Received: 20 June 2019 | Revised: 4 October 2019 | Accepted: 5 October 2019

DOI: 10.1111/ajt.15654

ORIGINAL ARTICLE

AJT

A gut microbiota score predicting acute graft-versus-host disease following myeloablative allogeneic hematopoietic stem cell transplantation

Lijie Han[1,2] | Ke Zhao[2] | Yuanyuan Li[3] | Haohao Han[1] | Lizhi Zhou[4] | Ping Ma[1] | Zhiping Fan[2] | Hui Sun[1] | Hua Jin[2] | Zhongxing Jiang[1] | Qifa Liu[2,5] | Jie Peng[6]

[1]Department of Hematology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China

[2]Department of Hematology, Nanfang Hospital, Southern Medical University, Guangzhou, China

[3]School of Foreign Languages, Henan University of Chinese Medicine, Zhengzhou, China

[4]Department of Biostatistics, School of Public Health, Southern Medical University, Guangzhou, China

[5]Guangdong Provincial Key Laboratory of Construction and Detection in Tissue Engineering, Southern Medical University, Guangzhou, China

[6]Department of Oncology, the Second Affiliated Hospital of Guizhou Medical University, Kaili, China

Correspondence
Jie Peng
Email: sank44@sina.com

Qifa Liu
Email: liuqifa628@163.com

Zhongxing Jiang
Email: jiangzx@zzu.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: 81770190, 81870144, 81600141, 81700176 and 2016A030310390; Sciences Foundation of Guizhou Province, Grant/Award Number: 20185579-X; National Primary Research Plan, Grant/Award Number: 2017YFA105500 and 2017YFA105504; Natural Science Foundation of Guangdong Province; Medical Science and Technology Planning Project of Henan Province, Grant/Award Number: 201403227

Although studies have reported that intestinal microbiota are associated with acute graft-versus-host disease (aGVHD), they lacked a satisfactory method for predicting aGVHD. We collected stool and blood samples at day 15 posttransplant from 150 patients from two centers who underwent myeloablative conditioning allogeneic hematopoietic stem cell transplantation (allo-HSCT). Stool microbiota were detected by 16S ribosomal RNA gene sequencing; inflammatory factors and T lymphocytes were detected by multiplex immunoassays and flow cytometry, respectively. A gut microbiota score (GMS) from a LASSO (least absolute shrinkage and selection operator) model was developed and validated to predict aGVHD. In the discovery cohort, the GMS could predict II-IV aGVHD (area under the receiver operating characteristic [ROC] curve [AUC] = 0.904, $P < .0001$). Furthermore, the validation model was consistent with the discovery set (AUC = 0.887, $P < .0001$). Regulatory T/T-helper 17 (Treg/Th17) cells ratio in the low GMS subgroup was higher compared with the high GMS ($P = .012$), and the validation set is consistent with the discovery set ($P = .003$). In addition, high cytokine levels were associated with high GMS. In conclusion, the GMS at neutrophil engraftment could predict aGVHD, and it was a potential and novel method. The GMS was associated with the inflammatory factor and Treg/Th17 balance.

Abbreviations: aGVHD, acute graft-versus-host disease; allo-HSCT, allogeneic hematopoietic stem cell transplantation; AUC, area under the curve; BMT, bone marrow transplantation; BuCY, busulfan and cyclophosphamide; CMV, cytomegalovirus; CR, complete remission; CsA, cyclosporin A; DLI, donor lymphocyte infusion; GEP, gene expression profiling; GMS, gut microbiota score; IFI, invasive fungal infection; LASSO, least absolute shrinkage and selection operator; LDA, linear discriminant analysis; LEfSe, linear discriminant analysis effect size; MMF, mycophenolate; MTX, methotrexate; MUD, matched unrelated donor; OUT, operational taxonomic unit; PBSCT, peripheral blood stem cell to transplantation; PCR, polymerase chain reaction; ROC, receiver operating characteristic; TBI, total body irradiation.

Lijie Han, Ke Zhao, Yuanyuan Li, Haohao Han, and Lizhi Zhou contributed equally to this work.

KEYWORDS
graft survival, graft-versus-host disease (GVHD), graft-versus-leukemia (GVL)/graft versus tumor, hematology/oncology, immunosuppression/immune modulation, translational research/science

1 | INTRODUCTION

Acute graft-versus-host disease (aGVHD) remains a major cause of death for patients who undergo allogeneic hematopoietic stem cell transplantation (allo-HSCT).[1,2] Approximately one-third of patients fail to response to initial corticosteroid therapy, and survival is poor among the remaining patients.[2-4] Mortality is associated with clinical severity in current grading systems.

Once aGVHD occurs, serious complications, such as infections and organ dysfunction, may appear subsequently, resulting in high mortality of recipients after allo-HSCT.[2-4] So it is necessary to predict aGVHD occurrence to take proper measures preemptively for patients. Recently, gut microbiota serving as a biomarker for predicting aGVHD have been studied.[5] One new report from Seattle has suggested that the gut microbiota at neutrophil engraftment could predict severe aGVHD after allo-HSCT.[5] However, because of the heterogeneity of patients from different cohorts, models from previous studies were developed in single institutions and not validated in other centers; thus, it is pivotal to build a predictive model for aGVHD that can be validated by other centers.

Machine learning is an emerging and promising field that involves the reduction of large high-dimensional features from databases of gene expression profiling (GEP), aberrant DNA methylation, and cytokine genes, predictive of GVHD and prognosis.[6-8] This algorithm can effectively prevent overfitting, particularly that caused by too many variables. In previous studies, we used the machine learning method named "LASSO" to solve the high-dimensional features and built models to precisely predict clinical outcome in other cancers, such as hepatocellular carcinoma.[9,10] Therefore, we speculated that using a LASSO model based on 16S ribosomal RNA (rRNA) gene sequencing of gut microbiota could precisely predict aGVHD occurrence. Many machine learning models have a prominent drawback, in that it is difficult to explain performance, particularly when constructing clinical predictive models. Therefore, it is necessary to understand the associated model mechanism based on the gut microbiota and inflammation in aGVHD.

In this study, we prospectively collected stool and blood samples from patients undergoing myeloablative conditioning allo-HSCT at day 15 ± 1 (neutrophil engraftment) posttransplant from two medical centers. Stool microbiota and the inflammatory factor levels in the blood were examined by 16S rRNA gene sequencing and ProcartaPlex multiplex immunoassays, respectively. aGVHD was determined by a retrospective review of clinical data. We used the LASSO method and ultimately developed a model based on gut microbiota to predict the probability of aGVHD. This algorithm defines a gut microbiota score (GMS) with distinct risk that may eventually prove useful as a predictor for aGVHD. We also uncovered an association between the GMS, inflammatory factors, immune status, and aGVHD, which was valuable for better understanding of the potential mechanism of this model.

2 | PATIENTS AND METHODS

2.1 | Patients and study design

An original cohort of 178 patients who underwent myeloablative allo-HSCT was recruited from Nanfang Hospital, Southern Medical University, and the First Affiliated Hospital of Zhengzhou University of Chinese between December 2016 and May 2018. Three patients died before engraftment, 19 patients failed to examine fecal or blood specimens successfully, and six patients (including three with hematologic relapse and three patients only with Minimal Residual Disease positive) received donor lymphocyte infusion (DLI) within 100 days posttransplant were excluded from this study. So the final analytical data included 150 cases (Figure 1). This study was approved by the ethical committee of Nanfang Hospital, Southern Medical University, and the First Affiliated Hospital of Zhengzhou University, China. After approval of the study, consent was obtained from the participants for biospecimen collection and analysis. This study was conducted in accordance with the Declaration of Helsinki.[11]

The discovery cohort comprised 102 patients from Nanfang Hospital, and the validation cohort consisted of an independent series of 48 patients from the First Affiliated Hospital of Zhengzhou University. This project was approved by the institutional review board.

2.2 | Samples

Stool and blood samples were collected from patients who underwent allo-HSCT at preconditioning, at day 0 and day 15 posttransplantation. The stool and plasma samples were tagged and stored at −80°C until retrieved for DNA extraction or detection, respectively. T lymphocyte subsets in blood were directly examined by flow cytometry.[11] If there were more than 6 hours between collection and disposition, the samples were discarded.

2.3 | Conditioning and GVHD prophylaxis

Three myeloablative conditioning regimens were used for the patients including the following: two standard myeloablative regimens (BuCY: busulfan + cyclophosphamide; TBI + CY: total body irradiation + CY), and a sequential intensified regimen (fludarabine + cytarabine plus TBI + Cy

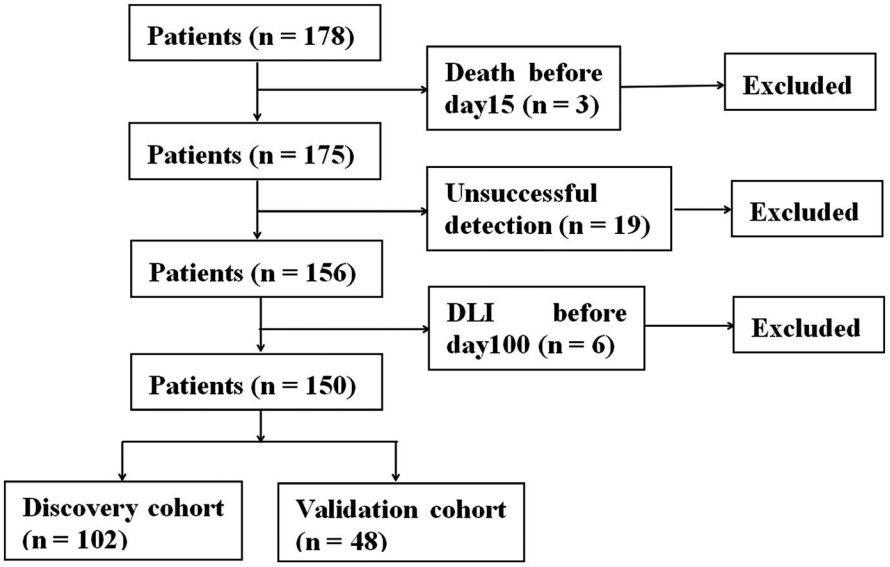

FIGURE 1 Diagram of patient groups enrolled in this study

+etoposide).[11-13] The selection of conditioning regimens was based on disease type and status at transplantation. Generally, patients with acute myelogenous leukemia in complete remission (CR) received BuCY, and those with acute lymphoblastic leukemia in CR received TBI + CY, whereas those in non-CR were given the intensified regimen. In addition, some high-risk patients were also given the intensified regimen.[11-13]

