FULL METHODS

Patients

Eligible patients were required to be between 18-80 years old and have newly diagnosed stage 1-4 aGVHD of the lower GI tract. Biopsy confirmation of GI aGVHD was required. However, patients could be enrolled and receive the first dose of F-652 prior to the biopsy results being available. Patients whose biopsy results returned with evidence of active infection or otherwise did not support a diagnosis of GI GVHD were removed from the trial and replaced, although they continued to be evaluated for toxicity. Patients could not have received previous systemic immunosuppressive therapy for the treatment of aGVHD except for a maximum of five days of previous corticosteroid therapy, nor could they have received corticosteroid therapy for a non-GVHD indication at a prednisone-equivalent dose > 0.5 mg/kg/day within seven days of enrollment. Previous use of topical corticosteroids, including poorly absorbed oral agents, was allowed.

Patients with an absolute neutrophil count (ANC) less than 500/μL, a serum creatinine level $\geq$ 3 mg/dL, uncontrolled infections, or medical conditions requiring vasopressor and/or mechanical ventilation support at the time of enrollment were excluded. Also ineligible were patients who had received a donor lymphocyte infusion, unless as part a of their originally planned transplant therapy (and not for persistent/recurrent disease), patients with concurrent chronic GVHD (i.e. overlap syndrome), those with a history of an epithelial malignancy (including melanoma or any carcinoma), those with a history of mantle cell lymphoma or anaplastic large cell lymphoma (due to reports of the IL-22 receptor being aberrantly expressed on cells from these tumor types[1,2]), and patients with a history of psoriasis (since IL-22 has been suggested to be involved in the pathogenesis of this disease[3,4]). Full details of inclusion and exclusion criteria are listed in **Supplemental Table 1**.

Study design and treatment

This was an open-label, single-cohort, multicenter phase 2 study sponsored by Evive Biotechnology (Shanghai) Ltd (formerly Generon [Shanghai] Corporation Ltd). Enrollment commenced in

May 2016 and was completed in March 2019. Patients were followed until day 56 for efficacy analyses, safety was also monitored until day 56 (or approximately 28 days following the last dose of F-652), and monitoring continued until day 365 for survival evaluation. The study was conducted at Memorial Sloan Kettering Cancer Center (MSK), University of Texas MD Anderson Cancer Center, and City of Hope National Medical Center (clinicaltrials.gov study NCT02406651). The protocol and informed consents were approved by the institutional review boards of all participating institutions. All patients signed informed consents in accordance with the Declaration of Helsinki.

F-652 was administered at a dose of 45 μg/kg intravenously weekly for 4 doses with concurrent systemic corticosteroid treatment. The dose and schedule of F-652 administration were based on the results of a phase 1 study performed in healthy volunteer subjects[5]. The maximum dose administered in that study was well tolerated, and it was selected for use in this trial. At the time of study enrollment here, all patients received prednisone (or IV equivalent) at a dose of 2 mg/kg/day, which was mandated to be continued for a minimum of three days following the first dose of IL-22. Corticosteroids were then allowed to be tapered per institutional practice, although a minimum dose of 0.25 mg/kg/day of prednisone (or methylprednisolone equivalent dose) was required to still be present by day 28 of therapy. Patients who had progression of aGVHD after seven days of therapy (prior to dose #2) or no response/stable symptoms after 14 days of therapy (prior to dose #3) were discontinued from treatment and considered to be treatment failures. Prior to each treatment dose, patients were required to meet criteria of ANC $\geq$ 500/mm^3, a serum creatinine $\leq$ 3 mg/dL, and all non-hematologic toxicities attributed as probably or definitely related to F-652 were required to be resolved to $\leq$ grade 1 or to the patient's baseline condition.

Study endpoints and evaluations

The objectives of the study were to determine the safety and PK of F-652 treatment, as well as its efficacy in combination with systemic corticosteroids for the treatment of lower GI aGVHD at day 28. Patients who received $\geq$ 1 dose of F-652 were evaluable for safety and toxicity. AEs were graded using the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.03.

Treatment response was assessed for lower GI aGVHD and overall aGVHD on day 28 of therapy[6] according to previously defined staging and response criteria[7]. Treatment responses were categorized as complete response (CR), very good partial response (VGPR), partial response (PR), and treatment failure. GI aGVHD CR was defined as the absence of any symptoms related to the lower GI tract. VGPR was defined as near resolution of lower GI aGVHD symptoms, justifying a continued tapering of corticosteroids. This category included patients tolerating food or enteral feeding and having predominantly formed stools with no overt GI bleeding or abnormal cramping. PR was defined as the improvement of at least one stage in the severity of lower GI aGVHD. Treatment failure was defined by the absence of improvement or progression of lower GI aGVHD. Patients who required additional intervening therapy for GI aGVHD symptoms were considered treatment failure even if improvement in GVHD symptoms was observed. Overall aGVHD response evaluation followed the same criteria, although mixed response (MR) was included in the response assessment and defined as improvement in one or more organs without improvement in others. MR was considered treatment failure.

Secondary study objectives included lower GI aGVHD treatment response at day 56 and overall aGVHD responses at days 28 and 56, and OS at 1 year after first infusion of F-652. In a subset of patients, the study also included the exploratory objectives of post-treatment GI histology evaluation at day 28 or end of treatment (EoT), peripheral blood analyses including biomarkers and cytokines, and stool microbiota assessment before and after IL-22 therapy with F-652.

