Computational and Structural Biotechnology Journal 19 (2021) 2928–2937

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

Gut lactate-producing bacteria promote CD4 T cell recovery on Anti-retroviral therapy in HIV-infected patients

Wei Lyu [a,1], Qingren Meng [b,c,1], Jingfa Xiao [b], Jing Li [d], Jian Wang [b], Zhifeng Qiu [a], Xiaojing Song [a], Hua Zhu [e], Changjun Shao [b], Yanan Chu [b], Qian Zhou [f], Taisheng Li [a], Routy Jean-Pierre [g], Jun Yu [b,h], Yang Han [a,*], Yu Kang [b,*]

[a] Department of Infectious Disease, Peking Union Medical College Hospital, & Center for AIDS Research, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, China
[b] CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, & China National Center for Bioinformation, Beijing 100101, China
[c] School of Medicine, Southern University of Science and Technology, Shenzhen 518055, China
[d] Department of Rheumatology and Immunology, Peking University People's Hospital, Beijing 100044, China
[e] Institute of Laboratory Animal Science, Chinese Academy of Medical Sciences (CAMS) Comparative Medical Center, Peking Union Medical College (PUMC), Beijing 100021, China
[f] CAS Key Laboratory of Quantitative Engineering Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, People's Republic of China
[g] Chronic Viral Illnesses Service and Division of Hematology, McGill University Health Centre, Montreal, QC, Canada
[h] University of Chinese Academy of Sciences, No.19 Yuquan Road, Shijingshan District, Beijing 100049, China4i9

ARTICLE INFO

Article history:
Received 22 October 2020
Received in revised form 2 May 2021
Accepted 9 May 2021
Available online 11 May 2021

Keywords:
Lactic acid bacteria
Metagenome
HIV
Immune recovery

ABSTRACT

Anti-retroviral therapy (ART) effectively suppresses viral replication in HIV-infected patients, however CD4 + cell restoration to normal value is not achieved by 15–20% of patients who are called immune non-responders. Gut microbiota composition has been shown to influence host immunity. Herein, to identify intestinal microbial agents that may influence the CD4 recovery in HIV-infected patients, we utilized a "Quasi-paired cohort" method to analyze intestinal metagenome data from immunological responders (IRs) and immunological non-responders (INRs). This method identified significant enrichment for *Streptococcus* sp. and related lactate-producing bacteria (LAB) in IRs. In a validation cohort, positive correlations between the abundance of these LAB and the post-ART CD4 + recovery was observed, and a prediction model based on these LAB performed well in predicting immune recovery. Finally, experiments using a germ-free mouse model of antibody-induced CD4 + cell depletion showed that supplementation with a lactate-producing commensal *Streptococcus thermophilus* strongly promoted CD4 recovery. In conclusion, our study identified a group of LAB that was associated with enhanced immune recovery in post-ART HIV-infected patients and promotes CD4 + cell restoration in a mouse model. These findings favour supplementation of LAB commensal as a therapeutic strategy for CD4 + cell count improvement in HIV-infected patients.

1. Introduction

The initiation of anti-retroviral treatment (ART) can effectively suppress viral replication in HIV-infected patients and restore the CD4 + cell counts to normal value (>500/mm^3) in the majority of patients, and define them as immunological responders (IRs) [1]. However, even after years of ART, CD4 + cell counts remain below 500/mm^3 in 15–20% of patients who are called immunological non-responders (INRs) [2]. In INRs, CD4 + cell count is not influenced by switching or intensifying the ART-regimen. Factors associated with low CD4 recovery include late initiation of ART, nadir CD4 cell count, pre-treatment viral load, older age, and markers of inflammation [3,4]. The inflammatory conditions are characterized as increased CD8 + T cell count [5], low CD4/CD8 ratio (<1), elevated plasma IL-6, CRP, which indicates a higher risk for development of non-AIDs morbidity and mortality [6].

* Corresponding authors.
E-mail addresses: hanyang_pumch@126.com (Y. Han), kangy@big.ac.cn (Y. Kang).
[1] These authors contribute equally.

https://doi.org/10.1016/j.csbj.2021.05.021

Recently, gut microbial composition is emerging as a new factor associated with inflammation and CD4 recovery in HIV infection [7]. Gut dysbiosis, namely dysregulation of intestinal microbiota, and increased gut permeability lead to enhanced systemic inflammation and are commonly seen in chronic conditions such as obesity and aging [8,9]. In HIV-infected patients, gut-associated lymphoid tissue is the major site where HIV depletes CD4 + cells [10], and several lines of evidence suggest the depletion of gut CD4 + T cells is associated with gut dysbiosis [11–13], microbial translocation and systemic inflammation [14,15]. In contrast to optimal viral control with ART, gut dysbiosis and systemic inflammation persist especially in INRs, increasing their risk of developing inflammatory non-AIDS comorbidities such as cardiovascular disease, diabetes mellitus, liver steatosis and cancer [16]. Thus, an emerging research priority is to discover strategies to rectify the gut dysbiosis for treatment of HIV-infected patients in addition to ART. However, which components of the intestinal microbiota and how they influence the immune homeostasis are not fully understood.

Here, we first compared the shotgun-sequenced intestinal metagenome data of IR and INR patients on two-year of ART from our previous study [17] and identified a group of lactic acid bacteria (LAB) significantly enriched in the IR samples from the dataset. Then the role of these lactate-producers in CD4 + cell restoration was further confirmed by the metagenome data of a validation cohort in which treatment-naïve HIV-infected patients were recruited and followed-up after 12 months on ART. Finally, an animal model of CD4 + cell depletion was developed by infused anti-CD4 antibody in germ-free mice, and supplementary *Streptococcus thermophilus*, a LAB commensal species, was able to promote CD4 + cell recovery, especially for naïve CD4 + cells. These findings suggest a previously unidentified potential of a group of LAB taxa in promoting immune recovery.

2. Results

2.1. A group of lactate-producing taxa correlated with favourable CD4 recovery

To identify species among gut microbiota composition that can influence a post-ART CD4 recovery in HIV-infected patients, we re-analysed the metagenomic data obtained in our previous study [17] for IR and INR groups. In brief, IR (immunological responder) and INR (immunological non-responder) were defined as HIV-patients whose CD4 + T cell count > 350/mm^3 and < 350/mm^3 after at least two years of ART, respectively. For the samples collected, *i.e.,* after two years of ART, we found no significant differences in the alpha-diversity (*i.e.*, Shannon and richness indexes) between the IR and INR groups (**Fig. S1a-d**), and the principal coordinate analysis (PCoA) did not effectively separate IR and INR patients (**Fig. S1e**). Given the many confounding factors which greatly elevate the inter-individual metagenomic diversity and prevent the identification of causative species or metabolic features from metagenomic data [18,19], we then utilized our newly-developed analytical method – "quasi-paired cohort" which is powerful in control of inter-individual diversity and successful in identifying causative microbiome features of autism [20] (Fig. 1a, refer to Methods and Fig. S2 for detail).

Using this method, we reconstituted a quasi-cohort of 64 pairs of IR and INR microbiota from our published metagenomic datasets (Table S1) [17] and identified 33 and 25 intestinal species enriched in the IR and INR groups, respectively (Wilcoxon signed rank test for paired samples, $p < 0.005$, Fig. 1**b**, Table S2). Out of these species, 22/33 species of IR group and 10/25 species of INR group respectively exhibit positive and negative correlations with the CD4/CD8 ratio, the informative prognostic marker for immune recovery (Fig. 2**a**). Notably, 6 out of the 22 IR-associated species were lactic acid bacteria (LAB), including *Streptococcus infantis, Granulicatella* unclassified, *Lactococcus garvieae*, *Lactobacillus sanfranciscensis, Leuconostoc lactis,* and *Gemella* unclassified, whereas none LAB species were found in the INR-associated species (Fig. 2a). These LAB taxa, most of which belong to Order Lactobacillales, share the common ability to produce lactate and often exhibit morphological and physiological similarities.

