JCI insight

Inflammation-type dysbiosis of the oral microbiome associates with the duration of COVID-19 symptoms and long COVID

John P. Haran,[1,2,3] Evan Bradley,[1,3] Abigail L. Zeamer,[2,3] Lindsey Cincotta,[1] Marie-Claire Salive,[1] Protiva Dutta,[1] Shafik Mutaawe,[1] Otuwe Anya,[1] Mario Meza-Segura,[2] Ann M. Moormann,[4] Doyle V. Ward,[2,3] Beth A. McCormick,[2,3] and Vanni Bucci[2,3]

[1]Department of Emergency Medicine, [2]Department of Microbiology and Physiological Systems, [3]Program in Microbiome Dynamics, and [4]Department of Medicine, University of Massachusetts Medical School, Worcester, Massachusetts, USA.

In the COVID-19 pandemic, caused by SARS-CoV-2, many individuals experience prolonged symptoms, termed long-lasting COVID-19 symptoms (long COVID). Long COVID is thought to be linked to immune dysregulation due to harmful inflammation, with the exact causes being unknown. Given the role of the microbiome in mediating inflammation, we aimed to examine the relationship between the oral microbiome and the duration of long COVID symptoms. Tongue swabs were collected from patients presenting with COVID-19 symptoms. Confirmed infections were followed until resolution of all symptoms. Bacterial composition was determined by metagenomic sequencing. We used random forest modeling to identify microbiota and clinical covariates that are associated with long COVID symptoms. Of the patients followed, 63% developed ongoing symptomatic COVID-19 and 37% went on to long COVID. Patients with prolonged symptoms had significantly higher abundances of microbiota that induced inflammation, such as members of the genera *Prevotella* and *Veillonella*, which, of note, are species that produce LPS. The oral microbiome of patients with long COVID was similar to that of patients with chronic fatigue syndrome. Altogether, our findings suggest an association with the oral microbiome and long COVID, revealing the possibility that dysfunction of the oral microbiome may have contributed to this draining disease.

Introduction

The oral cavity holds the second largest microbial community in the human body, after the gut, with over 1000 species of commensal bacteria residing therein (1). Dysbiosis or disrupted homeostasis caused by an imbalance in the microflora in the oral cavity has been linked to many other systemic inflammatory or infectious diseases (2). There is mounting evidence that links oral bacterial species to systemic diseases including pneumonia (1, 3, 4). Bacteria in the oral cavity may promote respiratory infections either directly via aspiration or indirectly by enzyme production that may hinder pathogen clearance, promote lung colonization, or alter respiratory epithelial immune responses (5).

SARS-CoV-2 is responsible for the current COVID-19 pandemic. This pandemic began in early 2020 and has caused over half a million deaths in the United States alone (6). Building upon the body of evidence that the microbiome plays a role in the regulation of innate and adaptive immunity to viral infections (7, 8), studies done early in the pandemic have demonstrated a connection between an altered gut microbiome and the severity of COVID-19 (9, 10). Additionally, among patients with COVID-19 there has been a large number of coinfection cases with organisms that originate from the oral cavity (11). Recently, decreased oral microbiome diversity and increased dysbiotic species abundances have been identified as predictive of COVID-19 (12). This has raised the possibility of using the oral microbiome to diagnose SARS-CoV-2 infection; however, studies linking the observed dysbiotic oral microbiota to disease outcomes have been lacking. Also lacking is evidence that this COVID-related microbiome, which occurs early in the disease process, is predictive of key outcomes such as symptom duration.

Most hospitalized patients have persistent, long-lasting symptoms that can take weeks to resolve (13) and negatively affect health-related quality of life (14). Symptoms persisting greater than 4 weeks after an acute infection are called ongoing symptomatic COVID-19, as characterized by The British National Institute for

Health and Care Excellence (15). Symptoms lasting even longer, 8–12 weeks or greater (16), and symptoms characterized by fatigue, headache, dyspnea, and anosmia (17, 18) are termed long-lasting COVID-19 symptoms (long COVID). Long COVID does not currently have a strict definition (19). At the 10-week mark after SARS-CoV-2 infection, more than 50% of patients with long COVID suffer profound fatigue (20). Increasing age, body mass index, and female gender are known to associate with long COVID (16). It is currently unknown why most people recover fully within 2 to 3 weeks and others experience symptoms for weeks or months longer (21). There is evidence, however, of persistently perturbed inflammatory pathways long after the acute SARS-CoV-2 infection has subsided (22).

Given the emerging associations between the human microbiome and SARS-CoV-2 infection and the unknown driver for patients with COVID-19 suffering from long-lasting symptoms, we sought to explore if oral microbiome dysbiosis associates with ongoing symptoms among patients with COVID-19 after hospitalization. Accordingly, we enrolled a cohort of patients with COVID-19 who tested positive for SARS-CoV-2 infection by PCR from an emergency department in the United States, collected oral swabs early in the disease course, and followed them for 4-week and 10-week symptom resolution outcomes. We analyzed oral microbiome composition by shotgun metagenomic sequencing. Our findings uniquely describe how dysbiosis of the oral microbiome may have played a pivotal role in lengthening symptom duration, leading to the long COVID syndrome.

Results

Patient population. From a prospective sampling of 164 patients presenting with COVID-19 symptoms over a 9-month period, 84 (51.2%) tested positive by PCR for SARS-CoV-2. Of these patients, 27 were successfully contacted for follow-up at both 4 weeks and 10 weeks (Figure 1). The average age was 62.6 (SD 12.5) with 70.4% men, 66.7% White, 7.4% African American, and 25.9% Hispanic. Among the cohort for high-risk medical comorbidities, 16 (59.3%) had hypertension, 8 (29.6%) diabetes, and 5 (18.5%) chronic obstructive pulmonary disease. Neither the medical comorbidities nor the patients' Charlson Comorbidity Index (CCI) scores differed by symptom duration outcome (Table 1). None of these patients lived in the same household. All of these patients were admitted to the hospital, with 4 (14.8%) admitted to the ICU. The average hospital length of stay was 8.3 days (SD 7.7), with 85.2% requiring oxygen and 25.9% getting advanced oxygen delivery by high flow or positive airway pressure. Two patients were intubated with an endotracheal tube.

Symptom duration. The average length of symptom duration was 45.8 days (SD 30.4), with 14 patients (51.9%) experiencing continuation of symptoms after 4 weeks from disease onset, and 10 patients (37.0%) experiencing symptoms longer than 10 weeks. The symptoms that lasted the longest were respiratory in nature (81.5% cough or short of breath) followed by fatigue (55.6%), gastrointestinal symptoms (14.8%), confusion or "brain fog" (22.2%), and ageusia or anosmia (14.8%). Brain fog is a symptom more recently linked to long COVID and characterized by a lack of clear memory or an ability to focus (23, 24). There were no significant differences in demographics, medical history, or hospital treatments among the 2 outcome categories (Table 1). However, among patients with symptoms lasting longer than 10 weeks, fatigue and brain fog were the most prominent symptoms that lasted the longest duration.