Cyclosporin A (CsA) and methotrexate (MTX) (on days +1, +3, and +6) were administered to patients who underwent matched sibling donor transplants for GVHD prophylaxis. CsA + MTX +ATG (ATG, thymoglobulin; Genzyme, Cambridge) was administered to patients who underwent a MUD (matched unrelated donor) transplant. CsA + MTX +ATG + mycophenolate (MMF) was used for patients who underwent a haploidentical donor transplant for GVHD.[11-13]

2.4 | Infection prophylaxis and treatment

At our institution, oral sulfamethoxazole and norfloxacin were used in all cases for infection prophylaxis.[11,12] Oral sulfamethoxazole was administered from 7 days before conditioning to 1-year posttransplant. Although norfloxacin was used for 7 days until conditioning start, ganciclovir was given for the prophylaxis and treatment of cytomegalovirus (CMV) infections, and acyclovir was administered for other viruses. Antifungal agents were used for fungal infection prophylaxis. Fluconazole (0.3 g/d) or itraconazole (0.4 g/kg/d) was administered up to 60 days posttransplant to patients with no history of invasive fungal infection (IFI), and those with a history of IFI received voriconazole (0.4 g/d), itraconazole (0.4 g/d), caspofungin (50 mg/d), or ambisome (2 mg/kg/d) intravenously. Oral voriconazole or itraconazole was substituted for intravenous treatment when the peripheral white blood cell count was greater than 2.0×10^9/L, and it was discontinued 90 days posttransplant.

Generally, patients were given imipenem, or it was combined with an amikacin as a first-line antibiotic for fever during neutropenia. Vancomycin or piperacillin/tazobactam was used as a second-line antibiotic. Other antibiotics were variably administered, and these were used for a minority of patients. For instance, tigecycline was administered occasionally for noneffective second-line antibiotics for bacterial infections.

2.5 | Clinical metadata

All clinical data including aGVHD, neutrophil engraftment, and prophylactic or therapeutic antibiotics were determined by retrospective review of clinical charts by individuals blinded to the microbiota results of the participants. aGVHD was defined according to the 1994 Consensus Conference on aGVHD Grading and graded from I to IV grades,[14] which included the whole of aGVHD onset before day 100 posttransplantation, whereas late-onset aGVHD (that is beyond day 100 posttransplant), or relapsed aGVHD, or DLI-associated aGVHD was not considered. Neutrophil engraftment was defined as the third day an absolute neutrophil count was greater than 0.5×10^9/L posttransplant. Based on aGVHD, the study groups were categorized into grades 0-I aGVHD (non-aGVHD) and grades II-IV aGVHD (aGVHD).

2.6 | 16S rRNA gene sequencing for fecal specimens

For each fecal specimen, DNA was extracted and purified, and the V3 to V4 region of the 16S rRNA gene was polymerase chain reaction (PCR)–amplified using modified universal bacterial primers.[11] Microbiome DNA concentrations were measured using a real-time quantitative PCR assay targeting the V3-V4 region of the 16S rRNA gene. Purified PCR products were sequenced with the Hiseq2500 PE250 platform.[11] Sequence data were compiled and processed using mothur version 1.31.2 is a software of process and analysis for 16S rRNA gene. Sequence data were screened and filtered for quality and then aligned to the full-length 16S rRNA gene using the SILVA reference alignment as a template. Sequences were grouped into operational taxonomic units (OTUs) of 97% similarity.[11]

2.7 | Microbial diversity and different microbiota taxa analysis

Microbial diversity was estimated by the inverse Simpson index, an ecological estimate of diversity calculated to represent the reciprocal of the expected probability of randomly selected bacterial sequences belonging to the same OTU.[15] A nonparametric test (Mann-Whitney) was used to compare the statistical significance of the groups. Phylogenetic classification at the family level was analyzed based on a naive Bayesian classification scheme and Greengenes reference database.[16] A nonparametric factorial Kruskal-Wallis rank-sum test was used to identify different microbiota taxa to detect features that were significantly different and in abundance between the groups. Linear Discriminant Analysis (LDA) Effect Size (LEfSe) analysis was performed using LEfSe software. The nonparametric factorial Kruskal-Wallis rank-sum test was used in LEfSe analysis to detect characteristics between the aGVHD and non-aGVHD groups. The effect sizes of the identified characteristics were then analyzed with an LDA model.[11]

2.8 | Development and validation of a gut microbiota score (GMS) for predicting aGVHD

As reported in previous studies,[17-19] least absolute shrinkage and selection operator (LASSO) is a powerful algorithm for solving the problem of regression with high-dimensional predictors. In this study, the LASSO logistic regression model was used to choose the most important gut microbiota from the discovery cohort. For the sum of the absolute values of the parameters being bound by a constant, a log partial likelihood subject is minimized via using LASSO logistic regression:

$$\widehat{\beta} = \text{argmin}\ell(\beta), \text{subject to } \sum |\beta j| \le s,$$

where $\widehat{\beta}$ is the acquired parameters, $\ell(\beta)$ is the log partial likelihood of the logistic regression model, $s > 0$ is a constant LASSO method shrink coefficient, and some coefficients were reduced to zero using the absolute constraint. Hence, the LASSO algorithm can be used to reduce and select features. In our study, the standardized constraint parameters were set as to −2.835, and LASSO selected 20 nonzero coefficients. Then, a logistic regression model was obtained with its outcome being the odds ratio for individuals. Finally, a gut microbiota score (or GMS) formula was defined based on selected features from a total of 42 gut microbiota. A GMS for predicting aGVHD was then constructed via the gut microbiota. Using receiver operating characteristic (ROC) analysis, we estimated the performance of the GMS model in the discovery and independent validation sets. The optimal cut-off value for GMS was determined using the Youden index (YI) in the prediction aGVHD.

2.9 | Inflammatory factors

Levels of interleukin-1β (IL-1β), IL-6, IL-17A, and tumor necrosis factor-alpha (TNF-α) in plasma were detected with the ProcartaPlex multiplex immunoassay kit according to the manufacturer's protocol (eBioscience Corporation). Plates were read with the Luminex 200™ system (Luminex Corporation) and analyzed using ProcartaPlex Software (eBioscience Corporation).

2.10 | T lymphocyte subsets

T lymphocyte subsets were detected as described previously.[11] Th17 cells were examined using the Intracellular Staining Kit (BD Pharmingen). Cells were incubated for 5 hours with phorbol-12-myristate-13-acetate (50 ng/mL) plus ionomycin (2.5 μg/mL, all reagents from Sigma Chemical) to stimulate IL-17A production, and the samples were supplemented with Golgistop (0.7 μL/mL) during the last 4 hours to trap proteins in the cytoplasm. The proportion of T lymphocyte subsets (Treg cells: CD45+CD3+CD4+CD25+Foxp3+; Th17 cells: CD45+CD3+CD4+CD8−IL-17A+) in peripheral blood was analyzed by flow cytometry.[11]

2.11 | Statistical analysis

All statistical analyses were conducted using R statistical software (https://www.r-project.org/, version 3.5.1) and Graph Pad Prism (version 7.0). Considering the competing risks of GVHD and relapse, cumulative incidence curves in a competing risk setting were applied to calculate probabilities of aGVHD and chronic GVHD (cGVHD). Groups were compared with the Gray test.[20] Probabilities of overall survival (OS), disease-free survival (DFS), and relapse were evaluated using the Kaplan-Meier estimate. The correlations of the gut microbiota were analyzed in discovery and validation cohorts using the "heat map" package. The "glmnet" package was used to perform the LASSO algorithm. Scatter dot and box plots indicate median and 95% confidence intervals (CIs). Statistical methods including Mann-Whitney *U* and correlation analysis were applied in the data analysis of GMS. Univariate and multivariate analyses of aGVHD were performed using the "rms" package. A two-sided $P < .05$ was considered statistically significant.

