RESEARCH ARTICLE

Gut microbiome functionality might be associated with exercise tolerance and recurrence of resected early-stage lung cancer patients

Andrea Marfil-Sánchez[1☯], Bastian Seelbinder[1☯], Yueqiong Ni[1], Janos Varga[2], Judit Berta[2], Virag Hollosi[2], Balazs Dome[2,3,4], Zsolt Megyesfalvi[2,3,4], Edit Dulka[5], Gabriella Galffy[5], Glen J. Weiss[6‡], Gianni Panagiotou[1‡*], Zoltan Lohinai[2‡]

1 Systems Biology and Bioinformatics, Leibniz Institute for Natural Product Research and Infection Biology, Hans Knöll Institute, Jena, Germany, 2 National Koranyi Institute of Pulmonology, Budapest, Hungary, 3 Department of Thoracic Surgery, National Institute of Oncology-Semmelweis University, Budapest, Hungary, 4 Division of Thoracic Surgery, Department of Surgery, Comprehensive Cancer Center, Medical University of Vienna, Vienna, Austria, 5 County Hospital of Torokbalint, Torokbalint, Hungary, 6 MiRanostics Consulting, Oro Valley, Arizona, United States of America

☯ These authors contributed equally to this work.
‡ GJW, GP and ZL are joint senior authors on this work.
* gianni.panagiotou@leibniz-hki.de

Abstract

Impaired exercise tolerance and lung function is a marker for increased mortality in lung cancer patients undergoing lung resection surgery. Recent data suggest that the gut-lung axis regulates systemic metabolic and immune functions, and microbiota might alter exercise tolerance. Here, we aimed to evaluate the associations between gut microbiota and outcomes in lung cancer patients who underwent lung resection surgery. We analysed stool samples, from 15 early-stage lung cancer patients, collected before and after surgical resection using shotgun metagenomic and Internal Transcribed Spacer (ITS) sequencing. We analysed microbiome and mycobiome associations with post-surgery lung function and cardiopulmonary exercise testing (CPET) to assess the maximum level of work achieved. There was a significant difference, between pre- and post-surgical resection samples, in microbial community functional profiles and several species from *Alistipes* and *Bacteroides* genus, associated with the production of SCFAs, increased significantly in abundance. Interestingly, an increase in VO_2 coincides with an increase in certain species and the "GABA shunt" pathway, suggesting that treatment outcome might improve by enriching butyrate-producing species. Here, we revealed associations between specific gut bacteria, fungi, and their metabolic pathways with the recovery of lung function and exercise capacity.

Introduction

The ethology or progression of lung cancer has been linked to numerous factors, including the gut microflora [1]. The microbiome can control epithelial cell proliferation/ differentiation, nutrition, detoxification, metabolism, and hormonal homeostasis. Gut microbiota has been implicated in tumour development in cancers [2, 3]. It has been shown that several chronic lung diseases, including cancer may be linked to a dysbiotic airway microbiota and frequently occur in conjunction with gastrointestinal disorders [4].

Previous studies showed that exercise training can alter the abundance of butyrate in the gut through an increase in the relative abundance of butyrate-producing species [5, 6]. Butyrate is a short-chain fatty acid (SCFA) that can regulate proliferation of epithelial cells in the gut, improve the integrity of the gut barrier, and alter the immune system and gene expression of the host [6]. A decrease in the abundance of Firmicutes has been repeatedly reported in lung cancer patients [2, 7] and has been previously associated with dysbiosis of gut metabolism, resulting in lower SCFAs concentration [2]. In contrast, a higher ratio of Firmicutes to Bacteroidetes bacteria has been show to correlate with maximal oxygen uptake during exercise (VO_2) [8]. Barton *et al*. [9] showed that athletes have greater faecal SCFA concentrations and altered abundance of bacterial pathways related to the biosynthesis of amino acids and the metabolism of carbohydrates. Others found that exercise training can alter the gut microbiota, at the taxonomic and functional level, in obese and non-obese humans [10] and reduce the expression of genes associated to metabolism of fructose and amino acids [11]. Moreover, exercise training may affect the gut mucus layer integrity, which is involved in the prevention of microbial adherence to the gut epithelium and acts as a substrate for bacteria of the mucosa-associated microflora. Increased heat shock protein response that prevents tight junction breakdown between epithelial cells [12] and improves resilience of the gut barrier [13, 14] has been observed in trained athletes. Others showed that physical activity decreases the risk for colorectal cancer by 24% [15].

Cardiopulmonary exercise testing (CPET) is a non-invasive, physiological test that provides a comprehensive overview of the pulmonary response to exercise and allows for the evaluation of the body's metabolic state, functional capacity, and impairment, through the assessment of submaximal and peak exercise responses. The metabolic changes that characterize health status and tumour recurrence have a stronger correlation with exercise tolerance, as measured by CPET, than with resting pulmonary and cardiac function testing, thus resulting in an increase in the use of CPET in patient management. Among many parameters, CPET measures the oxygen uptake (VO_2) peak that is used in predicting postoperative overall condition and pulmonary complications [16, 17].

In this study, our primary aim was to characterize the taxonomy and functionality of the gut microbiomes' metabolic interaction in lung cancer patients who underwent lung resection surgery. Our secondary aim was to describe the taxonomy changes in the same setting according to postoperative disease recurrence. We focus on the alterations of microbiome functionality to understand the complex gut-lung axis' metabolic interactions.

Materials and methods

Study population

From 2018 to 2019, we initially screened a cohort of 98 lung cancer patients from the National Koranyi Institute of Pulmonology and County Hospital of Torokbalint. From these, 20 were in early-stage and had their tumours surgically resected. Ultimately, 15 cases were suitable for the CPET testing and underwent pulmonary rehabilitation. A flow chart of study participants and

excluded patients is shown in S1 Fig. Baseline and followed-up samples were collected from those patients. Clinicopathological data collected included gender, age, smoking history, chemo- and radiotherapy treatments, type of operation, and overall survival (OS) or disease-free survival. Tumour, node, metastases stage according to the Union for International Cancer Control (8th edition [18]), and age at the time of diagnosis were recorded. The study and all treatments were conducted in accordance with the guidelines of the National Comprehensive Cancer Network and the Helsinki Declaration of the World Medical Association. The study was approved by the national level ethics committee (Hungarian Scientific and Research Ethics Committee of the Medical Research Council (ETTTUKEB- 50302-2/2017/EKU)). Informed consent was obtained for all patients. Stool samples were collected before (prior to surgery within 14 days) and 12 months after lung resection surgery. To assess a patient's overall condition, we performed a CPET test one year post-surgical resection to coincide with the time of follow-up stool sample collection. Major characteristics of the patient cohort are displayed in S1 Table. Patients received no neo-adjuvant therapy, and selected cases were treated with post-operative systemic chemotherapy with a platinum-based doublet regimen therapy (S1 Table).

Pulmonary function (PF) and CPET testing

We used spirometry at rest one year after surgery to assess key lung function parameters including {percentage of reference value} (Forced Expiratory Volume (FEV1) {%}), total Lung Capacity (TLC{%}), residual volume (RV{%}).

We performed CPET that begins with mild followed by intense exercise on an upright bicycle. Patients breathing through a mouthpiece. We measured the ventilation and respiratory gas parameters during exercise using oxygen and carbon dioxide gas analysers. Respiratory volumes were computed by integrating the air flow signals over the time of inspiration and expiration. The CPET test lasted until the maximal effort workload achieved for approximately 10 minutes.