Oral microbiome composition predicted ongoing symptomatic COVID-19. We set out to explore the associations of oral microbiome composition with the symptoms of ongoing symptomatic COVID-19. To do this we profiled the oral microbiome of subjects with acute COVID-19 infection using shotgun metagenomic sequencing (see Methods). Microbial species abundances were determined by running Metaphlan3 (25). We estimated microbiome α diversity by calculating Shannon diversity index (26). We started by applying unsupervised learning methods, such as principal coordinate analysis (PCoA) and t-distributed stochastic neighbor embedding (t-SNE) and, as expected, found that interindividual variability overwhelmingly accounted for the majority of the information in the data (Supplemental Figure 1; supplemental material available online with this article; https://doi.org/10.1172/jci.insight.152346DS1). PERMANOVA analysis on samples classified according to COVID-19 symptoms duration was not statistically significant ($P < 0.05$). We then applied random forest classification (RFC; refs. 27, 28) to identify microbiome and clinical features associated with ongoing disease. Feature selection was performed using the Boruta algorithm on 5-fold cross-validated data and then running RFC using the union of the selected Boruta features on the same 5-fold cross-validated data to estimate model performance (29). We compared classification accuracy for different models that were trained only on demographics and clinical data; only on microbiome species abundances; only on Shannon diversity; on demographics, clinical data, and Shannon diversity; on demographics, clinical data, microbiome species, and Shannon diversity; and on clinical data, microbiome species, and Shannon diversity (Figure 2A).
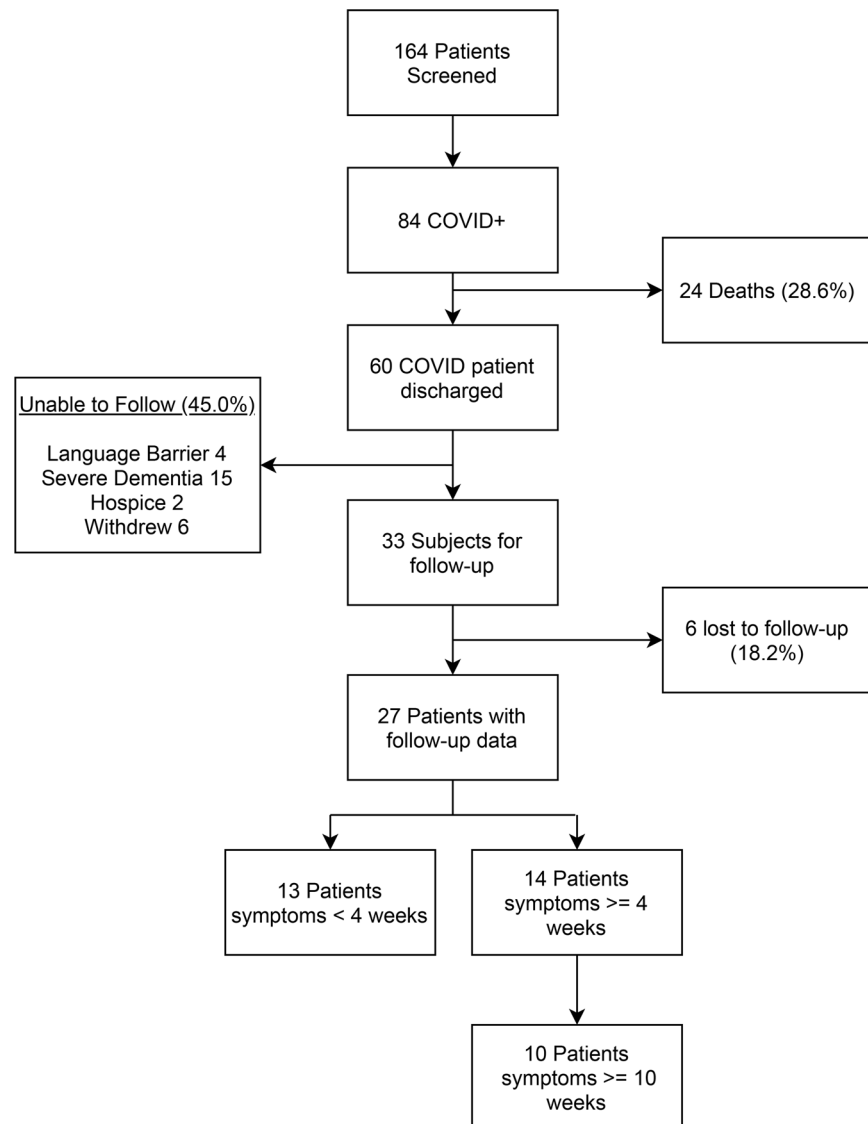
Figure 1. Study enrollment flow chart.

Each model was run starting from 10 different random seeds to calculate appropriate performance statistics. The mean F1 score, the harmonic mean of precision and recall, was used to select the top-performing model for a given outcome. The best model — clinical data and microbiome species and Shannon diversity — performed with a mean F1 score of 0.751 (Figure 2A).

Specific microbial members had the greatest contribution to correctly classifying samples. We detected both bacterial and eukaryotic organisms in the oral microbiome analysis, with only bacteria demonstrating associations with the outcomes. We examined the 19 bacterial species, whose abundances were associated with ongoing symptomatic COVID-19, and 2 clinical covariates based on their median RFC-estimated permutated importance score over the 10 RFC pipeline iterations (Figure 2, B and C). The findings from the model indicate that both viral load and Shannon diversity were of moderate importance, whereas specific microbiome members contributed the most to correct sample prediction. In particular, 2 of the 3 top predictors (*Veillonella dispar* and *Veillonella infantium*) as well as 2 other species associated with ongoing symptomatic COVID-19 belong to the genus *Veillonella*. Members of this genus are gram-negative anaerobic coccus that can cause infection in humans (30). Specifically, *V. infantium* has been found in the bronchoalveolar lavage fluid of the patients with COVID-19, suggesting it is a significant coinfectious agent (31). Other pathobionts (organisms that can coexist or cause disease under certain circumstances), such as *Solobacterium moorei* (32, 33), *Streptococcus infantis* (34), and *Rothia dentocariosa* (35), were in higher abundances in patients with ongoing

Table 1. Demographics, hospital treatments, and symptoms by outcome category

Patient characteristics[A]	Early symptom resolution (n = 13)	Ongoing symptomatic COVID-19 (n = 4)	Long COVID (n = 10)	P
Demographics and medical				
Age (mean [SD]) (yr)	62.3 (14.3)	63.8 (13.5)	62.5 (10.9)	0.98
Male, n (%)	11 (84.6)	3 (75.0)	5 (50.0)	0.19
Female, n (%)	2 (15.4)	1 (25.0)	5 (50.0)	0.19
White, n (%)	9 (69.2)	2 (50.0)	7 (70.0)	0.75
African American, n (%)	1 (7.7)	1 (25.0)	0 (0.0)	0.27
Hispanic, n (%)	3 (23.1)	2 (25.0)	3 (30.0)	0.93
Smoker, n (%)	4 (30.8)	2 (50.0)	3 (30.0)	0.75
CCI (mean [SD])	4.1 (3.1)	1.75 (1.5)	3.2 (2.2)	0.31
Hypertension, n (%)	9 (69.2)	1 (25.0)	6 (60.0)	0.29
Diabetes, n (%)	6 (46.2)	0 (0.0)	2 (20.0)	0.15
Chronic obstructive lung disease, n (%)	1 (7.7)	1 (25.0)	3 (30.0)	0.37
BMI (mean [SD])	30.2 (6.4)	39.3 (5.3)	31.5 (4.8)	0.77
ICU admission, n (%)	2 (15.4)	1 (25.0)	1 (10.0)	0.77
Remdesivir, n (%)	5 (38.5)	4 (100.0)	6 (60.0)	0.09
Clinical trial, n (%)	4 (30.8)	1 (25.0)	1 (10.0)	0.49
Longest lasting symptoms				
Fatigue, n (%)	6 (46.2)	1 (25.0)	8 (80.0)	0.11
Respiratory symptoms, n (%)	10 (76.9)	3 (75.0)	9 (90.0)	0.68
GI symptoms, n (%)	3 (23.1)	0 (0.0)	1 (10.0)	0.45
Fever, n (%)	2 (15.4)	0 (0.0)	0 (0.0)	0.31
Ageusia/anosmia, n (%)	3 (23.1)	0 (0.0)	1 (10.0)	0.45
Confusion/"brain fog," n (%)	0 (0.0)	1 (25.0)	5 (50.0)	0.017
Duration of symptoms days (mean [SD])	18.8 (11.5)	47.8 (5.4)	80.1 (10.7)	<0.001

χ^2 Test was used to compare categoric variables and analysis of variance for continuous variables. Advanced O_2, if patients received oxygen beyond nasal cannula (i.e., high flow, continuous positive airway pressure); clinical trial, if patient received therapy as part of a clinical trial; GI, gastrointestinal; CCI, Charlson Comorbidity Index. [A]Data are presented as the number (%), unless otherwise specified.

symptomatic COVID-19. Interestingly, *S. infantis* has been found to be enriched in fecal samples from patients with COVID-19 (9) and *R. dentocariosa* was predictive of SARS-CoV-2 presence in hospital rooms (36).

In addition to being implicated in coinfection, the *Veillonella* species is also known to produce a large amount of LPSs (37). Another pattern that emerges from these data is that the higher abundances of other LPS-producing species are predictive of ongoing symptomatic COVID-19. Five members of the *Prevotella* genus are positively associated with ongoing symptomatic COVID-19 in our analysis. *Prevotella* exhibits increased inflammatory properties (38) and has been thought to be a clinically important pathobiont involved in promoting chronic inflammation (39, 40). Other proinflammatory species such as *Leptotrichia wadei* (12) are also in higher abundances in patients with a longer symptom duration.