Study Definitions

The International Bone Marrow Transplant Registry classification guided the aGVHD grading, except grades A-D were labeled grades I-IV. GVHD with purely acute features was graded accordingly, even if it occurred after day 100[8]. Relapse was defined as recurrence of hematologic malignancy after allo-HCT, whereas TRM was defined as death from any cause in continued remission. Overall survival (OS) and progression-free survival (PFS) were defined following standard criteria, and causes of death were described according to the Copelan algorithm[9].

Blood biomarkers

Plasma blood samples were prospectively collected prior to treatment initiation (baseline) and 28 days after initiation of treatment (or EoT if prior to day 28) in a subset of patients (n=24), and three patients did not have available samples. Samples were stored per institutional guidelines and batch-shipped to Viracor-IBT Laboratories at Lee's Summit, MO for analysis. The assays for quantification of REG3α and suppressor of tumorigenicity 2 (ST2) were sandwich enzyme-linked immunosorbent assays (ELISAs) performed in a microtiter plate format. The results from each biomarker were used to calculate an algorithm result[10].

IL-22 pharmacokinetics, IL-22 binding protein, and pharmacodynamics

Serial blood samples for PK evaluation and determination of serum concentrations of IL-22 were obtained at multiple timepoints on dosing days and during follow-up. Specifically, sampling obtained on dosing days 0 (dose #1) and 21 (dose #4) included samples 1, 8, 72, and 168 hours after F-652 administration. Frozen samples were batch-shipped to Syneos Health Clinical Lab (Princeton, NJ) for analysis. PK parameters were estimated using standard non-compartmental analytical procedures (WinNonlin v5.2, Pharsight Corporation, USA). IL-22 binding protein (IL-22BP) was assessed with the human IL-22BP DuoSet ELISA (DY1087-05: R&D systems) in accordance with the manufacturer's protocol. C-reactive protein (CRP) was tested in a subset of patients before and after F-652 infusion as a potential *in vivo* pharmacodynamic activity marker since IL-22 has the capacity to induce acute-phase proteins including CRP[11].

Cytokine and chemokine measurements

Plasma samples were prospectively collected at baseline and day 28 post-treatment (or EoT if prior to day 28) in a subset of 14 patients. Samples were stored per institutional guidelines. Multiple cytokine and chemokine biomarkers were analyzed with a Th17 bead-based multiplex assay using Luminex technology. Specifically, the Milliplex MAP human Th17 magnetic bead assay from EMD Millipore was

utilized. The Median Fluorescent Intensity (MFI) data was analyzed using a 5-parameter logistic or spline curve-fitting method for calculating analyte concentrations in samples. The panel included tumor necrosis factor (TNF)-α, IL-27, IL-5, IL-1β, IL-17α, IL-9, IL-12p70, IL-15, IL-10, interferon (IFN)-γ, and C-C motif chemokine ligand 20 (CCL20).

Fecal Samples and Microbiota

Stool samples were collected prospectively at baseline and day 28 (+/- 4 days) post-treatment (or EoT if prior to day 28) in a subset of patients at each center using similar procedures. DNA extraction, PCR-amplification of genomic 16S ribosomal RNA V4/V5 regions, and sequencing were performed in a central laboratory at MSK. Microbial alpha-diversity is a value that summarizes a microbiota community according to the count of unique taxa and how evenly their frequencies are distributed. This value was computed using the reciprocal Simpson's index (S)[12]. Patients were stratified according to treatment response (responder vs. non-responder). Additionally, a cohort of 27 allo-HCT patients with lower GI aGVHD treated with corticosteroids alone was retrospectively selected to serve as a microbiome comparator group (**Supplemental Table 2**). Recipients of unmodified allografts were chosen at random for this comparator cohort contingent on having fecal samples in the MSK institutional microbiome biobank collected approximately at the time of aGVHD onset and at day 28 post-therapy. All patients provided written consent for biospecimen collection.

Statistical Analyses

Differences in categorical variables were evaluated using Chi-square or Fisher's exact test as appropriate, while differences in categorical and continuous variables were evaluated by using Wilcoxon rank-sum test. Cumulative incidence functions were used to estimate GVHD and TRM. The competing risks for each outcome were death or relapse for GVHD, and relapse for TRM. PFS and OS were estimated using the Kaplan-Meier method. A Simon's two-stage optical design was implemented to distinguish between an unpromising response rate of 35% and a promising response rate of 60% for day-28 treatment

response, which was utilized for determination of the sample size. The type I and type II errors were both set at 0.10.

Safety was evaluated by the incidence of AEs, and severity and type of AEs. AEs were tabulated according to the Medical Dictionary for Regulatory Activities (MedDRA) by system organ class. PK parameters were estimated using the standard non-compartmental analytical procedures (WinNonlin v5.2, Pharsight Corporation, USA). The actual times of sample collection were used in the estimation of the PK parameters. PK comparisons based on response status were estimated by a Wilcoxon rank-sum test. Changes in cytokine levels from baseline to post-treatment time points were evaluated using a Wilcoxon signed-rank test. Associations between CCL20 and IL-22 PK levels were assessed using the Spearman correlation method. Comparison of REG3α and ST2 levels before and after treatment were performed using a Wilcoxon matched-pairs test.