We performed an incremental exercise test on an electronically braked cycle ergometer (Ergoline-900, Marquette) that begin with 3 minutes rest and 3 minutes constant pedalling at 20 W, work rate was increased from 5, 10, or 15 W/min in ramp profile. Patient pedalling speed was 60 rpm. Pulmonary ventilation (VE) and gas exchange (VO_2 and carbon dioxide output (VCO_2)) were measured breath-by-breath by a mass flow sensor and exercise metabolic measurement system (Vmax 29c, SensorMedics). During the test, we monitored heart rate (HR) using a 12-lead ECG (Cardiosoft, SensorMedics) and oxygen saturation by pulse oximetry (SatTrak, SensorMedics). We measured CPET parameters and lung function parameters including forced expiratory volume in one-second workload (Watt), maximal ventilation during exercise (VE max), ventilatory equivalents for oxygen (VE/VO_2), Oxygen pulse (O_2/HR), ventilatory equivalents for carbon dioxide (VE/VCO_2).

DNA extraction from stool samples

All samples were processed by Novogene. DNA extraction, library preparation and sequencing was done as described previously [19]. Stool samples were thoroughly mixed with 900 μL of CTAB lysis buffer. All samples were incubated at 65°C for 60 min before being centrifuged at 12000×g for 5 min at 4°C. Supernatants were transferred to fresh 2 mL microcentrifuge tubes and 900 μL of phenol:chloroform:isoamyl alcohol (25:24:1, pH = 6.7; Sigma-Aldrich) was added for each extraction. Samples were mixed thoroughly prior to being incubated at room temperature for 10 min. Phase separation occurred by centrifugation at 12,000×g for 15 min at 4°C, and the upper aqueous phase was re-extracted with a further 900 μL of phenol:chloroform:isoamyl alcohol. Next, samples were centrifuged at 12,000×g for 10 min at 4°C, and the

upper aqueous phases were transferred to fresh 2 mL microcentrifuge tubes. The final extraction was performed with 900 μL of chloroform:isoamyl alcohol (24:1), and layer separation occurred by centrifugation at 12,000×g for 15 min at 4˚C. Precipitation of DNA was achieved by adding the upper phase from the last extraction step to 450 μL of isopropanol (Sigma-Aldrich) containing 50 μL of 7.5M ammonium acetate (Fisher). Samples were incubated at–20˚C overnight. Samples were centrifuged at 7500×g for 10 min at 4˚C, and supernatants were discarded. Finally, DNA pellets were washed three times in 1 mL of 70% (v/v) ethanol (Fisher). The final pellet was air-dried and re-suspended in 200 μL of 75mM TE buffer (pH = 8.0; Sigma-Aldrich).

Library preparation and sequencing

Sequencing library was generated based on Illumina technologies and followed manufacturer's recommendations. Index codes were added to each sample. Briefly, the genomic DNA was randomly fragmented to a size of 350 bp, then DNA fragments were narrowly size selected with sample purification beads. The selected fragments were then end polished, A-tailed, and ligated with adapter. These fragments were filtered with beads again and amplified by PCR reaction. At last, the library was analysed for size distribution and quantified using real-time PCR. The library was then to be sequenced on an Illumina platform Novaseq 6000 (Novogene) with paired-end reads of 150 bp.

The concentration of genomic DNA for ITS2 sequencing was determined by Qubit and the DNA quality was checked on the gel. 200 ng of DNA was used as input for PCR reaction with corresponding primer set specifically binding to different hypervariable regions. Each primer set had a unique barcode. PCR product was then run on the gel and DNA fragment with the proper amplification size was cut and purified. The purified PCR product was then used as template for library preparation. The PCR products were pooled together with equal amount and then end polished, A-tailed, and ligated with the adapter. These fragments were filtered with beads again. After PCR reaction, the library was analysed for size distribution and quantified using real-time PCR. The library was then to be sequenced on Hiseq2500.

Quality control

Quality control of raw reads was performed using the Sunbeam 2.1 pipeline [20]. First, cutadapt [21] version 2.8 was used to remove universal adapter sequences. Next, trimmomatic [22] version 0.36 was used to perform Illumina-specific adapter trimming, window quality trimming (Q5 over 25 nt), and 3' and 5' clipping (Q<6). Resulting reads shorter than 36nt were removed. Decontamination of human reads was performed by mapping quality-controlled reads with BWA [23] version 0.7.17 against a masked human reference genome (GRCh38-89). Masking of low entropy regions was performed using BBmask. Reads with 99% coverage and >97% identity to the human reference were removed.

Taxonomic and functional annotation

Taxonomic annotation was performed using MetaPhlAn2 [24] version 2.7.7 with default settings, generating taxonomic relative abundances.

The PIPITS pipeline [25] version 2.4 was used for taxonomy annotation of fungal Internal Transcribed Spacer (ITS) with default parameters including quality filtering, read-pair merging, ITS2 extraction and chimera removal. Remaining reads were binned based on 97% similarity as operational taxonomic units and aligned to the UNITE fungi database using Mothur classifier [26].

Functional annotation was performed using HUMAnN2 [27] version 0.11.2. In the pipeline, the reads were mapped to the MetaCyc database for pathway annotation, and the UniRef90 database to estimate gene family abundances. These abundances were aggregated into MetaCyc pathway abundances using MinPath based on the MetaCyc database.

Microbial community diversity analysis

Alpha diversity indices detailing microbial community composition within samples were calculated using the R packages vegan [28] and fossil [29]. For estimating beta diversity reflecting community dissimilarities, Bray-Curtis distances were calculated using R package vegan [28].

Co-abundance network

For bacterial co-abundance network reconstruction, the species relative abundance table was split into before and after surgical resection samples, and they were processed independently with SparCC [30].

Correlation analyses with lung function parameters

Partial Spearman's correlations adjusted for Chronic obstructive pulmonary disease (COPD) and cancer type were determined between significantly different bacterial features (species and pathways), fungal species abundances, and lung function parameters. For visualization only bacterial features with significant correlations ($P<0.05$) and fungal species with significant correlations ($P<0.05$) and absolute correlation coefficient > 0.65 were selected.

Prediction of VO$_2$, recurrence and overall survival (OS)

Prediction of VO$_2$, recurrence and OS was performed by calculating microbial balances [31]. A working implementation is available https://github.com/UVic-omics/selbal. A microbial balance is a special kind of log-contrast. Briefly, let X = (X$_1$,X$_2$,. . .,X$_k$) be a composition with k components. For two disjoint subsets of k$_+$ and k$_-$ parts, respectively, the balance B of the component sets X$_+$ and X$_-$ is determined by the formula:

$$B\left(X_{+}, X_{-}\right) \propto \frac{1}{k_{+}} \sum_{i \in I_{+}} \log X_{i} - \frac{1}{k_{-}} \sum_{j \in I_{-}} \log X_{j}$$

Feature selection for X$_+$ and X$_-$ was performed using the function "selbal.cv", which implements an iterative cross-validation procedure to (i) identify the optimal number of components (C$_{opt}$) to be included in the balance and (ii) estimate the importance of the selected components. Only features (bacterial species and MetaCyc pathways) that changed significantly after surgery were used for feature selection.

Since "selbal.cv" is a forward selection process where components are included sequentially at every step, we have a sequence of balances, ranging from C = 2 to C = 20 components. Once C$_{opt}$ has been determined, we apply "selbal.cv" to the whole data set, with the number of taxa C$_{opt}$, and obtain the global balance.

In order to estimate the robustness of the global balance and the importance of selected components, all balances with C$_{opt}$ components obtained in the cross-validation process are retrieved and compared to the global balance, obtaining the relative frequencies of the different balances and the proportion of times that each component was included in a balance. The more often a component was selected, the higher its importance.

Statistical analysis

Testing for significant differences in alpha diversity between before and after surgical resection was performed using Wilcoxon signed-rank test. To test for significant differences in the microbial composition between before and after surgical resection, beta-diversity was analysed using permutational multivariate analysis of variance (PERMANOVA) as implemented in the function adonis from R package vegan [28].