Dysbiotic inflammatory-type oral microbiome associated with the development of long COVID syndrome. We repeated our machine learning–based analysis described above to predict long COVID outcome from microbial abundance and clinical covariates. RFC was not able to capture any signal in the data for models that lacked microbiome information (i.e., only on demographics and clinical data; only on Shannon diversity; and on demographics, clinical data, and Shannon diversity; Figure 2A). The top-performing RFC for long COVID was the one trained on clinical data and microbiome species, resulting in an F1 score on 0.615 (Figure 3A). From the modeling, we identified 29 different bacterial species whose abundances were associated with long COVID (Figure 3, B and C). Similar to ongoing symptomatic COVID-19, multiple *Veillonella* species were associated with long COVID. Several of the top-predicting species (4 out of 29) belong to the genus *Actinomyces*. *Actinomyces* cause actinomycosis, a rare infectious disease in which bacteria can spread to the respiratory tract causing inflammation (41). As with ongoing symptomatic COVID-19, multiple *Prevotella* species (38) are associated with long COVID. *Prevotella* species are overrepresented in patients with COVID-19 and are thought to produce proteins that can promote SARS-CoV-2 infection and increase clinical severity of
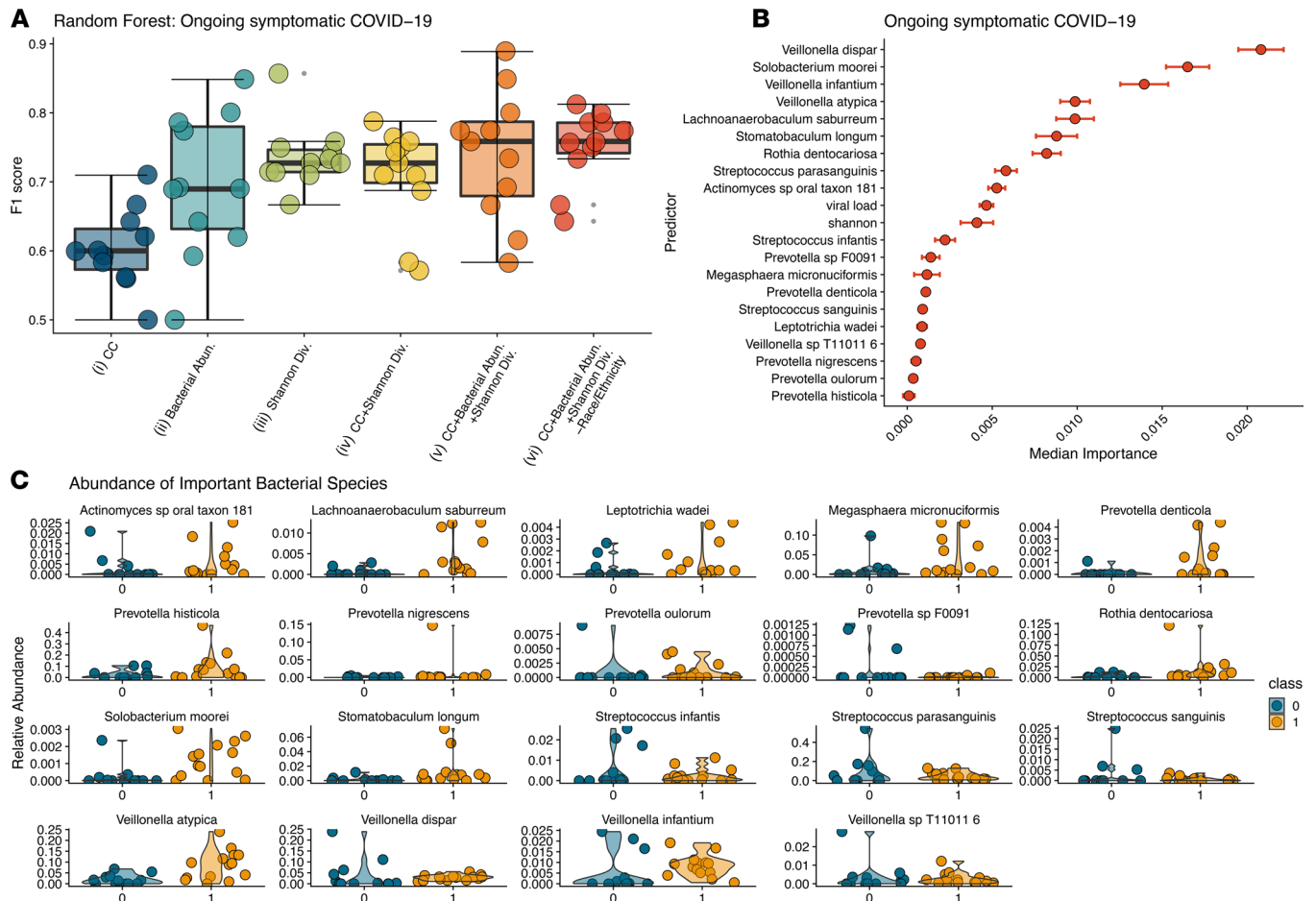
Figure 2. Bacterial abundances predict ongoing symptomatic COVID-19. RFC to identify predictors of ongoing symptomatic COVID-19 using 6 different combinations of data modalities. (**A**) F1 scores for the different RFC models trained on different sets of covariates. Box plot represents the median and IQR. (**B**) Ranking of forest predictors based on median permutated variable importance for the top-performing model. (**C**) Relative abundances for each bacteria found to be important in predicting ongoing symptomatic COVID-19 from the top-performing RFC model (on clinical data, microbiome species, and Shannon diversity). Violin plots showing the distribution of relative abundance for microbes in each patient with symptoms less than 4 weeks and 4 weeks or longer. 0 indicates "no" and 1 indicates "yes ongoing symptomatic COVID-19." CC, clinical covariates; Abn, abundances; Div, diversity; RFC, forest classification modeling.

COVID-19 (42). Additional species known to cause infections such as the *S. anginosus* group bacterial species, which has been reported to be particularly important in the pathogenesis of respiratory infections (43), and *Gemella sanguinis,* which has been shown to cause bloodstream infections in patients with COVID-19 (44) were also found to be associated with long COVID.

Inflammatory metabolic pathways associated with ongoing symptomatic COVID-19 and long COVID states. Building upon the taxonomy analysis, we explored the metabolic pathways and their association with ongoing symptomatic COVID-19 and long COVID states using HUMAnN3 (45). For each outcome we again performed RFC analysis and compared classification accuracy for different trained models: demographics, clinical data, and relative pathway abundances and only relative pathway abundances. For both patients with ongoing symptomatic COVID-19 and patients with long COVID, the top-performing model was (ii) only relative pathway abundances, producing an F1 score of 0.814 and 0.689, respectively (Figure 4A and Figure 5A). We identified greater than 40 metabolic gene pathways whose abundances were associated with both ongoing symptomatic and long COVID (Figure 4B and Figure 5B). The top 15 predictors indicate a striking proinflammatory pattern.

For ongoing symptomatic COVID-19, there are 5 pathways involved in the biosynthesis of branched-chain amino acids that are reduced in patients with longer symptoms (Figure 4, B and C). These include the superpathway of L-isoleucine I (MetaCyc PWY-3001), L-isoleucine biosynthesis III (PWY-5103), superpathway of branched amino acids (BRANCHED-CHAIN-AA-SYN-PWY), L-valine (VALSYN-PWY), and L-isoleucine (ILEUSYN-PWY) biosynthesis pathways (ref. 46; Figure 4C). Branched amino acids have been
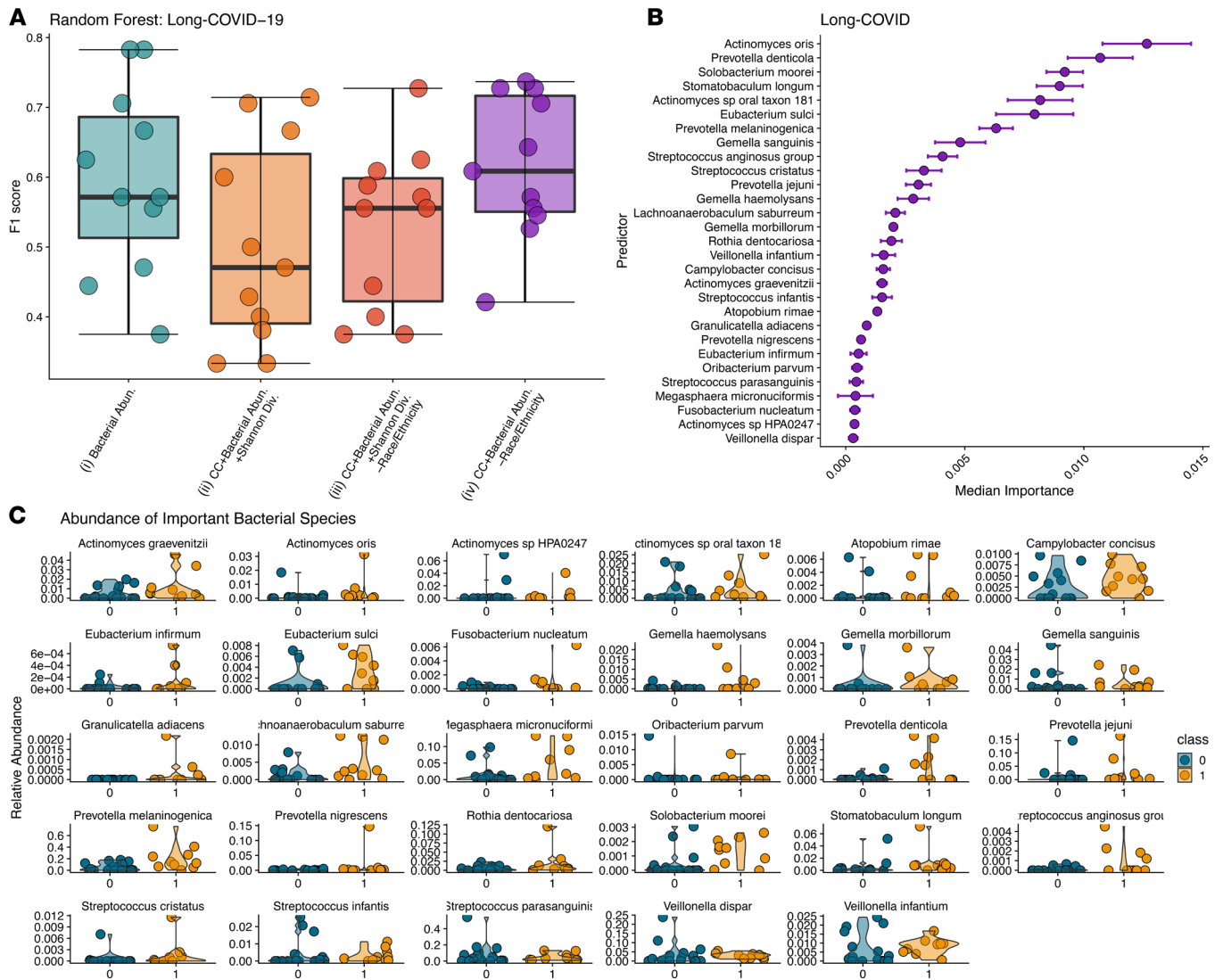
Figure 3. Bacterial abundances can predict long COVID. RFC modeling to predict long COVID. (**A**) F1 scores for all subsets of trainable RFC models. (**B**) Ranking of top 29 predictors associated with long COVID based on median permutated variable importance from the top-performing model (on demographics, clinical data, and Shannon diversity). (**C**) Relative abundances for each bacteria identified by model (on demographics, clinical data, and Shannon diversity) as important for predicting long COVID are presented as violin plots. Long COVID is presented as orange plots. CC, clinical covariates; Abn., abundances; Div., diversity. RFC, forest classification modeling.