REFERENCES

1. Gelebart P, Zak Z, Dien-Bard J, Anand M, Lai R. Interleukin 22 signaling promotes cell growth in mantle cell lymphoma. *Transl Oncol*. 2011;4(1):9-19.
2. Bard JD, Gelebart P, Anand M, Amin HM, Lai R. Aberrant expression of IL-22 receptor 1 and autocrine IL-22 stimulation contribute to tumorigenicity in ALK+ anaplastic large cell lymphoma. *Leukemia*. 2008;22(8):1595-1603.
3. Ekman AK, Bivik Eding C, Rundquist I, Enerback C. IL-17 and IL-22 Promote Keratinocyte Stemness in the Germinative Compartment in Psoriasis. *J Invest Dermatol*. 2019;139(7):1564-1573 e1568.
4. Lowes MA, Suarez-Farinas M, Krueger JG. Immunology of psoriasis. *Annu Rev Immunol*. 2014;32:227-255.
5. Tang KY, Lickliter J, Huang ZH, et al. Safety, pharmacokinetics, and biomarkers of F-652, a recombinant human interleukin-22 dimer, in healthy subjects. *Cell Mol Immunol*. 2019;16(5):473-482.
6. Levine JE, Logan B, Wu J, et al. Graft-versus-host disease treatment: predictors of survival. *Biol Blood Marrow Transplant*. 2010;16(12):1693-1699.
7. Martin PJ, Bachier CR, Klingemann HG, et al. Endpoints for clinical trials testing treatment of acute graft-versus-host disease: a joint statement. *Biol Blood Marrow Transplant*. 2009;15(7):777-784.
8. Rowlings PA, Przepiorka D, Klein JP, et al. IBMTR Severity Index for grading acute graft-versus-host disease: retrospective comparison with Glucksberg grade. *Br J Haematol*. 1997;97(4):855-864.
9. Copelan E, Casper JT, Carter SL, et al. A scheme for defining cause of death and its application in the T cell depletion trial. *Biol Blood Marrow Transplant*. 2007;13(12):1469-1476.
10. Levine JE, Braun TM, Harris AC, et al. A Prognostic Score for Acute Graft-Versus-Host Disease Based on Biomarkers: A Multicenter Study. *Lancet Haematol*. 2015;2(1):e21-e29.
11. Liang SC, Nickerson-Nutter C, Pittman DD, et al. IL-22 induces an acute-phase response. *J Immunol*. 2010;185(9):5531-5538.
12. Magurran A. Measuring biological diversity: Malden, MABlackwell Publishing; 2004.

Supplemental Table 1. Subject Inclusion and Exclusion Criteria

Subject Inclusion Criteria

- Patients 18 years or older.
- Newly diagnosed stage 1-4 lower GI aGVHD with a minimum stool volume >500 mL/day (determined by the maximum stool output in the preceding 3 days) following allogeneic HCT using bone marrow or peripheral blood stem cells, or cord blood.
- Subjects are willing to undergo a biopsy to confirm lower GI aGVHD. Biopsy results are not needed to initiate treatment. However, if aGVHD is not confirmed histologically, treatment with F-652 will be discontinued.
- Female subjects of childbearing potential who: agree to practice 2 effective methods of contraception at the same time from the time of signing the informed consent form (ICF) through 90 days after the last dose of study drug, OR are postmenopausal for at least 1 year before the screening visit, OR are surgically sterile, OR agree to practice true abstinence when this is in line with the preferred and usual lifestyle of the subject (periodic abstinence and withdrawal are not acceptable methods of contraception).
- Male subjects, even if surgically sterilized (i.e. status post-vasectomy) must agree to one of the following: agree to practice effective barrier contraception during the entire study treatment period and through 90 days after the last dose of study drug, or agree to practice true abstinence when this is in line with the preferred and usual lifestyle of the subject (periodic abstinence and withdrawal are not acceptable methods of contraception).
- Have adequate renal function (Serum creatinine $\leq$ 3 mg/dL).
- ANC $\geq$ 500/mm^3.
- Show evidence of a personally signed and dated informed consent document indicating that the subject (or legally acceptable representative) has been informed of all pertinent aspects of the trial.

Subject Exclusion Criteria

- Evidence of relapse or progression of hematologic malignancy at the time of study enrollment.
- Active uncontrolled infection. Subjects with a controlled infection receiving definitive therapy for 48 hours prior to enrollment are eligible.
- Subjects requiring vasopressors or mechanical ventilation.
- Subjects who have received previous systemic corticosteroids for the treatment of acute GI GVHD for longer than 5 days. Subjects who were treated with systemic corticosteroids for aGVHD for a prior allogeneic HCT $\geq$ 12 months ago are eligible.
- Subjects who received any corticosteroid therapy (for non-GVHD) at doses > 0.5 mg/kg/day prednisone (or IV equivalent) within 7 days prior to the onset of GVHD therapy.
- Subjects who developed aGVHD after unplanned donor lymphocyte infusion.
- Subjects with chronic GVHD features (i.e., acute/chronic GVHD overlap syndrome or classical chronic GVHD).
- History of psoriasis.
- History of epithelial malignancies including melanoma or any carcinomas.
- History or diagnosis of mantle cell lymphoma or anaplastic large cell lymphoma.
- Subject is pregnant or breast-feeding.
- Evidence of current uncontrolled cardiovascular conditions, including uncontrolled hypertension, uncontrolled cardiac arrhythmias, symptomatic congestive heart failure, unstable angina, or myocardial infarction within the past 6 months.
- The subject or guardian is unable to give informed consent or unable to comply with the treatment protocol including appropriate supportive care, follow-up, and research tests.
- The subject has tested positive for the Clostridium difficile toxin within 7 days of study entry.
- Cytotoxic, biologic, or investigational agents are not permitted throughout the study. These include, but are not limited to, anti-thymocyte globulin, alemtuzumab, rituximab, photopheresis, and thalidomide. Subjects who participated in any other investigational drug trial or had exposure to any other investigational agent, device or procedure, within 4 weeks prior to screening and throughout the entire trial, except for trials of investigational drugs administered prophylactically for GVHD or CMV post allogeneic HSCT. In this exception, the other investigational drug must be discontinued upon enrolling (i.e., screening/sign ICF) into this study.
- Any serious medical or psychiatric illness that could, in the Investigator's opinion, potentially interfere with the completion of treatment according to this protocol.