Bacterial species and MetaCyc pathways were filtered by 10% prevalence across all samples and their relative abundances were used for statistical comparisons of before vs. after surgery. Differentially abundant features were identified by the Wilcoxon signed-rank test and were considered significantly different if $P < 0.05$ after FDR correction using the Benjamini-Hochberg procedure.

Species-species correlation coefficients were estimated as the average of 20 inference iterations and 100 permutations were used for the pseudo P-value calculation. A correlation was considered significant if pseudo $P < 0.05$.

Partial Spearman's correlations between lung function and CPET parameters with bacterial and fungal features, adjusted for Chronic obstructive pulmonary disease (COPD) and cancer type, were performed using R package ppcor [32] and were considered significant if $P < 0.05$.

Results

Community taxonomic and functional diversity before and one year after surgical resection

We analysed data from 15 lung cancer patients who underwent lung resection surgery. Our cohort includes a balanced set of men (n = 7) and women (n = 8), and patients with stage I (n = 4), stage II (n = 8) and stage IIIA (n = 3) disease. Tumours recurred in six patients. Detailed patients' characteristics are shown in S1 Table.

These subjects had an overall good performance status (Eastern Cooperative Group Performance Status 0) at diagnosis.

The structure and function of the gut microbiome were assessed using shotgun metagenomic sequencing. In total, 337 bacterial species were detected. First, we compared the overall composition of the gut microbiome communities pre- versus post-surgical resection. There were no significant changes in bacterial species alpha diversity over time (Shannon $P = 0.98$, Simpson $P = 0.56$, Chao1 $P = 0.53$; Wilcoxon signed-rank test) (Fig 1A). Bacterial species beta diversity did not show significant differences either, but indicated a trend (Bray-Curtis dissimilarity, $P = 0.074$, $R^2 = 4.8\%$, PERMANOVA) (Fig 1B). We further compared the species' beta diversity taking into account the different stages of disease and chemotherapy treatment and found no significant differences (S2A and S2B Fig). Alpha and beta diversity comparisons of recurrent and non-recurrent patients showed no significant differences either, but indicated a trend with non-recurrent patients having higher alpha diversity at both, pre- and post-surgical resection (S2C and S2D Fig).

We used shotgun metagenomic sequencing to further examine the variation of gut bacterial functions. In total, 498 MetaCyc pathways were retrieved. We did not observe significant differences in MetaCyc alpha diversity evenness and richness after surgical resection (Shannon $P = 0.28$, Simpson $P = 0.6$, Chao1 $P = 0.49$; Wilcoxon signed-rank test) (Fig 1C). However, we found a significant difference in bacterial functional profiles (Bray-Curtis dissimilarity) from post-surgical samples compared with pre-resection samples ($P = 0.047$, $R^2 = 6.4\%$, PERMANOVA) (Fig 1D).
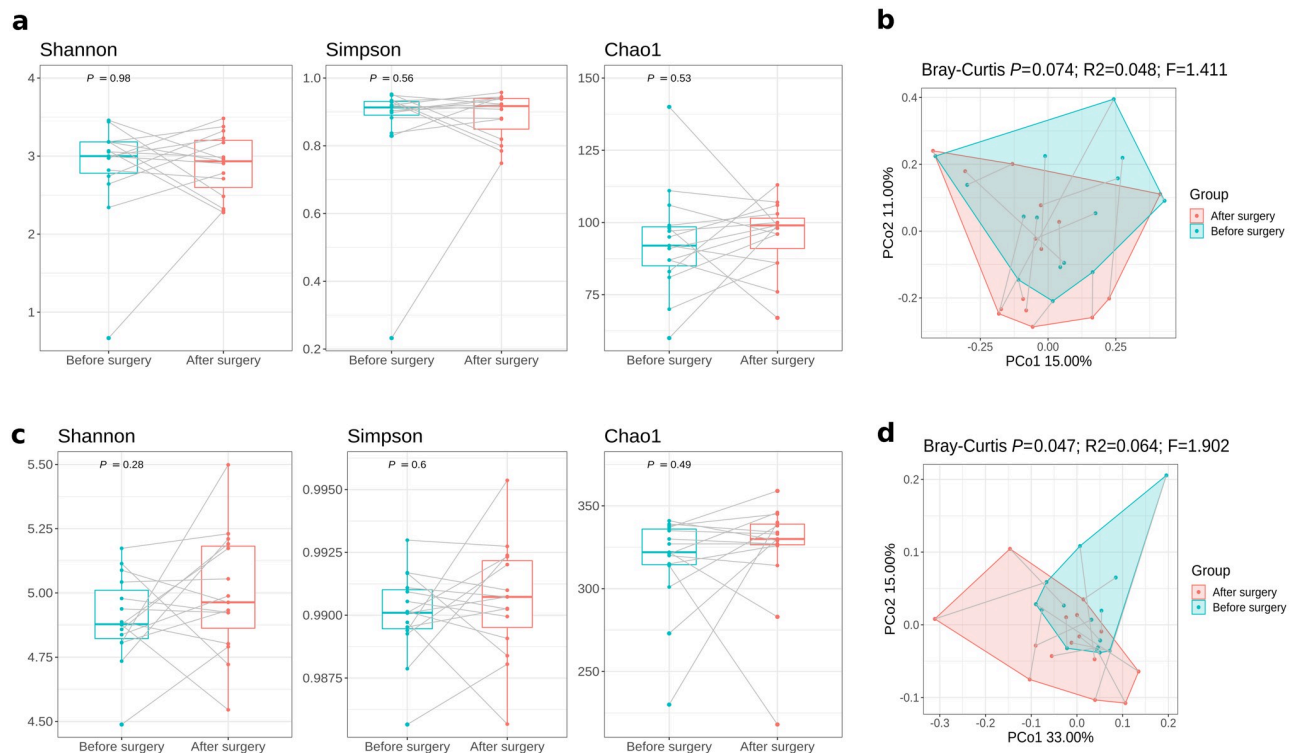
Fig 1. Comparison of the gut microbiota composition before and after surgical resection. (A, C) Boxplots, with median (centrelines), first and third quartiles (box limits) and 1.5x interquartile range (whiskers), showing alpha diversity Shannon, Simpson, and Chao1 indices at (A) bacterial species and (C) bacterial MetaCyc pathways. Gray lines connect samples of the same patient from before and after surgical resection. (B, D) Principal Coordinate Analysis plot based on Bray-Curtis distances before and after surgical resection at (B) bacterial species and (D) bacterial MetaCyc pathways. Gray lines connect the measurement of a patient before and after surgical resection.

https://doi.org/10.1371/journal.pone.0259898.g001

Alterations of gut microbial species, co-abundance network and functions one year after surgical resection

Next, we focused on taxonomic changes at the species level. We investigated which prevalent species changed significantly in relative abundance post-surgical resection. Thirty-two bacterial species showed significant changes after surgical resection compared to pre-surgical resection ($P<0.05$, Wilcoxon signed-rank test) (S2 Table). Of these, 15 were enriched pre-surgical resection, and 17 were enriched post-surgical resection (Fig 2A). We observed a significant decrease in the abundance of atypical opportunistic pathogens, such as *Klebsiella pneumoniae* [33], and *Odoribacter splanchnicus* [34]. These species are usually harmless within the gut of their host, but cause infection outside this niche. Other opportunistic pathogens like *Sutterella wadsworthensis* [35] decreased in abundance. Several species from *Alistipes* and *Bacteroides* genus, related to a healthy microbiome, increased significantly in abundance. These species have been associated with the production of SCFAs [36, 37].

