, UK. [74]Tumour Immunogenomics and Immunosurveillance Laboratory, University College London Cancer Institute, London, UK. [75]Department of Molecular Oncology and Immunology, the Netherlands Cancer Institute, Amsterdam, Netherlands. [76]Oncode Institute, Utrecht, Netherlands. [77]Experimental Oncology, Institute for Oncology and Radiology of Serbia, Belgrade, Serbia. [78]Immune Regulation and Tumour Immunotherapy Group, Cancer Immunology Unit, Research Department of Haematology, University College London Cancer Institute, London, UK. [79]Centre for Medical Image Computing, Department of Medical Physics and Biomedical Engineering, University College London, London, UK. [80]Department of Medical Physics and Bioengineering, University College London Cancer Institute, London, UK. [81]Department of Medical Physics and Biomedical Engineering, University College London, London, UK. [82]Institute of Nuclear Medicine, Division of Medicine, University College London, London, UK. [83]Institute of Structural and Molecular Biology, University College London, London, UK. [84]University College London, London, UK. [85]Department of Radiology, University College London Hospitals, London, UK. [86]UCL Respiratory, Department of Medicine, University College London, London, UK. [87]Department of Thoracic Surgery, University College London Hospital NHS Trust, London, UK. [88]Lungs for Living Research Centre, UCL Respiratory, University College London, London, UK. [89]Department of Thoracic Medicine, University College London Hospitals, London, UK. [90]University College London Hospitals, London, UK. [91]The Institute of Cancer Research, London, UK. [92]The University of Texas MD Anderson Cancer Center, Houston, TX, USA. [93]Independent Cancer Patients' Voice, London, UK. [94]University Hospital Southampton NHS Foundation Trust, Southampton, UK. [95]Department of Oncology, University Hospital Southampton NHS Foundation Trust, Southampton, UK. [96]Academic Division of Thoracic Surgery, Imperial College London, London, UK. [97]Royal Brompton and Harefield Hospitals, Guy's and St Thomas' NHS Foundation Trust, London, UK. [98]Department of Histopathology, Royal Brompton and Harefield Hospitals, Guy's and St Thomas' NHS Foundation Trust, London, UK. [99]National Heart and Lung Institute, Imperial College London, London, UK. [100]Royal Surrey Hospital, Royal Surrey Hospitals NHS Foundation Trust, Guilford, UK. [101]University of Surrey, Guilford, UK. [102]Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK. [103]Liverpool Heart and Chest Hospital, Liverpool, UK. [104]Princess Alexandra Hospital, The Princess Alexandra Hospital NHS Trust, Harlow, UK. [105]School of Cancer Sciences, University of Glasgow, Glasgow, UK. [106]Cancer Research UK Beatson Institute, Glasgow, UK. [107]Queen Elizabeth University Hospital, Glasgow, UK. [108]NHS Greater Glasgow and Clyde, Glasgow, UK. [109]Pathology Department, Queen Elizabeth University Hospital, NHS Greater Glasgow and Clyde, Glasgow, UK. [110]Golden Jubilee National Hospital, Clydebank, UK.

Methods

Participant cohorts

TRACERx cohort. TRACERx is a UK-wide prospective multicenter study of participants with primary NSCLC that aims to define evolutionary trajectories for lung cancer through multiregional and longitudinal tumor sampling[17](NCT01888601). The study was approved by an independent Research Ethics Committee (13/LO/1546). All participants provided written informed consent. Participants are followed for up to 5 years from the point of primary diagnosis, through surgical resection to cure, cancer progression(s) and death. The study collects longitudinal clinical, epidemiological and imaging data as well as multiregional tumor tissue samples. The study protocol with inclusion/exclusion criteria has been published previously[18].

For the body composition and cachexia study, participants were included if a preoperative abdominal CT incorporating the third lumbar vertebra and performed within 3 months of primary surgical tumor resection, was available. For the delta cohort, all participants with available abdominal CT at the time of disease recurrence were included. Overall, 651 participants with a pre-surgery baseline CT were included, of which 188 had a relapse with corresponding CT scan.

BCLS cohort. Participants from the US cancer cohort in this analysis are part of the ongoing BLCS, a multi-institutional epidemiology cohort study at Massachusetts General Brigham (MGB) and the Dana-Farber Cancer Institute (DFCI). Inclusion criteria for this analysis were pathology-confirmed diagnosis of lung cancer with available abdominopelvic CT or positron emission tomography/CT scans within 4 months before and the 2 months after diagnosis, and relevant, a priori-defined clinical covariates available. Participants were excluded if imaging or clinical data were incomplete or missing. The study was approved by the institutional review boards of all institutions and the requirement for written informed consent was waived. Information regarding smoking status was prospectively collected in the TRACERx study and was also available for all participants in the BLCS cohort. Participant demographics are provided in Table 1.

ARCHER1009 cohort. ARCHER1009 was a randomized phase 3 study for participants with advanced NSCLC who were randomized to the epidermal growth factor receptor inhibitors dacomitinib and erlotinib[50]. This study received ethical approval from the Pfizer Institutional Review Board and was conducted in accordance with the Declaration of Helsinki (NCT01360554). For the cachexia study, 164 participants with available plasma and body weight data at two time points within 6 months were included. Classification of weight loss was similar to published schemes[6,70] and defined as weight stable/gain, >0–5% weight loss or >5% weight loss. Any prior treatment (chemotherapy, radiation or surgery) must have been completed at least 2 weeks before randomization at the start of the study.

Body composition and body weight measurements

For the TRACERx cohort, SAT, VAT and SKM were quantified as areas (cm^2) localized to the third lumbar vertebra (L3) level. CT scans were generated according to standard local protocols at the participating sites (Supplementary Table 12). L3 selection, image segmentation and tissue area quantification were conducted via an automated deep learning-based pipeline (Data Analysis Facilitation Suite (DAFS), Voronoi Health Analytics[16,71–73]). To measure SAT, VAT and SKM areas at the L3 level, analyses were run with the 'avg-L3mid[3]' command, which measures across three slices above and below the midst of L3 to increase data accuracy (Supplementary Fig. 8). The following Hounsfield unit (HU) boundaries were used: For SAT, −190 to −30 HUs; for VAT, −150 to −50 HUs; and for SKM, −29 to 150 HUs. Accurate L3 selection and segmentation quality was manually inspected for all participants by an experienced investigator (medical oncologist with over 10 years of experience). To this end, each CT annotation and segmentation was inspected via a sagittal, coronal and axial view of each scan using the 'quickcheck' option; mis-annotations were corrected using the CAST (CT Annotation and Segmentation Tool) feature from DAFS. To confirm the highly reproducible nature of the algorithm, 60 CT scans were rerun twice through the automated annotation, segmentation and measurements steps. Using the precision metrics previously proposed by Arribas et al., ≤0.01% variance was observed between the runs (Supplementary Table 13)[74]. Absolute (cm^2) change between baseline and relapse was calculated for SAT, VAT and SKM for each participant.

For the independent BLCS cohort, baseline body composition was measured by a previously developed automated system[75,76], which produced body composition measurements at the L3 level for SAT, VAT and SKM in cm^2. Manual quality control was conducted for all images in the BLCS cohort by an experienced investigator (radiologist with over 10 years of experience).

Body weight data in the TRACERx, BCLS and ARCHER1009 cohorts were collected from medical records in routine health care settings; scale calibrations and type of clothing was conducted according to local standards. For the TRACERx cohort, body weight changes were assessed between baseline and relapse; for the ARCHER1009 cohort, body weight changes were assessed between time of enrollment and end of study. BMI-adjusted weight change was calculated as follows, similar to previous publications[77]:

	BMI (kg/m^2)				
Weight loss (%)	**≥28**	**25–27.9**	**22–24.9**	**20–21.9**	**<20**
±2.4	0	0	1	1	3
2.5–5.9	1	2	2	2	3
6–10.9	2	3	3	3	4
11–14.9	3	3	3	4	4
≥15	3	4	4	4	4

Whole-exome and RNA sequencing

WES was performed on DNA purified from tumor tissue with matched germline DNA from whole blood, as described previously[18,78]. RNA was extracted from primary tumor tissue, and downstream analyses were conducted as reported previously; where applicable, multiregional gene expression data were handled as the average per tumor or adjusted by a linear mixed-effects model, also accounting for histology and sex (Martinez-Ruiz et al, 2023). P value < 0.01 and absolute log fold change >1 was used to identify differentially expressed genes. GSEA was done by pre-ranking the genes by fold-change values and using the fgsea R package (v1.22.0) and using the hallmark gene set from the Molecular Signatures Database (v.7.4) with a minimal size of 15 and maximal size of 500. GISTIC2.0 was used to analyze copy number alterations in tumor tissues according to the cachexia and non-cachexia groups[38].

A list of candidate cachexia genes, that is, genes possibly associated with cachexia, was generated by reviewing supporting literature, including GWAS on obesity and cachexia, because both extremes can share common metabolic perturbations, as well as murine and *Drosophila* models (see Supplementary Table 9).

Plasma proteomics

Plasma proteomes were profiled using the Olink Explore 3072 platform at Bioxpedia following the standard Olink-certified protocol[43]. TRACERx plasma samples and control samples were plated on three 96-well plates that were processed in one batch. For data analysis, protein expression as $\log_2$ of normalized protein expression was used. Comparisons of protein expression between body composition groups were made by Welch two-sample t-tests with Benjamini–Hochberg correction. Data that did not pass the Olink-specified quality-control

metrics were excluded from the analysis. R packages OlinkAnalyze v3.1.0 and ggplot2 v3.3.6 were used for data analysis and visualization.