3 | RESULTS

3.1 | Clinical characteristics

The baseline characteristics of patients in this study are shown in Table 1. The median age was 33 years (range: 14-56 years). Most donors (74.7%) were male; No. 2-3 HLA-mismatched patients accounted for 16.0% of the patients, 4-5 mismatched patients were 34.0%, and 0-1 mismatched patients were 50.0%. The total dose of ATG was 7 and 7.5 mg/kg in MUD and haploidentical transplant patients, respectively. Intensified conditioning was administered to 30.7% of the patients, and 53.3% patients had β-lactam antibiotics prescribed, 65.3% recipients possessed a high inverse Simpson index at engraftment, and 42.0% cases experienced II-IV aGVHD (Table 1). The graft source

TABLE 1 Characteristics of patients in the discovery and validation sets

Variable	All patients no. (n = 150)	Discovery set (n = 102)	Validation set (n = 48)	*P* value
Sex				
Female	57 (38.0%)	41 (40.2%)	16 (33.3%)	.419
Male	93 (62.0%)	61 (59.8%)	32 (66.7%)	
Age (y)				
≤33	70 (46.7%)	43 (42.2%)	27 (56.3%)	.107
>33	80 (53.3%)	59 (57.84%)	21 (43.8%)	
Donor sex				
Female	38 (25.3%)	23 (22.5%)	15 (31.3%)	.253
Male	112 (74.7%)	79 (77.5%)	33 (68.7%)	
HLA-mismatched				
0-1	75 (50.0%)	56 (54.9%)	19 (39.6%)	.214
2-3	24 (16.0%)	15 (14.7%)	9 (18.7%)	
4-5	51 (34.0%)	31 (30.4%)	20 (41.7%)	
Underlying disease				
ALL	56 (37.3%)	42 (41.2%)	14 (29.2%)	.331
AML	87 (58.0%)	55 (53.9%)	32 (66.7%)	
MDS	7 (4.7%)	5 (4.9%)	2 (4.1%)	
BUCY				
Yes	78 (52.0%)	54 (54.9%)	24 (50.0%)	.737
No	72 (48.0%)	48 (45.1%)	24 (50.0%)	
Intensified conditioning				
Yes	46 (30.7%)	31 (30.4%)	15 (31.3%)	.915
No	104 (69.3%)	71 (69.6%)	33 (68.7%)	
Graft source, No.				
PBSCT	83 (55.3%)	60 (58.8%)	23 (47.9%)	.210
BMT+PBSCT	67 (44.7%)	42 (41.2%)	25 (52.1%)	
Antibiotics (%)				
β-Lactam	80 (36.7%)	63 (39.4%)	17 (29.3%)	.229
Vancomycin	55 (25.2%)	36 (22.5%)	19 (32.8%)	
Amikacin	83 (38.1%)	61 (38.1%)	22 (37.9%)	
Bloodstream infection				
Yes	28 (18.7%)	17 (16.7%)	11 (22.9%)	.360
No	122 (81.3%)	85 (83.3%)	37 (77.1%)	
Inverse Simpson index				
≥2	98 (65.3%)	67 (65.7%)	31 (64.6%)	.895
<2	52 (34.7%)	35 (34.3%)	17 (35.4%)	
GVHD II-IV				
Yes	63 (42.0%)	43 (42.2%)	20 (41.7%)	.955
No	87 (58.0%)	59 (57.8%)	28 (58.3%)	

P value is derived from the difference between the discovery data set and the validation data set in either the clinical characteristics.

ALL, acute lymphoblastic leukemia; AML, acute myelogenous leukemia; BMT, bone marrow transplantation; BuCY, busulfan and cyclophosphamide; GVHD, graft-versus-host-disease; HID, haploidentical donor; HLA, Human Leukocyte Antigen; MSD, matched sibling donor; MNC, mononuclear cell; MDS, myelodysplastic syndrome; PBSC, peripheral blood stem cell.

was peripheral blood stem cell (PBSC) or bone marrow transplantation (BM) + PBSC in the discovery (58.8% or 41.2%) and validation sets (47.9% or 52.1%), respectively. None of the variables were found to be significantly different in the two cohorts ($P > .05$).

3.2 | GVHD, survival, and relapse

The median day of aGVHD onset was day 20 (10-62) posttransplant; aGVHD occurred in 5 cases before day 14 posttransplant (samples collected time) characterized by the symptom of aGVHD gradually and developed obviously, not significant at onset. The overall cumulative incidences of grade II-IV aGVHD by day +100 posttransplant were 42.2% (37.3%-47.1%) and 41.7% (34.6%-48.8%), respectively, for the discovery and validation cohorts ($P = .926$, Figure 2A). Similarly, the cumulative incidences of grade III-IV aGVHD were 12.7% (9.4%-16.0%) and 18.8% (13.2%-24.4%), respectively, for the discovery and validation sets ($P = .340$, Figure 2B).

The median follow-up was 17.5 months (range: 2-24 months). The overall 2-year cumulative incidence of cGVHD posttransplant was 43.4% (37.5%-49.3%) and 35.2% (27.4%-43.0%), respectively, for the discovery and validation cohorts ($P = .459$, Figure S1A), and that of extensive cGHVD was 11.5% (8.0%-15.0%) and 14.7% (9.1%-20.3%) ($P = .646$, Figure S1B).

The cumulative incidences for OS, DFS, and relapse are shown in Figure S2. No significant difference in the 2-year cumulative OS and DFS posttransplant was identified between the discovery and validation sets (OS: 77.1% [72.7%-81.5%] vs 73.8% [67.2%-80.4%], $P = .698$; DFS: 68.3% [63.5%-73.1%] vs 71.8% [65.1%-78.5%], $P = .622$; respectively, Figure S2A,B). The 2-year cumulative incidence of relapse posttransplant was 15.0% (11.1%-18.9%) and 11.7% (6.8%-16.6%), respectively, for the discovery and validation sets ($P = .612$, Figure S2C), including 3 cases that relapsed within day 100 posttransplant for the whole cohorts.

3.3 | aGVHD-associated gut microbiota at engraftment in the discovery and validation cohorts

The different richness of the microbiota in fecal samples was shown in Figure 3. More microbiota in stool was observed in the non-aGVHD group than in the aGVHD group in the two independent cohorts. These results indicate that the taxonomic composition of the fecal microbiota was less complex in the aGVHD group with fewer distinct members in the discovery and validation cohorts.

3.4 | Most microbiota were consistent between the two cohorts

As shown in Figure 4A,B, the gut bacteria abundances were consistent between the two cohorts. In the two heat maps generated from the discovery and validation sets, the blue color indicates a positive correlation

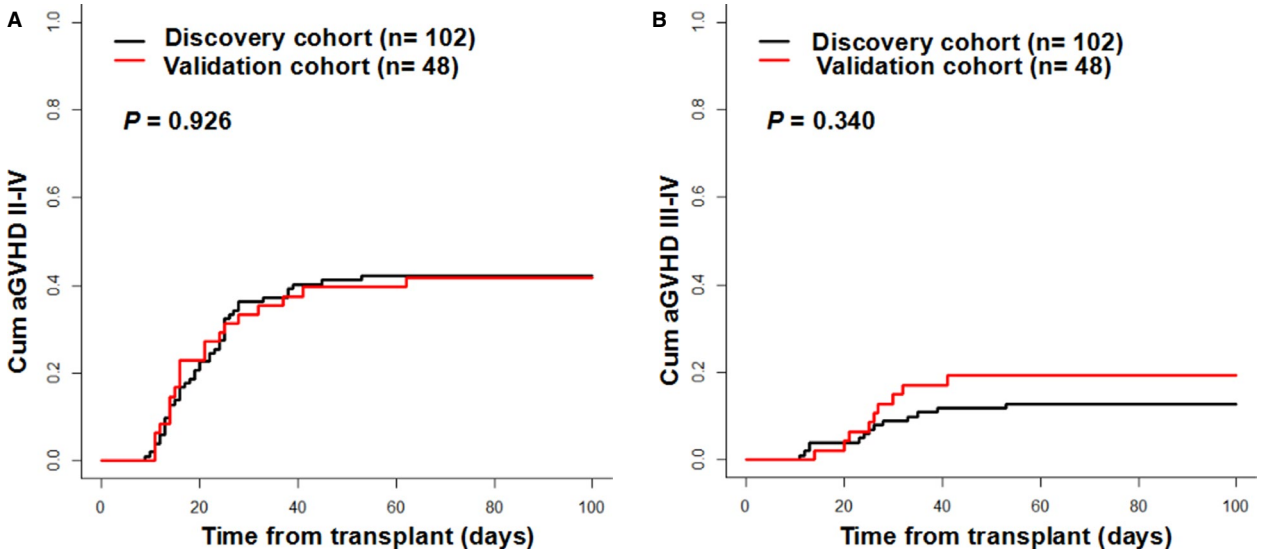

FIGURE 2 Cumulative incidence of aGVHD. A, The cumulative incidence of grade II-IV aGVHD by day +100 posttransplant was 42.2% and 41.7% for the discovery and validation cohorts, respectively (*P* = .926). B, The incidence of grade III-IV aGVHD by day +100 posttransplant was 12.7% and 18.8% for the discovery and validation cohorts, respectively (*P* = .340) [Color figure can be viewed at wileyonlinelibrary.com]