Next, we investigated specific contributions by each of the 32 outcome-associated taxa to the counts of a variety of T cell subsets, a metric that can be used to assess HIV status or prognosis. For context, the count of CD4 + cells, and its subsets including CD4 + CD28 +(functional subsets), naïve CD4+ (CD4 + CD45RA +), homing naïve CD4+ (CD4 + CD45RA + CD62L +), and memory CD4 + cells (CD4 + CD45RO +), are used clinically for positive prognoses of post-ART immune restoration; whereas the count of CD8 + cells, and its subsets consisting of CD8 + CD38 + and CD8 + DR + cells, are markers of activated inflammation and poor prognosis. Specifically, the CD4/CD8 ratio can be a highly informative and reliable marker for prognosis prediction [21–23].

Using Spearman's correlation coefficient (rho), we found that the 22 IR-associated species were positively correlated with CD4 + cell counts and other positive prognosis predictors at the end of two-year ART, whereas these species were often negatively correlated with CD8 + cell count and other negative prognosis predictors (Fig. 2a). In contrast, the ten INR-associated species showed inverse correlations to these prognostic markers, though not as strong as those enriched in IR groups (Fig. 2a). In particular, two LAB species, *Streptococcus infantis* and *Gemella* unclassified, showed the strongest correlations (rho value > 0.6) with the CD4/CD8 ratio; together, these species exhibited the most pronounced positive relationships with CD4 + cell counts and all other positive indicators including CD4 + CD28+, naïve CD4+, homing naïve CD4+, and memory CD4 + cells, whereas a negative correlation to CD8 + cell counts (Fig. 2a). Thus, the high proportion of LAB in IR-associated species and the strongest correlations to prognostic markers they exhibit suggest a role of lactate-producing bacteria in promoting immune recovery among post-ART patients.

To further compare the potential of intestinal lactate production between IR and INR groups, we quantified the abundance of the key enzymes from canonical lactate-producing pathways in their microbiota. Lactate dehydrogenases (LDH) is the key enzyme that transforms pyruvate to lactate in anaerobic glycolysis, and its two forms, L-LDH and D-LDH produce L- and D-lactate, respectively. We used annotations from the analytical tool of HUMAnN2 [24] to calculate the abundance of metabolic pathways and enzymes in metagenome data, and we found that the abundance of D-LDH (EC1.1.1.28), as read count per million reads (CPM), was significantly elevated in the IR group, in contrast to L-LDH (EC1.1.1.27) which showed no difference between groups (Wilcoxon rank sum test, $p < 0.05$, Fig. 2**b**). Furthermore, the "mixed acid fermentation" pathway which only generates lactate via D-LDH (EC1.1.1.28) was significantly enriched in the IR patients when compared to INR in the *quasi-paired* cohort (Wilcoxon signed rank test, $p < 0.001$, Fig. 2**c**). As D-LDH is rather specific for microbes [25] and essentially unidirectional in producing lactate in contrast to L-LDH that catalyze the reverse reaction to pyruvate more efficiently [26], only the abundance of D-LDH is considered as an indicator for the microbial lactate production in the gut. The elevated abundance of D-LDH and related pathway supports the enrichment for lactate-producing bacteria observed in the IR microbiota samples and implies an increased local microbial lactate production.

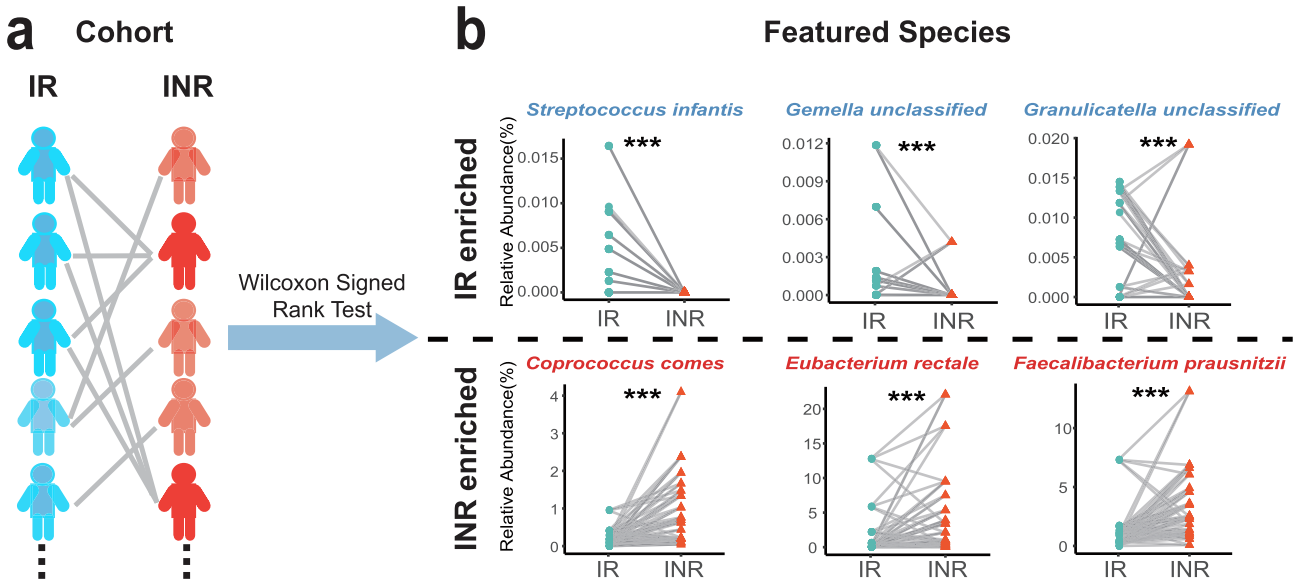


Fig. 1. The principle of "quasi-paired cohort". a. The scheme of the method "quasi-paired cohort" that "paired" patients of similar metagenomic profile. b. Representative species enriched in IR and INR groups identified with the "quasi-paired cohort" approach. Wilcoxon Signed-Rank test, *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