For the orthogonal GDF15 validation, GDF15 was measured in TRACERx plasma samples using the Elecsys GDF15 immunoassay (Roche) at baseline ahead of surgery and at diagnosis of relapse. Plasma samples from the ARCHER1009 samples were obtained at the end of study and human GDF15 was measured using ELISA (R&D Systems).

Statistical analyses

All statistical analyses were performed in R v4.1.1. GDF15 plasma levels were $\log_{10}$ transformed. Tests involving correlations were done using stat_cor from the R package ggpubr (v04.0) with linear regression and Spearman's rank-order correlation. Categorical comparisons were made by analysis of variance, student's *t*-test or Wilcoxon test. *P* values were two sided and considered as statistically significant if below 0.05. Data structuring, plotting and analysis in R was also done by ggplot2 (v3.3.6), cowplot (v1.1.1), table1 (v1.4.2), gg.gap (v1.3), data.table (v1.14.2), readxl (v1.4.0), xlsx (v0.6.5), ggsci (v2.9), ggsignif (v0.6.3), ggdendro (v0.1.23), scales (v1.2.0), survminer (v0.4.9), survival (v3.4.0), GenomicRanges (v1.46.1), ggvenn (v0.1.9), EnhancedVolcano (v1.10.0), fuzzyjoin (v0.1.6), tidyr (v1.2.0) and dplyr (v1.0.8). For LCSS, an event was death from lung cancer; and all other participants were censored at the date last seen alive or death from other causes (competing risk analyses, in which death from other causes were treated as a competing event, did not produce materially different associations). For OS, an event was death from any cause; and all other participants were censored at the date last seen alive.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The WES and RNA-seq data (from the TRACERx study) used during this study have been deposited at the European Genome–phenome Archive (EGA), which is hosted by The European Bioinformatics Institute (EBI) and the Centre for Genomic Regulation (CRG) under the accession codes EGAS00001006494 (WES) and EGAS00001006517 (RNA-seq); access is controlled by the TRACERx data access committee. Details on how to apply for access are available on the linked page.

The Olink dataset, de-identified body composition and body weight as well as clinical outcome data from TRACERx and BLCS are provided in a Zenodo repository (https://doi.org/10.5281/zenodo.7617516), together with plasma GDF15 data from ARCHER1009.

Code availability

Code to reproduce the figures is provided in a Zenodo repository (https://doi.org/10.5281/zenodo.7617516).

References

70. Dewys, W. D., Costa, G. & Henkin, R. I. Clinical parameters related to anorexia. *Cancer Treat. Rep.* **65**, 49–52 (1981).
71. Popuri, K., D. Cobzas, N. Esfandiari, et al. Body Composition Assessment in Axial CT Images Using FEM-Based Automatic Segmentation of Skeletal Muscle. *IEEE Trans. Med. Imaging* **35**, 512–520 (2016).
72. Ma, D., V. Chow, K. Popuri, et al. Comprehensive Validation of Automated Whole Body Skeletal Muscle, Adipose Tissue, and Bone Segmentation from 3D CT images for Body Composition Analysis: Towards Extended Body Composition. Preprint at https://arxiv.org/abs/2106.00652 (2021).
73. Cespedes Feliciano, E. M., Popuri, K., Cobzas, D. et al. Evaluation of automated computed tomography segmentation to assess body composition and mortality associations in cancer patients. *J. Cachexia Sarcopenia Muscle* **11**, 1258–1269 (2020).
74. Arribas, L., Sabaté-Llobera, A., Domingo, M.C. et al. Assessing dynamic change in muscle during treatment of patients with cancer: Precision testing standards. *Clin. Nutr.* **41**, 1059–1065 (2022).
75. Bridge, C. P., Rosenthal, M., Wright, B. et al. In *OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis*. (eds. Stoyanov, D. et al.) Fully-Automated Analysis of Body Composition from CT in Cancer Patients Using Convolutional Neural Networks (Springer International Publishing, 2018).
76. Magudia, K., Bridge, C. P., Bay, C. P. et al. Population-Scale CT-based Body Composition Analysis of a Large Outpatient Population Using Deep Learning to Derive Age-, Sex-, and Race-specific Reference Curves. *Radiology* **298**, 319–329 (2021).
77. Martin, L., Senesse, P., Gioulbasanis, I. et al. Diagnostic Criteria for the Classification of Cancer-Associated Weight Loss. *J. Clin. Oncol.* **33**, 90–99 (2014).
78. Rosenthal, R., Cadieux, E. L., Salgado, R. et al. Neoantigen-directed immune escape in lung cancer evolution. *Nature* **567**, 479–485 (2019).

Acknowledgements

The TRACERx study (NCT01888601) is sponsored by University College London (UCL/12/0279) and has been approved by an independent Research Ethics Committee (13/LO/1546). TRACERx is funded by Cancer Research UK (C11496/A17786) and coordinated through the Cancer Research UK and UCL Cancer Trials Centre, which has a core grant from CRUK (C444/A15953). We gratefully acknowledge the participants and relatives who participated in the TRACERx study. We thank all site personnel, investigators, funders and industry partners that supported the generation of the data within this study. In particular, we acknowledge the support of Scientific Computing, the Advanced Sequencing Facility and Experimental Histopathology departments at the Francis Crick Institute. This work was also supported by the Cancer Research UK Lung Cancer Centre of Excellence, the CRUK City of London Centre Award (C7893/A26233) and the UCL Experimental Cancer Medicine Centre. This work was supported by the Francis Crick Institute that receives its core funding from Cancer Research UK (CC2041), the UK Medical Research Council (CC2041) and the Wellcome Trust (CC2041). In particular, we acknowledge the support of Scientific Computing, the Advanced Sequencing Facility and Experimental Histopathology departments at the Francis Crick Institute. We also acknowledge the help of M. Angelova, R. Bentham, E. Colliver, E. Gronroos, J. Rane and R. Zaidi at the Francis Crick Institute and UCL Cancer Institute, in proofreading the manuscript, the team at Bioxpedia (Aarhus, Denmark) in processing the plasma samples for the Olink assay, and we acknowledge the CANCAN consortium. For the purpose of open access, the author has applied a CC BY public copyright license to any author accepted manuscript version arising from this submission. This work was also supported by the Cancer Research UK Lung Cancer Centre of Excellence and the CRUK City of London Centre Award (C7893/A26233) as well as the UCL Experimental Cancer Medicine Centre. The authors thank the study participants and investigators participating in the BLCS study, which has been approved by the Committees on the use of human participants in research at MGB and the Harvard School of Public Health. The authors further acknowledge financial support from the National Institutes of Health (NIH) (D.C.: NIH (NCI) 5U01CA209414; H.A.: NIH-USA U24CA194354, NIH-USA U01CA190234, NIH-USA U01CA209414 and NIH-USA R35CA22052, R.M.: 5U01CA209414, M.R.: U01CA210171, U10CA180821 and R01CA255184), the European Union - European Research Council (H.A.: 866504) and the German Research Foundation (J.W.: 6406/2-1). The ARCHER1009 study was funded by

Pfizer. Figures 2a and 3c were created with BioRender.com. O.A.S. is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 467697427. T.K. is supported by the JSPS Overseas Research Fellowships Program (202060447). M.J.-H. has received funding from CRUK, NIH National Cancer Institute, IASLC International Lung Cancer Foundation, Lung Cancer Research Foundation, Rosetrees Trust, UKI NETs, the National Institute for Health Research (NIHR) and NIHR UCLH Biomedical Research Centre. A.M.F. is supported by Stand Up To Cancer (SU2C-AACR-DT23-17). T.B.K.W. is supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001169), the UK Medical Research Council (FC001169) and the Wellcome Trust (FC001169) as well as the Marie Curie ITN Project PLOIDYNET (FP7-PEOPLE-2013, 607722), Breast Cancer Research Foundation (BCRF), Royal Society Research Professorships Enhancement Award (RP/EA/180007) and the Foulkes Foundation. C.M.-R. is supported by the Rosetrees Trust (M630). M.R. is supported by the Hale Family Research Center for Pancreatic Cancer at DFCI and the Lustgarten Foundation. N.M. is a Sir Henry Dale Fellow, jointly funded by the Wellcome Trust and the Royal Society (grant number 211179/Z/18/Z) and receives funding from Cancer Research UK, Rosetrees and the NIHR BRC at University College London Hospitals and the CRUK University College London Experimental Cancer Medicine Centre. S.O.R. is supported by the Medical Research Council MRC.MC.UU.12012.1, a Wellcome Senior Investigator Award 214274/Z/18/Z and the NIHR Cambridge Biomedical Research Centre. C.S. is a Royal Society Napier Research Professor (RSRP\R\210001). C.S. is supported by the Francis Crick Institute that receives its core funding from Cancer Research UK (CC2041), the UK Medical Research Council (CC2041), and the Wellcome Trust (CC2041). For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. C.S. is funded by Cancer Research UK (TRACERx (C11496/A17786), PEACE (C416/A21999) and CRUK Cancer Immunotherapy Catalyst Network); Cancer Research UK Lung Cancer Centre of Excellence (C11496/A30025); the Rosetrees Trust, Butterfield and Stoneygate Trusts; NovoNordisk Foundation (ID16584); Royal Society Professorship Enhancement Award (RP/EA/180007); National Institute for Health Research (NIHR) University College London Hospitals Biomedical Research Centre; the Cancer Research UK-University College London Centre; Experimental Cancer Medicine Centre; the Breast Cancer Research Foundation (US); and The Mark Foundation for Cancer Research Aspire Award (Grant 21-029-ASP). This work was supported by a Stand Up To Cancer-LUNGevity-American Lung Association Lung Cancer Interception Dream Team Translational Research Grant (Grant Number: SU2C-AACR-DT23-17 to S.M. Dubinett and A.E. Spira). Stand Up To Cancer is a division of the Entertainment Industry Foundation. Research grants are administered by the American Association for Cancer Research, the Scientific Partner of SU2C. CS is in receipt of an ERC Advanced Grant (PROTEUS) from the European Research Council under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 835297).