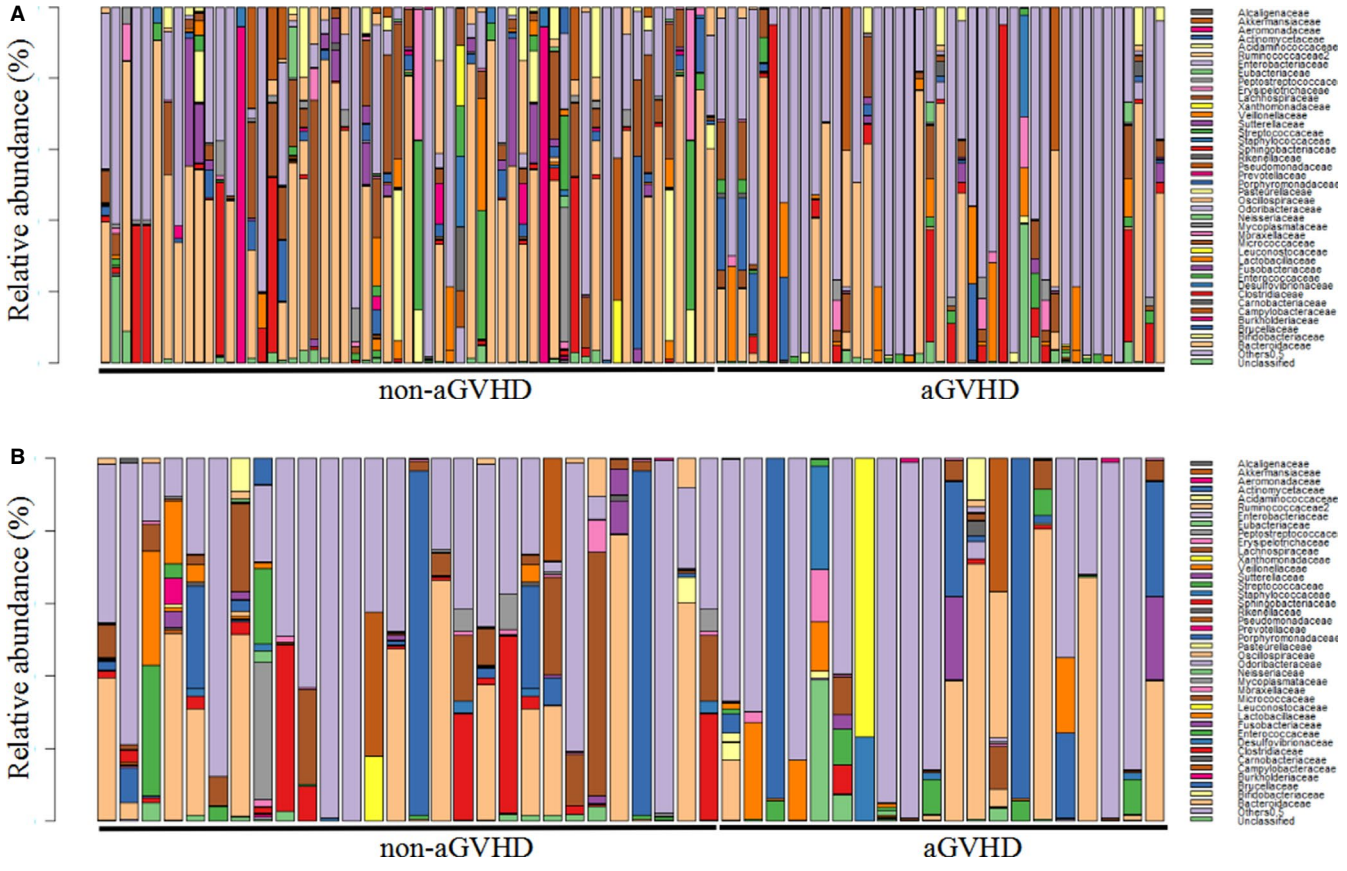

FIGURE 3 Intestinal ecosystems of patients at engraftment. Intestinal ecosystems of patients with aGVHD and non-aGVHD in the discovery set (n = 102) (A) and validation set (n = 48) (B). Each rank is a study subject, which represents the phylogenetic composition of each subject. The relative abundance of intestinal bacterial taxa is shown. aGVHD, acute graft-versus-host disease [Color figure can be viewed at wileyonlinelibrary.com]

and the red color indicates a negative correlation. The intensities of the colors represent the strength of the correlations. For example, the relative abundances of Lachnospiraceae and Peptostreptococcaceae were negatively correlated with aGVHD, whereas the relative abundance of Enterobacteriaceae was positively correlated with aGVHD. The results of the discovery set were consistent with that of the validation set. These results reveal that gut microbiota might interact with each other, and losing this balance would lead to the occurrence of aGVHD.

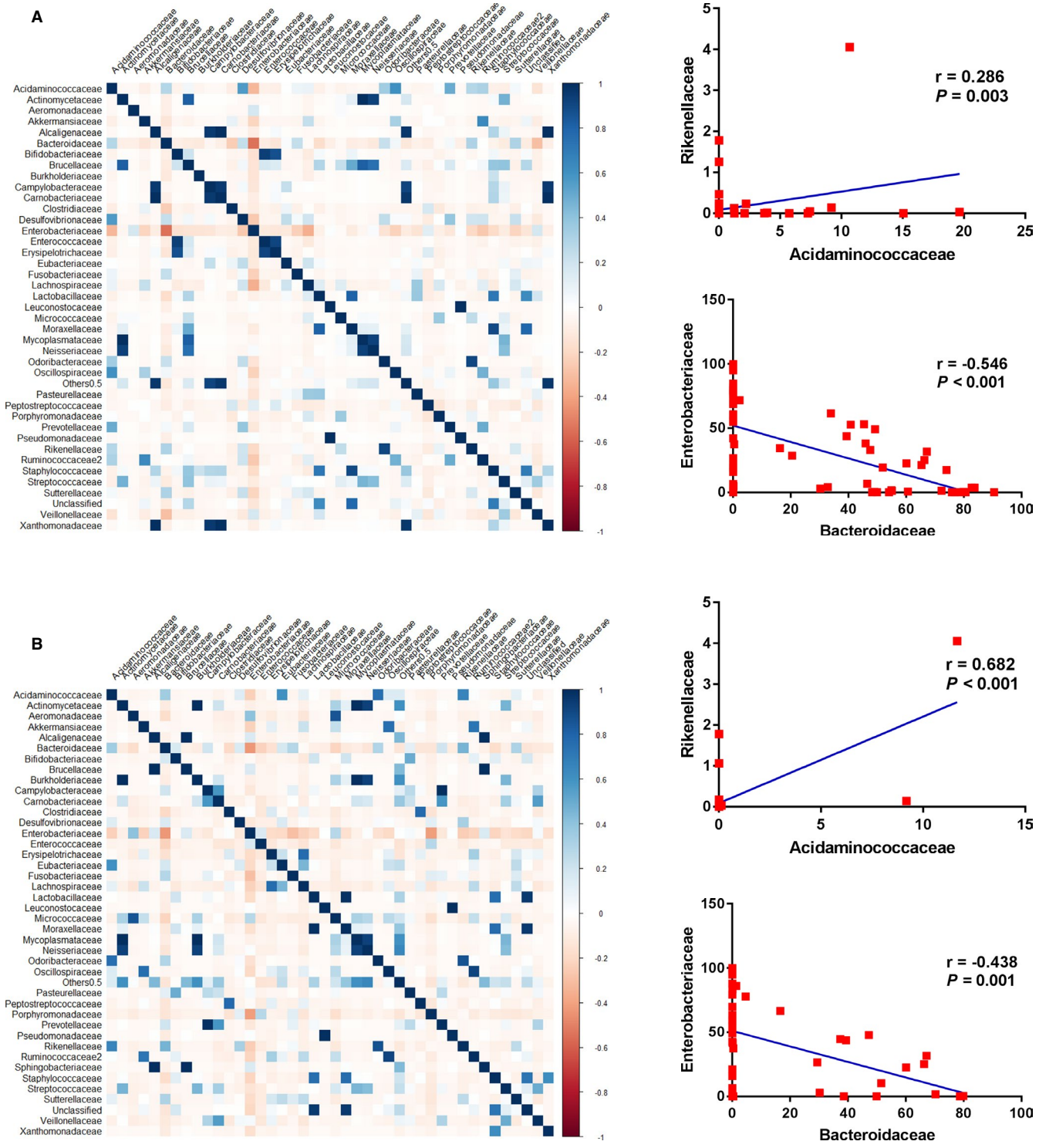

FIGURE 4 The correlation heatmaps of intestinal microbiota at engraftment. The correlation heatmaps of relative abundance of 42 intestinal bacteria are shown in the discovery set (A) and the validation set (B). The blue color represents positive correlation (0~1) and the red color represents negative correlation (−1~0). The depth of the color is related to the power of correlation. The associations of relative abundance of 4 intestinal bacteria are shown [Color figure can be viewed at wileyonlinelibrary.com]

3.5 | Gut microbiota score for predicting for aGVHD

To determine suitable predictors for aGVHD, the microbiota at engraftment posttransplant was evaluated. Feature selection and LASSO coefficient analysis of the 42 features of intestinal bacteria were performed. Based on 10-fold cross-validation via minimum criteria, 20 coefficients were chosen as the vertical line shown in the plot (Figure 5). Gut microbiota score (or GMS) calculation formula was shown in Table 2.

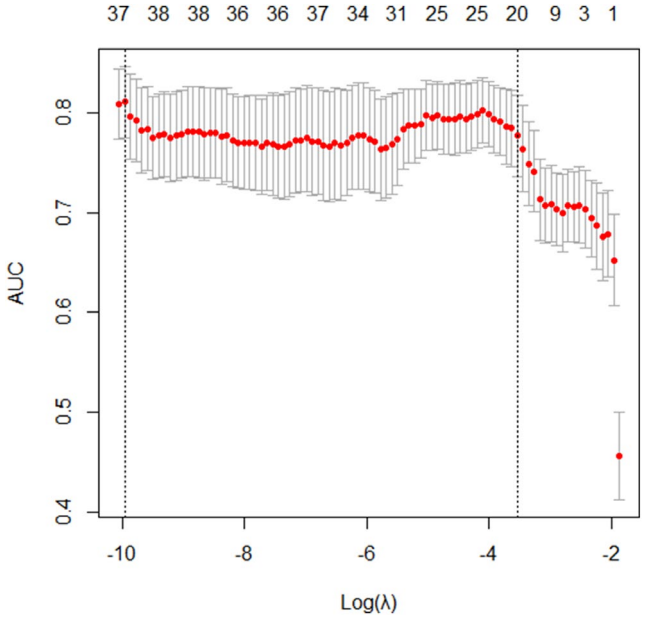

FIGURE 5 Tuning parameters for intestinal microbiota selection in the LASSO regression model. Feature selection and LASSO coefficient analysis of the 42 features of intestinal bacterial were performed. Based on 10-fold cross-validation via minimum criteria, the 20 coefficients were chosen as the vertical line presented in the plot. LASSO, least absolute shrinkage and selection operator [Color figure can be viewed at wileyonlinelibrary.com]

GMS is associated with inverse Simpson index, and the GMS could be a predictor for aGVHD. Inverse Simpson index was higher in the low GMS group compared with the high GMS group both in the discovery set (P = .001, Figure 6A) and in the validation set (P < .001, Figure 6B). Although not all of the bacteria could predict aGVHD, the inverse Simpson index and GMS of the organisms could predict subsequent aGVHD (Figure 6C,D). As determined by the area under the curve (AUC) in ROC plots, the inverse Simpson index and GMS of microbiota were determined. The results demonstrated that the AUC of the inverse Simpson index and GMS could be predictors for the subsequent development of aGVHD grades II-IV (AUC = 0.622, 95% CI: 0.505-0.739, P = .035 and AUC = 0.904, 95% CI: 0.848-0.960, P < .001, respectively; Figure 6C). Furthermore, the model for the GMS in the validation set also demonstrated a high AUC for predicting aGVHD and was consistent with the discovery set (AUC = 0.887, 95% CI :0.797-0.977, P < .001; Figure 6D). The AUC of the inverse Simpson index was 0.673 (95% CI: 0.520-0.826, P = .043, Figure 6D). Together, the efficacy of the GMS was significantly superior to the inverse Simpson index both in the discovery (P < .001) and validation sets (P = .005) (Figure 6C,D).