We estimated relationships among gut microbes by constructing co-abundance networks based on bacterial species relative abundance (using SparCC [30]) for pre- and post-surgical resection groups, respectively. For evaluating the impact of surgery on the microbial community structure, we focused only on correlations between significant species that changed either from positive to negative or vice versa (Fig 2B). Overall, co-occurrence relationships were observed within and between phylum Actinobacteria, Bacteroidetes, Firmicutes, and Proteobacteria. We found more correlations changing from negative to positive post-surgical
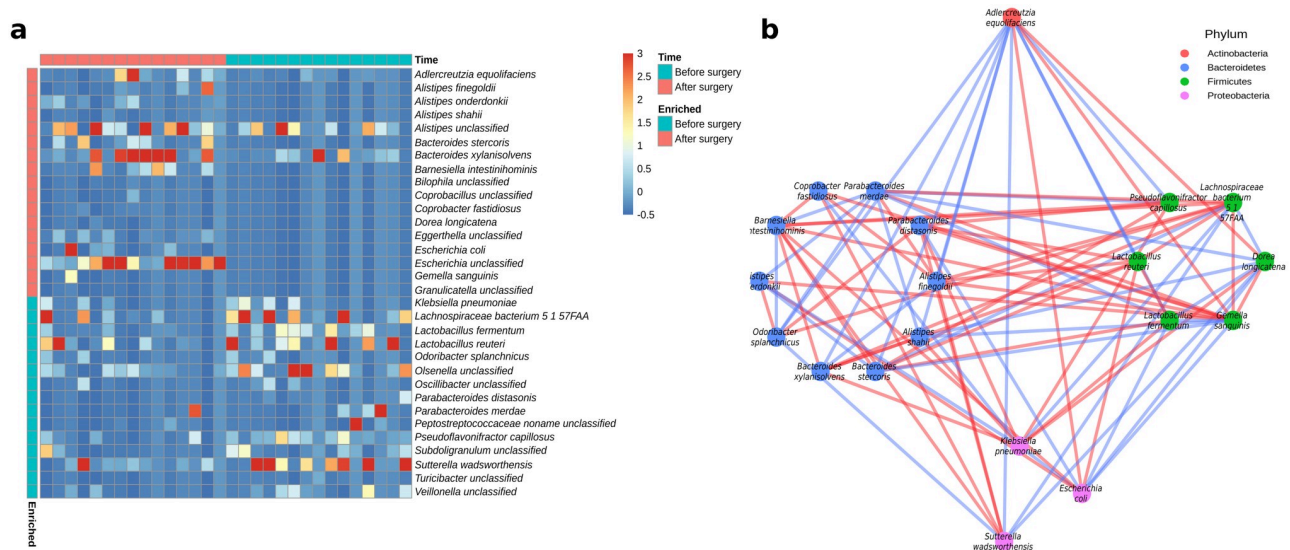
Fig 2. Taxonomic analysis of the gut microbiome in response to surgical resection. (A) Heatmap of differentially abundant bacterial species (*P*<0.05, Wilcoxon signed-rank test) before and after surgical resection. Red and blue in the far-left column indicates increased and decreased relative abundance, respectively. (B) Co-abundance network of bacterial species using SparCC [30]. Only correlations between differentially abundant bacterial species (*P*<0.05, Wilcoxon signed-rank test) that changed direction were used for network construction. The nodes are coloured based on their affiliated phyla. Edge colour indicates either correlations that changed from positive to negative (blue) or from negative to positive (red).

https://doi.org/10.1371/journal.pone.0259898.g002

resection (47) than from positive to negative (37). The species that showed the most changing correlations were *Gemella sanguinis*, *Adlercreutzia equolifaciens*, *Lachnospiraceae bacterium 5 1 57FAA*, *Parabacteroides merdae*, *K. pneumoniae* and *Barnesiella intestinihominis*.

Furthermore, 78 MetaCyc pathways differed significantly post-surgical resection compared to pre-surgical resection (*P*<0.05, Wilcoxon signed-rank test) (S3 Table) with 22 pathways enriched pre-surgical resection and 56 pathways post-surgical resection. The pathways with the highest increase were mostly related to the generation of precursor metabolites and energy and the biosynthesis of cofactors, electron carriers, and vitamins, whereas the pathways with the highest decrease were mostly related to core bacterial functions such as nucleoside and nucleotide biosynthesis (Fig 3).

Associations of bacteria and their function with lung function and CPET parameters