Author contributions

O.A.S. generated and analyzed the body composition and survival data from TRACERx, designed and conducted the genomic, transcriptomic and proteomic bioinformatic analyses and wrote the manuscript. J.W. generated and analyzed the body composition and survival data from BLCS and wrote the manuscript. M. Skrzypski generated body composition data in TRACERx. J.M.L. and T.K. assisted with data analyses and manuscript writing. F.Z. and A.C.K. assisted with data analysis. A.M.F., T.B.K.W., C.M.R., C.P., J.R.M.B., A.H., M.A.B. and M. Sokac assisted with bioinformatic analyses. S.C. and D.M.B. contributed and analyzed the GDF15 data from ARCHER1009. S.V., N.M., C.N.L., P.P., A.T. and S.W. assisted with sample collection and sample processing. N.J. assisted with statistical survival analyses. R.S. gave feedback on the analyses. C.P.B., D.C., R.M., C.B. and M.R. generated and contributed the BLCS data. N.S. and P.W. oversaw the measurements of GDF15 in TRACERx. Y.L. and N.P. contributed to the candidate cachexia gene list. K.P. and M.F.B. developed the DAFS body composition software and assisted with its implementation in TRACERx. N.M. assisted with the bioinformatic analyses. A.H. helped oversee the running of the TRACERx study and the survival analyses in the cachexia sub-study. S.O.R. guided the GDF15 data analyses and interpretation. N.J.B. assisted with the bioinformatic analyses and wrote the manuscript. H.A. supervised the analyses of the BLCS dataset. M.J.H. and C.S. jointly designed and supervised the study and helped write the manuscript. All authors reviewed and revised the manuscript.

Competing interests

O.A.S.: Advisory Board (AstraZeneca, AbbVie, Ascentage, Gilead, Janssen and Roche), speaker honoraria (Adaptive, AstraZeneca, AbbVie, BeiGene, Eli Lilly, Gilead, Janssen and Roche), research funding (BeiGene, AbbVie, Janssen and Roche). J.M.L. receives research funding from the NIHR. A.M.F. is a named inventor on a patent application to determine methods and systems for tumor monitoring (PCT/EP2022/077987). M.A.B. has consulted for Achilles Therapeutics. S.C.: employment (Pfizer). M.R. receives research funding from the NIH, Lustgarten Foundation, Stand Up to Cancer Foundation and the Hale Family Center for Pancreatic Cancer at DFCI. N.S.: consultancy and/or speaker honoraria for Abbott Laboratories, Afimmune, Amgen, AstraZeneca, Boehringer Ingelheim, Eli Lilly, Hanmi Pharmaceuticals, Janssen, Merck Sharp & Dohme, Novartis, Novo Nordisk, Pfizer, Roche Diagnostics and Sanofi; research grants for AstraZeneca, Boehringer Ingelheim, Roche Diagnostics and Novartis. P.W.: grant income from Roche Diagnostics, AstraZeneca, Boehringer Ingelheim and Novartis and speaker's fees from Novo Nordisk outside the submitted work. K.P. and M.F.B. actively direct Voronoi Health Analytics Incorporated, a Canadian corporation that sells commercial licenses for the DAFS software. R.H.M.: Advisory Board (ViewRay, AstraZeneca), consulting (Varian Medical Systems, Sio Capital Management), honorarium (Novartis, Springer Nature), ownership (Health-AI) and research funding (ViewRay). N.M. has stock options in and has consulted for Achilles Therapeutics and holds a European patents relating to targeting neoantigens (PCT/EP2016/ 059401), identifying patient response to immune checkpoint blockade (PCT/ EP2016/071471), determining HLA LOH (PCT/GB2018/052004), predicting survival rates of patients with cancer (PCT/GB2020/050221). D.M.B.: employment (Pfizer). S.O.R.: consultancy (Pfizer, AstraZeneca, Northsea and Third Rock Ventures). N.J.B. is a co-inventor of a patent to identify responders to cancer treatment (PCT/GB2018/051912), and a patent to predict HRD deficiency (US10190160B2). H.J.W.J.A. is scientific advisor and shareholder of Onc.AI, Love Health, Health-AI and Sphera, and receives consulting fees from BMS and Editas Medicine (all outside the presented work). M.J.-H. is a CRUK Career Establishment Awardee and has received funding from CRUK, IASLC International Lung Cancer Foundation, Lung Cancer Research Foundation, Rosetrees Trust, UKI NETs, NIHR and NIHR UCLH Biomedical Research Centre. M.J.-H. has consulted and is a member of the Scientific Advisory Board and Steering Committee, for Achilles Therapeutics, has received speaker honoraria from Astex Pharmaceuticals, Oslo Cancer Cluster and Pfizer, and is co-inventor on a patent PCT/US2017/028013 relating to methods for lung cancer detection. C.S. acknowledges grant support from AstraZeneca, Boehringer-Ingelheim, Bristol Myers Squibb, Pfizer, Roche-Ventana, Invitae (previously Archer Dx Inc - collaboration in minimal residual disease sequencing technologies), and Ono

Pharmaceutical. C.S. is an AstraZeneca Advisory Board member and Chief Investigator for the AZ MeRmaiD 1 and 2 clinical trials and is also Co-Chief Investigator of the NHS Galleri trial funded by GRAIL and a paid member of GRAIL's Scientific Advisory Board. He receives consultant fees from Achilles Therapeutics (also SAB member), Bicycle Therapeutics (also a SAB member), Genentech, Medicxi, Roche Innovation Centre – Shanghai, Metabomed (until July 2022), and the Sarah Cannon Research Institute. C.S. had stock options in Apogen Biotechnologies and GRAIL until June 2021, and currently has stock options in Epic Bioscience, Bicycle Therapeutics, and has stock options and is co-founder of Achilles Therapeutics. C.S. is an inventor on a European patent application relating to assay technology to detect tumour recurrence (PCT/GB2017/053289), the patent has been licensed to commercial entities and under his terms of employment with C.S due a revenue share of any revenue generated from such license(s). C.S. holds patents relating to targeting neoantigens (PCT/EP2016/059401), identifying patient response to immune checkpoint blockade (PCT/EP2016/071471), determining HLA LOH (PCT/GB2018/052004), predicting survival rates of patients with cancer (PCT/GB2020/050221), identifying patients who respond to cancer treatment (PCT/GB2018/051912), US patent relating to detecting tumour mutations (PCT/US2017/28013), methods for lung cancer detection (US20190106751A1) and both a European and US patent related to identifying insertion/deletion mutation targets (PCT/GB2018/051892) and is co-inventor to a patent application to determine methods and systems for tumour monitoring (PCT/EP2022/077987). C.S. is a named inventor on a provisional patent protection related to a ctDNA detection algorithm.