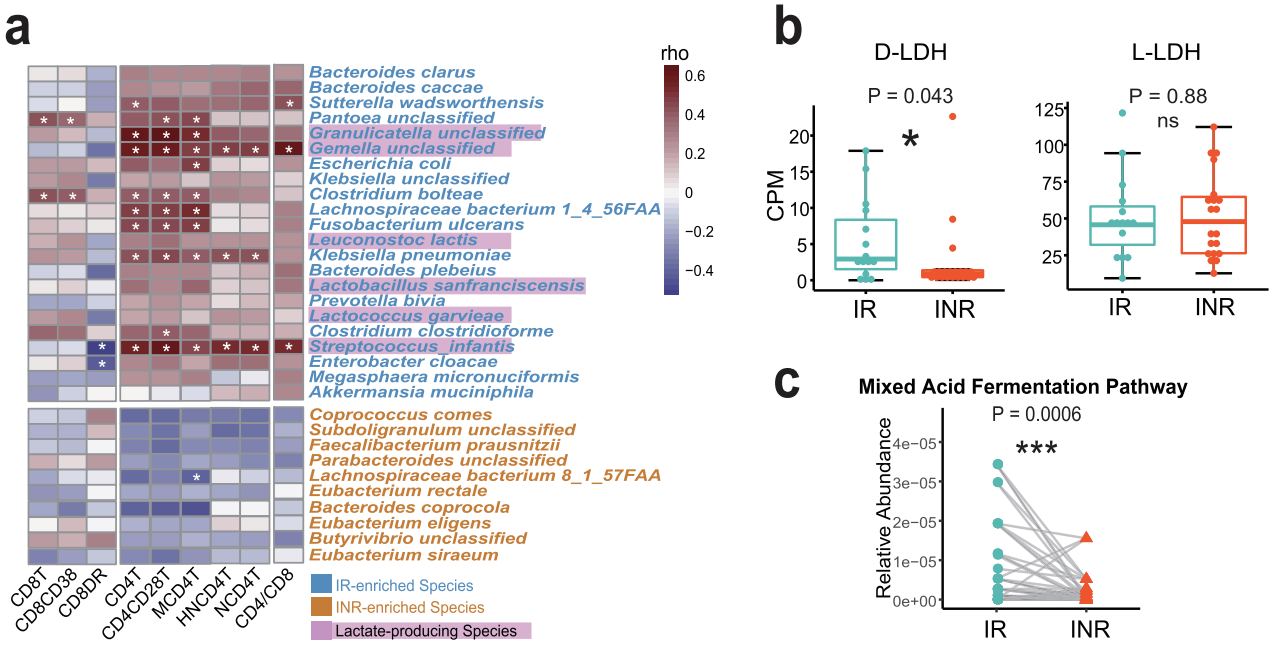


Fig. 2. Enrichment of lactate-producing bacteria and lactate-dehydrogenase in IR group. a. Heat map of species enriched in IR and INR groups and their correlations to prognostic immunological markers. Colors indicate the ρ value of Spearman's coefficient, and asterisk (*) denotes strong correlations of $\rho \geq 0.4$ or $\rho \leq -0.4$. Species highlighted with a purple background are typical lactic acid bacteria (LAB). Positive prognostic markers include CD4TC (TC, short for T cell count), CD4CD28TC, MCD4TC (Memory CD4 + T cell), HNCD4TC (Homing naïve CD4 + T cell), NCD4TC (Naive CD4 + T cell), and the most informative marker of CD4/CD8 (the ratio of CD4 + T cell count to CD8 + T cell count). Negative prognostic markers include CD8TC, CD8CD38TC, , CD8DRTC (CD8 + DR + T cell). **b.** Comparison of the abundance of D-LDH and L-LDH in the microbiota of IR and INR patients. LDH, lactate-dehydrogenase; CMP, read count per million reads. **c.** The relative abundance of the pathway "mixed acid fermentation" in the microbiota of IR and INR patients. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.2. The abundance of the IR-associated LAB performs well in predicting the post-ART CD4 recovery

To further validate the role of the LAB clusters in immune recovery, we recruited a validation cohort of 26 treatment-naïve HIV-infected patients for longitudinal observation over one-year after ART initiation. T cell subsets and viral load were monitored (Table S3), and faecal samples before and after ART were collected to perform shotgun-sequencing of metagenome. Several weeks after the initiation of ART, all the patients achieved complete viral suppression, and CD4 + cell count of 16/26 patients reached 350/mm^3 at the end of one-year of ART. However, the alpha- and beta-diversity of their stool samples did not change after ART, except for a mild decrease in richness (the number of species) was observed (Fig. S3). Then each species we identified in the first cohort was tested for correlations to the growth of post-ART CD4 + cell count (indicated as the ratio of After/Before), but none reached a strong correlation.

It is known that interacting species in microbiota are organized into functional groups and confer metabolic functions as a whole [27,28]. Therefore, we established co-abundance associations for all the species present in all samples and found nine co-abundance clusters in which species exhibit strict co-variance (rho > 0.8, Table S4). Interestingly, both the top two largest clusters (Cluster I & II) contained a large proportion of the LAB (Fig. 3**a**). Specifically, in Cluster I, 9/13 species were LAB, while in Cluster II, 4/5 species were LAB, primarily consisting of *Streptococcus* spp., *Gemella* spp., and *Granulicatella* spp., which are also identified as IR-associated species in the first cohort. The total abundance of Cluster I and Cluster II species was much higher in IR than in INR patients, and especially so for Cluster II, which reached statistical significance (Fig. 3**b**). Furthermore, the total abundances of the two clusters also exhibited positive correlations to CD4 + T cell counts and other CD4 + T cell subsets similarly to those IR-associated species (Fig. S4), which implies that the two LAB clusters may function as a whole to contribute to the immune recovery.

When further tested with the validation cohort, the total abundance of species in Clusters I and II (average of Before and After samples) strongly correlated to the growth rate (after/before ratio) of CD4 + cell count and that naïve CD4 + and CD4 + CD28 + cells on ART (Fig. 3**c**), while the abundance of other clusters did not show strong correlations to the post-ART change of CD4 + cell and its subsets (Table S5). This result suggests a dependence of post-ART immune recovery on the abundance of the species in the two LAB clusters. However, the abundance of the species in the two clusters, show little correlations with the change of CD8 + cell count after ART as well as that of its subsets (Fig. S5). We further tested the correlation between the initial abundance of the LAB clusters before ART and changes of T cell subsets after ART, and found them less strong than those of averaged abundance (Fig. S5).

All the patients in this cohort obtained a favourable recovery of CD4 count back to > 200/mm^3 after the one-year ART, and we found that patients of pre-ART CD4 count < 200/mm^3 and CD4/CD8 ratio < 0.5 often achieved a higher After/Before ratio of CD4 count. Notably, the abundance of both clusters of LAB were higher in these patients and possibly predicted their favourable prognosis (Fig. S6). Comparison between Before and After-ART samples showed that the abundance of the two LAB clusters significantly decreased after ART (Fig. 3d). Meanwhile, six out of the nineteen species whose relative abundance significantly decreased after ART were LAB, including *Bifidobacterium longum*, *Granulicatella* unclassified, and four species of *Streptococcus* spp. (Fig. S7), indicating that ART medicines are unfavorable for intestinal LAB species.

As the abundance of the two LAB clusters exhibited strong correlations to post-ART CD4 + cell restoration, we further evaluate their predictive value by constructing a random forest classifier with their abundance. The model was first trained and tested to discriminate IR and INR samples in the first cohort, and evaluated with ROC curve analysis. The area under curve (AUC) achieved 81% with 1000 bootstraps which indicated that the abundance of the species in the two LAB clusters accurately described the deviations between IR and INR (Fig. 4**a**). Interestingly, the species of *Granulicatella* unclassified, *Streptococcus infantis,* and *Gemella* unclassified which have also been identified as IR-associated species are among the top five species that contribute most to the model (Fig. 4**b**). Next, the model was used to predict the post-ART outcome for the validation cohort. The classifier score of each subject inferred from the model indicates the possibility of a sample to be correctly classified as IR samples in contrast to INR sample. This score, which we named as immune promotion score, could be regarded as a comprehensive indicator to represent the promoting effects conferred by the species in the two LAB clusters. According to the immune promotion score, patients were divided into quartiles, and a conspicuous trend of higher CD4 + cell growth can be observed in quartiles of higher scores (Fig. 4**c**), which further confirms the prognostic value of these LAB species in predicting immune recovery.

2.3. Streptococcus thermophilus promotes CD4 + cell recovery in a mouse model

To confirm the effects of LAB on CD4 recovery, we developed a mouse model of transient CD4 + lymphopenia by injecting intraperitoneal an anti-CD4 antibody as previously described [29]. Germ-free C57BL/6 mice were used in the experiment, and the effects of intestinal LAB on immune recovery were investigated by intragastric gavage of bacteria before CD4 + depletion and observed for 30 days. As *Streptococcus* species exhibited the strongest associations to favourable prognosis, we selected *Streptococcus thermophiles* (ST, n = 4), a common fermentation species in yogurt, as a representative to validate the effects of LAB and compared to *Eubacterium bakeri* (EB, n = 4), a butyrate-producing commensal. Negative controls (Ctrl, n = 4) of CD4 + cell depletion but no treatment and mock controls (Mock, n = 4) without CD4 + cell depletion or any other treatment were investigated in parallel (Fig. 5**a**). Mice administrated with bacteria showed no signs of illness or infection. The depletion of CD4 + cells in all experimental groups was confirmed by flow cytometry three days after the injection of antibody, and 16 s rDNA amplification and sequencing of faecal samples confirmed the inoculation of both species and showed no signs of contamination of other species three weeks after bacterial gavage. The faecal concentration of lactic acid was a little but not significantly elevated in ST mice than in EB mice (averagely 6.5 μM/g vs. 4.6 μM/g), whereas not detected in control groups.