One year post-surgery, we investigated the significant associations between differentially abundant taxonomic and functional features with lung function and CPET parameters (FEV1%, TLC%, RV%, workload (Watt), VE, VO_2, O_2/HR, VE/VCO_2, VE/VO_2) adjusting for cofounders such as COPD and cancer type. In total, 12 bacterial species were significantly correlated (*P*<0.05) with lung function and CPET parameters. From these species, six were positively correlated with parameters: *Veillonella unclassified* (O_2/HR), *P. merdae* (RV%), *L. bacterium 5 1 57FAA* (TLC% and RV%), *Granulicatella unclassified* (VO_2), *G. sanguinis* (VO_2) and *Eggerthella unclassified* (RV%), whereas four species showed negative correlations: *Dorea longicatena* (Specific Airway Conductance (sGaw%)), *Coprobacillus unclassified* (VE and O_2/HR), *Bilophila unclassified* (VE and O_2/HR), *and B. intestinihominis* (workload and O_2/HR). *Parabacteroides distasonis* was positively correlated with VE/VO_2 and negatively correlated with workload while *Alistipes onderdonkii* was positively correlated with sGaw% and VE/VCO_2 and negatively correlated with workload (Fig 4A). We also investigated the relations between
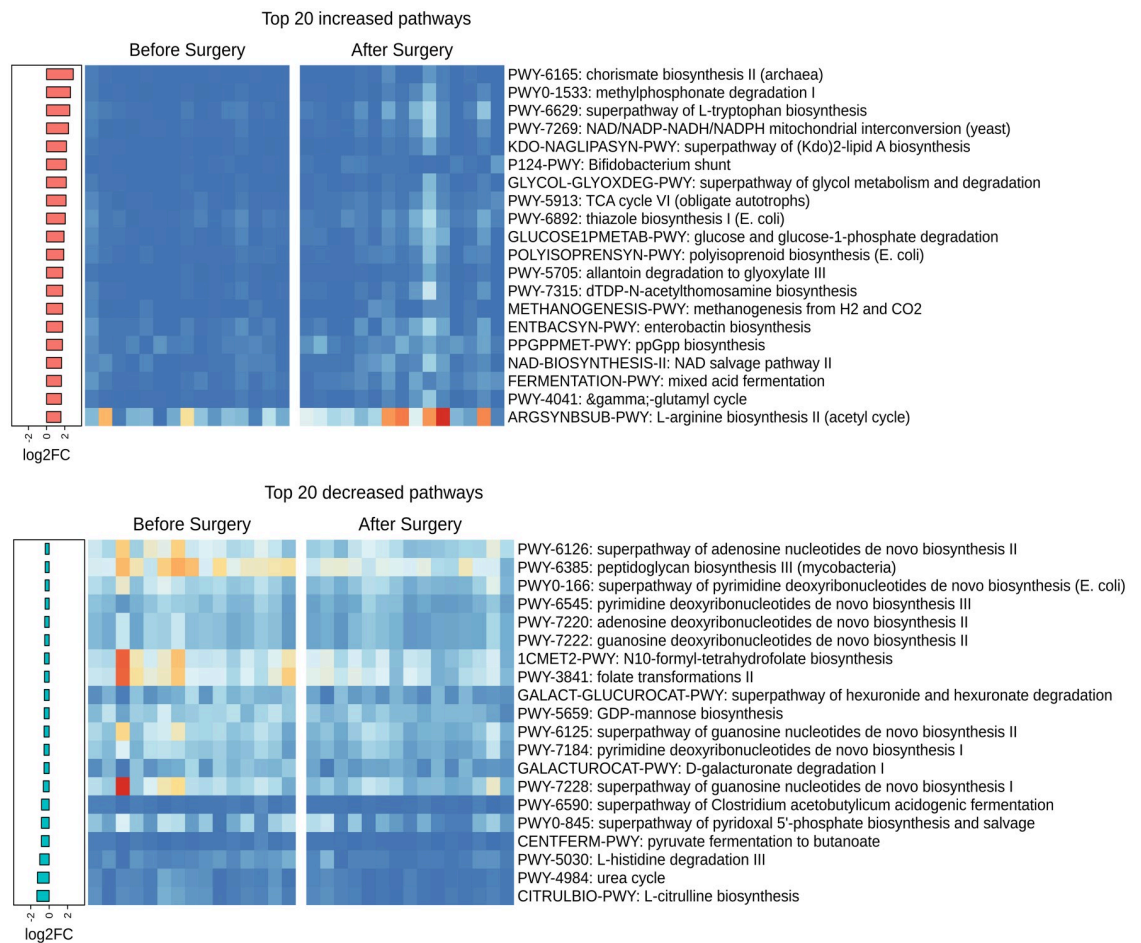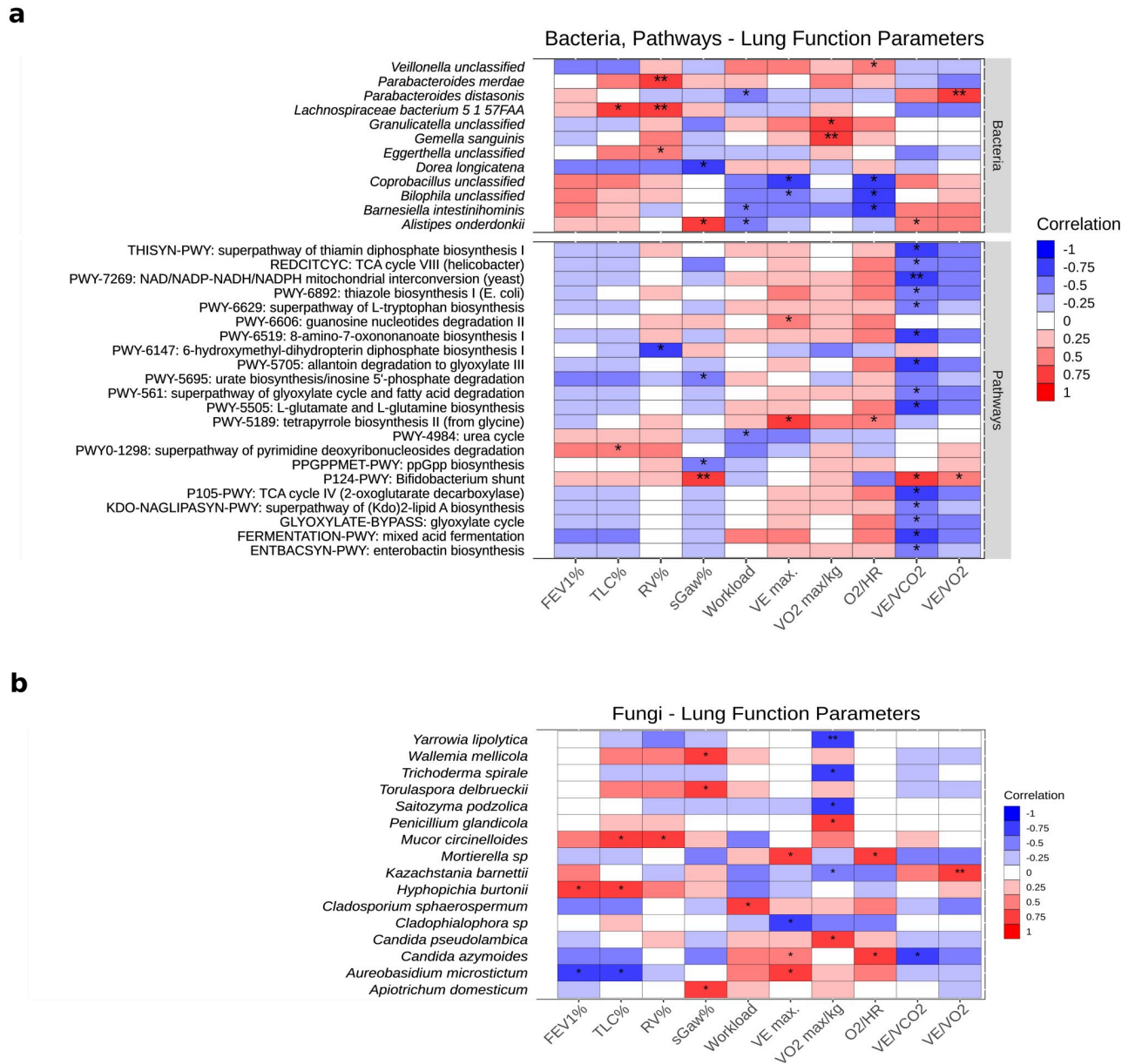
Fig 3. Functional analysis of the gut microbiome in response to surgical resection. Heatmaps of the top 20 increased (top) and decreased (bottom) differentially abundant bacterial MetaCyc pathways ($P < 0.05$, Wilcoxon signed-rank test). Histogram (left panel) shows the pathways ranked by variation (log2 fold-change after surgical resection).

post-surgery changes in the abundance of species with post-surgery changes in lung function parameters (FEV, FEV1% and FEV/FVC). We identified 10 bacterial species, including species correlated with CPET parameters, such as *P. distasonis* and *B. intestinihominis* significantly correlated with changes in lung function parameters (S3A Fig).

Of the significantly differentially abundant MetaCyc pathways, four were positively associated with at least one of TLC%, sGaw%, VE, O_2/HR, VE/VCO_2, and VE/VO_2, including pathways for generation of precursor metabolites and energy, nucleoside and nucleotide degradation, and tetrapyrrole biosynthesis. Eighteen significantly differentially abundant MetaCyc pathways showed a negative correlation with lung function parameters (RV%, sGaw %, workload, and VE/VCO_2), including pathways responsible for compound degradation, utilization, and assimilation (N = 3), cofactor, electron carrier, and vitamin biosynthesis (N = 4), amino acid biosynthesis (N = 2), compound biosynthesis (N = 4), and generation of precursor metabolites and energy (N = 5) (Fig 4A) (S4 Table). Post-surgery changes in the abundance of 4 pathways, related to amino acid biosynthesis, cofactor, electron carrier, and vitamin biosynthesis, and compound degradation, utilization and assimilation, were positively associated with post-surgery changes in lung function parameters (S3A Fig). In contrast, post-surgery

Fig 4. Correlations between taxonomic and functional profiles and CPET. (A) Heatmap of partial Spearman's rank correlation analysis between bacterial species and bacterial MetaCyc pathways versus lung function parameters adjusting for COPD and cancer type. Only differentially abundant species and pathways ($P<0.05$, Wilcoxon signed-rank test) were used. (B) Heatmap of partial Spearman's rank correlation analysis between fungal species versus lung function parameters adjusting for COPD and cancer type. (A, B) Cell colour indicates either negative correlation (blue) or positive correlation (red). Only species and pathways with significant correlations ((A) $P<0.05$; (B) $P<0.05$, absolute correlation coefficient > 0.65) are shown (*$P<0.05$, **$P<0.01$, ***$P<0.001$).

changes in the abundance of two pathways were negatively associated with post-surgery changes in lung function parameters (S3A Fig).