Additional information

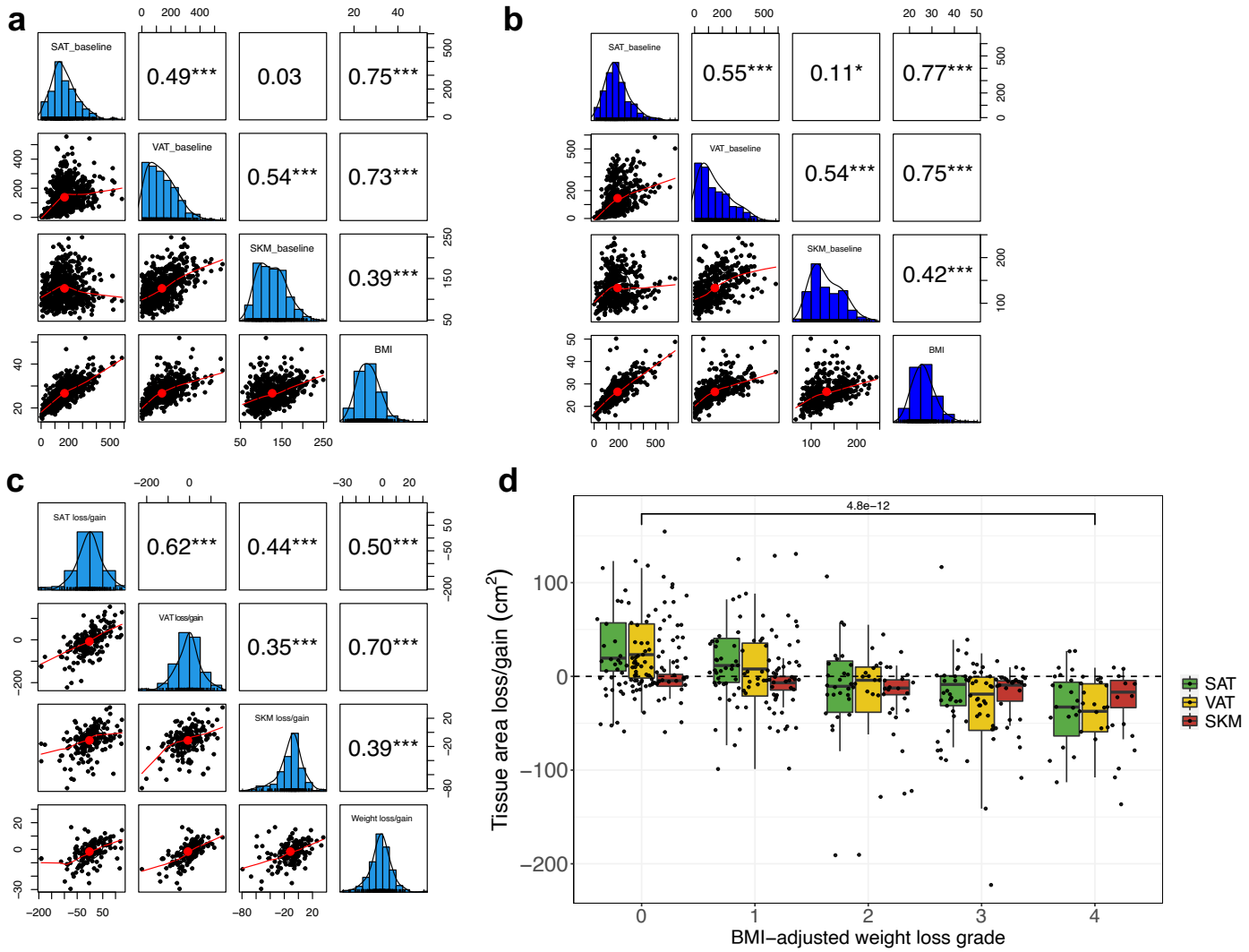
Extended data is available for this paper at https://doi.org/10.1038/s41591-023-02232-8.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41591-023-02232-8.

Correspondence and requests for materials should be addressed to Mariam Jamal-Hanjani or Charles Swanton.

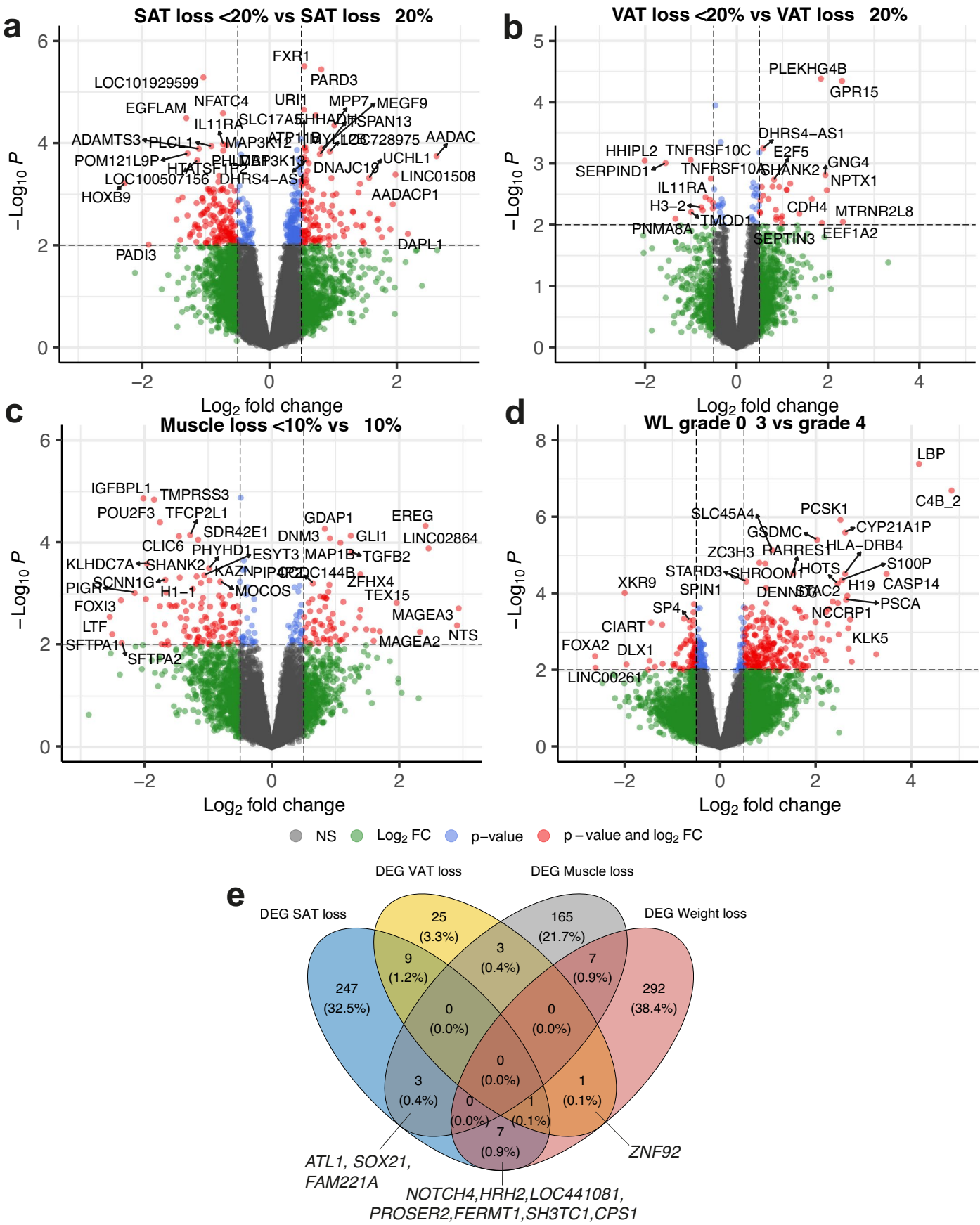
Peer review information *Nature Medicine* thanks Andrew Hoy and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Joao Monteiro, in collaboration with the *Nature Medicine* team.

https://doi.org/10.1038/s41591-023-02232-8



Extended Data Fig. 1 | Correlation between body composition metrics and body weight. a Spearman's correlation in the TRACERx cohort between subcutaneous adipose tissue (SAT), visceral adipose tissue (VAT), skeletal muscle (SKM) and body mass index (BMI) at primary diagnosis. **b** Spearman's correlation in the BLCS cohort between SAT, VAT, SKM and BMI at primary diagnosis. **c** Spearman's correlation between loss/gain of SAT, VAT, SKM and body weight between primary diagnosis and first relapse. **d** Losses and gains in cm^2 of SAT (green), VAT (yellow) and SKM (red) according to BMI-adjusted weight loss grade 0 to 4 (n=146). Bracket indicates p-value from two-sided Wilcoxon test; box plots represent lower quartile, median and upper quartile, whiskers extend to a maximum of 1.5 × IQR beyond the box. Points indicate individual data points. * indicates p-value <0.05, *** indicates p-value <0.001.