Supplemental Table 2. Antibiotic exposure prior to, during, and after F-652 treatment in MSK patients (n = 14)

Time Period	Antibiotic	Patients (n)
30 days prior to F-652	amikacin	1
	azithromycin	1
	aztreonam	2
	ceftriaxone	1
	ciprofloxacin	3
	meropenem	3
	metronidazole	1
	piperacillin-tazobactam	7
	sulfamethoxazole-trimethoprim	1
	vancomycin - oral	3
	vancomycin - intravenous	5
	none	2
During F-652 course	amikacin	1
	ampicillin	2
	ampicillin-sulbactam	1
	azithromycin	4
	ceftriaxone	1
	ciprofloxacin	1
	meropenem	1
	metronidazole	1
	penicillin	5
	piperacillin-tazobactam	3
	sulfamethoxazole-trimethoprim	1
	vancomycin - intravenous	4
	none	3*
30 days following completion of F-652 course	ampicillin	2
	azithromycin	4
	cefepime	3
	ceftazidime	1
	ceftriaxone	3
	ciprofloxacin	2
	levofloxacin	3
	linezolid	1
	meropenem	2

	metronidazole	1
	penicillin	5
	piperacillin-tazobactam	4
	sulfamethoxazole-trimethoprim	3
	vancomycin - oral	2
	vancomycin - intravenous	2
	none	3*

* 13/14 patients experienced antibiotic exposure during the time period encompassing F-652 treatment and 30 days prior. 14/14 patients experienced antibiotic exposure during the time period encompassing the F-652 treatment course, 30 days prior, and 30 days after.

Supplemental Table 3. Microbiome control cohort demographics (n=27)

Characteristic	Patients
Median age (range)	55 (19-66)
Male gender, n (%)	18 (67%)
Diagnosis, n (%)	
Acute leukemia/ MDS	13 (48%)
Lymphoma	8 (30%)
CML/ MM/ other	6 (22%)
MA conditioning, n (%)	
TBI-based	2 (7%)
Chemotherapy-based	5 (19%)
RI conditioning, n (%)	
TBI-based	10 (37%)
Chemotherapy-based	10 (37%)
Donor, n (%)	
MRD	3 (11%)
MUD/MMUD*	16/ 6 (82%)
Haploidentical	2 (7%)
HLA-match, n (%)	
8/8	19 (70%)
7/8	1 (4%)
< 7/8	7 (26%)
Stem cell source, n (%)	
BM	4 (15%)
PBSC	19 (70%)
Cord blood	4 (15%)
GVHD prophylaxis, n (%)	
CNI +/- MTX +/- sirolimus +/- MMF	19 (70%)
CNI/MMF	4 (15%)
MMF/sirolimus	1 (4%)
PTCy/ MMF +/- CNI +/- sirolimus	3 (11%)
LGI GVHD stage, n (%)	
Stage 1	9 (33%)
Stage 2	4 (15%)
Stage 3-4	14 (52%)
Treatment response	
Responder	15 (56%)
Non-responder	12 (44%)

Abbreviations: n, number; MDS, myelodysplastic syndrome; CML, chronic myeloid leukemia; MM, multiple myeloma; MA, myeloablative; TBI, total body irradiation; RI, reduced intensity; MRD, matched-related donor; MUD, matched-unrelated donor; MMUD, mismatched-unrelated donor; HLA, human leukocyte antigen; BM, bone marrow; PBSC, peripheral blood stem cell; GVHD, graft-versus-host disease; CNI, calcineurin inhibitor; MTX, methotrexate; MMF, mycophenolate mofetil; PTCy, post-transplant cyclophosphamide; LGI, lower gastrointestinal.
*Includes 4 patients who received cord blood grafts

SUPPPLEMENAL FIGURE LEGENDS

Supplemental Figure 1. CONSORT-style flow chart of study subjects. Target enrollment for efficacy evaluation was 27 patients with biopsy-proven lower GI aGVHD.

Supplemental Figure 2. GVHD treatment response according to lower GI clinical stage. Treatment responses were observed in all clinical stages, including stage 3-4 disease.

Supplemental Figure 3. Plasma cytokine levels in recipients of F-652 and systemic corticosteroids. Cytokine levels were measured at enrollment prior to aGVHD treatment (Pre) and after completion of treatment (Post).