shown to act as antiinflammatory agents (47, 48) with orally administered L-isoleucine and L-leucine exhibiting antiinflammatory activities (49). Four of 15 of the top pathways involve synthesis of molecules with antiinflammatory effects and are lower in patients with ongoing symptomatic COVID-19. These include the top predictor, Polyisoprenoid (50), whose biosynthesis has also been identified as significantly decreased in inflammatory conditions such as Crohn's disease (51). Tetrapyrrole (52) and farnesol (53) also have antiinflammatory effects. Conversely, 3 pathways for biosynthesis of proinflammatory molecules are increased in patients with ongoing symptomatic COVID-19: dTDP-L-rhamnose (DTDPRHAMSYN-PWY; ref. 54), pyrimidine (PWY-6545; ref. 55), and purine (P164 PWY; ref. 56) deoxyribonucleotides. Finally, both O-antigen building block biosynthesis (OANTIGEN-PWY), an important step in the LPS biosynthetic pathway (57), and the superpathway of phospholipid biosynthesis (PHOSLIPSYN-PWY), important in LPS production (58, 59), are higher among patients with ongoing symptomatic COVID-19. Similar patterns emerge with the long COVID analysis, which share 6 predictors with the ongoing symptomatic COVID-19 analysis.

Proinflammatory molecule synthesis was higher among patients with long COVID relative to those without as well as reduced branch-chain amino acid and antiinflammatory molecule biosynthesis (Figure 5C). Additional proinflammatory molecule biosynthesis is noted, with chorismite (PWY-6163; ref. 60),
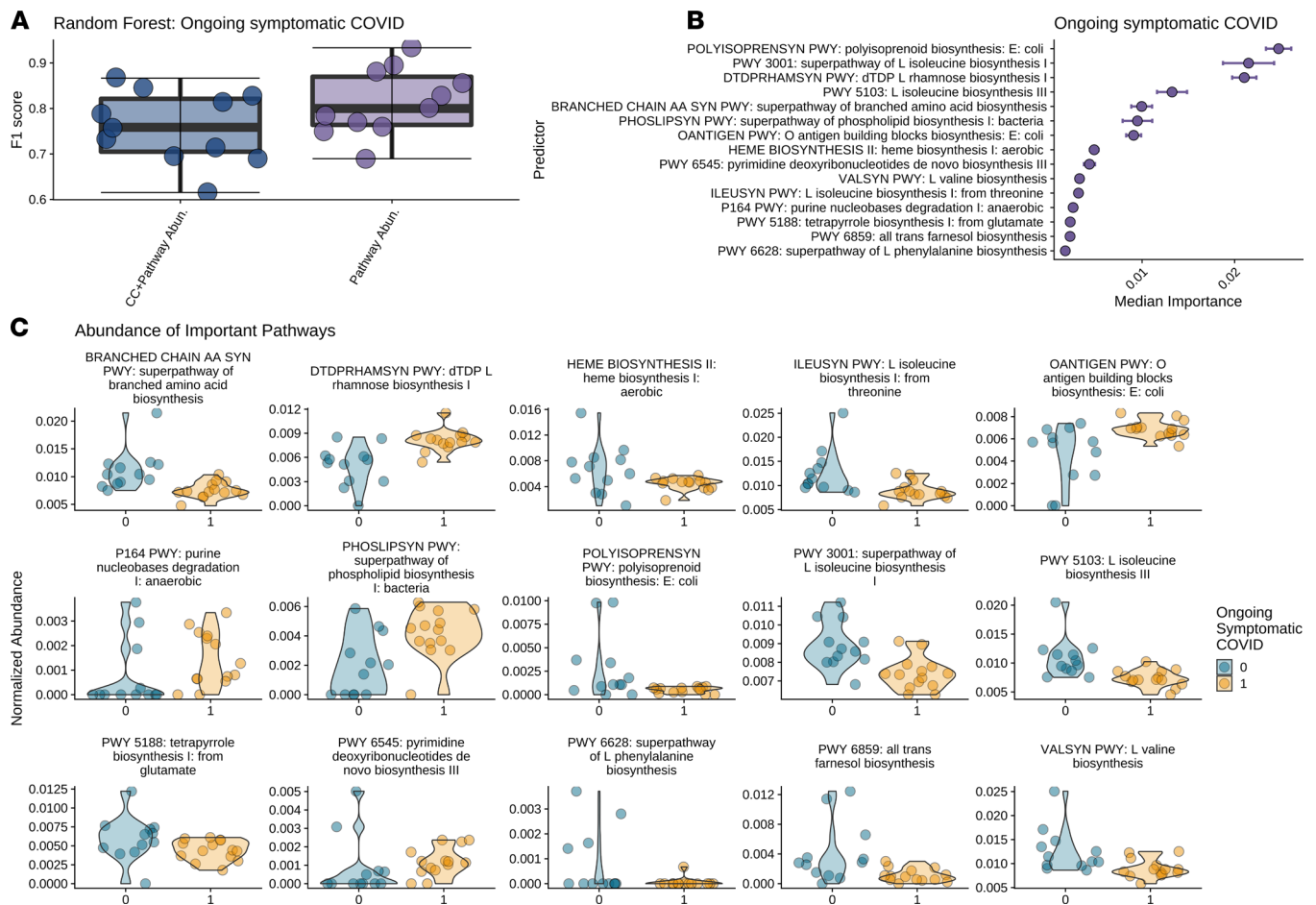
Figure 4. Bacterial metabolic pathways involving inflammation are significantly associated with ongoing symptomatic COVID-19. Results from RFC modeling to predict ongoing symptomatic COVID-19 and long COVID from HUMAnN3 pathway abundances. (**A**) F1 scores for demographics, clinical covariates, and pathway abundances and only on pathway abundances. (**B**) Ranking of forest predictors based on median permutated variable importance from the top-performing model, pathways only, for each outcome. (**C**) Relative pathway abundances for each pathway found to be important in predicting ongoing symptomatic COVID-19 and long COVID, respectively, by RFC modeling using only pathway abundances. We report violin plots showing the distribution of the relative abundance of pathways in patients with symptoms less than 4 weeks (blue) and 4 weeks or longer (yellow). RFC, forest classification modeling.

colanic acid (COLANSYN-PWY; ref. 61), and NAD biosynthesis (PWY-241; ref. 62) all being higher among the patients with long COVID.