https://doi.org/10.1038/s41591-023-02232-8

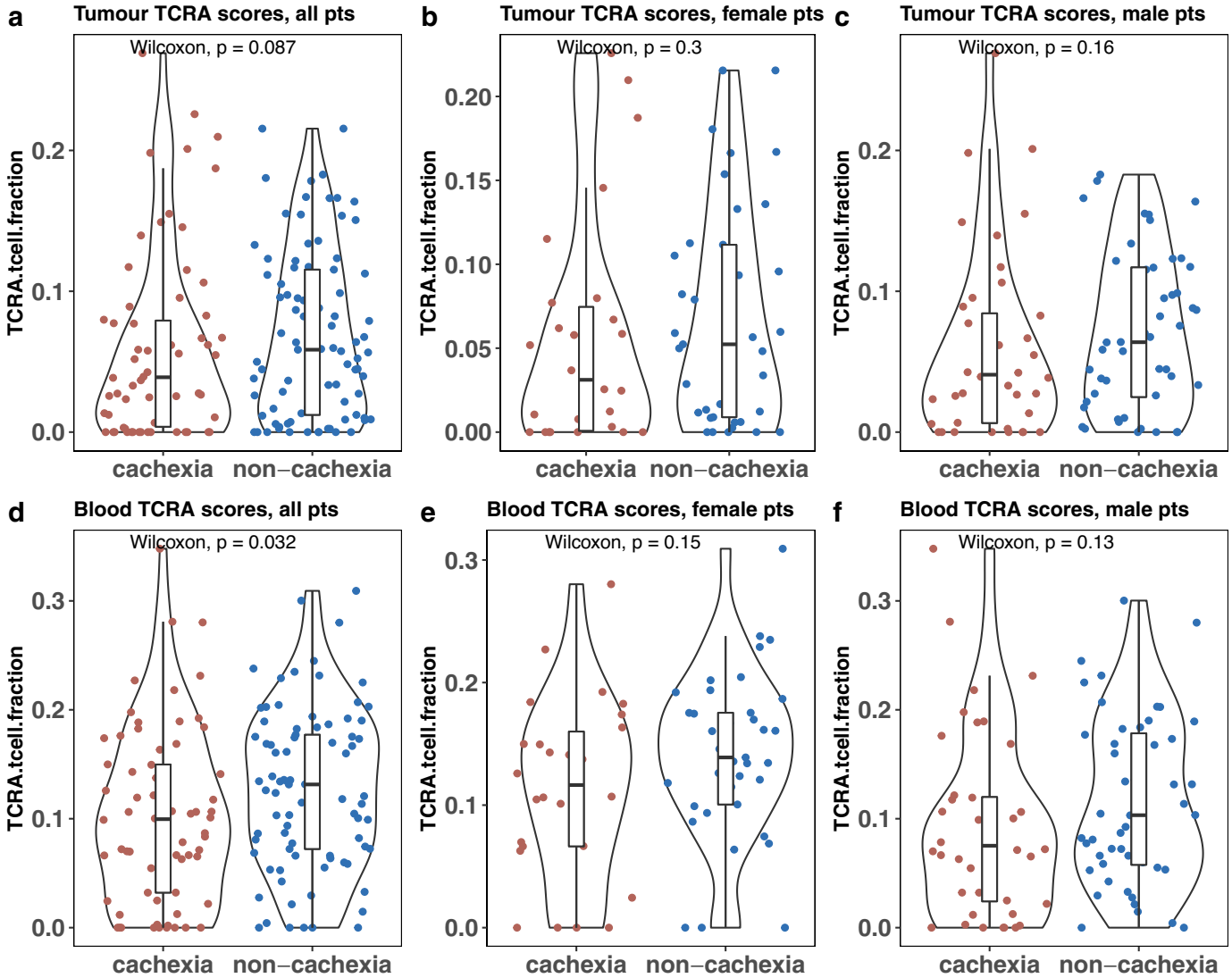


Extended Data Fig. 2 | See next page for caption.

https://doi.org/10.1038/s41591-023-02232-8

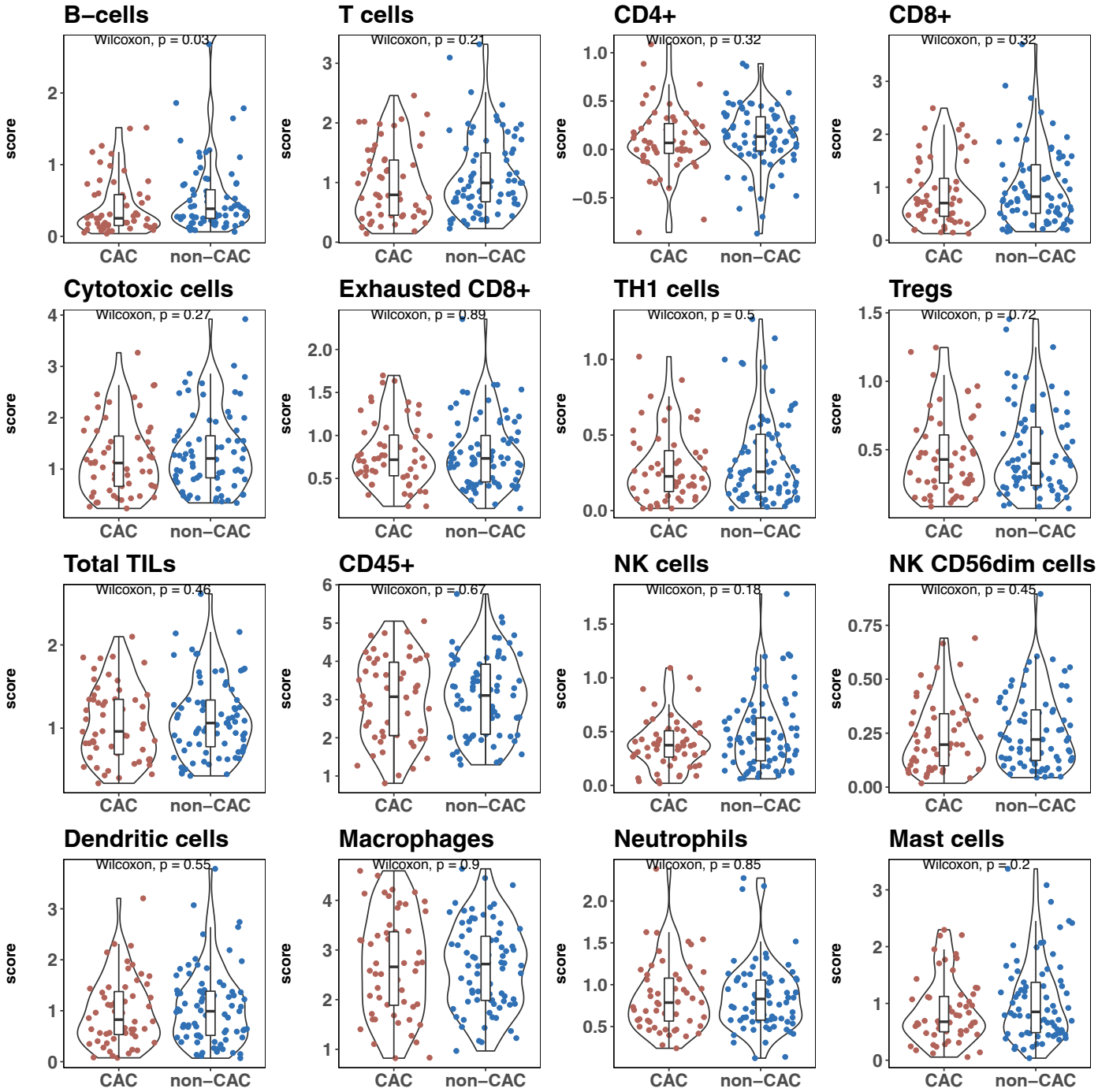
Extended Data Fig. 2 | Differential gene expression according to SAT, VAT, SKM and body weight loss. tumor differential gene expression between patients with **a** SAT loss <20% versus ≥20%, **b** VAT loss <20% versus ≥20%, **c** Muscle loss <10% versus ≥10%, **d** BMI adjusted weight loss grade 0-3 versus 4, all adjusted for number of tumour regions, sex, and histology. **e** Overlap of differentially expressed genes (DEG) between the ≥20% SAT, ≥20% VAT, ≥10% SKM and grade 4 weight loss groups. P values from moderated two-sided t-test without adjusting for multiple testing.

https://doi.org/10.1038/s41591-023-02232-8



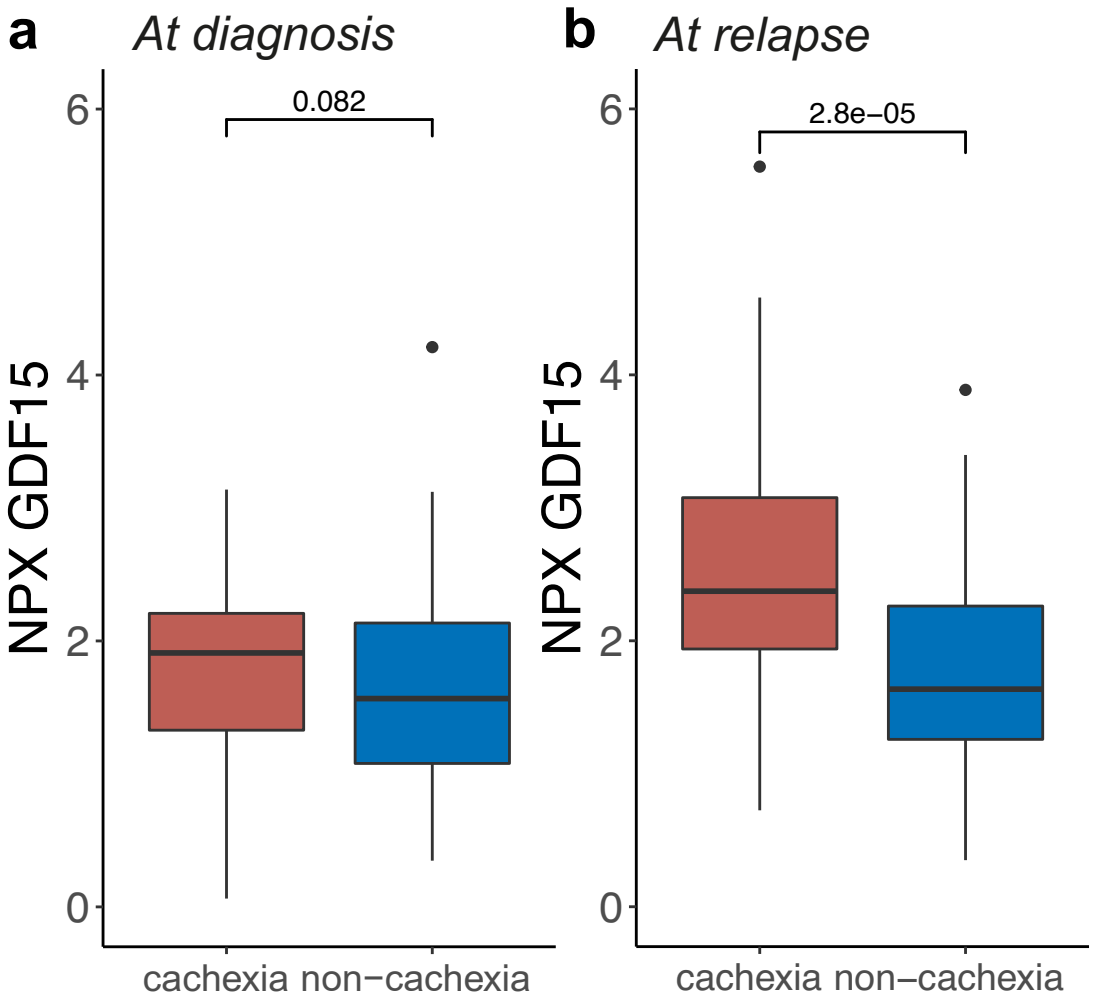
Extended Data Fig. 3 | TCRA scores according to CAC and non-CAC groups at diagnosis. a Tumour TCRA scores according to CAC (n=66) and non-CAC status (n=85). **b** Tumour TCRA scores according to CAC and non-CAC in female patients (n= 26 and 37, respectively) and **Cc** male patients (n=40 and 48, respectively). **d** Blood TCRA scores according to CAC (n=66) and non-CAC status (n=85). **e** Blood TCRA scores in blood in female patients (n= 26 and 37, respectively) and **Ff** male patients (n=40 and 48, respectively). P-values from Wilcoxon tests are two-sided and not adjusted for multiple testing. Box plots represent lower quartile, median and upper quartile; whiskers extend to a maximum of 1.5 × IQR beyond the box. Points indicate individual data points.