Consistently, the GMS was identified to be a predictor for the development of aGVHD grades III-IV. The ROC curve showed the predictive performance of the GMS in the discovery and validation sets for estimating grade III-IV aGVHD (Figure 6E,F). The results of ROC analysis showed high predictive performance of the GMS in predicting III-IV aGVHD (discovery cohort, AUC = 0.808, P < .001; validation cohort, AUC = 0.860, P = .001, respectively). The AUC of the inverse Simpson index was 0.704 (0.516-0.893) and 0.763 (0.610-0916) in the two cohorts, respectively. The predictive performance of GMS was higher than the inverse Simpson index, both in the discovery and validation sets, although no difference was reached (P = .346 and .316, respectively, Figure 6E,F).

TABLE 2 Gut microbiota score (GMS) calculation formula

Gut microbiota features	Coefficients	*P* value
Intercept	8.084e-01	.331
Lachnospiraceae	−1.046e-01	.010*
Peptostreptococcaceae	−2.923e-01	.021*
Enterobacteriaceae	8.598e-03	.370
Ruminococcaceae	6.512e-02	.458
Actinomycetaceae	−1.128e+01	.343
Akkermansiaceae	1.610e-03	.983
Alcaligenaceae	−2.107e+02	.432
Bacteroidaceae	−2.016e-02	.075
Bifidobacteriaceae	−3.489e-01	.076
Burkholderiaceae	−1.226e-01	.943
Carnobacteriaceae	3.554e-04	.999
Desulfovibrionaceae	−1.507e+00	.313
Micrococcaceae	8.105e-01	.239
Moraxellaceae	1.205e+02	.227
Mycoplasmataceae	−9.860e+01	.535
Odoribacteraceae	1.443e+00	.507
Prevotellaceae	−1.470e+01	.829
Pseudomonadaceae	−3.220e+01	.326
Veillonellaceae	1.312e-02	.705
Xanthomonadaceae	6.072e-02	.763

*P < .05.

In the discovery cohort, the cumulative incidence of II-IV aGVHD was lower for the low GMS subgroup than for the high GMS subgroup according to the GMS (21.7% [16.7%-26.7%] vs 84.8% [78.6%-91.0%], P < .001, Figure 7A); in the validation cohort, the incidence of II-IV aGVHD was consistent for the two subgroups (21.4% [13.6%-29.2%] vs 70.0% [59.8%-80.2%], P = .001, Figure 7B). Similarly, the incidence of III-IV aGVHD for the low GMS subgroup was also lower compared with the high GMS subgroup in the discovery cohort (4.3% [1.8%-6.8%] vs 30.3% [22.3%-38.3%], P < .001, Figure 7C); consistently, the incidence of III-IV aGVHD in the validation cohort was 3.6% (0.1%-7.1%) and 40.0% (29.0%-51.0%), respectively, for the low and high GMS subgroups (P = .002, Figure 7D).

3.6 | Risk factors for aGVHD

Based on univariate analysis (Table 3), we found that intensified conditioning, β-lactam antibiotics, the inverse Simpson index, and the GMS

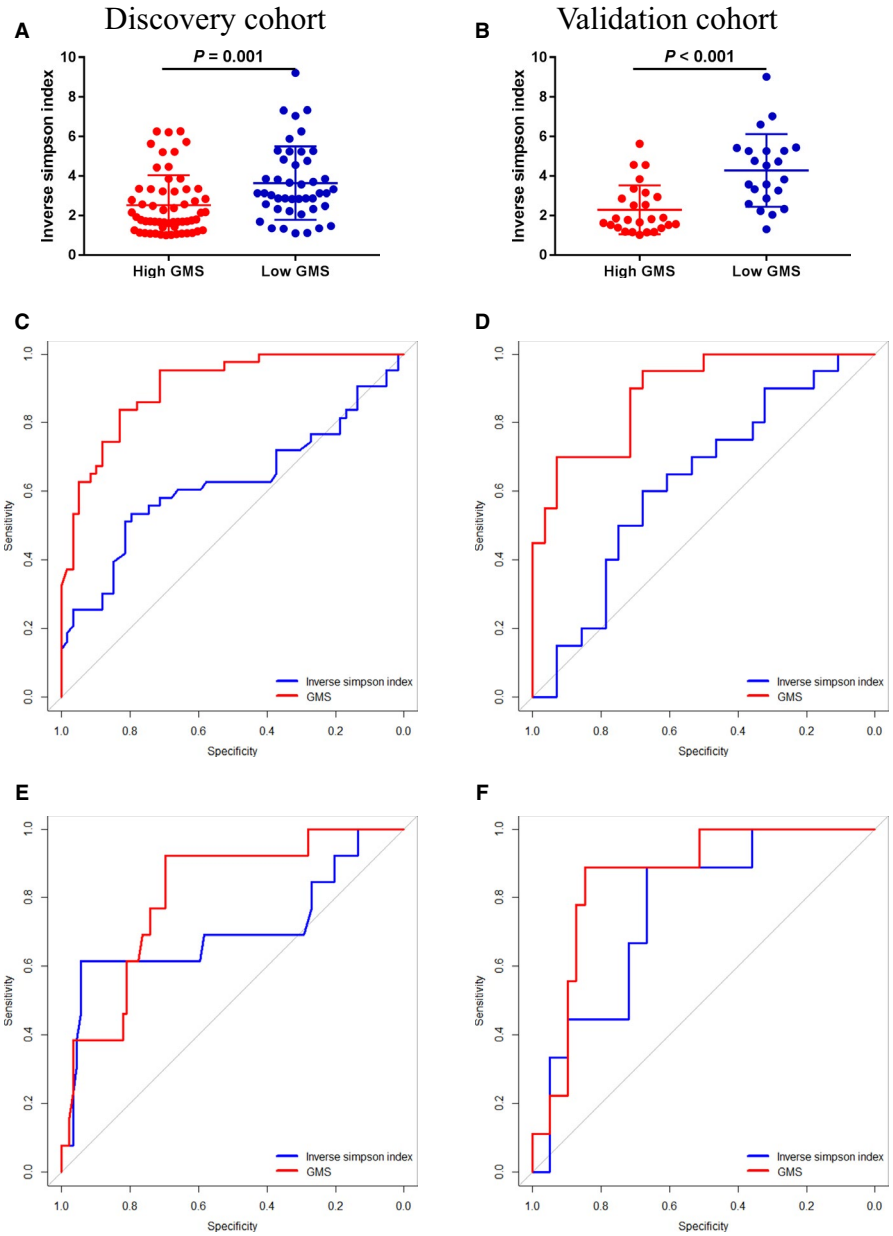

FIGURE 6 The GMS was associated with inverse Simpson index and could predict aGVHD. Inverse Simpson index was higher in the low GMS group compared with the high GMS group both in discovery set (P = .001, 4A) and validation set (P < .001, 4B). Receiver operating characteristic (ROC) curves show the predictive performance of GMS and inverse Simpson index for estimating grade II-IV aGVHD (C and D). The results of the bootstrap (n = 2000) test for the two ROC curves indicate that the AUC of the sum of GMS was significantly higher than those of the inverse Simpson index both in discovery set (0.904 vs 0.622, P < .001) and validation set (0.887 vs 0.673, P = .005). Similarly, ROC curves of the GMS showed the predictive performance in the discovery and validation cohorts for estimating grade III-IV aGVHD, respectively (AUC = 0.808 and 0.860; P < .001 and = .001, respectively; E and F). AUC, area under the ROC curve; GMS, gut microbiota score; ROC, receiver operating characteristic [Color figure can be viewed at wileyonlinelibrary.com]

are significantly associated with aGVHD (P = .003, .001, .002, and <.001, respectively; HR = 3.483 [1.546-8.153], 4.182 [1.760-10.691], 0.269 [0.111-0.627], and 2.976 [1.972-5.102], respectively). Because of the significant correlation between the GMS and inverse Simpson index, we separated the two variables to analyze independent risk factors combined with intensified conditioning and β-lactam antibiotics in the discovery cohort.

According to the logistic regression model for multivariate analysis of II-IV aGVHD in Table 4, high GMS, intensified conditioning, and β-lactam antibiotics were independent risk factors for II-IV aGVHD (P < .001, .042, and .047; HR = 2.612 [1.699-4.634], 1.500 [1.054-4.850], and 2.016 [1.018-6.838], respectively). Based on other factors incorporated in the multivariate analysis, we found that low inverse Simpson index, intensified conditioning, and β-lactam antibiotics were also independent risk factors for II-IV aGVHD (P = .017, .016 and .019; HR = 0.318 [0.122-0.799], 2.957 [1.224-7.386], and 3.092 [1.223-8.266], respectively). Patient age, donor gender, number of HLA mismatches, and vancomycin (intravenous) administration were not found to be independent risk factors for II-IV aGVHD (P > .05).