Discussion

Many patients recovering from SARS-CoV-2 infection have symptoms that last long after the acute infection has run its course and our study highlights this same phenomenon. Over one third of our cohort had symptoms lasting longer than 10 weeks and thus entered the long COVID stage. Fatigue and "brain fog" were the longer lasting, most prominent symptoms among these patients. In an attempt to better understand both patients with ongoing symptomatic COVID-19 and patients with long COVID, we investigated potential clinical and microbiome associations with these disorders. Our modeling identified (a) microbial associations that are known to promote inflammation via LPS production or other mechanisms, (b) reduction of antiinflammatory metabolic pathways, (c) pathobionts known to cause pulmonary infections, and (d) microbiota previously shown to have associations with COVID-19. Thus, our work begins to shed light on the hypothesis that the oral microbiome composition may influence the duration of COVID-19 symptoms.

Patients with longer COVID-19 symptoms had dysbiotic, inflammatory-type oral microbiome. The oral microbiome has been shown to closely associate with SARS-CoV-2 coinfections in the lungs (11) and the oral–lung aspiration axis is a key factor leading to many respiratory infectious processes (63). We hypothesized that the oral microbiome might associate with the duration of postacute infection symptoms presented in ongoing
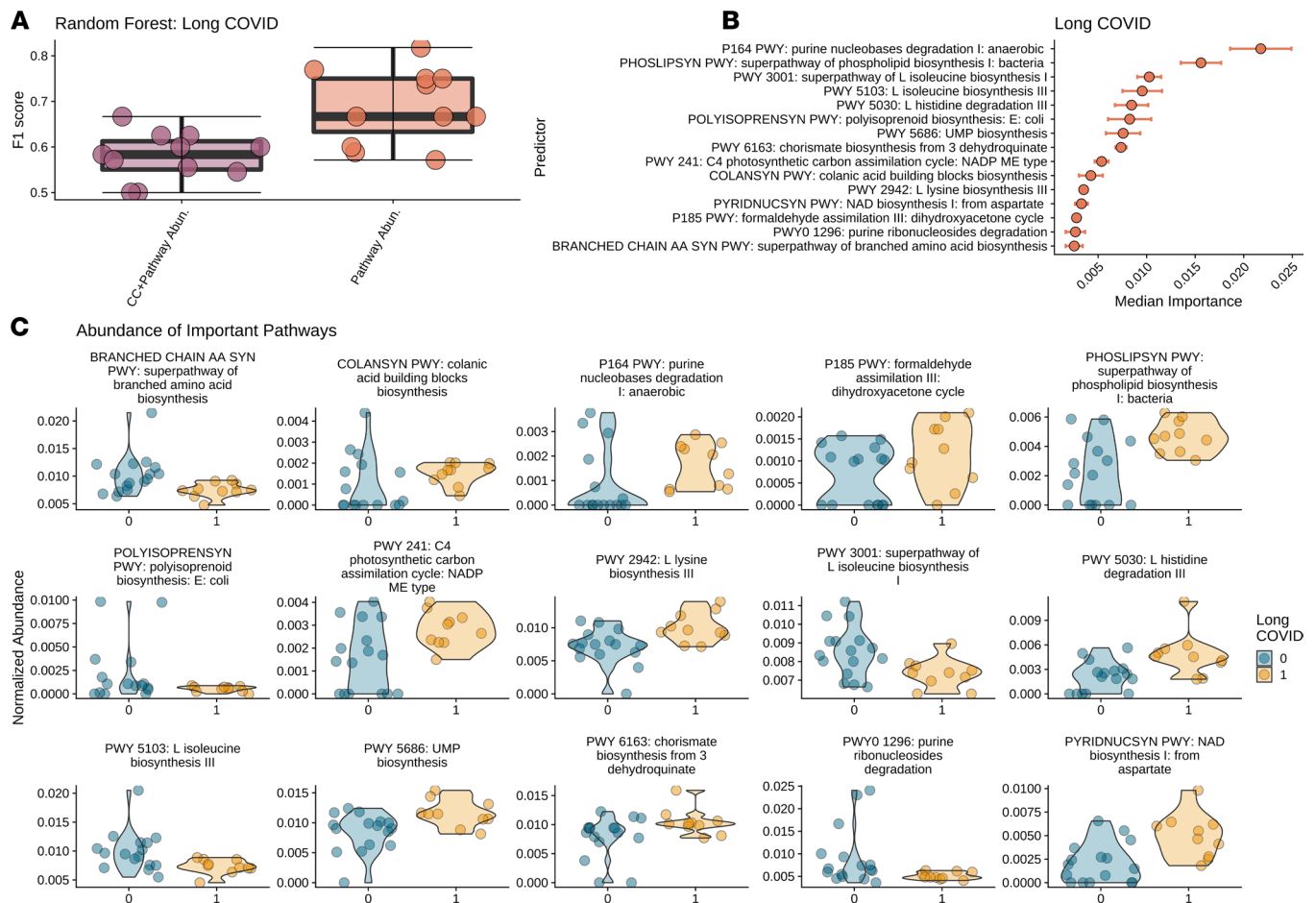
Figure 5. Bacterial metabolic pathways involving inflammation are significantly associated with long COVID. Results from RFC modeling to predict ongoing symptomatic COVID-19 and long COVID from HUMAnN3 pathway abundances. (**A**) F1 scores for (i) demographics, clinical covariates, and pathway abundances and only on pathway abundances. (**B**) Ranking of forest predictors based on median permutated variable importance from the top-performing model, pathways only, for each outcome. (**C**) Relative pathway abundances for each pathway found to be important in predicting long COVID, respectively, by RFC modeling using only pathway abundances. We report violin plots showing the distribution of the relative abundance of pathways in patients with symptoms less than 10 weeks (blue) and 10 weeks or longer (yellow). RFC, forest classification modeling.

symptomatic COVID-19 and long COVID states (64). Our findings extend previous work demonstrating how specific members of the genera *Prevotella* and *Veillonella* were distinctive in the oral microbiota of patients with COVID-19 (65). *Prevotella* species have been overrepresented in COVID-19 patient populations (42), whereas members of both the *Prevotella* and the *Veillonella* genera have been found in the bronchoalveolar lavage fluid of the patients with COVID-19 (31). Members of the *Prevotella* genus are thought to produce proteins that can promote SARS-CoV-2 infection and increase clinical severity of COVID-19 (42) and have previously been tied to systemic diseases, including low-grade systemic inflammation (38). The increased abundances of these 2 genera on the tongue have also been associated with an increased risk of death due to pneumonia in older, frail patients (66, 67). Finally, both genera induce inflammatory responses. *Veillonella* species have shown a strong capacity to induce IL-6 (68), whereas *Prevotella* strains primarily activate TLR-2 and enhance the expression of inflammatory cytokines, including IL-23 and IL-1 (69, 70). Other proinflammatory microbiota were identified in our analysis that also associated with longer disease symptoms such as *L. wadei* (12), *S. moorei* (71), and multiple *Actinomyces* species (41).

Metabolic pathways associated with the production of proinflammatory molecules were increased in abundance, whereas pathways associated with the production of antiinflammatory molecules were decreased in patients presenting with ongoing and long COVID symptoms. One of the top predictors, and thus demonstrating the strongest association in our data with both ongoing symptomatic COVID-19 and long COVID, was polyisoprenoid biosynthesis. Polyisoprenoid expresses antiinflammatory activity (50) and is significantly

decreased in inflammatory conditions such as Crohn's disease (51). Among the top predictors in our analysis was a reduced abundance of genes involved in the production of branched amino acids. Branched amino acids have long been shown to act as antiinflammatory agents (47, 48). Evidence is accumulating to support the hypothesis that systemic chronic inflammation contributes to the symptomatic progression to long COVID (22, 72). Given that changes in the microbiome composition can result in chronic inflammation and metabolic dysfunction (73), it is possible that the proinflammatory microbiome profiles we observe here could have played a pivotal role in this disease process.

LPS-producing bacteria may have promoted inflammation and have driven COVID-19 symptom duration. LPSs is an outer membrane component of gram-negative bacteria and can also be released in vesicles (74). Vesicle-associated LPS can have proinflammatory effects on host immune systems (75). Microbiome-derived LPS causes systemic inflammation (76, 77) and can even induce cognitive impairment and neuroinflammation (78, 79). Increases in LPS-producing bacteria, such as *Leptotrichia*, have been demonstrated in the oral cavity of patients with COVID-19 and are thought to be involved in the inflammatory response (12). Our analysis reveals higher abundances of many LPS-producing bacteria in patients with longer lasting symptoms. For example, *Veillonella* species, known to produce large amounts of LPSs (37), are present in increased abundances in our patients with COVID-19 with longer lasting symptoms. Increases in species such as *V. dispar*, *V. infantium*, and *V. atypica* are top predictors of ongoing symptomatic COVID-19, whereas *V. infantium* is found in higher abundances among patients with long COVID. Other LPS-producing species such as *L. wadei* (12) and *M. micronuciformis* (80) are also found to be in increased abundances. Additionally, our metabolic pathway analysis revealed an association with important steps in LPS biosynthesis and ongoing symptomatic COVID-19 and long COVID states. It is possible that LPS production may be a marker of other risk factors rather than a direct causal contributor. This would be critical to investigate in future work; however, this evidence points toward the important association of inflammation and long symptom disease states.