When sacrificed four weeks after CD4 + cell depletion, both groups of bacterial inoculation showed higher recovery than the negative control mice, showing increased CD4 + cell (%), naïve CD4 + cell (%) in CD4 + cells, as well as decreased CD8 + cell (%) in blood (Fig. 5**b,5c**). This result implies that both species promote the recovery of CD4 + cells while consequently decrease the proportion of CD8 + cells. Specifically, the effect of lactate-producing *S. thermophiles* on CD4 + recovery was much stronger than *E. bakeri* that the proportions of T cell subsets in ST group were almost the levels of mock group (Fig. 5b), indicating a superior effect than that of butyrate-producing *E. bakeri* on CD4 + restoration. Similar effects were also observed in T cell subsets collected in mesenteric lymph node (LN) and spleen (Fig. S8). However, the influences of both species on the proportions of some CD8 + cell subsets, *i.e.* effective CD8 + cells and CD8 + CD38 + cells in blood, mesenteric LN and spleen were not obvious (Fig. S9), so were their effects on γδT cells. These results indicated that the promoting effects of *S. thermophiles* seemed to be more directed to CD4 + cells than to CD8 + cells and other T cell subsets, consistent to the immunological effects of LAB species observed in the validation cohort. Besides, we investigated the development of lymphatic follicles in the lamina propria , the submucosal lymph tissue of colon. The lymphatic follicles were greatly enlarged under the stimulation of *S. thermophiles*, even exceeding the size of those in Mock mice, and the CD4 + cells stained by immunohistochemistry showed much stronger signals in ST mice when compared with EB and Ctrl mice (Fig. 5**d**).

3. Discussions

In this study, we reported an association between a group of intestinal LAB species and the enhanced CD4 + cell restoration on ART in HIV-infected patients. Gut microbes are commonly found to have immunomodulatory effects [30], and a range of individual or groups of microbes has been demonstrated to be able to stimulate or inhibit a specific subset of immune cells [31–33]. The micro-

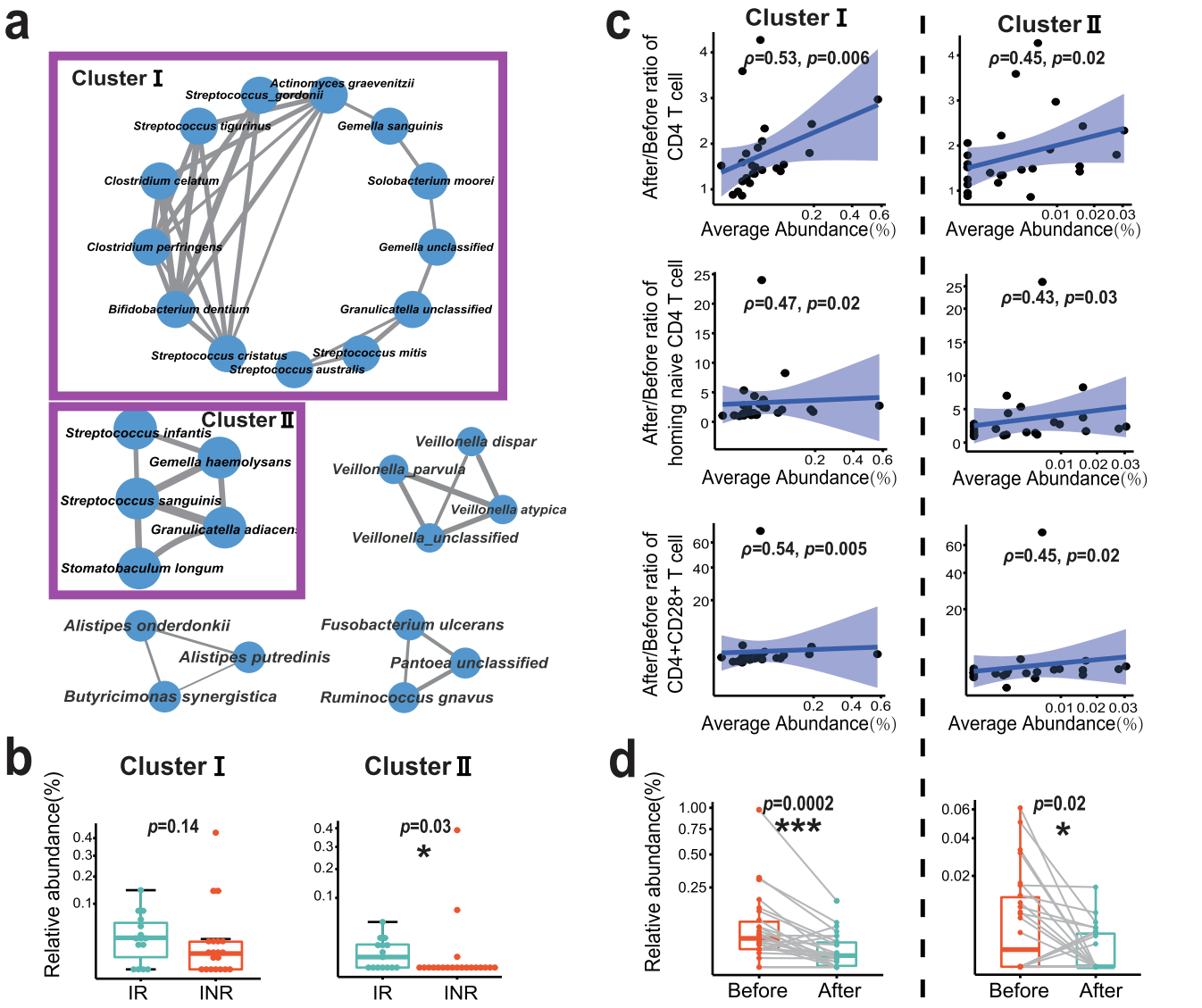
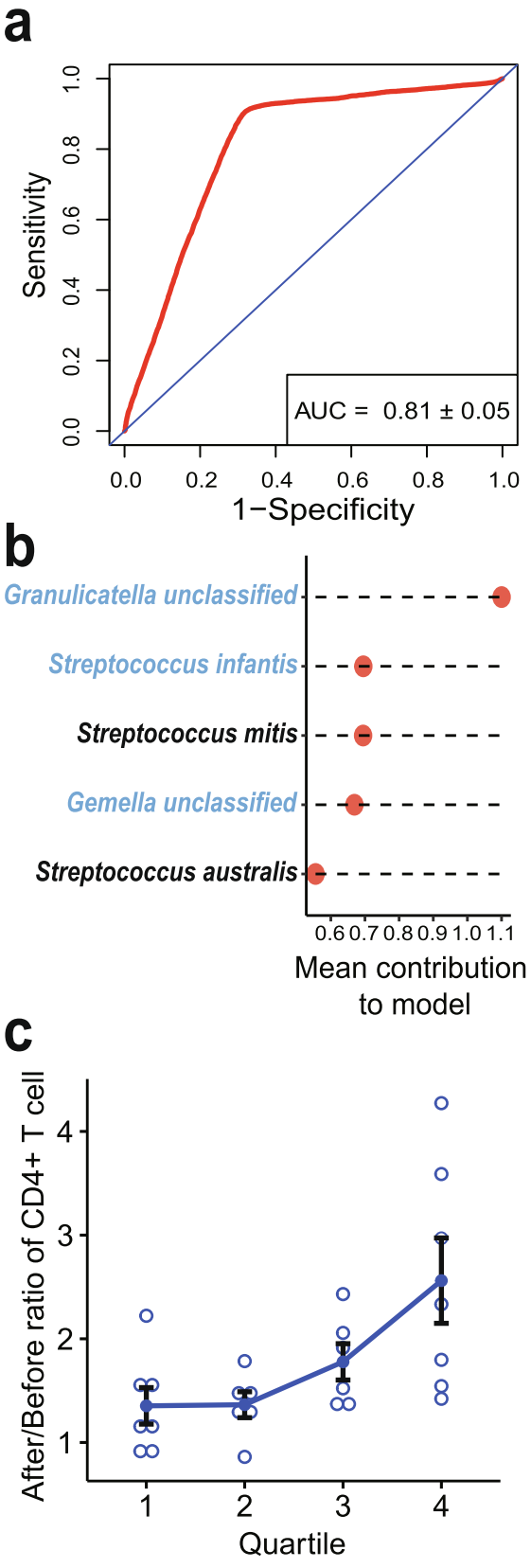