We compared the abundance of aerobic and anaerobic species before and after surgery and found a strong, significant decrease in the abundance of anaerobes after surgical resection (r = 68.9%, P = 0.0054, Wilcoxon signed-rank test) (S4A Fig). When stratifying for recurrent and non-recurrent patients separately, recurrent patients showed a very strong, significant

decrease in the abundance of anaerobes (r = 89.9%, P = 0.031, Wilcoxon signed-rank test) (S4B Fig). Non-recurrent showed only a trend (r = 54.5%, P = 0.15, Wilcoxon signed-rank test). Given the difference in effect, these findings imply a potential role of anaerobe bacteria in tumour suppression.

At last, we tried to predict VO_2, tumour recurrence, and overall survival (OS) from either (a) bacterial or (b) MetaCyc pathway relative abundances (S5 Fig, S5 Table). We only looked at features with significant change post-surgical resection. We selected features whose microbial balance [31] has a high linear correlation to target response. For VO_2, the balance of only six bacterial species positively correlates with a Pearson correlation coefficient of r = 0.78 (P = 6E-4). Here, the balance was termed to be the ratio of *A. equolifaciens* and *Alistipes* abundances (both increased significantly after surgery) to *P. distasonis*, *P. merdae* (both decreased significantly after surgery) and *G. sanguinis* abundances. The balance of 7 MetaCyc pathways correlates even better with r = 0.85 (P = 7E-5). Functions for carbohydrate utilization, fatty acid oxidation and arginine production were selected. Interesting, an increase in VO_2 coincides with an increase in the pathway "GABA shunt", in which 4-aminobutyrate (GABA) is produced from glutamate and further metabolized to succinate. Improvements in exercise tolerance and treatment outcome were associated with butyrate producing species. Succinate is involved in several metabolic processes. Importantly, it is involved in the control of reactive oxygen species and tumorigenesis [38]. In contrast, functions for metabolizing glucose and fucose may have a negative impact on VO_2. In contrast to previous findings [8], the F/B ratio was not correlated or predictive of VO_2. However, the ratio of Proteobacteria against Euryarchaeota and Actinobacteria correlated with r = 0.56 (p = 0.032), implying a potential link between members of these phyla and VO_2.

For prediction of tumour recurrence vs non-recurrence, patients can be accurately classified by a balance of three bacteria (CV-AUC = 95%, Acc: 100%). *O. splanchnicus* is hereby associated with tumour free outcome, whereas *G. sanguinis* and *Olsenella* were associated with tumour recurrence. A balance of two MetaCyc functions was also predictive of recurrence (CV-AUC: 90%, Acc: 100%). Hereby, L-alanine biosynthesis was associated with positive outcome. Arginine and polyamine biosynthesis were predicted to increase tumour recurrence.

For OS, a balance of 5 MetaCyc pathways correlated well (r = 87%, P = 2e-5). Interestingly, guanosine nucleotide de novo synthesis pathways were associated with improved OS, while adenosine and pyrimidine nucleotide de novo synthesis to decreased OS.

Associations of fungal species and lung function and CPET parameters

Besides the bacterial composition, we also attempted to associate the mycobiome with the lung function parameters. We built high-quality libraries for ITS2 sequencing for samples collected post-surgical resection (N = 15) and estimated the relative fungal abundance using the PIPITS pipeline [25]. In total, we identified 124 genera and 189 species. From these, 16 fungal species showed strong significant correlations ($P < 0.05$, absolute correlation coefficient > 0.65) with lung function and CPET parameters (FEV1%, TLC%, sGaw%, Workload, VE, VO_2, and O_2/HR) (Fig 4B). *Penicillium glandicola* and *Candida pseudolambica* were positively correlated with VO_2, whereas *Yarrowia lipolytica*, *Trichoderma spirale*, and *Saitozyma podzolica* showed negative correlations (S5 Table).

Discussion

CPET is a non-invasive method to test the overall condition of lung cancer patients who underwent lung resection surgery. CPET provides a comprehensive assessment of the exercise response and reflects the metabolic interactions of different organ systems. Shotgun

metagenomic sequencing allows taxonomic and functional annotation of the microbiome, thus it is the most comprehensive method for microbiome characterization.

Early metabolic changes might predict postoperative physical condition and outcome earlier than radiographic changes. The gut-lung axis plays a critical role in metabolic functional changes, and therefore, we hypothesized that the gut microbiome might be associated with different workloads and other CPET parameters.

Previous studies have shown that colorectal cancer patients have a disturbed gut microbial composition with a low abundance of species producing butyrate, such as species from the *Roseburia* and *Lachnospiraceae* genera, which may alter gene expression in healthy and cancerous cells [39]. In colorectal cancer cells, a dysfunction of the mitochondria contributes to an accumulation of butyrate in the cytosol and inhibition of histone deacetylases, resulting in the downregulation of proliferation and apoptosis pathways [39]. These changes may contribute to a decrease in tumour size and the probability of metastasis.

In our study, we observed post-surgical increases in (a) SCFA producing bacterial species and (b) pathways involved in carbohydrate, alcohol metabolism, and vitamin B production. These findings suggest a possible relationship between bacterial communities capable of metabolizing carbohydrates and alcohols into chemical compounds. The compounds involved in these pathways are considered important (a) to improve cell membrane barrier, (b) to act as precursor metabolites for human hormones such as serotonin and melatonin, and (c) for vitamin B_1 and B_3 production. Furthermore, we predicted a negative association between *O. splanchnicus* and tumor recurrence. *O. splanchnicus* is a potent butyrate producing species [40] by fermentation of L-Lysing (BioCyc P163-PWY [41]) and integral part of a healthy gut microbiome [42]. To improve CPET VO_2 and supress tumour recurrence, enriching the host microbiome with butyrate producing species and decreasing monosaccharide metabolizing species may improve treatment outcome.

From our study, we conclude that *A. equolifaciens*, *Alistipes*, *L. bacterium 5 1 57FAA* and *P. merdae* may be beneficial species for improved recovery of overall physical condition and lung capacity. These species were positively correlated with several lung function parameters or predictive of VO_2. In contrast, *G. sanguinis* was associated with decreased VO_2 and increased tumor recurrence. *D. longicatena* was associated with a decrease in lung function parameters, suggesting a negative impact of these two bacterial species on the recovery of lung capacity. Both *L. bacterium 5 1 57FAA* and *P. merdae* were negatively correlated with *D. longicatena*, implying that these two species may compete with *D. longicatena* for gut colonization and could therefore improve health post-surgery. While *O. splanchnicus* was not directly associated to improved lung function, a potential effect on tumour suppression should be worth further investigation. We observed a significant decrease in relative abundance of this bacterial species after resection. Like other SCFA producers, high fibber diets are required for the fermentation, proliferation and hence SCFA production. In the future, it should be worth to investigate if SCFA and other metabolites produced by *O. splanchnicus* are taken up by their human host and how these may influence tumour tissue.

Our study's limitation includes that in the real-life setting even though initial case numbers of recruited patients were high, the final numbers of patients with overall good clinical condition and available samples were low. To our best knowledge, the sample size is comparable to other studies of comparably unique conditions. Accordingly, this study did not allow us to draw robust conclusions on outcome-related biomarkers; however, it enabled us to meet our primary aim to study metabolic interactions. Another limitation is that the causality of metabolic and microbiome changes is not clear and follow-up studies with larger cohorts are needed.

We conclude that there are bacterial metabolic pathways that might be associated to increased oxygen uptake and exercise tolerance in patients one year after lung resection surgery. To our knowledge, this is the first study on microbiome functionality correlations with CPET, a highly specialized stress test with parameters that provide a unique detailed opportunity to study not only taxa but also metabolic interactions. Also, restoring specific bacteria might provide future therapeutic targets. Additionally, our findings highlight the potential of examining dynamic function parameters compared with traditional static metrics such as basic lung function and radiographic assessment in order to assess different organ systems' metabolic interactions. In this unique setting, the gut microbiota provided useful information on associations of exercise tolerance. We hypothesize that outcomes after lung resection surgery might be associated with distinct metabolic pathways that need to be confirmed in larger datasets. We offer critical postoperative metabolic and microbiome taxa and functional changes that are associated with distinct patient physical conditions and hopefully provide a reasonable basis for future studies aiming to increase patient outcomes.