https://doi.org/10.1038/s41591-023-02232-8



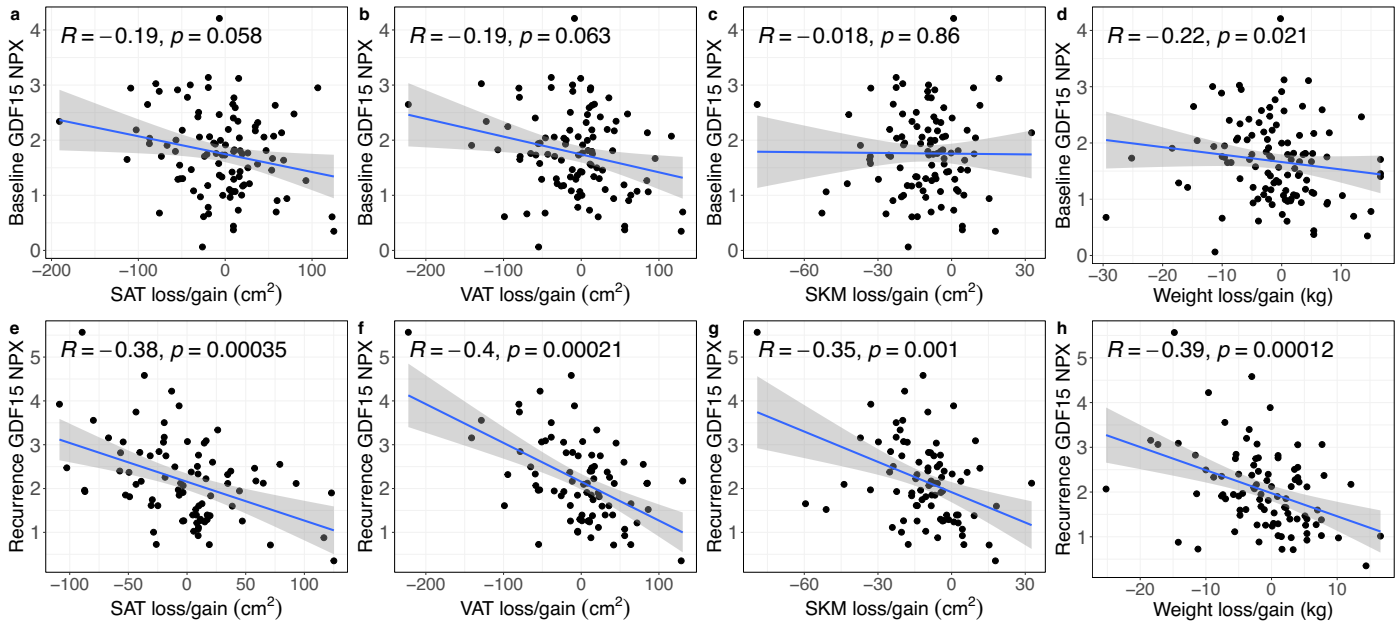
Extended Data Fig. 4 | Danaher scores according to CAC and non-CAC groups. P-values are from two-sided Wilcoxon tests and not adjusted for multiple comparisons. No significant difference was observed between cachexia (CAC, n=56) and non-cachexia (non-CAC, n=77) groups after Bonferroni correction. Box plots represent lower quartile, median and upper quartile; whiskers extend to a maximum of 1.5 × IQR beyond the box. Points indicate individual data points.

https://doi.org/10.1038/s41591-023-02232-8



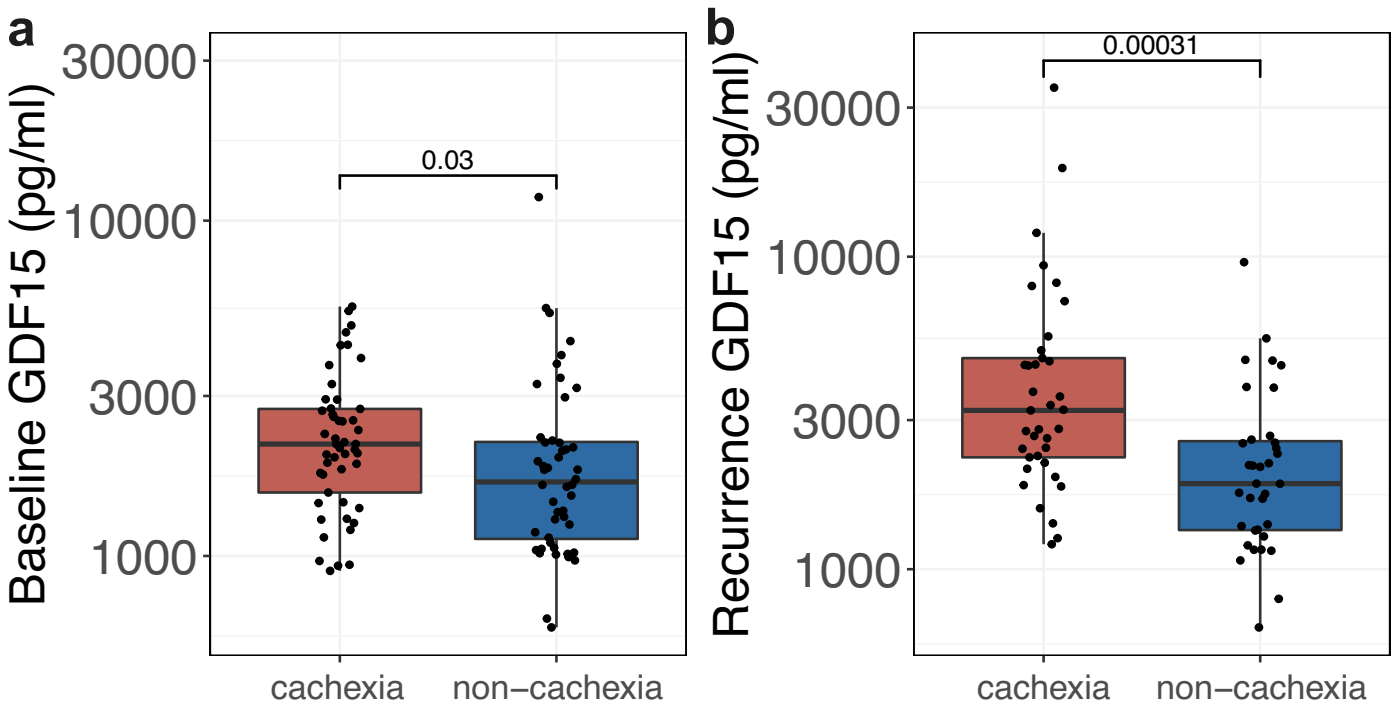
Extended Data Fig. 5 | Normalized plasma protein expression (NPX) of GDF15 in the cachexia versus non-cachexia groups. GDF15 protein expression in the CAC group (cachexia, red) and the non-CAC group (non-cachexia, blue) at cancer diagnosis (n=110) (**a**) and cancer relapse (n=110) (**b**). Two-sided Wilcoxon test. Box plots represent lower quartile, median and upper quartile, whiskers extend to a maximum of 1.5 × IQR beyond the box. Points indicate outliers. Y-axis represents log10 scale.