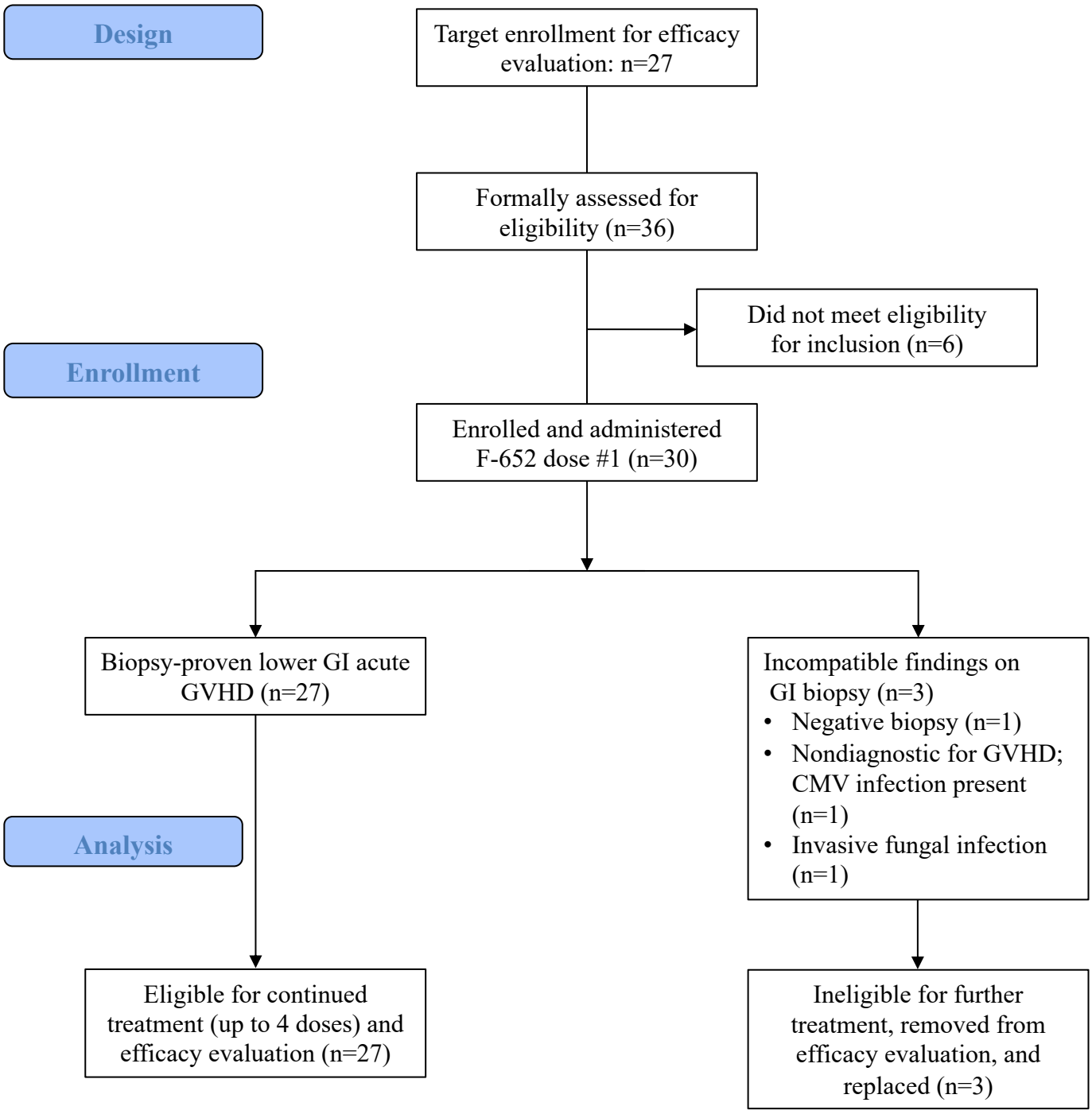
Supplemental Figure 4. Peak F-652 PK levels did not correlate with treatment response. (A) Drug PK levels one hour following the first dose of rhIL-22 in patients who went on to become treatment responders or nonresponders at day 28, showing no significant difference. (B) Drug PK levels eight hours following the first dose of rhIL-22 in patients who went on to become treatment responders or nonresponders at day 28, showing no significant difference.

Supplemental Figure 5. IL-22BP serum concentrations before and after treatment with F-652. (A) Baseline evaluation from samples taken at the time of enrollment showed no significance differences in IL-22BP serum concentrations between responders and nonresponders. (B) Post-treatment evaluation showed no significance differences in IL-22BP serum concentrations between responders and nonresponders.

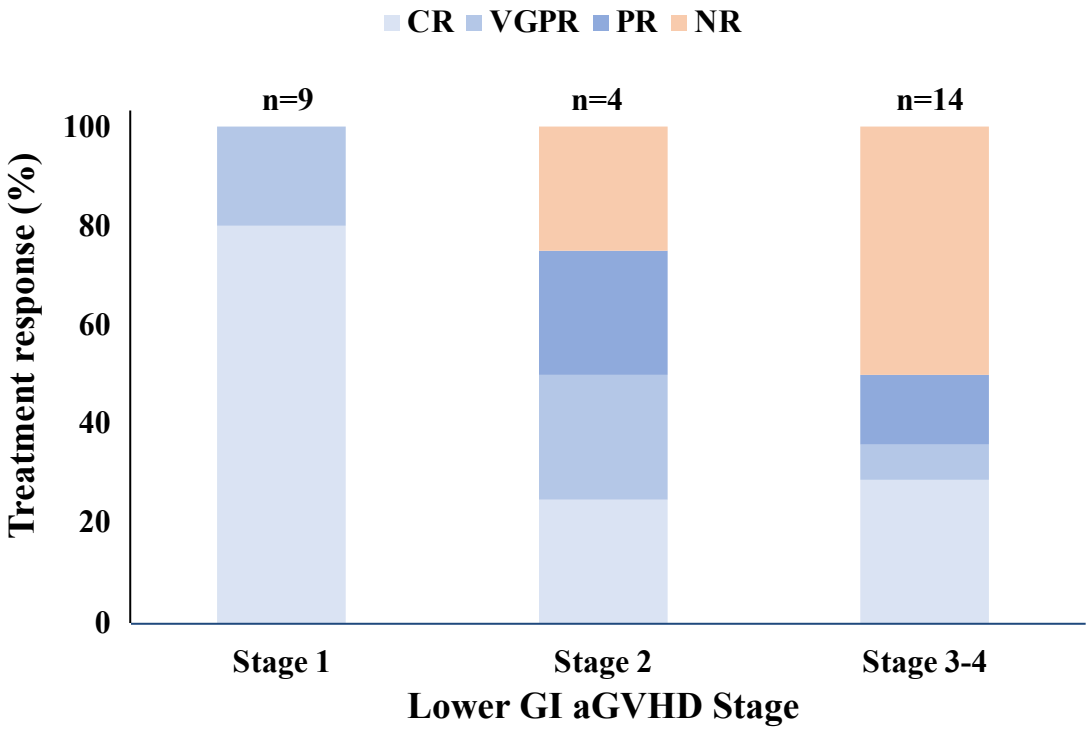
Supplemental Figure 6. PCA plots of the enteric flora before and after GVHD treatment. (A) Abundance of *Blautia* shown on the PCA plots of responders and nonresponders before and after GVHD treatment, indicating enrichment of *Blautia* in the previously identified cluster of F-652 responders. (B)