Myalgic encephalomyelitis/chronic fatigue syndrome linking to long-term COVID-19 symptoms through oral microbiome dysbiosis. There has been a growing concern that patients with COVID-19 with long-term sequelae resemble patients with myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS; ref. 81). These 2 conditions share some of the same symptoms, especially fatigue and cognitive impairment (17, 82). ME/CFS is a condition characterized by chronic fatigue, lasting at least 6 months, that impairs one's ability to perform daily activities and typically has additional impairments in memory and concentration (83). This syndrome is also closely linked to chronic inflammation as the driver of these patients' symptoms (84). The link to long-term symptoms is not unique to COVID-19 because patients with both SARS-CoV-1 and Middle East respiratory syndrome have also suffered from long-term sequelae in the previous epidemics (85).

ME/CFS has been hypothesized to be linked to infectious agents and microbiome dysbiosis has specifically been described in this syndrome through either the presence of pathobionts or microbial species that promote chronic inflammation (86). The gut microbiome has been shown to have reduced diversity and altered composition in patients with ME/CFS (87), and viral-induced microbiome changes are also thought to play a pivotal role (88). Clinical trials targeting the gut microbiome have shown promise in treating ME/CFS (89). Interestingly, patients with ME/CFS have been shown to have altered dysbiotic oral microbiomes characterized by increased abundances in the genera *Leptotrichia*, *Prevotella*, and *Fusobacterium* (90). Using whole genome sequencing, we have shown many species belonging to these genera are increased abundance in both patients with ongoing symptomatic COVID-19 and patients with long COVID. Specifically, top-predicting species *L. wadei*, *P. sp F0091*, *P. denticola*, *P. nigrescens*, *P. histicola*, and *P. oulorum* in the ongoing symptomatic COVID-19 group and *P. denticola*, P. *melaninogenica*, *P. jejuni*, *P. nigrescens*, and *F. nucleatum* in the long COVID group were all present in higher abundances in patients suffering from longer lasting symptoms. These finding add intriguing evidence of a possible link between patients with ME/CFS and patients with COVID-19 suffering from longer lasting symptoms related to inflammation in the oral microbiome.

Strengths and limitations. This study has several notable strengths and limitations. This study is limited in the number of patients enrolled and followed for symptom duration outcomes. A more robust cohort would allow deeper investigation of preexisting medical conditions and medications that might shape the oral microbiome composition. Larger cohorts would also include a more diverse patient set involving those treated as outpatients and more intensive care unit admissions. Generalization of our findings would need to be performed in a more diverse patient population. This limitation is balanced by our application of whole genome sequencing, which provides greater resolution than 16S rRNA gene sequencing used in many of the previous microbiome investigations (91). We also applied RFC which enable us to

include both clinical and microbiome data in our modeling (27, 28). This modeling approach has significant advantages compared with traditional classification techniques. Because it is agnostic to model structure (e.g., nonparametric regression), it does not need to meet common assumptions underlying classical regression techniques, and is able to intrinsically perform permutated ranked feature selection (29). We also have the advantage of collecting samples at the time of diagnosis before medical treatments that may alter the microbiome composition.

Conclusions. In conclusion, the oral microbiome of patients with prolonged symptoms falling under the ongoing symptomatic COVID-19 or long COVID states demonstrated a dysbiotic pattern of increased pathobionts, an increase in inflammation-inducing and LPS-producing microbiota, and a reduction of metabolic pathways known to have antiinflammatory properties. Although this work needs further validation, it supports the tenet that the microbiome may have played a role in prolonging symptom duration among COVID-19 through promotion of inflammation. The microbiome may therefore hold the key to better understanding the postinfection prolonged syndromes now facing patients after they recover from acute infection and provide a way to predict and subsequently act upon and prevent the development of long COVID.

Methods

Study setting and population. This prospective cohort consists of patients presenting to an emergency department located in central Massachusetts from April 2020 through February 2021. We enrolled patients who presented with symptoms consistent with a COVID-19 infection, but we analyzed only those who tested positive by PCR for SARS-CoV-2 and could contact for follow-up. We defined symptoms of COVID-19 based on the CDC guidelines (92).

Data collection. We collected baseline factors that included demographics, medical history, and presenting disease duration, and symptomatology. Comorbidity was assessed at baseline using the CCI, a widely used instrument designed to measure the burden of medical diseases and predict mortality (93). Patients were then followed through their hospital course for treatment types and length of stay. After discharge from the hospital, subsequent healthcare visits were recorded through the medical record. Patients were contacted by phone after 4 weeks of total symptoms after discharge and then again a second time if they were experiencing ongoing symptoms, after 10 weeks. Patients were categorized as symptoms lasting longer than 4 weeks and symptoms lasting longer than 10 weeks for analysis. Patients were also queried as to the type of symptoms that lasted the longest. Patients were excluded from follow-up if they died, were unable to communicate in English, had severe dementia, were in hospice, or withdrew themselves from the study.

Sample collection and processing. Oropharyngeal samples were collected using OMNIgene•ORAL collection kits (OMR-120, DNAgenotek). Briefly, the posterior oropharynx was swabbed for 30 seconds and then the swab was inserted into a tube with a DNA/RNA stabilization buffer. Samples were heated to 65°C–70°C for 1 hour to inactivate SARS-CoV-2 (94) and stored frozen. Nucleic acids were extracted by first thawing samples and then treating with 5ul Proteinase K (P8107S, New England Biolabs) for 2 hours at 50°C. DNA and RNA was then extracted using ZymoBIOMICS DNA/RNA Miniprep Kits (R2002, Zymo Research) as per the manufacture's protocol.

Sequence processing and analysis. Metagenomic DNA sequencing libraries were constructed using the Nextera XT DNA Library Prep Kit (FC-131-1096, Illumina) and sequenced on a NextSeq500 Sequencing System as 2 × 150 nucleotide paired-end reads. Shotgun metagenomic reads were first trimmed and quality filtered to remove sequencing adapters and host contamination using Trimmomatic (95) and Bowtie2 (96), respectively, as part of the KneadData pipeline (https://bitbucket.org/biobakery/kneaddata). As in our previous work (28, 97), metagenomic data were profiled for microbial taxonomic abundances and microbial metabolic pathways using Metaphlan3 (25) and HUMAnN3 (45), respectively. The total number of microbial and contaminant reads recovered as presented in Supplemental Table 1.

SARS-CoV-2 viral load quantification. PCR was performed using the ViiA 7 Real-Time PCR System (Applied Biosystems) and the GoTaq Probe 1-Step RT-qPCR System (Promega, A6120). The primer-probe set N1 (2019-nCoV_N1-F: 5′-GACCCCAAAATCAGCGAAAT-3′; 2019-nCoV_N1-R: 5′-TCTGGTTACTGC-CAGTTGAATCTG-3′; 2019-nCoV_N1-P: 5′-FAM-ACCCCGCATTACGTTTGGACC-BHQ1-3′) designed by the CDC were obtained from Integrated DNA Technologies (10006713) and used at concentrations of 500 nM and 125 nM, respectively (98). Eluted RNA (5 μl) were used to prepare 20 μl PCR reactions. Cycling conditions were as indicated by the CDC: 45°C for 15 minutes, 95°C for 2 minutes, followed by 45 cycles of 95°C for 3 seconds and 55°C for 30 seconds (98). Cycle threshold values were converted into viral RNA copies based

on a standard curve prepared from 4-fold serial dilutions of known quantities (1.0×10^6 to 2.44×10^2 viral copies) of a SARS-CoV-2_N positive control plasmid (10006625, Integrated DNA Technologies). The lower limit threshold for positive detection in our study was 244 viral copies per reaction. Viral load was calculated as the number of genome copies per milliliter of transport media to resuspend tongue swabs. The assay was run in triplicate for each sample and 3 nontemplate wells were included as negative controls.

Data availability. Data relating to the metagenomic sequencing that support the findings of this study have been uploaded to the NCBI BioProject (https://www.ncbi.nlm.nih.gov/bioproject/) and are available for download via accession PRJNA735193 under the title "Oral Microbiome associated with Coronavirus disease 2019 (COVID-19)."