https://doi.org/10.1038/s41591-023-02232-8



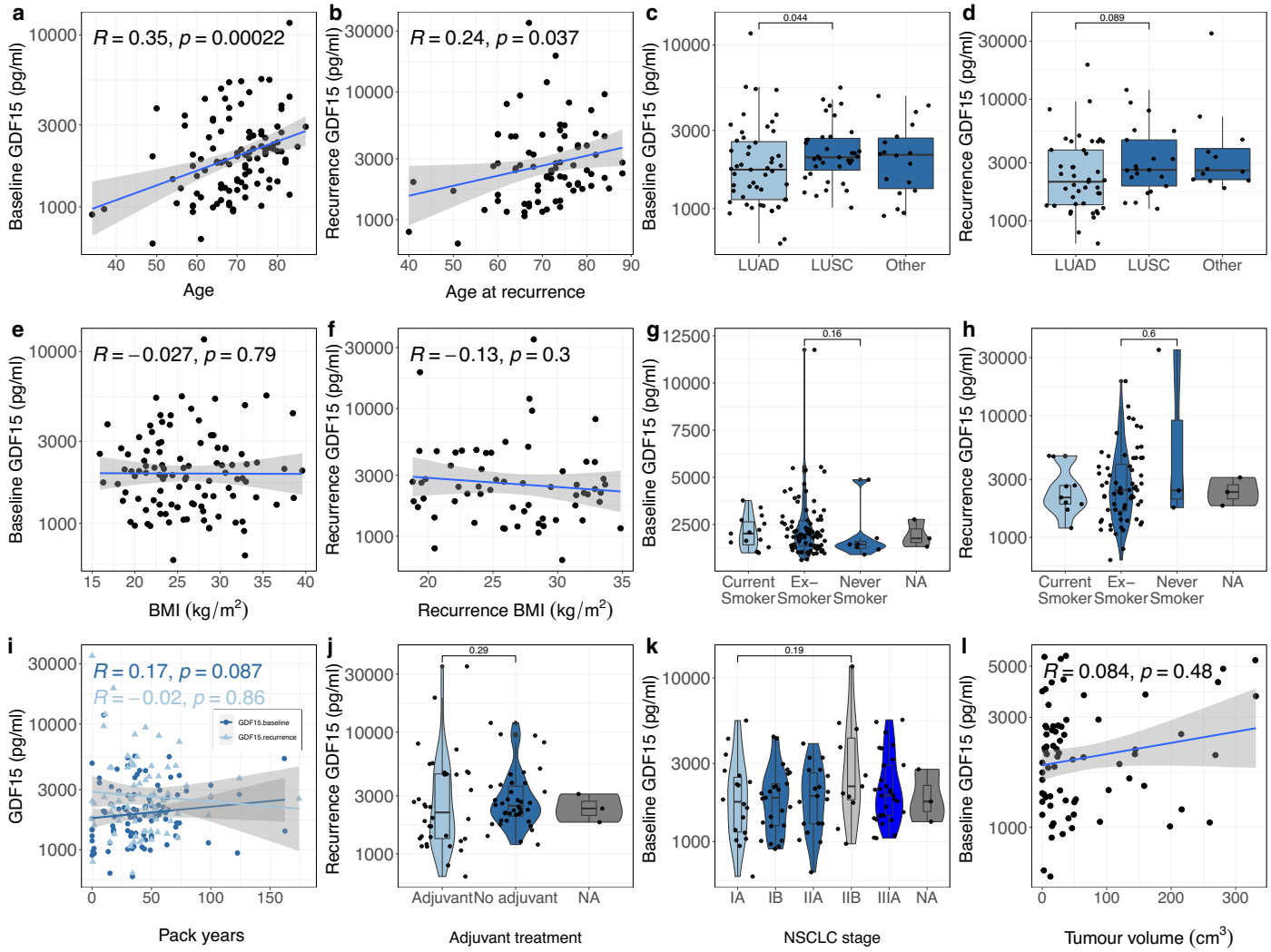
Extended Data Fig. 6 | Correlation between normalized plasma GDF15 expression (NPX) and changes in body composition and body weight. **a-d** Spearman's correlation between body composition at diagnosis (a-C, n=103) and body weight at diagnosis (d, n=104) and plasma GDF15. **e-h** Spearman's correlation between body composition at relapse (e-g, n=61) and body weight at relapse (h, n=76). Grey error bands represent the 95% confidence interval of the fitted linear model. Y-axes represent log10 scales.

https://doi.org/10.1038/s41591-023-02232-8



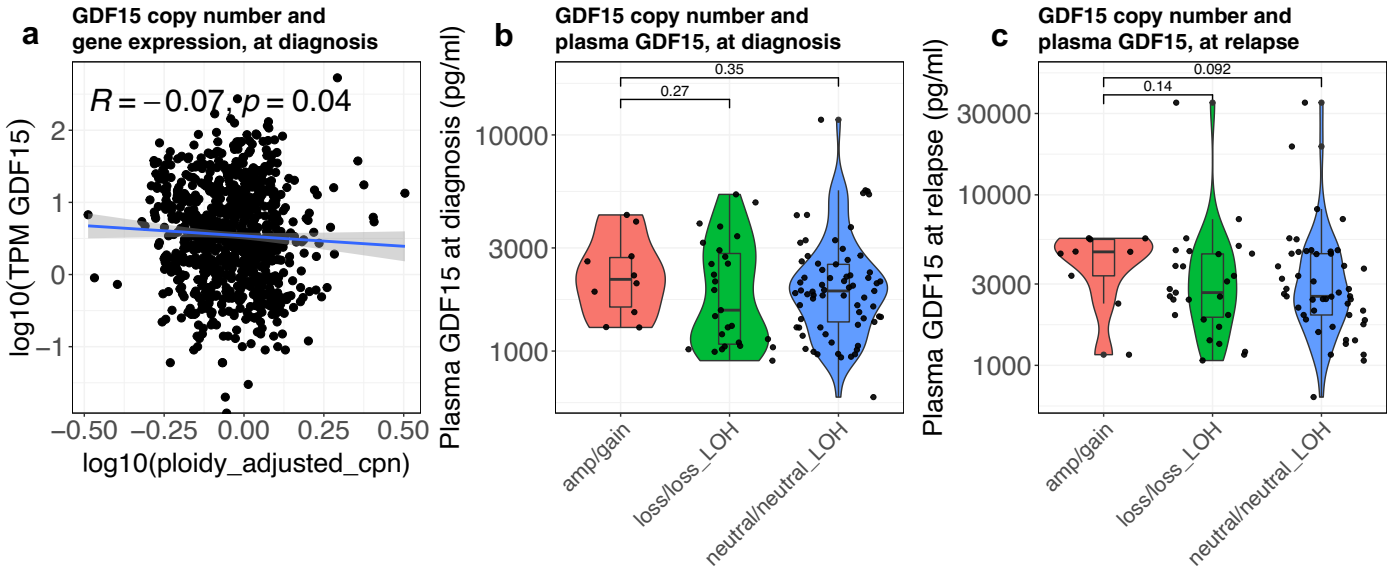
Extended Data Fig. 7 | Plasma GDF15 levels in patients with versus without cachexia. Plasma GDF15 levels in the CAC group (cachexia, red) and the non-CAC group (non-cachexia, blue) at cancer diagnosis (**a**, n=96) and cancer relapse (**b**, n=71). Two-sided Wilcoxon test. Box plots represent lower quartile, median and upper quartile, whiskers extend to a maximum of 1.5 × IQR beyond the box. Points indicate individual data points. Y-axes represent log10 scales.

https://doi.org/10.1038/s41591-023-02232-8



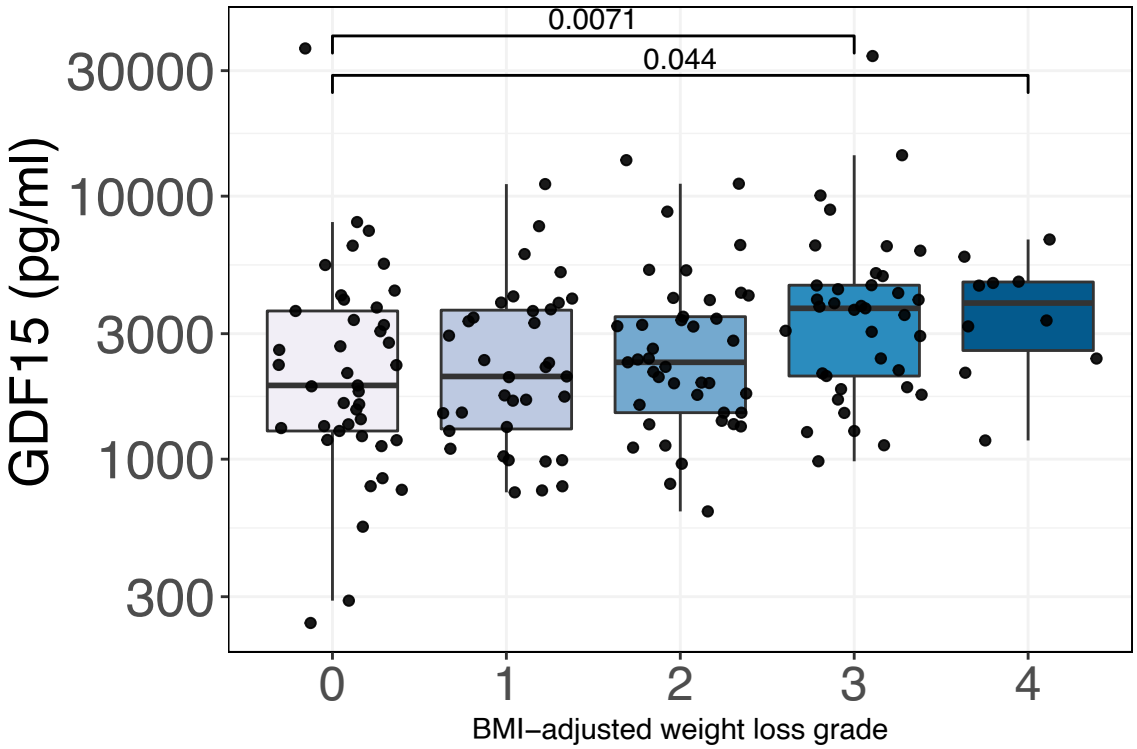
Extended Data Fig. 8 | Correlation of GDF15 plasma levels at diagnosis and at relapse with clinical features. a, **b** Spearman's correlation between age and plasma GDF15 at diagnosis (a, n=107) and relapse (b, n=79). **c**, **d** Correlation between histology and plasma GDF15 at diagnosis (c, n=107) and relapse (d, n=79), two-sided Wilcoxon test. **e**, **f** Spearman's correlation between body mass index (BMI) and plasma GDF15 at diagnosis (e, n=107) and relapse (f, n=79). **g**, **h** Correlation between smoking status and plasma GDF15 at diagnosis (g, n=107) and relapse (h, n=79), two-sided Wilcoxon test. **i** Spearman's correlation between smoking pack years and plasma GDF15 (at diagnosis n=107, at relapse n=79). **j** Correlation between adjuvant therapy and plasma GDF15 at relapse (n=79), two-sided Wilcoxon test. **k** Correlation between lung cancer stage and plasma GDF15 at diagnosis (n=107), two-sided Wilcoxon test. **l** Spearman's correlation between tumour volume and plasma GDF15 at diagnosis (n=73). P-values not adjusted for multiple comparisons. Grey error bands represent the 95% confidence interval of the fitted linear model. Box plots represent lower quartile, median and upper quartile, whiskers extend to a maximum of 1.5 × IQR beyond the box. Points indicate individual data points. Y-axis represents log10 scales. NSCLC, non-small cell lung cancer; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma.