Fig. 3. Two clusters of co-occurrent LAB species and their correlations to post-ART immune recovery. a. The top five co-occurrence clusters of species in HIV gut microbiota. The thickness of lines indicates rho value of Spearman's coefficient. **b.** Comparison of total abundance of species in cluster I (left), cluster II (right) between IR and INR patients. Wilcoxon rank sum test, *, $p < 0.05$. **c.** The correlations of the total abundance of species in cluster I (left) and cluster II (right) to the post-ART growth rate of CD4 + T cells and its subsets of patients in the validation cohort. The growth rate was represented as the After/Before ratio of cell count, and the abundance is averaged between before and after ART. The strength of the correlations was evaluated by the Spearman rank-sum test with the rho value labelled above. **d.** Change of the total abundance of species in cluster I (left) and cluster II (right) after ART. Wilcoxon signed-rank test, *, $p < 0.05$; ***, $p < 0.001$.

bial immunomodulatory effect is phylogenetically independent and varies from species to species, possibly due to their differences in antigenicity, aggressiveness, and the metabolites they produced [30,34]. As most of the species related to the immune recovery in HIV-infected patients are LAB, and quantification of the abundance of D-LDH gene and related pathway demonstrated significant enrichment in IR patients, we therefore speculate an elevated intestinal lactic acid concentration contributes to the immune-promotion effect. Lactate is a signal molecule that substantially influences the gene expression profile in a wide range of cell types, including immune cells [35]. However, its effects on immunity and inflammation are still controversial [36,37]. As the immune recovery-related LAB we have identified represent as a small proportion of intestinal lactate-producers, the contribution from a group of specific species might overwhelm that from others. However, the dissection of the lactate effects from that of specific LAB species might be rather difficult since the fecal concentration of lactic acid is often undetectable due to high efficiency in gut re-absorption and local utilization by enterocytes and bacteria [38].

Another obstacle for clarifying the role of LAB on immune recovery after ART is the selection of an appropriate animal model. The pathogenesis of HIV infection and changes upon ART are rather complicated, and there is no available model to fully recapitulate the complex and multifactorial process of CD4 depletion in HIV and recovery following ART. We finally selected the model of CD4 depletion by anti-CD4 antibodies to dissect the pure effect of LAB on the recovery of CD4 count in no additional conditions during the HIV infection and ART, such as "leaky gut", chronic inflammation, or latent HIV infection. Therefore, the LAB effects on CD4 count we observed are not confined to HIV infection, and might be potentially generalized into other immune inhibitory conditions such as post-antineoplastic chemotherapy. Although LAB are generally beneficial for health, clinical trials with edible LAB, including *Lactobacillus spp*. [39], have not achieved success in improving post-ART immune recovery for HIV patients. Nevertheless, results from our study pointed to a group of less-noticed LAB, *i.e., Streptococcus* spp. *Gemella* spp. and *Granulicatella* spp., which correlated to a better prognosis, and the favourable

effects of *Streptococcus thermophilus* in our animal experiment deserve further clinical investigations and in comparison to other commonly applied LAB.



Many LAB species have been utilized as probiotics, and studies have reported the effects of LAB species on pathogen resistance and ameliorating inflammation in intestine [40,41]. Our study also demonstrated the effects of *S. thermophiles*, a common fermentative LAB in yogurt, on the restoration of CD4 + T cells, which suggests a new strategy in dealing with the CD4 + lymphopenia. As revealed in our study, the extent of CD4 + cell restoration is dependent on the abundance of the group of LAB. These findings imply that the supplementation of LAB probiotics especially of *Streptococcus* species which are vulnerable to standard ART might be a promising strategy to improve the CD4 + cell lymphopenia. However, long-term intake of LAB probiotic should be cautiously prescribed due to the complex interactions between probiotics and host [42]. For example, LAB often produce both L-lactate and D-lactate, and overdose absorption of gut D-lactate might lead to lactic acidosis as the D-lactate would not ready to be utilized or metabolized by the host, and probiotic strains depleted of the D-LDH gene might be more suitable for long-term users.

In conclusion, our study uncovered a group of intestinal LAB that closely related to the enhanced post-ART CD4 + T cell restoration. These primary findings provide a novel avenue for investigations in microbiota-related mechanisms underlying the poor immunological recovery in HIV-infected patients and for attempts in rectifying their immune abnormalities via tuning the intestinal microbiota components.

4. Materials and Methods

4.1. Ethical compliance

All experimental protocols were approved by the Ethics Committee of the Peking Union Medical College Hospital (reference numbers: JS1617) and registered in ClinicalTrials.gov (ID: NCT04297501). The study design complied with all relevant ethical regulations and aligned with the Declaration of Helsinki. All participants gave their informed consent.

4.2. Samples collection and fecal DNA extraction

The information of IR and INR HIV patients in the cohort of treated patients was retrieved from our previous study [17]. Briefly, this cohort included HIV patients who experienced suppressive ART for over 2 years and were recruited from the Department of Infectious Diseases, Peking Union Medical College Hospital, China from December 2015 to September 2016. IRs (immunological responders) and INRs (immunological non-responders) were defined as CD4 + T-cell counts $\geq$ 350 cells/mm^3 and < 350 cells/mm^3 after 2 years of ART. For the validation cohort, 26 newly diagnosed HIV patients and naïve of anti-retroviral treatment were recruited and followed until one year after ART in the same Hospital from March 2018 to July 2019. In accordance with the exclusion criteria in the treated cohort, subjects who have used antibiotics, antacid, probiotics, or prebiotics or have experienced diarrhea or digestive symptoms within the previous one month before sample

Fig. 4. A prediction model based on the abundance of species in Cluster I and II. **a.** Performance of the prediction model evaluated with ROC. AUC, area under curve. **b.** Top five of species contributors to the prediction model. Species in blue colour are also identified as species enriched in IR patients. **c.** The growth rate of CD4 + cell (After/Before ratio of CD4 + cell count) in each quartile of the patients in the validation cohort which was divided according to their immune promotion score. The score of each patient was inferred from the prediction model with the abundance of species in Cluster I and II. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