Statistics. To determine similarity in oral microbiome samples among the patients with COVID-19 and to associate microbiome features to duration of symptom outcomes, we started by performing traditional unsupervised correspondence analysis (PCoA and t-SNE). Because most of the signal from the unsupervised analysis was accounted by interindividual variability, we then decided to run supervised machine learning models. We built a RFC pipeline to predict either ongoing symptomatic COVID-19 or long COVID from a given data subset. One sample failed the sequencing run, and thus 26 samples were included in our modeling. The first step of our pipeline used the feature selection algorithm Boruta on 5-fold cross-validated data to estimate model performance (29). The permutated variable importance from each RFC was also calculated. Each model was run starting from 10 different random seeds to calculate performance metrics. F1 score, the harmonic mean of precision and accuracy, was used to select the top-performing model for each outcome. *P* values of less than 0.05 were considered significant.

Study approval. This prospective cohort study was approved by the IRB at the University of Massachusetts Medical School. Written informed consent was received from all study participants prior to inclusion in the study.

Author contributions

JPH, BAM, AM, and EB conceived and led the study. JPH, EB, and PD supervised the conduct of the study and data collection. PD, MCS, SM, and OA managed the clinical data, including quality control. LC OA, and MMS handled the sample collection and storage. DVW managed sample extraction and sequencing and performed metagenomic profiling. ALZ and VB provided statistical advice on study design and analyzed the data. JPH and EB wrote the manuscript with input from all authors. JPH composed the first draft of the majority of the manuscript and was responsible for incorporation of all authors edits. Accordingly, JPH was assigned the first author slot.

Acknowledgments

Address correspondence to: John P. Haran, 55 Lake Avenue North, Worcester, Massachusetts 01655, USA. Phone: 508.450.8688; Email: john.haran@umassmed.edu.

1. Dewhirst FE, et al. The human oral microbiome. *J Bacteriol*. 2010;192(19):5002–5017.
2. Willis JR, Gabaldón T. The human oral microbiome in health and disease: from sequences to ecosystems. *Microorganisms*. 2020;8(2):308.
3. Seymour GJ, et al. Relationship between periodontal infections and systemic disease. *Clin Microbiol Infect*. 2007;13(suppl 4):3–10.
4. Awano S, et al. Oral health and mortality risk from pneumonia in the elderly. *J Dent Res*. 2008;87(4):334–339.
5. Scannapieco FA. Role of oral bacteria in respiratory infection. *J Periodontol*. 1999;70(7):793–802.
6. CDC. COVID Data Tracker. https://covid.cdc.gov/covid-data-tracker/#datatracker-home. Accessed September 2, 2021.
7. Trompette A, et al. Dietary fiber confers protection against flu by shaping Ly6c⁻ patrolling monocyte hematopoiesis and CD8⁺ T cell metabolism. *Immunity*. 2018;48(5):992–1005.
8. Belkaid Y, Harrison OJ. Homeostatic immunity and the microbiota. *Immunity*. 2017;46(4):562–576.

9. Zuo T, et al. Depicting SARS-CoV-2 faecal viral activity in association with gut microbiota composition in patients with COVID-19. *Gut*. 2021;70(2):276–284.

10. Yeoh YK, et al. Gut microbiota composition reflects disease severity and dysfunctional immune responses in patients with COVID-19. *Gut*. 2021;70(4):698–706.

11. Bao L, et al. Oral microbiome and SARS-CoV-2: beware of lung co-infection. *Front Microbiol*. 2020;11:1840.

12. Ren Z, et al. Alterations in the human oral and gut microbiomes and lipidomics in COVID-19. *Gut*. 2021;70(7):1253–1265.

13. Tenforde MW, et al. Symptom duration and risk factors for delayed return to usual health among outpatients with COVID-19 in a multistate health care systems network — United States, March-June 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69(30):993–998.

14. Garrigues E, et al. Post-discharge persistent symptoms and health-related quality of life after hospitalization for COVID-19. *J Infect*. 2020;81(6):e4–e6.

15. National Institute for Health and Care Excellence. COVID-19 rapid guideline: managing the long-term effects of COVID-19. https://www.nice.org.uk/guidance/ng188. Accessed September 2, 2021.

16. Sudre CH, et al. Attributes and predictors of long COVID. *Nat Med*. 2021;27(4):626–631.

17. Marshall M. The lasting misery of coronavirus long-haulers. *Nature*. 2020;585(7825):339–341.

18. Baig AM. Chronic COVID syndrome: need for an appropriate medical terminology for long-COVID and COVID long-haulers. *J Med Virol*. 2021;93(5):2555–2556.

19. Brodin P. Immune determinants of COVID-19 disease presentation and severity. *Nat Med*. 2021;27(1):28–33.

20. Townsend L, et al. Persistent fatigue following SARS-CoV-2 infection is common and independent of severity of initial infection. *PLoS One*. 2020;15(11):e0240784.

21. Greenhalgh T, et al. Management of post-acute covid-19 in primary care. *BMJ*. 2020;370:m3026.

22. Doykov I, et al. 'The long tail of Covid-19' — the detection of a prolonged inflammatory response after a SARS-CoV-2 infection in asymptomatic and mildly affected patients. *F1000Res*. 2020;9:1349.

23. Koralnik IJ, Tyler KL. COVID-19: a global threat to the nervous system. *Ann Neurol*. 2020;88(1):1–11.

24. Stefano GB. Historical insight into infections and disorders associated with neurological and psychiatric sequelae similar to long COVID. *Med Sci Monit*. 2021;27:e931447.

25. Beghini F, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *Elife*. 2021;10:e65088.

26. DeJong TM. A comparison of three diversity indices based on their components of richness and evenness. *Oikos*. 1975;26(2):222–227.

27. Hajjem A, et al. Mixed-effects random forest for clustered data. *J Stat Comput Simul*. 2014;84(6):1313–1328.

28. Haran JP, et al. Alzheimer's disease microbiome is associated with dysregulation of the anti-inflammatory P-glycoprotein pathway. *mBio*. 2019;10(3):e00632–19.

29. Wipperman MF, et al. Gastrointestinal microbiota composition predicts peripheral inflammatory state during treatment of human tuberculosis. *Nat Commun*. 2021;12(1):1141.

30. Cobo F, et al. Bacteremia caused by Veillonella dispar in an oncological patient. *Anaerobe*. 2020;66:102285.

31. Wu F, et al. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020;579(7798):265–269.

32. Liu WJ, et al. First case report of bacteremia caused by Solobacterium moorei in China, and literature review. *BMC Infect Dis*. 2019;19(1):730.

33. Pedersen RM, et al. Solobacterium moorei bacteremia: identification, antimicrobial susceptibility, and clinical characteristics. *J Clin Microbiol*. 2011;49(7):2766–2768.

34. Bek-Thomsen M, et al. Population diversity and dynamics of Streptococcus mitis, Streptococcus oralis, and Streptococcus infantis in the upper respiratory tracts of adults, determined by a nonculture strategy. *Infect Immun*. 2008;76(5):1889–1896.

35. Sadhu A, et al. Rothia dentocariosa endocarditis complicated by multiple cerebellar hemorrhages. *Diagn Microbiol Infect Dis*. 2005;53(3):239–240.

36. Marotz C, et al. Microbial context predicts SARS-CoV-2 prevalence in patients and the hospital built environment [preprint]. https://doi.org/10.1101/2020.11.19.20234229. Posted on medRxiv November 22, 2020.

37. Delwiche EA, et al. The veillonellae: gram-negative cocci with a unique physiology. *Annu Rev Microbiol*. 1985;39:175–193.

38. Larsen JM. The immune response to Prevotella bacteria in chronic inflammatory disease. *Immunology*. 2017;151(4):363–374.

39. Bellocchi C, Volkmann ER. Update on the gastrointestinal microbiome in systemic sclerosis. *Curr Rheumatol Rep*. 2018;20(8):49.

40. Kim D, Kim WU. Editorial: can Prevotella copri be a causative pathobiont in rheumatoid arthritis? *Arthritis Rheumatol*. 2016;68(11):2565–2567.

41. Valour F, et al. Actinomycosis: etiology, clinical features, diagnosis, treatment, and management. *Infect Drug Resist*. 2014;7:183–197.

42. Khan AA, Khan Z. COVID-2019-associated overexpressed Prevotella proteins mediated host-pathogen interactions and their role in coronavirus outbreak. *Bioinformatics*. 2020;36(13):4065–4069.

43. Noguchi S, et al. The clinical features of respiratory infections caused by the Streptococcus anginosus group. *BMC Pulm Med*. 2015;15:133.

44. Bonazzetti C, et al. Unexpectedly high frequency of enterococcal bloodstream infections in coronavirus disease 2019 patients admitted to an Italian ICU: an observational study. *Crit Care Med*. 2021;49(1):e31–e40.

45. Franzosa EA, et al. Species-level functional profiling of metagenomes and metatranscriptomes. *Nat Methods*. 2018;15(11):962–968.

46. Caspi R, et al. The MetaCyc database of metabolic pathways and enzymes — a 2019 update. *Nucleic Acids Res*. 2020;48(d1):D445–D453.