https://doi.org/10.1038/s41591-023-02232-8



Extended Data Fig. 9 | Copy number alterations of GDF15. a Ploidy-adjusted copy number compared to transcript-per-million GDF15 gene expression, at baseline, log10 transformed, Spearman's correlation, n=1050 regions from 348 patients. Grey error bands represent the 95% confidence interval of the fitted linear model. **b** Copy number events in relation to circulating GDF15 levels, at diagnosis (n=106), log10 transformed, p-values from two-sided Wilcoxon test, not adjusted for multiple testing. **c** Copy number events in relation to circulating GDF15 levels, at relapse (n=44), log10 transformed. P-values from two-sided Wilcoxon test, not adjusted for multiple testing. Box plots represent lower quartile, median and upper quartile; whiskers extend to a maximum of 1.5 × IQR beyond the box. Points indicate individual data points. Y-axes represent log10 scales.

https://doi.org/10.1038/s41591-023-02232-8



Extended Data Fig. 10 | Plasma GDF15 levels and BMI-adjusted weight loss in the ARCHER1009 cohort. GDF15 levels according to BMI-adjusted weight loss category in patients (n=164) treated in the ARCHER1009 trial. P values from two-sided Wilcoxon test, not adjusted for multiple testing. Box plots represent lower quartile, median and upper quartile, whiskers extend to a maximum of 1.5 × IQR beyond the box. Points indicate individual data points. Y-axis represents log10 scale.

nature portfolio

Corresponding author(s): Mariam Jamal-Hanjani
Charles Swanton

Last updated by author(s): Jan 17, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Clinical and pathological data is collected from patients during study follow up - this period is a minimum of five years. Data collection is overseen by the sponsor of the study (Cancer Research UK & UCL Cancer Trials Centre) and takes place in 19 hospitals across the United Kingdom. A centralised database called MACRO is used for this purpose. Recruitment started in April 2014 and is still ongoing (in London and Manchester). Survival data were last updated on 15 June 2021.
Data analysis	All code to reproduce the figures will be available in a Zenodo repository with DOI reference. R version 4.1.1 R packages for data and statistical analyses table1_1.4.2 gg.gap_1.3 data.table_1.14.2 readxl_1.4.0 xlsx_0.6.5 dplyr_1.0.8 tidyr_1.2.0 cowplot_1.1.1 ggsci_2.9 ggpubr_0.4.0 ggplot2_3.3.6 ggsignif_0.6.3

ggdendro_0.1.23
scales_1.2.0
fuzzyjoin_0.1.6
survminer_0.4.9
survival_3.4-0
GenomicRanges_1.46.1
ggvenn_0.1.9
fgsea_v1.22.0
EnhancedVolcano_1.10.0
OlinkAnalyze_v3.1.0
OlinkR_1.0.1

GISTIC v2.0.23 for copy number analysis
CIBERSORTx v2019 for cell type deconvolution
DAFS version 3.6.9 for body composition analyses

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The whole exome sequencing (WES) data and RNA sequencing (RNA-seq) data (in each case from the TRACERx study) used during this study have been deposited at the European Genome–phenome Archive (EGA), which is hosted by The European Bioinformatics Institute (EBI) and the Centre for Genomic Regulation (CRG) under the accession codes EGAS00001006494 (WES) and EGAS00001006517 (RNAseq); access is controlled by the TRACERx data access committee. Details on how to apply for access are available at the linked page. Olink data from TRACERx and GDF15 data from TRACERx and ARCHER1009 are deposited will be provided in a Zenodo repository linked in the manuscript. Hallmark genesets58 were downloaded from Molecular Signatures Database (MSigDB) (https://www.gsea-msigdb.org/gsea/msigdb/).

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Sex and gender were not considered in the study design. Self-reported sex was collected and used in the analyses.
Population characteristics	TRACERx cohort: 43.9% are females , 56.1% males; 92.5% are smokers or have a smoking history, 7.5% are never smokers; 20.1% of patients were diagnosed at stage IA, 21.4% at IB, 19.0% at IIA, 17.4% at IIB, 21.2% at IIIA and 0.9% at IIIB; 55.6% of diagnosed tumours were adenocarcinomas, 32.4% were squamous cell carcinomas and 10.0% were of other histological subtypes; 90.6% of the cohort is from a white ethnic background and the mean age of the patients was 69. BLCS cohort: 47.1% are females, 52.9% are male; 88.3 are smokers or have a smoking history, 11.7% are never smokers; 48.3% were stage I, 15% were stage II, 36.7% were stage III. 53.1% of the tumors were adenocarcinoma, 24.8% were squamous cell carcinoma. 97.9% were from white ethnic background, mean age of the patients was 65. ARCHER1009: 35% are females, 65% are male; 82% are smokers or had a smoking history, 19% are never smokers; all patients had locally advanced or metastatic NSCLC, 69% had adenocarcinoma, 27% squamous-cell carcinoma. Mean age was 63 years.
Recruitment	Samples for this study were chosen based on the availability of banked plasma and computer tomography (CT) scans. No other selection criteria were applied.
Ethics oversight	The TRACERx study was approved by the NRES Committee London with the following details: REC reference: 13/LO/1546 Protocol number: UCL/12/0279, IRAS project ID: 138871. ARCHER1009 was approved by the Pfizer Institutional Review Board, the BLCS was approved by the Committees on the Use of human Subjects in Research at the Massachusetts General Brigham (MGB) and the Harvard School of Public Health (IRB19-1612). All studies were conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from all study participants. (ClinicalTrials.gov Identifier: NCT01888601, NCT01360554; BLCS NIH 1U01CA209414-01A1).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Samples from the TRACERx cohort were chosen based on the availability of matched baseline and relapse plasma material as well as corresponding baseline and relapse computer tomography (CT) images for body composition analysis. Cases from the Boslon Lung Cancer Study were selected based on presence of stage I-III NSCLC and available CT imaging of third lumbar vertrebrae level. Samples from the ARCHER1009 cohort were selected based on availability of plasma.
Data exclusions	Please see study inclusion/exclusion criteria above.
Replication	GDF15 and Olink measurements were taken from distinct samples.
Randomization	Non-interventional study, not applicable
Blinding	Blinding is not relevant as this is an observational study. Patients were not allocated to any intervention and they were followed up and assessed as per routine practice. No biomarker results (tissue and bloods) are reported back to patients, so there is no likelihood of people changing their behaviours based on these findings. The laboratory analyses were all performed without knowing the outcome (DFS or survival) status of the patients, which represents a form of blinding.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Elecsys GDF15 immunoassay (F. Hoffmann-La Roche, Basel, Switzerland), Ref 08946779190 Human GDF-15 Quantikine ELISA kit (R&D Systems, Minneapolis, MN), Ref SGD150
Validation	Elecsys GDF15 immunoassay validated in >900 patients (O'Connor et al, JAMA, 2009) Human GDF-15 Quantikine ELISA kit validated in other studies (Breen et al, Cell Metabolism, 2020)

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	TRACERx Lung: https://clinicaltrials.gov/ct2/show/NCT01888601 ARCHER1009: https://clinicaltrials.gov/ct2/show/NCT01360554). BLCS: https://sites.sph.harvard.edu/blcs/
Study protocol	TRACERx Lung: The study protocol is available as a supplement to Jamal-Hanjani et al. 2017, https://doi.org/10.1056/NEJMoa1616288

ARCHER1009: The study protocol is available as a supplement to Ramalingam et al, 2014, https://doi.org/10.1016/S1470-2045(14)70452-8

BLCS: The Boston Lung Cancer Study is a cancer epidemiology cohort, details are provided on https://sites.sph.harvard.edu/blcs/.

Data collection

TRACERx: Clinical and pathological data is collected from patients during study follow up - this period is a minimum of five years. Data collection is overseen by the sponsor of the study (Cancer Research UK & UCL Cancer Trials Centre) and takes place in hospitals across the United Kingdom.
ARCHER1009: This study ran between 2011 and 2013 and enrolled patients with advanced or metastatic NSCLC to be randomized between dacomitinib or erlotinib. The study is completed and closed, primary readout was published in 2014.
BLCS: This cohort comprises lung cancer cases enrolled at Massachusetts General Hospital and the Dana-Farber Cancer Institute from 1992-present

Outcomes

TRACERx Lung: disease-free survival is a pre-defined primary endpoint determined by RECIST v1.1 criteria.
ARCHER1009: Primary endpoint was progression-free survival.