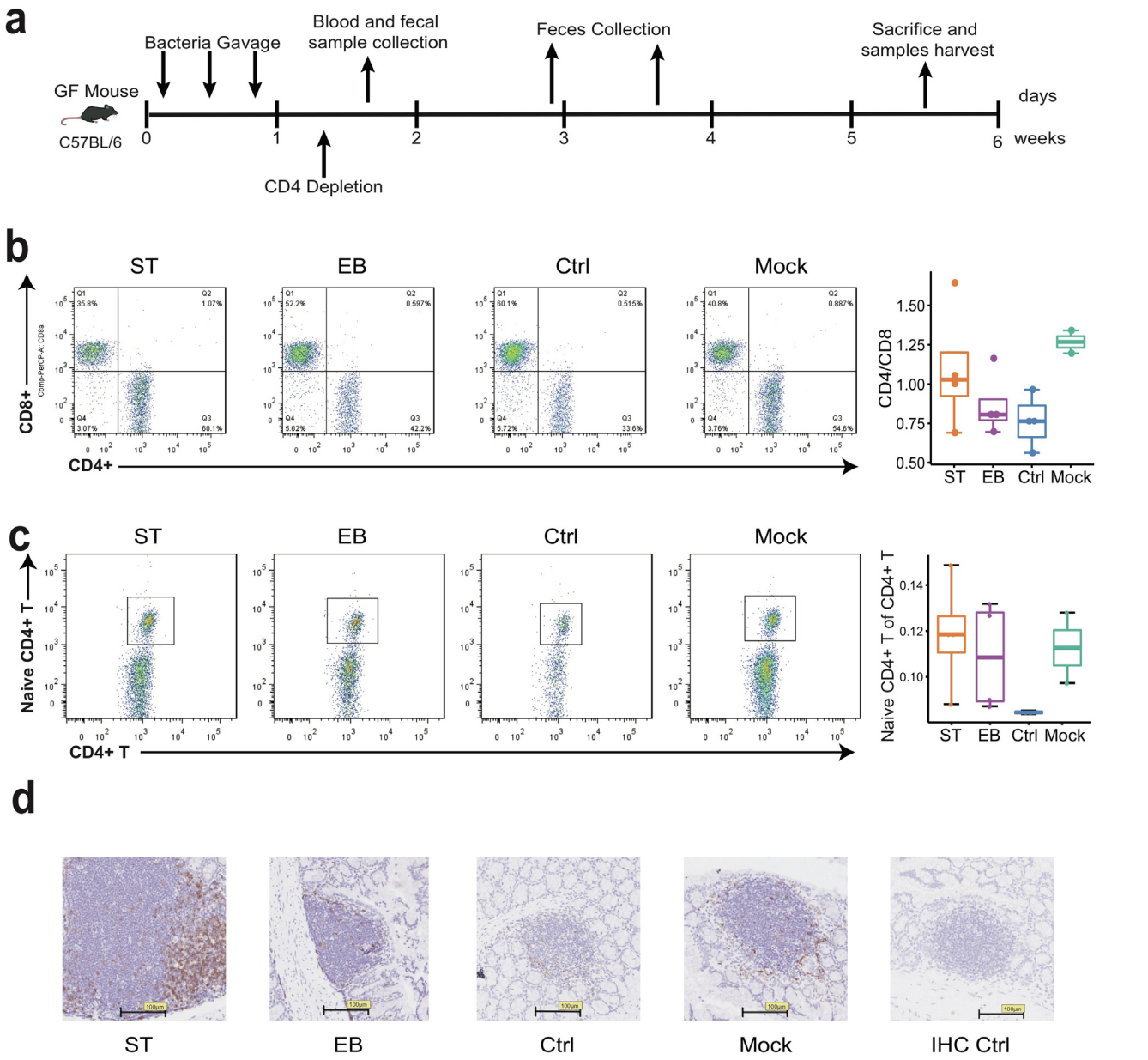


Fig. 5. Effects of a LAB species of *Streptococcus thermophilus* on the recovery of CD4 lymphopenia in a mouse model. a. Scheme of the experimental design for a probiotic-based preventive regimen using the mouse model of transient CD4 + cell depletion. **b-c.** Representative flow cytometry plots showing the proportion of CD4 + cells **and** CD8 + cells **(b)**, and Naïve CD4 T cells **(c)** in blood of the indicated groups. The proportions of each cell subset in all samples were plotted on the right. **d.** Representative microscopic photos of lymphatic follicles in the lamina propria . Brown cells are immunohistochemistry stained CD4 + T cells. Scale bar = 100 μm. ST, *Streptococcus thermophilus* gavage group (n = 4); EB, *Eubacterium Bakeri* gavage group (n = 4); Ctrl, negative control group (n = 4); and Mock, mock control group (n = 4). *, $p < 0.05$.

collection were excluded. In addition, the patients with active opportunistic infection, co-infection of HBV and HCV, or other comorbidities were also excluded from our cohorts.

For all participants, faecal samples before and one-year after ART were collected and placed in collection tubes in PSP® Spin Stool DNA Plus Kit (Stratec Co., Germany), and stored at − 80℃. The faecal DNA was extracted in the P2 plus laboratory in the hospital according to the manufacture's instruction. Besides, the counts of CD4 + and CD8 + T cell and their subsets were analyzed by flow cytometry as listed in Table S3.

4.3. DNA library construction and sequencing

The DNA libraries were constructed, quantified, and then sequenced on the Illumina HiSeq 4000 platform in Annoroad Gene Technology Co. (Beijing) to generate 150-bp paired-end reads, with an insert size around 350 bp. The raw reads were processed with KneadData package (https://bitbucket.org/biobakery/kneaddata), where low-quality reads, adapters, human DNA contamination and shorter reads<60 bp were removed by Trimmomatic [43] and Bowtie2 [44].

4.4. Bacteria taxonomy and pathway profiling

For the metagenomic data of each sample, the taxonomy profile and relative abundance of each species were estimated by performing MetaPhlAn2 [45] with default parameters, where qualified reads were assigned to clade-specific marker genes of over 7500 species identified from ~ 13500 bacteria genomes.

For each sample, the strains-level diversity of all its component species was analyzed using StrainPhlAn [46] with the default parameters. Based on the MetaPhlAn2 output of marker genes,

StrainPhlAn distinguishes strains by identifying sequence variations in marker genes per species. The phylogenetic trees of each species were further constructed with RAxML [47] and visualized with FigTree [http://tree.bio.ed.ac.uk/software/figtree/].

Depending on the output of MetaPhlAn2 and the pangenome database of ChocoPhlAn, software HUMAnN2 (The HMP Unified Metabolic Analysis Network 2) [24] that is based on the pathway database of MetaCyc [48], was used to identify the microbial pathways in each sample with default parameters except alignment e-value as 1e-5. The module humann2_renorm_table was utilized to estimate the relative abundance of each microbial pathways normalized as a proportion of total assigned reads and quantify the CPM (Counts Per Million) of each enzyme.

4.5. Construction of quasi-paired cohort

The metagenome data of IRs and INRs were downloaded from the dataset of a previous study [17] (No. SRP111623). The taxonomical profile of each sample was configurated as above described, and all samples were then plotted into the high-dimension space with the abundance of each species as a dimension. The similarity of a sample to its neighbours was evaluated with KNN, that it, the average Euclidean distances of the sample to its nearest k neighbours (k was the square root of sample size). Similarly, intragroup KNN and intergroup KNN were also calculated for each sample to evaluate its similarity to neighbours of own side and opposite side. Then three steps were taken to construct the quasi-paired cohort. First, to remove outliner and redundant samples which were defined as KNN > mean + SD (outliner) or < mean-SD (redundancy). Next, to identify boundary samples which were more similar to neighbors of the opposite group than own side, that is, intragroup KNN > intergroup KNN. Finally, to construct the quasi-paired cohort with pairs of boundary samples and one of its nearest k neighbors of the opposite group, and to remove redundant pairs (refer to Fig.S2 for the flowchart of "quasi-paired cohort").

4.6. Co-abundance species clusters

We first evaluated the pairwise correlations between species according to their abundance in IR and INR patients with Spearman rank sum test. Then a co-abundance network of species was constructed based on the correlations. When the cut-off of Spearman's coefficient (rho value) was set as > 0.8, clusters emerged in the MCODE module of Cytoscape [49] software with default parameters.

4.7. Random forest classifier and prediction model

The abundance of species in Cluster I and II were used to construct a random forest classifier with the packages of caret [https://cran.r-project.org/web/packages/caret/] and random Forest [https://cran.r-project.org/web/packages/randomForest/] in R. The model was trained with 40% of all IRs and INRs samples and tested with all samples with 1000 times of bootstrapping. The performance of the model was evaluated by AUROC with the R packages of ROCR [https://cran.r-project.org/web/packages/ROCR/] and pROC [https://cran.r-project.org/web/packages/pROC/]. For each species, the mean value of its contribution to the model across all bootstraps was calculated to indicate its importance in deviating IRs and INRs. Finally, the model was used to predict the recovery of CD4 + cell count after one-year of ART for patients in the validation cohort. For each patient, an opportunity score that classified the patient as IR or INR was inferred by the prediction model with the patient's total abundance of species in Cluster I and II, and the score was named as immune promotion score. Then patients were divided into quartiles according to their immune promotion score and compared for their CD4 T cell count growth after one-year of ART.