3.7 | Microbiota and Treg/Th17 balance

To investigate the association between the microbiota model and Treg/Th17 balance, counts of Treg and Th17 cells, and CD4+ T cells (C45+CD3+CD4+CD8- cells) in the blood were detected (Figure 8). At engraftment posttransplantation, the counts of Treg cells in the low GMS subgroup was higher than that in the high GMS subgroup in the discovery set (P = .027; Figure 8); however, the Th17 cells were lower in the low GMS subgroup compared with the high

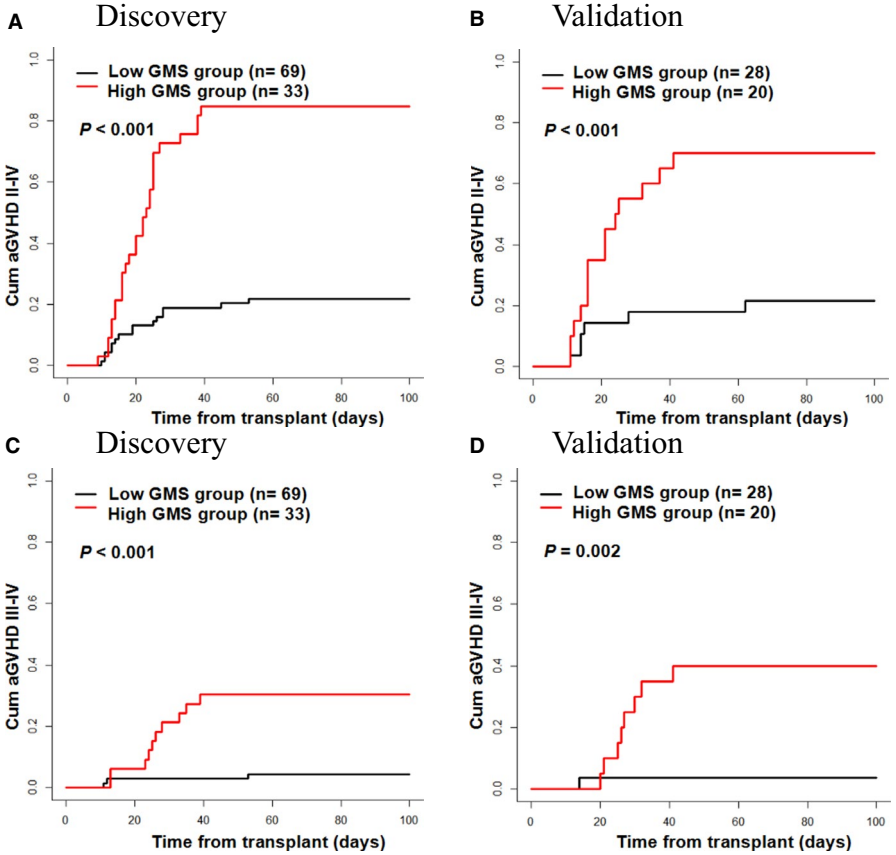

FIGURE 7 Cumulative incidence of aGVHD according to the GMS. A, The cumulative incidence of grade II-IV aGVHD by day +100 posttransplant was lower for the low GMS subgroup than for the high GMS subgroup in the discovery cohort (21.7% vs 84.8%, $P < .001$). B, The validation was consistent with the discovery cohort, the incidence of II-IV aGVHD for the low and high GMS subgroups was 21.4% and 70.0%, respectively ($P = .001$). C, The incidence of III-IV aGVHD was also lower for the low GMS subgroup than for the high GMS subgroup in the discovery cohort (4.3% vs 30.3%, $P < .001$). D, The incidence of III-IV aGVHD in the validation cohort was still consistent with the discovery cohort for the low and high GMS subgroups (3.6% vs 40.0%, $P = .002$) [Color figure can be viewed at wileyonlinelibrary.com]

GMS subgroup ($P = .012$). The ratio of Treg/Th17 cells in the low GMS subgroup was also higher than that in the high GMS subgroup ($P = .012$). In addition, the validation sets were consistent with the discovery sets (Figure 8).

3.8 | Microbiota and cytokines

To explore the association between the microbiota and inflammatory factors, the cytokines, IL-1β, IL-6, IL-17A, and TNF-α were compared in the low and high GMS subgroups (Figure 9). In the discovery sets, the results indicate that the levels of IL-1β, IL-6, IL-17A, and TNF-α were higher in the high GMS subgroup than those found in the low GMS subgroup ($P = .005$, .001, .015, and .003, respectively). Furthermore, the validation sets were also in accordance with the discovery sets (Figure 9).

4 | DISCUSSION

The gut microbiota is associated with the development of aGVHD, and loss of microbiota diversity is an independent risk factor for aGVHD.[15,21-23] In this study, we demonstrated that the constitution of the microbiota at engraftment following myeloablative allo-HSCT is associated with aGVHD. Furthermore, we used the LASSO algorithm to develop and validate a model in which the GMS could predict the subsequent occurrence and development of aGVHD. Of interest, we further reveal a related mechanism in the LASSO model and aGVHD.

In this and our prior studies, we demonstrated that both the diversity and constitution of gut microbiota at neutrophil engraftment are associated with aGVHD.[11] This study focused mainly on whether the microbiota might act as a biomarker for aGVHD. Recently, a few studies have indicated that the constitution of gut microbiota at neutrophil engraftment could be a predictor for the development of aGVHD. Weber et al[24] indicated that urinary 3-indoxyle sulfate, which is a metabolite of gut Clostridia, could be a biomarker for aGVHD. Golob[5] found that a gradient of 20 types of bacterial species (difference in the abundance sum of the positive correlates with minus the sum of the negative correlates) could predict (AUC = 0.83) severe aGVHD. The present study demonstrates that the constitution of the gradient of bacterial species could also be a predictive biomarker for the subsequent development of aGVHD.

Increasing numbers of studies have demonstrated that the gut microbiome plays an important role in the development of aGVHD.[21,23,25,26] However, the relevant microbiota found to be related to aGVHD was not consistent among different transplant centers.[5,15,22,25-28] It may be interpreted that gut microbiota are impacted by many factors, such as antibiotics, conditioning, diet, and geographic environment,[11,15,29-31] and our previous studies have demonstrated such phenomena.[11] The present study demonstrates that the microbiota negatively (eg, Lachnospiraceae and Peptostreptococcaceae) or positively (eg, Enterobacteriaceae) correlate with aGVHD occurrence. In this study, we first used the LASSO method to reduce the total bacterial species, and 20 types of bacterial species were ultimately chosen. Then, a robust logistic

TABLE 3 Univariate analysis of aGVHD in the discovery set

	Univariate analysis	
Variable	**HR (95% CI)**	***P* value**
Sex (male vs female)	1.240 (0.556-2.805)	.600
Age (y) (>33 vs ≤33)	1.021 (0.460-2.277)	.959
Donor sex (male vs female)	0.932 (0.365-2.428)	.884
HLA-mismatched (0 vs 2-3 vs 4-5)	1.392 (0.895-2.182)	.143
Underlying disease (ALL vs AML + MDS)	0.952 (0.476-1.893)	.888
BUCY (yes vs no)	0.696 (0.312-1.531)	.370
Intensified conditioning (yes vs no)	3.483 (1.546-8.153)	.003*
Graft source (PBSCT vs BMT+PBSCT)	1.697 (0.682-4.260)	.253
β-Lactam (yes vs no)	4.182 (1.760-10.691)	.001*
Vancomycin (yes vs no)	1.155 (0.505-2.627)	.730
Amikacin (yes vs no)	1.048 (0.470-2.356)	.907
Bloodstream infection (yes vs no)	2.251 (0.787-6.758)	.133
Inverse Simpson index (≥2 vs <2)	0.269 (0.111-0.627)	.002*
Gut microbiota score (high vs low)	2.976 (1.972-5.102)	<.001*

AML, stable disease; ALL, acute lymphoblastic leukemia; BMT, bone marrow transplantation; CR, complete response; CI, confidence interval; HR, hazard ratio; HLA, human lymphocyte antigen; MDS, myelodyplastic syndrome; NA, not available; PBSCT, peripheral blood stem cell.

*$P < .05$.

model was built, and λ was selected with the smallest cross-validation error. Based on analysis of discovery and independent validation sets, the GMS model demonstrated a highly precise prediction for aGVHD. These results indicate that combining a machine learning model and gut microbiota could serve as a useful tool for estimating the occurrence of aGVHD before treatment. in addition, our study also demonstrated that the GMS was associated with survival, which was consistent with previous studies showing that microbiota diversity was associated with survival following allo-HSCT[15,22]; and the present study indicated that the microbiota had no impact on cGVHD and relapse (data not shown).

To further clarify the mechanism associated with this model and aGVHD, we investigated the changes in specific immunological cells and cytokine levels. In our previous study, we found that gut microbiota influenced the Treg/Th17 balance,[11] consistent with recent research.[32,33] In the present study, we explored the relationship between GMS and immunity homeostasis. We further uncovered that a high GMS correlated with the Treg/Th17 balance and verified these findings in both the training and validation sets. Although 81 cases of this study were registered in a previously published paper,[11] the present study indicated that the relationship between microbiota and aGVHD was verified by two different centers. In addition, our results revealed that the GMS was associated with the IL-1β, IL-6, IL-17A, and TNF-α levels at engraftment. It is notable that these inflammatory factors were associated with the Treg/Th17 balance. Based on the above results, we suggested that these inflammatory factors, mostly generated from patients with a high GMS, might result in a subsequent occurrence of aGVHD by influencing the Treg/Th17 balance.

Many clinical factors have been reported to affect the gut microbiota during allo-HSCT, including patient age, conditioning regimens, antibiotics, and diet.[11,15,29,30] In this study, in addition to being consistent with previous studies, we also found that the GMS was an independent risk factor for aGVHD. Taur et al[15] reviewed the microbiota diversity in myeloablative and nonmyeloablative conditioning regimens and found that the diversity in the myeloablative regimen was lower compared with that in the nonmyeloablative regimen, and the loss in diversity resulted in increased transplant-related mortality, including GVHD-related mortality. This research demonstrated that conditioning would result in destruction of the mucosal barrier and inflammatory factor release.[11,15,34,35] In this study, our results indicate that the microbiota was perturbed due to intensified myeloablative conditioning and antibiotics, resulting in increased inflammatory factors and influencing the Treg/Th17 balance, thus promoting the development of aGVHD.