Microbial α-diversity shown on the PCA plots of responders and nonresponders before and after GVHD treatment, indicating increased diversity in the previously identified cluster of rhIL-22/F-652 responders.

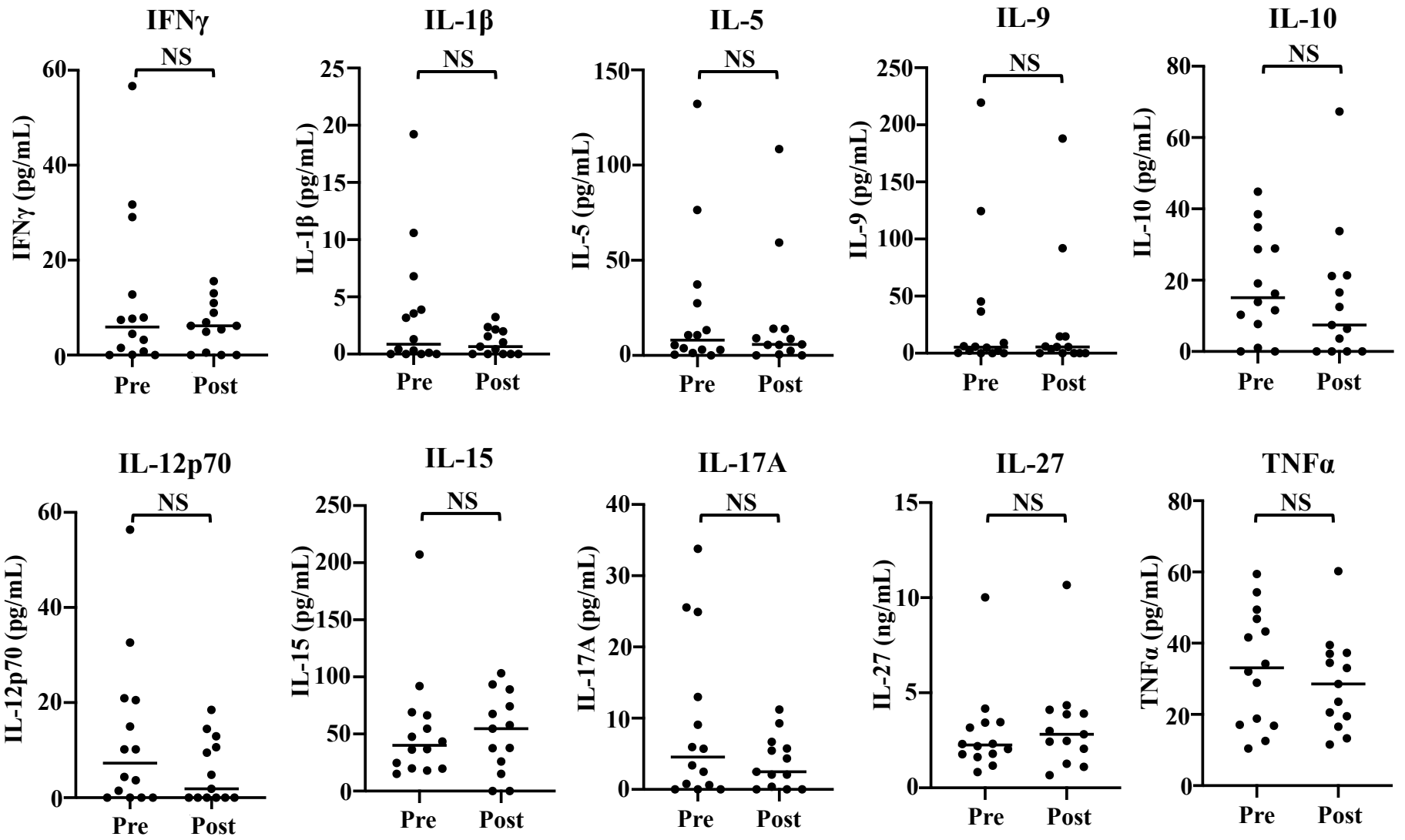
Protocol Patient Flow Diagram



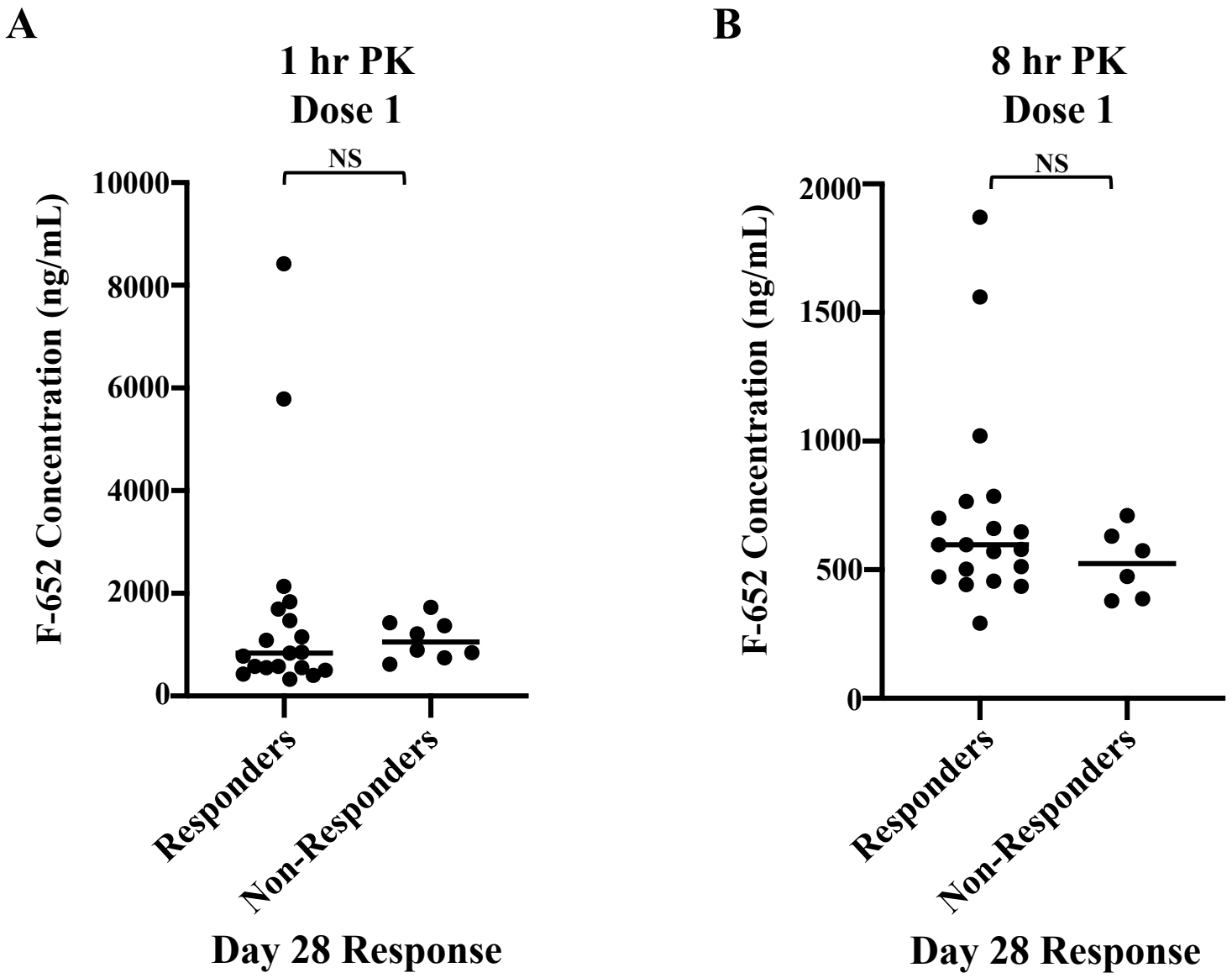
Supplemental Figure 1