4.8. Animal model of CD4 cell depletion in germ-free mice

All experiment protocols were approved by the Animal Care and Use Committee of the institute (ZH18001) and were carried out in accordance with the guidelines of the institute. Male and female C57BL6J germ-free mice were bred and fed under germ-free conditions in Institute of Laboratory Animal Science, Chinese Academy of Medical Sciences (CAMS, Beijing) until 4 weeks old for experiments. Animal model of CD4 + cell depletion was constructed by two boluses (with a three-day interval) of intraperitoneal injection of 50 μg anti-CD4 antibody (clone GK1.5, BioXCell) as previously described [29].

We randomized sixteen C57BL6J germ-free mice into four treatment groups (M:F = 1:1). They were *Streptococcus thermophilus* gavage group (ST, n = 4), *Eubacterium Bakeri* gavage group (EB, n = 4), negative control group (Ctrl, n = 4), and mock control group (Mock, n = 4). During the week before the anti-CD4 antibody injection, three doses of bacterial cell suspensions (1×10^8 cell/dose) were given for mice in ST and EB groups by gavage. Mice in Mock and Ctrl groups were given equal volume saline for control. Then, all mice were depleted CD4 + cells except those in the Mock group. Faecal samples of all mice were collected in the first and second week after the last antibody injection for 16 s rDNA sequencing to confirm inoculation and detect contamination. Faecal concentration of lactic acid was measured by spectrophotometric method with the lactic acid assay kit (Cat. No. BC2230, Solarbio, China). After faecal DNA extraction as described above, V3-V4 region of the 16S rDNA was amplified and sequenced by Sanger's method. Four weeks after the first antibody injection, all mice were sacrificed. Samples of blood, spleen, and mesenteric lymph nodes (MLN) were obtained from all mice, sieved through a 70 μm cell strainer (Corning) in RPMI 1640 medium with 10% FBS and single-cell suspensions (10^6 cells/100 μl) were prepared for flow cytometry. Colon tissues were fixed in formalin for pathological and immunohistochemistry assay.

4.9. Flow cytometry and immunohistochemistry assay

Flow cytometry analysis follows a scheme showed in Fig. S9. Cell surface markers were first stained, and the cells were then fixed and permeabilized with an intracellular staining buffer set (Thermo Fisher Scientific) following the manufacturer's protocol and stained with intracellular or intranuclear markers. Antibodies (Table S6) were purchased from eBiosciences (Thermo Fisher Scientific). Memory CD4 cells were defined as CD4 + CD44+, naïve CD4 cells as CD4 + CD62L+, effective CD8 cells as CD8 + CD38 + H-2 kb+, γδT cells as CD3 + TCR gamma/delta + . Flow cytometry was performed using FACSAriaTMII (BD Biosciences) and the data was analyzed using FlowJo v10.0.7 software (Tree Star Inc., Ashland, OR, USA).

The colon tissues were embedded in paraffin, sectioned and stained with hematoxylin and eosin (H&E). A microscopic assessment of lymphatic follicles was performed, as well as immunohistochemistry assay of CD4+, CD8+, and Treg (CD4 + CD25 + FOXP3 +) cells with corresponding antibodies (Table S6).

4.10. Statistics analysis

All statistical analyses were performed with R software. Wilcoxon Signed Rank test was used to identify enriched species and pathways for IR and INR patients in the quasi-paired cohort, as well as species significantly changed after ART. The difference in

enzymes abundance between IRs and INRs and difference in T cell subsets proportions between mice groups were tested by Wilcoxon Rank sum test. The correlation among species, or between species abundance and clinical markers of T cell subsets were evaluated by Spearman rank sum test. When multiple hypotheses were considered simultaneously, *p*-values were adjusted to control the false discovery rate with the method described previously [50]. The figures were plotted with the R package of ggplot2 [https://cran.r-project.org/web/packages/ggplot2/] and ggpubr [https://cran.r-project.org/web/packages/ggpubr/].

5. Data availability

Extended metagenomic data of this study are available at the URL https://bigd.big.ac.cn/ described in the project CRA002425.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank to Prof. Liping Zhao in School of Environmental and Biological Sciences, Rutgers, The State University of New Jersey who provided constructive suggestions to this project.

Funding

This study was supported by the National Natural Science Foundation of China (NSFC, 31671350, 31970568), National Science and Technology Major Project (2018ZX10712001-018-002), Programs of the Chinese Academy of Sciences (QYZDY-SSW-SMC017; Y8YZ02E001), and CAMS Innovation Fund for Medical Science (CIFMS, 2017-12 M-2-005), the Chinese National Key Technologies R&D Program for the 13th Five-year Plan (Grant 2017ZX10202102-004-001), CAMS Innovation Fund for Medical Sciences (CAMS-12 M) (2017-I2M-1-014).

Author contributions

L.W. and H.Y., conceived the project and were in charge of participants enrolment, sample collection, and clinical assay; M.Q., performed the experiment, data analysis, and draft the manuscript; L.J., conducted animal experiments and flow cytometry; Z.H., performed the animal experiments; Q.Z. and S.X, participated in sample collection and clinical assay, W.J., S.C., and C.Y., in charge of bacterial culture and animal experiments, J.X, Y.J, and R.JP., revised the manuscript; K.Y., conceived and designed the project, performed data analysis, wrote and edited the manuscript. All the authors have revised and approved the manuscript submission.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2021.05.021.