Future studies in similar unique datasets are needed to confirm our findings and the modulation of gut microbiota, including butyrate producing taxa to increase long-term benefit from lung resection surgery.

Supporting information

S1 Fig. Study design of surgically resected lung cancer patient cohort.
(PDF)

S2 Fig. Comparison of the gut microbiota composition pre- and post-surgery. (A, B) Principal Coordinate Analysis plot based on Bray-Curtis distances of (A) the different stages and (B) chemotherapy treatment. (C) Boxplots, with median (centrelines), first and third quartiles (box limits) and 1.5x interquartile range (whiskers), showing alpha diversity Shannon, Simpson, and Chao1 indices of recurrent and non-recurrent patients. (D) Principal Coordinate Analysis plot based on Bray-Curtis distances pre- and post-surgery of recurrent and non-recurrent patients.
(PDF)

S3 Fig. Correlations between taxonomic and functional profiles and CPET. (A) Heatmap of partial Spearman's rank correlation analysis between the fold-change of bacterial species and bacterial MetaCyc pathways versus the fold-change lung function parameters adjusting for COPD and cancer type. Only differentially abundant species and pathways ($P<0.05$, Wilcoxon signed-rank test) were used. (B) Heatmap of partial Spearman's rank correlation analysis between fungal species versus the fold-change of lung function parameters adjusting for COPD and cancer type. (A-B) Cell color indicates either negative correlation (blue) or positive correlation (red). Only species and pathways with significant correlations ($P<0.05$) are shown ($^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$).
(PDF)

S4 Fig. Prediction of VO2, tumour recurrence, and overall survival (OS). Prediction of VO_2, tumour recurrence, and overall survival (OS) from bacterial species (left), bacterial phyla (middle) or MetaCyc pathways (right) relative abundances.
(PDF)

S5 Fig. Abundance of aerobic and anaerobic species. (A, B) Boxplots, with median (centrerelines), first and third quartiles (box limits) and 1.5x interquartile range (whiskers), showing the abundance of aerobic and anaerobic species. Gray lines connect samples of the same

patient from before and after surgical resection.
(PDF)

S1 Table. Major patients' characteristics.
(XLSX)

S2 Table. Species Wilcoxon signed-rank test results.
(TSV)

S3 Table. Pathways Wilcoxon signed-rank test results.
(TSV)

S4 Table. Partial Spearman's correlation results of bacterial species and pathways with lung function and CPET parameters.
(TSV)

S5 Table. VO$_2$, recurrence and overall survival predictions.
(XLSX)

S6 Table. Partial Spearman's correlation results of fungal species with lung function and CPET parameters.
(TSV)

Author Contributions

Conceptualization: Glen J. Weiss, Gianni Panagiotou, Zoltan Lohinai.

Data curation: Judit Berta, Virag Hollosi, Balazs Dome, Edit Dulka, Gabriella Galffy.

Formal analysis: Andrea Marfil-Sánchez, Bastian Seelbinder, Yueqiong Ni.

Investigation: Andrea Marfil-Sánchez, Bastian Seelbinder, Janos Varga, Zsolt Megyesfalvi.

Resources: Judit Berta, Virag Hollosi, Balazs Dome, Edit Dulka, Gabriella Galffy.

Visualization: Andrea Marfil-Sánchez, Bastian Seelbinder.

Writing – original draft: Andrea Marfil-Sánchez, Bastian Seelbinder, Yueqiong Ni.

Writing – review & editing: Glen J. Weiss, Gianni Panagiotou, Zoltan Lohinai.

References

1. Zhuang H, Cheng L, Wang Y, Zhang Y-K, Zhao M-F, Liang G-D, et al. Dysbiosis of the Gut Microbiome in Lung Cancer. Front Cell Infect Microbiol. 2019; 9. https://doi.org/10.3389/fcimb.2019.00112 PMID: 31065547

2. Liu F, Li J, Guan Y, Lou Y, Chen H, Xu M, et al. Dysbiosis of the gut microbiome is associated with tumor biomarkers in lung cancer. Int J Biol Sci. 2019; 15: 2381–2392. https://doi.org/10.7150/ijbs.35980 PMID: 31595156

3. Zitvogel L, Galluzzi L, Viaud S, Vétizou M, Daillère R, Merad M, et al. Cancer and the gut microbiota: An unexpected link. Sci Transl Med. 2015; 7: 271ps1. https://doi.org/10.1126/scitranslmed.3010473 PMID: 25609166

4. Bingula R, Filaire M, Radosevic-robin N, Bey M, Berthon J, Bernalier-donadille A, et al. Desired Turbulence? Gut-Lung Axis, Immunity, and Lung Cancer. 2017; 2017. https://doi.org/10.1155/2017/5035371 PMID: 29075294

5. MATSUMOTO M, INOUE R, TSUKAHARA T, USHIDA K, CHIJI H, MATSUBARA N, et al. Voluntary Running Exercise Alters Microbiota Composition and Increases n-Butyrate Concentration in the Rat Cecum. Biosci Biotechnol Biochem. 2008; 72: 572–576. https://doi.org/10.1271/bbb.70474 PMID: 18256465

6. Peng L, Li Z-R, Green RS, Holzman IR, Lin J. Butyrate Enhances the Intestinal Barrier by Facilitating Tight Junction Assembly via Activation of AMP-Activated Protein Kinase in Caco-2 Cell Monolayers. J Nutr. 2009; 139: 1619–1625. https://doi.org/10.3945/jn.109.104638 PMID: 19625695

7. Zhang WQ, Zhao SK, Luo JW, Dong XP, Hao YT, Li H, et al. Alterations of fecal bacterial communities in patients with lung cancer. Am J Transl Res. 2018; 10: 3171–3185. PMID: 30416659

8. Durk RP, Castillo E, Márquez-Magaña L, Grosicki GJ, Bolter ND, Lee CM, et al. Gut Microbiota Composition Is Related to Cardiorespiratory Fitness in Healthy Young Adults. Int J Sport Nutr Exerc Metab. 2019; 29: 249–253. https://doi.org/10.1123/ijsnem.2018-0024 PMID: 29989465

9. Barton W, Penney NC, Cronin O, Garcia-Perez I, Molloy MG, Holmes E, et al. The microbiome of professional athletes differs from that of more sedentary subjects in composition and particularly at the functional metabolic level. Gut. 2017; gutjnl-2016-313627. https://doi.org/10.1136/gutjnl-2016-313627 PMID: 28360096

10. ALLEN JM, MAILING LJ, NIEMIRO GM, MOORE R, COOK MD, WHITE BA, et al. Exercise Alters Gut Microbiota Composition and Function in Lean and Obese Humans. Med Sci Sport Exerc. 2018; 50: 747–757. https://doi.org/10.1249/MSS.0000000000001495 PMID: 29166320

11. Munukka E, Ahtiainen JP, Puigbó P, Jalkanen S, Pahkala K, Keskitalo A, et al. Six-Week Endurance Exercise Alters Gut Metagenome That Is not Reflected in Systemic Metabolism in Over-weight Women. Front Microbiol. 2018; 9. https://doi.org/10.3389/fmicb.2018.02323 PMID: 30337914

12. Dokladny K, Moseley PL, Ma TY. Physiologically relevant increase in temperature causes an increase in intestinal epithelial tight junction permeability. Am J Physiol Liver Physiol. 2006; 290: G204–G212. https://doi.org/10.1152/ajpgi.00401.2005 PMID: 16407590