47. Lee JH, et al. Anti-inflammatory and anti-genotoxic activity of branched chain amino acids (BCAA) in lipopolysaccharide (LPS) stimulated RAW 264.7 macrophages. *Food Sci Biotechnol*. 2017;26(5):1371–1377.

48. Da Silva MS, et al. Whey protein hydrolysate and branched-chain amino acids downregulate inflammation-related genes in vascular endothelial cells. *Nutr Res*. 2017;38:43–51.

49. Saxena RN, et al. Anti-inflammatory and analgesic properties of four amino-acids. *Indian J Physiol Pharmacol*. 1984;28(4):299–305.

50. Pronin AV, et al. Plant polyisoprenoids and control of cholesterol level. *Arch Immunol Ther Exp (Warsz)*. 2014;62(1):31–39.

51. Klaassen MAY, et al. Anti-inflammatory Gut microbial pathways are decreased during Crohn's disease exacerbations. *J Crohns Colitis*. 2019;13(11):1439–1449.

52. Jelić D, et al. Porphyrins as new endogenous anti-inflammatory agents. *Eur J Pharmacol*. 2012;691(1-3):251–260.

53. Ku CM, Lin JY. Anti-inflammatory effects of 27 selected terpenoid compounds tested through modulating Th1/Th2 cytokine secretion profiles using murine primary splenocytes. *Food Chem*. 2013;141(2):1104–1113.

54. van der Beek SL, et al. Streptococcal dTDP-L-rhamnose biosynthesis enzymes: functional characterization and lead compound identification. *Mol Microbiol*. 2019;111(4):951–964.

55. Dimitrova P, et al. Restriction of de novo pyrimidine biosynthesis inhibits Th1 cell activation and promotes Th2 cell differentiation. *J Immunol*. 2002;169(6):3392–3399.

56. Barnes VM, et al. Acceleration of purine degradation by periodontal diseases. *J Dent Res*. 2009;88(9):851–855.

57. Han W, et al. Defining function of lipopolysaccharide O-antigen ligase WaaL using chemoenzymatically synthesized substrates. *J Biol Chem*. 2012;287(8):5357–5365.

58. Rothfield L, Pearlman M. The role of cell envelope phospholipid in the enzymatic synthesis of bacterial lipopolysaccharide. Structural requirements of the phospholipid molecule. *J Biol Chem*. 1966;241(6):1386–1392.

59. Emiola A, et al. Crosstalk between the lipopolysaccharide and phospholipid pathways during outer membrane biogenesis in Escherichia coli. *Proc Natl Acad Sci U S A*. 2016;113(11):3108–3113.

60. Rodrigues-Vendramini FAV, et al. Promising new antifungal treatment targeting chorismate synthase from Paracoccidioides brasiliensis. *Antimicrob Agents Chemother*. 2019;63(1):e01097–18.

61. Hanna A, et al. Role of capsular colanic acid in adhesion of uropathogenic Escherichia coli. *Appl Environ Microbiol*. 2003;69(8):4474–4481.

62. Gerner RR, et al. NAD metabolism fuels human and mouse intestinal inflammation. *Gut*. 2018;67(10):1813–1823.

63. Mammen MJ, et al. Oral-lung microbiome interactions in lung diseases. *Periodontol 2000*. 2020;83(1):234–241.

64. Dani M, et al. Autonomic dysfunction in 'long COVID': rationale, physiology and management strategies. *Clin Med (Lond)*. 2021;21(1):e63–e67.

65. Iebba V, et al. Profiling of oral microbiota and cytokines in COVID-19 patients. *Front Microbiol*. 2021;12:671813.

66. Asakawa M, et al. Tongue microbiota and oral health status in community-dwelling elderly adults. *mSphere*. 2018;3(4):e00332–18.

67. Kageyama S, et al. Relationships of variations in the tongue microbiota and pneumonia mortality in nursing home residents. *J Gerontol A Biol Sci Med Sci*. 2018;73(8):1097–1102.

68. van den Bogert B, et al. Immunomodulatory properties of Streptococcus and Veillonella isolates from the human small intestine microbiota. *PLoS One*. 2014;9(12):e114277.

69. Segal LN, et al. Enrichment of the lung microbiome with oral taxa is associated with lung inflammation of a Th17 phenotype. *Nat Microbiol*. 2016;1:16031.

70. Segal LN, et al. Enrichment of lung microbiome with supraglottic taxa is associated with increased pulmonary inflammation. *Microbiome*. 2013;1(1):19.

71. Hiranmayi KV, et al. Novel pathogens in periodontal microbiology. *J Pharm Bioallied Sci*. 2017;9(3):155–163.

72. Inciardi RM, et al. Coronavirus 2019 disease (COVID-19), systemic inflammation, and cardiovascular disease. *J Am Heart Assoc*. 2020;9(16):e017756.

73. Sommer F, Bäckhed F. The gut microbiota--masters of host development and physiology. *Nat Rev Microbiol*. 2013;11(4):227–238.

74. Farhana A, Khan YS, eds. *StatPearls*. StatPearls Publishing; 2021.

75. Bonnington KE, Kuehn MJ. Protein selection and export via outer membrane vesicles. *Biochim Biophys Acta*. 2014;1843(8):1612–1619.

76. Noailles A, et al. Systemic inflammation induced by lipopolysaccharide aggravates inherited retinal dystrophy. *Cell Death Dis*. 2018;9(3):350.

77. Salguero MV, et al. Dysbiosis of Gram-negative gut microbiota and the associated serum lipopolysaccharide exacerbates inflammation in type 2 diabetic patients with chronic kidney disease. *Exp Ther Med*. 2019;18(5):3461–3469.

78. Zhao J, et al. Neuroinflammation induced by lipopolysaccharide causes cognitive impairment in mice. *Sci Rep*. 2019;9(1):5790.

79. Lee JW, et al. Neuro-inflammation induced by lipopolysaccharide causes cognitive impairment through enhancement of beta-amyloid generation. *J Neuroinflammation*. 2008;5:37.

80. Shetty SA, et al. Comparative genome analysis of Megasphaera sp. reveals niche specialization and its potential role in the human gut. *PLoS One*. 2013;8(11):e79353.

81. Wostyn P. COVID-19 and chronic fatigue syndrome: is the worst yet to come? *Med Hypotheses*. 2021;146:110469.

82. [No authors listed]. Long COVID: let patients help define long-lasting COVID symptoms. *Nature*. 2020;586(7828):170.

83. Yancey JR, Thomas SM. Chronic fatigue syndrome: diagnosis and treatment. *Am Fam Physician*. 2012;86(8):741–746.

84. Komaroff AL. Inflammation correlates with symptoms in chronic fatigue syndrome. *Proc Natl Acad Sci U S A*. 2017;114(34):8914–8916.

85. O'Sullivan O. Long-term sequelae following previous coronavirus epidemics. *Clin Med (Lond)*. 2021;21(1):e68–e70.

86. Proal A, Marshall T. Myalgic encephalomyelitis/chronic fatigue syndrome in the era of the human microbiome: persistent pathogens drive chronic symptoms by interfering with host metabolism, gene expression, and immunity. *Front Pediatr*. 2018;6:373.

87. Giloteaux L, et al. Reduced diversity and altered composition of the gut microbiome in individuals with myalgic encephalomyelitis/chronic fatigue syndrome. *Microbiome*. 2016;4(1):30.

88. Newberry F, et al. Does the microbiome and virome contribute to myalgic encephalomyelitis/chronic fatigue syndrome? *Clin Sci (Lond)*. 2018;132(5):523–542.

89. Rao AV, et al. A randomized, double-blind, placebo-controlled pilot study of a probiotic in emotional symptoms of chronic fatigue syndrome. *Gut Pathog*. 2009;1(1):6.

90. Wang T, et al. Chronic fatigue syndrome patients have alterations in their oral microbiome composition and function. *PLoS One*. 2018;13(9):e0203503.

91. Ranjan R, et al. Analysis of the microbiome: advantages of whole genome shotgun versus 16S amplicon sequencing. *Biochem Biophys Res Commun*. 2016;469(4):967–977.

92. CDC. Symptoms. https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html. Updated 2/22/2021. Accessed September 2, 2021.

93. Austin SR, et al. Why summary comorbidity measures such as the charlson comorbidity index and elixhauser score work. *Med Care*.

2015;53(9):e65–e72.

94. Rabenau HF, et al. Stability and inactivation of SARS coronavirus. *Med Microbiol Immunol*. 2005;194(1-2):1–6.

95. Bolger AM, et al. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114–2120.

96. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):1–359.

97. Haran JP, et al. The nursing home elder microbiome stability and associations with age, frailty, nutrition and physical location. *J Med Microbiol*. 2018;67(1):40–51.

98. CDC. Test for COVID-19 Only. https://www.cdc.gov/coronavirus/2019-ncov/lab/virus-requests.html. Updated August 17, 2021. Accessed September 2, 2021.