TABLE 4 Multivariate analysis of aGVHD in the discovery set

	Model 1		Model 2	
Variable	**HR (95% CI)**	***P* value**	**HR (95% CI)**	***P* value**
Intensified conditioning (yes vs no)	2.957 (1.224-7.386)	.016*	1.500 (1.054-4.850)	.042*
β-Lactam (yes vs no)	3.092 (1.223-8.266)	.019*	2.016 (1.018-6.838)	.047*
Inverse Simpson index (≥2 vs <2)	0.318 (0.122-0.799)	.017*	NA	NA
Gut microbiota score (high vs low)	NA	NA	2.795 (1.825-4.931)	<.001*

Model 1: including intensified conditioning, β-lactam antibiotics, and inverse Simpson index. Model 2: including intensified conditioning, β-lactam antibiotics, and gut microbiota score (GMS).

CI, confidence interval; HR, hazard ratio; NA, not available.

*$P < .05$.

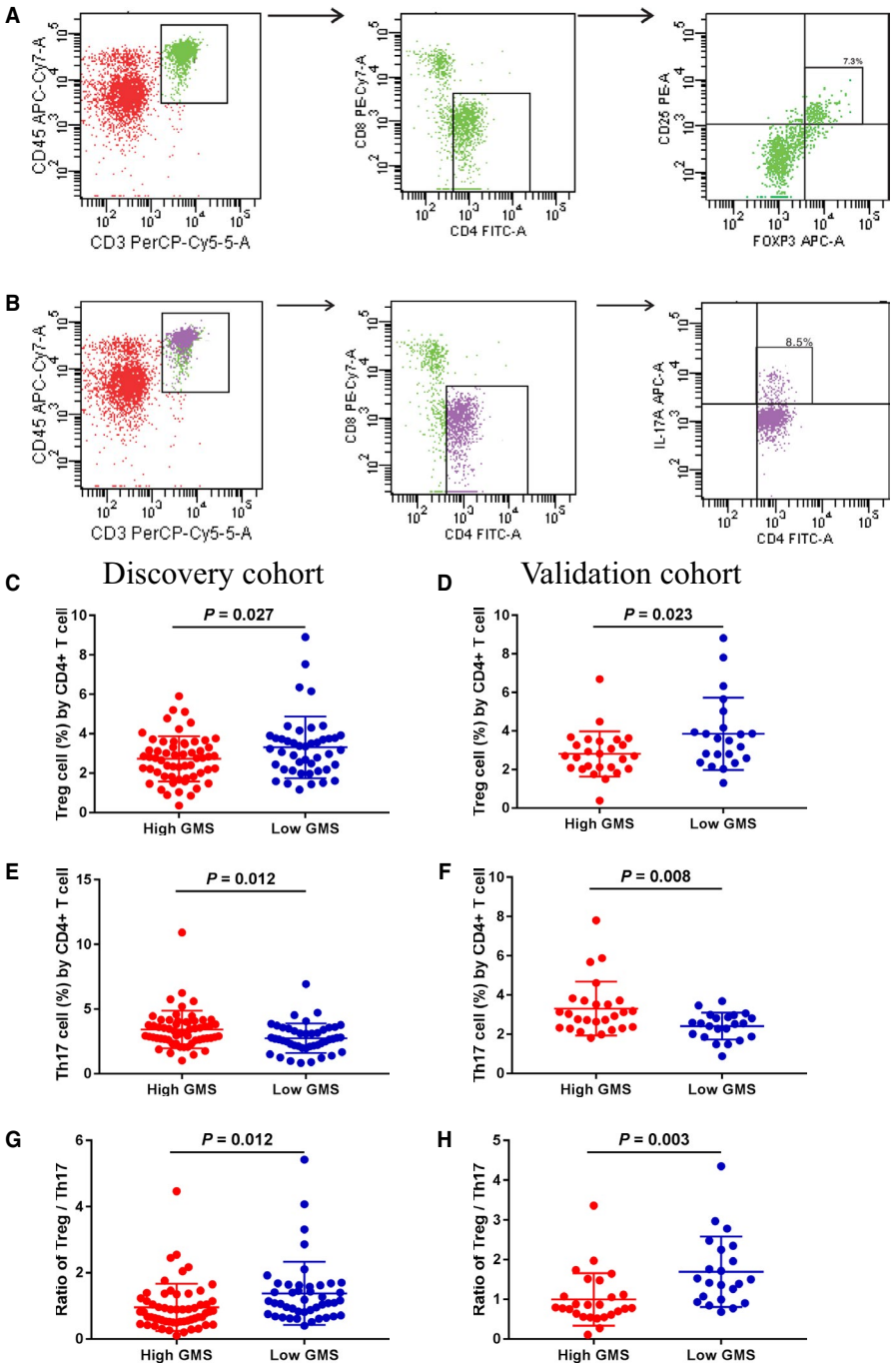

FIGURE 8 GMS was associated with T cell immune. The phenotype of regular T cells (Treg) (A) and IL-17-producing T cells Th17 (B) in peripheral blood mononuclear cells was examined by flow cytometric analysis. The percentages of Treg (C and D) and Th17 (E and F) cells were shown in the discovery and validation cohorts. The percentage of Treg cells was higher in the low GMS group than in the high GMS group both in the discovery set ($P = .034$) and validation set ($P = .010$). On the contrary, the percentage of Th17 cells was lower in the low GMS group than in the high GMS group both in the discovery set ($P = .011$) and validation set ($P = .006$). In addition, low GMS level was associated with high ratio of Treg/Th17 ratio (G and H) compared with the high GMS in the discovery set ($P = .001$, 4A) and validation set. GMS, Gut microbiota score [Color figure can be viewed at wileyonlinelibrary.com]

There are several limitations to our study. First, the patient sample size was relatively small, making this study more like a pilot study. Hence, a larger database of prospective ongoing and recent studies from several centers would be more powerful. Second, the mechanism underlying the microbiota in aGVHD in the study is superficial; thus, this will be studied further in the future. Third, more cytokines should be tested with proteomics applied to analyze the cytokine levels in subsequent studies.

In conclusion, this study demonstrated that the constitution of the gut microbiota at neutrophil engraftment following myeloablative allo-HSCT could be a predictive marker for developing aGVHD. The GMS based on the gut microbiota could be used to predict aGVHD, and the GMS was associated with the inflammatory factors and the Treg/Th17 balance. To more clearly understand the GMS as a predictive marker for aGVHD posttransplant, a deeper investigation into internal molecular mechanisms and pathways should be performed in the future.

ACKNOWLEDGMENTS

We thank Zhensheng Dong (Beijing Genomics Institute, Shenzhen, China) for analysis of the sequencing data from fecal samples as well as the staff of the Beijing Genomics Institute. This work was funded by the National Natural Science Foundation of China (Nos. 81770190, 81870144, 81600141, and 81700176), the Sciences Foundation of

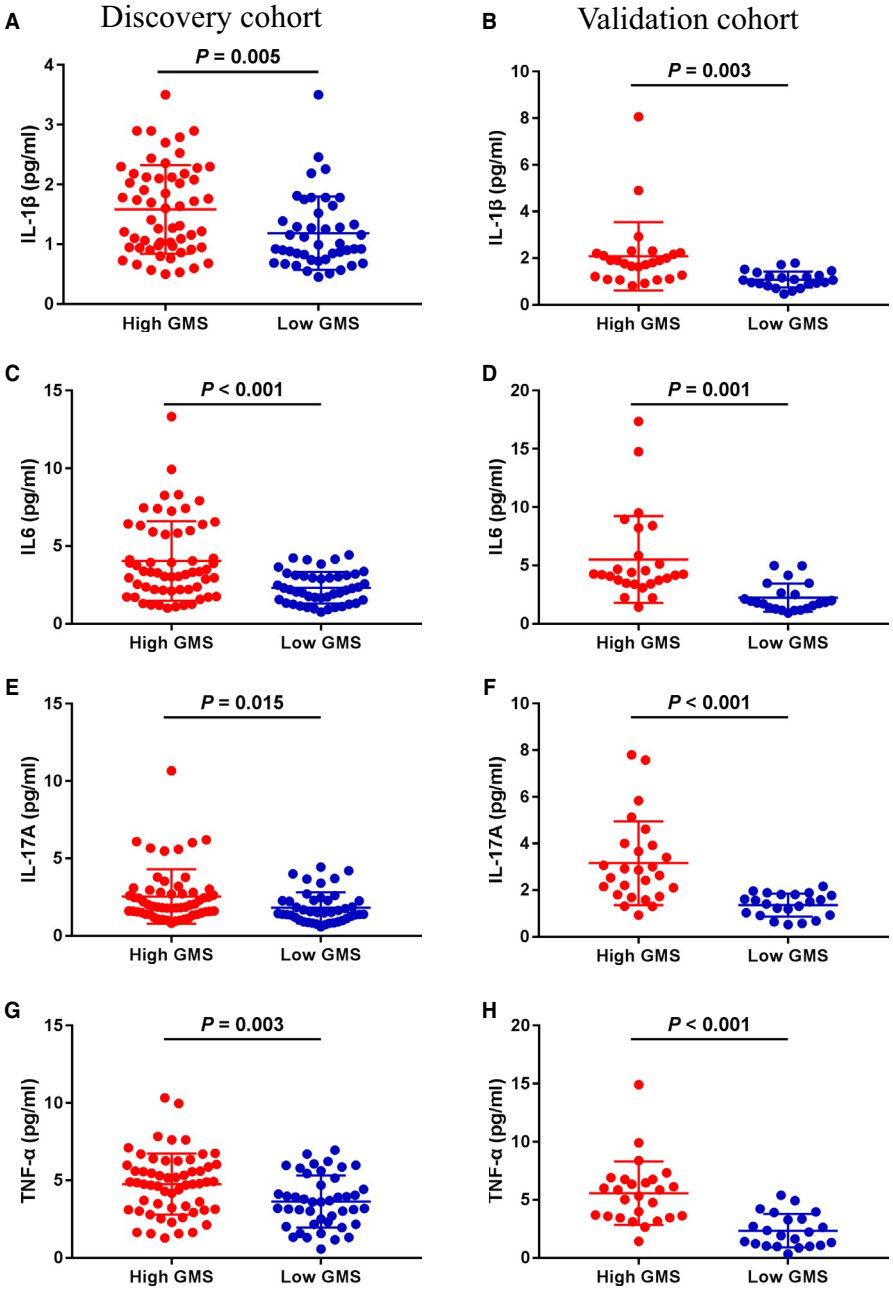

FIGURE 9 GMS was associated with cytokine levels in peripheral blood. The cytokine levels (IL-1β, IL-6, IL-17A, and TNF-α) were higher in the high GMS group than in the low GMS group both in the discovery set (P = .017, .01, .033, and .033, respectively) and validation set (P = .019, .012, .016, and .035, respectively). GMS, gut microbiota score; IL-1β, interleukin-1β; IL-6, interleukin-6; IL-17A, interleukin-17; TNF-α, tumor necrosis factor-α [Color figure can be viewed at wileyonlinelibrary.com]

Guizhou Province (Grant No. 20185579-X), the National Primary Research Plan (Nos. 2017YFA105500 and 2017YFA105504), the Natural Science Foundation of Guangdong Province (No. 2016A030310390), and Medical Science and Technology Planning Project of Henan Province (201403227).