13. Lira FS, Rosa JC, Pimentel GD, Souza HA, Caperuto EC, Carnevali LC, et al. Endotoxin levels correlate positively with a sedentary lifestyle and negatively with highly trained subjects. Lipids Health Dis. 2010; 9: 82. https://doi.org/10.1186/1476-511X-9-82 PMID: 20684772

14. Fehrenbach E, Niess AM, Schlotz E, Passek F, Dickhuth HH, Northoff H. Transcriptional and translational regulation of heat shock proteins in leukocytes of endurance runners. J Appl Physiol. 2000; 89: 704–10. https://doi.org/10.1152/jappl.2000.89.2.704 PMID: 10926657

15. Wolin KY, Yan Y, Colditz GA, Lee I-M. Physical activity and colon cancer prevention: a meta-analysis. Br J Cancer. 2009; 100: 611–616. https://doi.org/10.1038/sj.bjc.6604917 PMID: 19209175

16. Bolliger CT, Jordan P, Solèr M, Stulz P, Grädel E, Skarvan K, et al. Exercise capacity as a predictor of postoperative complications in lung resection candidates. Am J Respir Crit Care Med. 1995; 151: 1472–1480. https://doi.org/10.1164/ajrccm.151.5.7735602 PMID: 7735602

17. Morice RC, Walsh GL, Ali MK, Roth JA. Redefining the lowest exercise peak oxygen consumption acceptable for lung resection of high risk patients. Chest. 1996; 110: 161S.

18. Mirsadraee S. The 7th lung cancer TNM classification and staging system: Review of the changes and implications. World J Radiol. 2012; 4: 0. https://doi.org/10.4329/wjr.v4.i4.128 PMID: 22590666

19. Seelbinder B, Chen J, Brunke S, Vazquez-Uribe R, Santhaman R, Meyer A-C, et al. Antibiotics create a shift from mutualism to competition in human gut communities with a longer-lasting impact on fungi than bacteria. Microbiome. 2020; 8: 133. https://doi.org/10.1186/s40168-020-00899-6 PMID: 32919472

20. Clarke EL, Taylor LJ, Zhao C, Connell A, Lee JJ, Fett B, et al. Sunbeam: An extensible pipeline for analyzing metagenomic sequencing experiments. Microbiome. 2019; 7: 1–13. https://doi.org/10.1186/s40168-018-0604-3 PMID: 30606251

21. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011; 17: 10. https://doi.org/10.14806/ej.17.1.200

22. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014; 30: 2114–2120. https://doi.org/10.1093/bioinformatics/btu170 PMID: 24695404

23. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25: 1754–60. https://doi.org/10.1093/bioinformatics/btp324 PMID: 19451168

24. Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, et al. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat Methods. 2015; 12: 902–903. https://doi.org/10.1038/nmeth.3589 PMID: 26418763

25. Gweon HS, Oliver A, Taylor J, Booth T, Gibbs M, Read DS, et al. PIPITS: an automated pipeline for analyses of fungal internal transcribed spacer sequences from the I llumina sequencing platform. Bunce M, editor. Methods Ecol Evol. 2015; 6: 973–980. https://doi.org/10.1111/2041-210X.12399 PMID: 27570615

26. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Appl Environ Microbiol. 2009; 75: 7537–7541. https://doi.org/10.1128/AEM.01541-09 PMID: 19801464

27. Franzosa EA, McIver LJ, Rahnavard G, Thompson LR, Schirmer M, Weingart G, et al. Species-level functional profiling of metagenomes and metatranscriptomes. Nat Methods. 2018; 15: 962–968. https://doi.org/10.1038/s41592-018-0176-y PMID: 30377376

28. Dixon P. VEGAN, a package of R functions for community ecology. J Veg Sci. 2003; 14: 927–930. https://doi.org/10.1111/j.1654-1103.2003.tb02228.x

29. Vavrek MJ. fossil: Palaeoecological and palaeogeographical analysis tools. Palaeontol Electron. 2011; 14.

30. Friedman J, Alm EJ. Inferring Correlation Networks from Genomic Survey Data. PLoS Comput Biol. 2012; 8: 1–11. https://doi.org/10.1371/journal.pcbi.1002462 PMID: 22629235

31. Rivera-Pinto J, Egozcue JJ, Pawlowsky-Glahn V, Paredes R, Noguera-Julian M, Calle ML. Balances: a New Perspective for Microbiome Analysis. mSystems. 2018; 3: 53–71. https://doi.org/10.1128/mSystems.00053-18 PMID: 30035234

32. Kim S. ppcor: An R Package for a Fast Calculation to Semi-partial Correlation Coefficients. Commun Stat Appl Methods. 2015; 22: 665–674. https://doi.org/10.5351/CSAM.2015.22.6.665 PMID: 26688802

33. Holt KE, Wertheim H, Zadoks RN, Baker S, Whitehouse CA, Dance D, et al. Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in Klebsiella pneumoniae, an urgent threat to public health. Proc Natl Acad Sci U S A. 2015; 112: E3574–81. https://doi.org/10.1073/pnas.1501049112 PMID: 26100894

34. Göker M, Gronow S, Zeytun A, Nolan M, Lucas S, Lapidus A, et al. Complete genome sequence of Odoribacter splanchnicus type strain (1651/6). Stand Genomic Sci. 2011; 4: 200–9. https://doi.org/10.4056/sigs.1714269 PMID: 21677857

35. Brook I, Wexler HM, Goldstein EJC. Antianaerobic antimicrobials: spectrum and susceptibility testing. Clin Microbiol Rev. 2013; 26: 526–46. https://doi.org/10.1128/CMR.00086-12 PMID: 23824372

36. Parker BJ, Wearsch PA, Veloo ACM, Rodriguez-Palacios A. The Genus Alistipes: Gut Bacteria With Emerging Implications to Inflammation, Cancer, and Mental Health. Front Immunol. 2020; 11: 906. https://doi.org/10.3389/fimmu.2020.00906 PMID: 32582143

37. Rios-Covian D, Salazar N, Gueimonde M, de Los Reyes-Gavilan CG. Shaping the Metabolism of Intestinal Bacteroides Population through Diet to Improve Human Health. Front Microbiol. 2017; 8: 376. https://doi.org/10.3389/fmicb.2017.00376 PMID: 28326076

38. Tretter L, Patocs A, Chinopoulos C. Succinate, an intermediate in metabolism, signal transduction, ROS, hypoxia, and tumorigenesis. Biochim Biophys Acta—Bioenerg. 2016; 1857: 1086–1101. https://doi.org/10.1016/j.bbabio.2016.03.012 PMID: 26971832

39. Wang T, Cai G, Qiu Y, Fei N, Zhang M, Pang X, et al. Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. ISME J. 2012; 6: 320–329. https://doi.org/10.1038/ismej.2011.109 PMID: 21850056

40. Gomez-Arango LF, Barrett HL, McIntyre HD, Callaway LK, Morrison M, Dekker Nitert M. Increased Systolic and Diastolic Blood Pressure is Associated with Altered Gut Microbiota Composition and Butyrate Production in Early Pregnancy. Hypertension. 2016; 68: 974–981. https://doi.org/10.1161/HYPERTENSIONAHA.116.07910 PMID: 27528065

41. Karp PD, Billington R, Caspi R, Fulcher CA, Latendresse M, Kothari A, et al. The BioCyc collection of microbial genomes and metabolic pathways. Brief Bioinform. 2019; 20: 1085–1093. https://doi.org/10.1093/bib/bbx085 PMID: 29447345

42. Nagai F, Morotomi M, Watanabe Y, Sakon H, Tanaka R. Alistipes indistinctus sp. nov. and Odoribacter laneus sp. nov., common members of the human intestinal microbiota isolated from faeces. Int J Syst Evol Microbiol. 2010; 60: 1296–1302. https://doi.org/10.1099/ijs.0.014571-0 PMID: 19667375