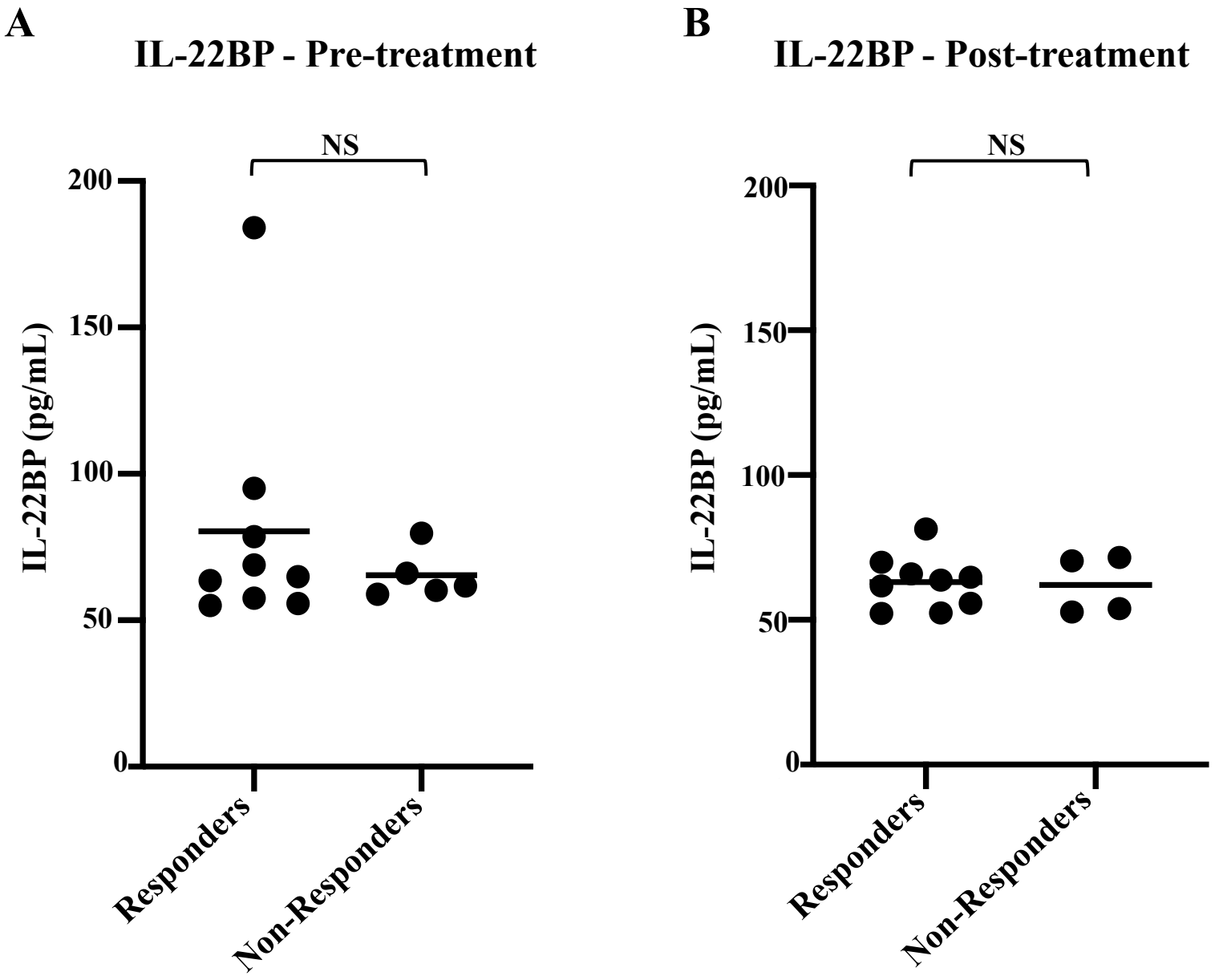
Supplemental Figure 2



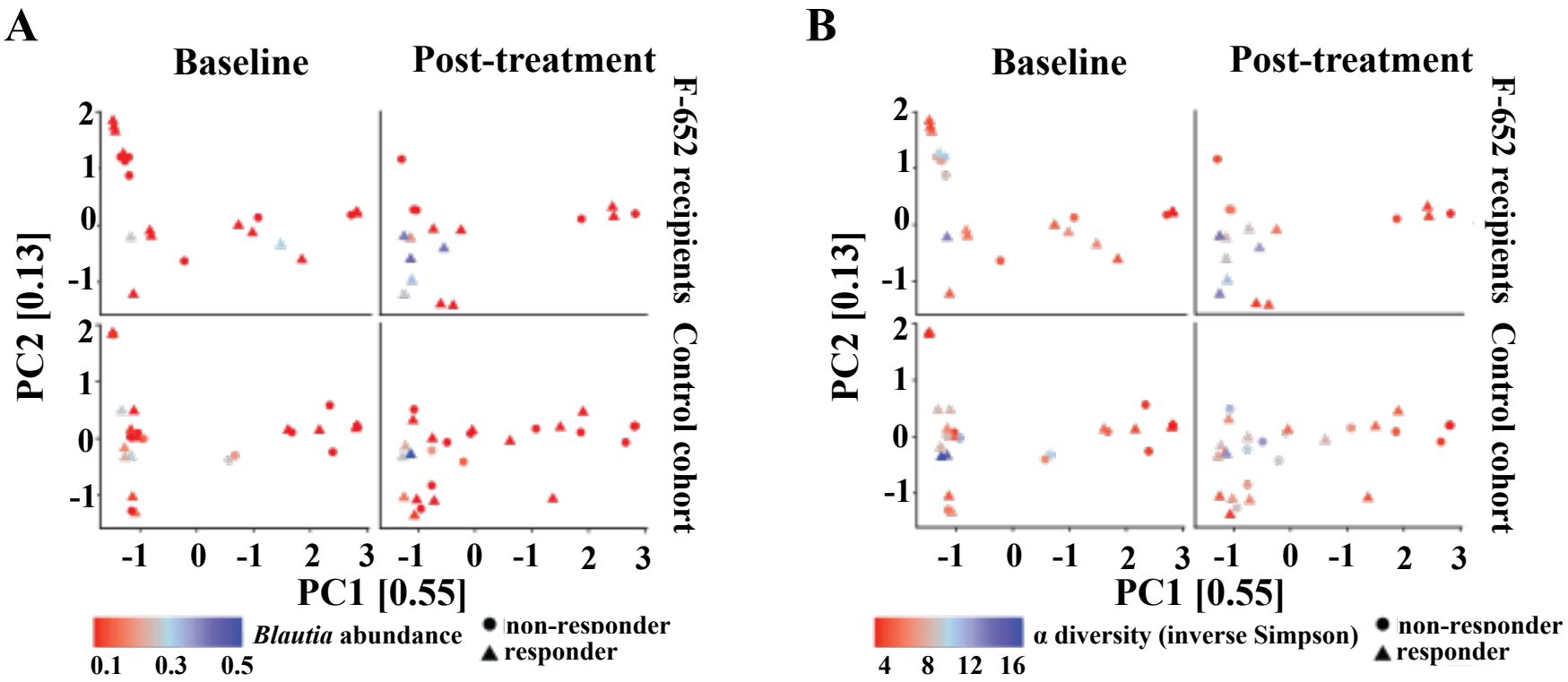
Supplemental Figure 3



Supplemental Figure 4



Supplemental Figure 5



Supplemental Figure 6