DISCLOSURE

The authors of this manuscript have no conflicts of interest to disclose as described by the *American Journal of Transplantation*.

AUTHOR CONTRIBUTIONS

L.H., J.P., Y.y.L., and Q.L. designed and wrote the paper; H.J. and K.Z. supervised the collection of samples and recorded the clinical data; P.M. and H.h.H. were responsible for DNA extraction from stool samples; Z.p.F., contributed to the examinations of the inflammatory factors and T lymphocyte subsets; H.S. and Z.J. assisted in the preparation and writing of the manuscript; L.Z. performed data analysis. All authors have read and approved the final manuscript.

DATA AVAILABILITY STATEMENT

I confirm that my article contains a Data Availability Statement even if no data are available. I confirm that I have included a citation for available data in my references section, unless my article type is exempt.

ORCID

Jie Peng https://orcid.org/0000-0001-8907-701X

REFERENCES

1. Anasetti C, Logan BR, Lee SJ, et al. Peripheral-blood stem cells versus bone marrow from unrelated donors. *N Engl J Med*. 2012;367(16):1487-1496.

2. Ferrara JL, Levine JE, Reddy P, Holler E. Graft-versus-host disease. *Lancet*. 2009;373(9674):1550-1561.
3. Satwani P, Freedman JL, Chaudhury S, et al. A multicenter study of bacterial blood Stream infections in pediatric allogeneic hematopoietic cell transplantation recipients: the role of acute gastrointestinal graft-versus-host disease. *Biol Blood Marrow Transplant*. 2017;23(4):642-647.
4. Brissot E, Labopin M, Stelljes M, et al. Comparison of matched sibling donors versus unrelated donors in allogeneic stem cell transplantation for primary refractory acute myeloid leukemia: a study on behalf of the Acute Leukemia Working Party of the EBMT. *J Hematol Oncol*. 2017;10(1):130.
5. Golob JL, Pergam SA, Srinivasan S, et al. Stool microbiota at neutrophil recovery is predictive for severe acute graft vs host disease after hematopoietic cell transplantation. *Clin Infect Dis*. 2017;65(12):1984-1991.
6. Martínez-Laperche C, Buces E, Aguilera-Morillo MC, et al. A novel predictive approach for GVHD after allogeneic SCT based on clinical variables and cytokine gene polymorphisms. *Blood Adv*. 2018;2(14):1719-1737.
7. Pellagatti A, Benner A, Mills KI, et al. Identification of gene expression-based prognostic markers in the hematopoietic stem cells of patients with myelodysplastic syndromes. *J Clin Oncol*. 2013;31(28):3557-3564.
8. Olk-Batz C, Poetsch AR, Nöllke P, et al. Aberrant DNA methylation characterizes juvenile myelomonocytic leukemia with poor outcome. *Blood*. 2011;117(18):4871-4880.
9. Peng J, Zhang J, Zhang Q, Xu Y, Zhou J, Liu L. A radiomics nomogram for preoperative prediction of microvascular invasion risk in hepatitis B virus-related hepatocellular carcinoma. *Diagn Interv Radiol*. 2018;24(3):121-127.
10. Peng J, Qi X, Zhang Q, et al. A radiomics nomogram for preoperatively predicting prognosis of patients in hepatocellular carcinoma. *Transl Cancer Res*. 2018;7(4):936-946.
11. Han L, Jin H, Zhou L, et al. Intestinal microbiota at engraftment influence acute graft-versus-host disease via the Treg/Th17 balance in Allo-HSCT recipients. *Front Immunol*. 2018;9:669.
12. Han LJ, Wang Y, Fan ZP, et al. Haploidentical transplantation compared with matched sibling and unrelated donor transplantation for adults with standard-risk acute lymphoblastic leukaemia in first complete remission. *Br J Haematol*. 2017;179(1):120-130.
13. Xuan LI, Huang F, Fan Z, et al. Effects of intensified conditioning on Epstein-Barr virus and cytomegalovirus infections in allogeneic hematopoietic stem cell transplantation for hematological malignancies. *J Hematol Oncol*. 2012;5:46.
14. Przepiorka D, Weisdorf D, Martin P, et al. Consensus Conference on Acute GVHD Grading. *Bone Marrow Transplant*. 1995;15(6):825-828.
15. Taur Y, Jenq RR, Perales M-A, et al. The effects of intestinal tract bacterial diversity on mortality following allogeneic hematopoietic stem cell transplantation. *Blood*. 2014;124(7):1174-1182.
16. Schloss PD, Westcott SL, Ryabin T, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol*. 2009;75(23):7537-7541.
17. Lim EL, Trinh DL, Ries RE, et al. MicroRNA expression-based model indicates event-free survival in pediatric acute myeloid leukemia. *J Clin Oncol*. 2017;35(35):3964-3977.
18. Taskesen E, Babaei S, Reinders MM, de Ridder J. Integration of gene expression and DNA-methylation profiles improves molecular subtype classification in acute myeloid leukemia. *BMC Bioinformatics*. 2015;16(Suppl 4):S5.
19. Tian XP, Huang WJ, Huang HQ, et al. Prognostic and predictive value of a microRNA signature in adults with T-cell lymphoblastic lymphoma. *Leukemia*. 2019;33(10):2454-2465.
20. Gooley TA, Leisenring W, Crowley J, Storer BE. Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. *Stat Med*. 1999;18(6):695–706.
21. Shallis RM, Terry CM, Lim SH. Changes in intestinal microbiota and their effects on allogeneic stem cell transplantation. *Am J Hematol*. 2018;93(1):122-128.
22. Weber D, Jenq RR, Peled JU, et al. Microbiota disruption induced by early use of broad-spectrum antibiotics is an independent risk factor of outcome after allogeneic stem cell transplantation. *Biol Blood Marrow Transplant*. 2017;23(5):845-852.
23. Shono Y, van den Brink M. Gut microbiota injury in allogeneic haematopoietic stem cell transplantation. *Nat Rev Cancer*. 2018;18(5):283-295.
24. Weber D, Oefner PJ, Hiergeist A, et al. Low urinary indoxyl sulfate levels early after transplantation reflect a disrupted microbiome and are associated with poor outcome. *Blood*. 2015;126(14):1723-1728.
25. Whangbo J, Ritz J, Bhatt A. Antibiotic-mediated modification of the intestinal microbiome in allogeneic hematopoietic stem cell transplantation. *Bone Marrow Transplant*. 2017;52(2):183-190.
26. Simms-Waldrip TR, Sunkersett G, Coughlin LA, et al. Antibiotic-Induced depletion of anti-inflammatory clostridia is associated with the development of graft-versus-host disease in pediatric stem cell transplantation patients. *Biol Blood Marrow Transplant*. 2017;23(5):820-829.
27. Jenq RR, Taur Y, Devlin SM, et al. Intestinal blautia is associated with reduced death from graft-versus-host disease. *Biol Blood Marrow Transplant*. 2015;21(8):1373-1383.
28. Ingham AC, Kielsen K, Cilieborg MS, et al. Specific gut microbiome members are associated with distinct immune markers in pediatric allogeneic hematopoietic stem cell transplantation. *Microbiome*. 2019;7(1):131.
29. Chen J, Ryu E, Hathcock M, et al. Impact of demographics on human gut microbial diversity in a US Midwest population. *PeerJ*. 2016;4:e1514.
30. Kasai C, Sugimoto K, Moritani I, et al. Comparison of the gut microbiota composition between obese and non-obese individuals in a Japanese population, as analyzed by terminal restriction fragment length polymorphism and next-generation sequencing. *BMC Gastroenterol*. 2015;15:100.
31. Lee S-E, Lim J-Y, Ryu D-B, et al. Alteration of the intestinal microbiota by broad-spectrum antibiotic use correlates with the occurrence of intestinal graft-versus-host disease. *Biol Blood Marrow Transplant*. 2019;25(10):1933-1943.
32. Riwes M, Reddy P. Microbial metabolites and graft versus host disease. *Am J Transplant*. 2018;18(1):23-29.
33. Varelias A, Ormerod KL, Bunting MD, et al. Acute graft-versus-host disease is regulated by an IL-17-sensitive microbiome. *Blood*. 2017;129(15):2172-2185.
34. Zeng MY, Inohara N, Nunez G. Mechanisms of inflammation-driven bacterial dysbiosis in the gut. *Mucosal Immunol*. 2017;10(1):18-26.
35. Raghunathan VM, Sheng I, Lim SH. Intestinal dysbiosis and allogeneic hematopoietic progenitor cell transplantation. *J Transl Med*. 2016;14(1):335.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Han L, Zhao K, Li Y, et al. A gut microbiota score predicting for acute graft-versus-host disease following myeloablative allogeneic hematopoietic stem cell transplantation. *Am J Transplant*. 2020;20:1014–1027. https://doi.org/10.1111/ajt.15654