References

[1] Ghosn J, Taiwo B, Seedat S, Autran B, Katlama C. Hiv. Lancet 2018;392(10148):685–97.
[2] Prabhu, S., J.I. Harwell, and N. Kumarasamy, Advanced HIV: diagnosis, treatment, and prevention. Lancet HIV, 2019. 6(8): p. e540-e551.
[3] Sereti, I., et al., Prospective International Study of Incidence and Predictors of Immune Reconstitution Inflammatory Syndrome and Death in People Living With Human Immunodeficiency Virus and Severe Lymphopenia. Clin Infect Dis, 2020. 71(3): p. 652-660.
[4] Battegay M, Nüesch R, Hirschel B, Kaufmann GR. Immunological recovery and antiretroviral therapy in HIV-1 infection. Lancet Infect Dis 2006;6(5):280–7.
[5] Serrano-Villar, S., et al., HIV-infected individuals with low CD4/CD8 ratio despite effective antiretroviral therapy exhibit altered T cell subsets, heightened CD8+ T cell activation, and increased risk of non-AIDS morbidity and mortality. PLoS Pathog, 2014. 10(5): p. e1004078.
[6] Lu W et al. CD4:CD8 ratio as a frontier marker for clinical outcome, immune dysfunction and viral reservoir size in virologically suppressed HIV-positive patients. J Int AIDS Soc 2015;18:20052.
[7] Ouyang J, Isnard S, Lin J, Fombuena B, Marette A, Routy B, et al. Metformin effect on gut microbiota: insights for HIV-related inflammation. AIDS Res Ther 2020;17(1). https://doi.org/10.1186/s12981-020-00267-2.
[8] Cox AJ, West NP, Cripps AW. Obesity, inflammation, and the gut microbiota. Lancet Diabetes Endocrinol 2015;3(3):207–15.
[9] Ferrucci L, Fabbri E. Inflammageing: chronic inflammation in ageing, cardiovascular disease, and frailty. Nat Rev Cardiol 2018;15(9):505–22.
[10] Veazey RS et al. Gastrointestinal tract as a major site of CD4+ T cell depletion and viral replication in SIV infection. Science 1998;280(5362):427–31.
[11] Bandera, A., et al., Altered gut microbiome composition in HIV infection: causes, effects and potential intervention. Curr Opin HIV AIDS, 2018. 13(1): p. 73-80.
[12] McHardy IH, Li X, Tong M, Ruegger P, Jacobs J, Borneman J, et al. HIV Infection is associated with compositional and functional shifts in the rectal mucosal microbiota. Microbiome 2013;1(1):26. https://doi.org/10.1186/2049-2618-1-26.
[13] Lozupone C, Li M, Campbell T, Flores S, Linderman D, Gebert M, et al. Alterations in the gut microbiota associated with HIV-1 infection. Cell Host Microbe 2013;14(3):329–39.
[14] Dillon SM, Frank DN, Wilson CC. The gut microbiome and HIV-1 pathogenesis: a two-way street. AIDS 2016;30(18):2737–51.
[15] Dillon SM, Lee EJ, Kotter CV, Austin GL, Dong Z, Hecht DK, et al. An altered intestinal mucosal microbiome in HIV-1 infection is associated with mucosal and systemic immune activation and endotoxemia. Mucosal Immunol 2014;7(4):983–94.
[16] Peterson TE, Baker JV. Assessing inflammation and its role in comorbidities among persons living with HIV. Curr Opin Infect Dis 2019;32(1):8–15.
[17] Lu W et al. Association Between Gut Microbiota and CD4 Recovery in HIV-1 Infected Patients. Front Microbiol 2018;9:1451.
[18] Fischbach MA. Microbiome: Focus on Causation and Mechanism. Cell 2018;174(4):785–90.
[19] Brüssow H. Problems with the concept of gut microbiota dysbiosis. Microb Biotechnol 2020;13(2):423–34.
[20] Zhang M, Chu Y, Meng Q, Ding R, Shi X, Wang Z, et al. A quasi-paired cohort strategy reveals the impaired detoxifying function of microbes in the gut of autistic children. Sci Adv 2020;6(43):eaba3760. https://doi.org/10.1126/sciadv.aba3760.
[21] McBride JA, Striker R, Coyne CB. Imbalance in the game of T cells: What can the CD4/CD8 T-cell ratio tell us about HIV and health?. PLoS Pathog 2017;13(11):e1006624. https://doi.org/10.1371/journal.ppat.1006624.
[22] Trickey, A., et al., CD4:CD8 Ratio and CD8 Count as Prognostic Markers for Mortality in Human Immunodeficiency Virus-Infected Patients on Antiretroviral Therapy: The Antiretroviral Therapy Cohort Collaboration (ART-CC). Clin Infect Dis, 2017. 65(6): p. 959-966.
[23] Hughes RA et al. Long terms trends in CD4+ cell counts, CD8+ cell counts, and the CD4+:CD8+ ratio. AIDS 2018;32(10):1361–7.
[24] Abubucker, S., et al., Metabolic reconstruction for metagenomic data and its application to the human microbiome. PLoS Comput Biol, 2012. 8(6): p. e1002358.
[25] Kang KP, Lee S, Kang SK. D-lactic acidosis in humans: review of update. Electrolyte Blood Press 2006;4(1):53–6.
[26] Tarmy EM, Kaplan NO. Kinetics of Escherichia coli B D-lactate dehydrogenase and evidence for pyruvate-controlled change in conformation. J Biol Chem 1968;243(10):2587–96.
[27] Zhang C, Yin A, Li H, Wang R, Wu G, Shen J, et al. Dietary Modulation of Gut Microbiota Contributes to Alleviation of Both Genetic and Simple Obesity in Children. EBioMedicine 2015;2(8):968–84.
[28] Claesson MJ, Jeffery IB, Conde S, Power SE, O'Connor EM, Cusack S, et al. Gut microbiota composition correlates with diet and health in the elderly. Nature 2012;488(7410):178–84.
[29] Fleming AL, Field EH, Tolaymat N, Cowdery JS. Age influences recovery of systemic and mucosal immune responses following acute depletion of CD4 T cells. Clin Immunol Immunopathol 1993;69(3):285–91.
[30] Geva-Zatorsky N, Sefik E, Kua L, Pasman L, Tan TG, Ortiz-Lopez A, et al. Mining the Human Gut Microbiota for Immunomodulatory Organisms. Cell 2017;168(5):928–943.e11.
[31] Ivanov II, Atarashi K, Manel N, Brodie EL, Shima T, Karaoz U, et al. Induction of intestinal Th17 cells by segmented filamentous bacteria. Cell 2009;139(3):485–98.
[32] An D, Oh S, Olszak T, Neves J, Avci F, Erturk-Hasdemir D, et al. Sphingolipids from a symbiotic microbe regulate homeostasis of host intestinal natural killer T cells. Cell 2014;156(1-2):123–33.
[33] Stein-Thoeringer CK, Nichols KB, Lazrak A, Docampo MD, Slingerland AE, Slingerland JB, et al. Lactose drives Enterococcus expansion to promote graft-versus-host disease. Science 2019;366(6469):1143–9.

[34] Cervantes-Barragan L et al. Lactobacillus reuteri induces gut intraepithelial CD4(+)CD8alphaalpha(+) T cells. Science 2017;357(6353):806–10.
[35] Haas R, Smith J, Rocher-Ros V, Nadkarni S, Montero-Melendez T, D'Acquisto F, et al. Lactate Regulates Metabolic and Pro-inflammatory Circuits in Control of T Cell Migration and Effector Functions. PLoS Biol 2015;13(7):e1002202. https://doi.org/10.1371/journal.pbio.100220210.1371/journal.pbio.1002202.g00110.1371/journal.pbio.1002202.g00210.1371/journal.pbio.1002202.g00310.1371/journal.pbio.1002202.g00410.1371/journal.pbio.1002202.g00510.1371/journal.pbio.1002202.g00610.1371/journal.pbio.1002202.g00710.1371/journal.pbio.1002202.s00110.1371/journal.pbio.1002202.s00210.1371/journal.pbio.1002202.s00310.1371/journal.pbio.1002202.s00410.1371/journal.pbio.1002202.s00510.1371/journal.pbio.1002202.s00610.1371/journal.pbio.1002202.s00710.1371/journal.pbio.1002202.s00810.1371/journal.pbio.1002202.s00910.1371/journal.pbio.1002202.s01010.1371/journal.pbio.1002202.s011.
[36] Pucino V, Certo M, Bulusu V, Cucchi D, Goldmann K, Pontarini E, et al. Lactate Buildup at the Site of Chronic Inflammation Promotes Disease by Inducing CD4 (+) T Cell Metabolic Rewiring. Cell Metab 2019;30(6):1055–1074.e8.
[37] Brooks GA. The Science and Translation of Lactate Shuttle Theory. Cell Metab 2018;27(4):757–85.
[38] Flint HJ, Duncan SH, Scott KP, Louis P. Links between diet, gut microbiota composition and gut metabolism. Proc Nutr Soc 2015;74(1):13–22.
[39] Tenore, S.B., et al., Immune effects of Lactobacillus casei Shirota in treated HIV-infected patients with poor CD4+ T-cell recovery. AIDS, 2020. 34(3): p. 381-389.
[40] Iraporda C, Errea A, Romanin DE, Cayet D, Pereyra E, Pignataro O, et al. Lactate and short chain fatty acids produced by microbial fermentation downregulate proinflammatory responses in intestinal epithelial cells and myeloid cells. Immunobiology 2015;220(10):1161–9.
[41] Mokoena MP. Lactic Acid Bacteria and Their Bacteriocins: Classification, Biosynthesis and Applications against Uropathogens: A Mini-Review. Molecules 2017;22(8):1255. https://doi.org/10.3390/molecules22081255.
[42] Suez J, Zmora N, Segal E, Elinav E. The pros, cons, and many unknowns of probiotics. Nat Med 2019;25(5):716–29.
[43] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;30(15):2114–20.
[44] Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9(4):357–9.
[45] Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, et al. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat Methods 2015;12(10):902–3.
[46] Truong DT, Tett A, Pasolli E, Huttenhower C, Segata N. Microbial strain-level population structure and genetic diversity from metagenomes. Genome Res 2017;27(4):626–38.
[47] Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 2014;30(9):1312–3.
[48] Caspi R, Altman T, Dreher K, Fulcher CA, Subhraveti P, Keseler IM, et al. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res 2012;40(D1): D742–53.
[49] Su G, Morris JH, Demchak B, Bader GD. Biological network exploration with Cytoscape 3. Curr Protoc Bioinformatics 2014;47(1). https://doi.org/10.1002/0471250953.2014.47.issue-110.1002/0471250953.bi0813s47.
[50] Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J Roy Stat Soc: Ser B (Methodol) 1995;57(1):289–300.