SUPPLEMENTARY APPENDIX

Role of interferon-γ in immune-mediated graft failure after allogeneic hematopoietic stem cell transplantation

Pietro Merli,[1] Ignazio Caruana,[1] Rita De Vito,[2] Luisa Strocchio,[1] Gerrit Weber,[1] Francesca Del Bufalo,[1] Vanessa Buatois,[3] Paolo Montanari,[3] Maria Giuseppina Cefalo,[1] Angela Pitisci,[1] Mattia Algeri,[1] Federica Galaverna,[1] Concetta Quintarelli,[1] Valentina Cirillo,[1] Daria Pagliara,[1] Walter Ferlin,[3] Maria Ballabio,[3] Cristina De Min[3] and Franco Locatelli[1,4]

[1]Bambino Gesù Children's Hospital, Department of Pediatric Hematology/Oncology, Cellular and Gene Therapy, Rome, Italy; [2]Bambino Gesù Children's Hospital, Department of Laboratories, Pathology Unit, Rome, Italy; [3]Novimmune SA, Geneva, Switzerland and [4]Department of Pediatrics, Sapienza, University of Rome, Rome, Italy

 doi:10.3324/haematol.2019.216101

Received: January 7, 2019.

Accepted: February 18, 2019.

Pre-published: February 21, 2019.

Correspondence: *PIETRO MERLI* - pietro.merli@opbg.net

Supplemental data

Materials and Methods

Patients

Patients treated with a monoclonal antibody against IFNγ, TNFα or IL6 in the previous 8 weeks before HSCT were not included in the study.

Cytokine profile

Streptavidin Gold plates (MSD, Rockville, MD) were pre-washed with PBS-Tween 0.05% (wash buffer). For IFNγ-detection, plates were coated with 25 µL of biotinylated anti-hIFNγ clone 1-D1K (Mabtech, Nacka Strand, Sweden) at 0.5 µg/mL in PBS-BSA 1%-Tween 0.05% and incubated 1 hour at 22°C with agitation at 600 rpm. The calibration standard (recombinant human IFNγ, R&D Systems, Minneapolis, MN) ranging from 20 to 100,000 pg/mL was prepared in pooled human serum (Seralab, West Sussex, UK) and incubated 1 hour at 22°C. Plates were then washed and 25 µL of standard and samples added in duplicate per well. After 1-hour incubation at 22°C with agitation at 600 rpm, plates were washed and 25 µL of the detection anti-hIFNγ antibody 1-D1K Sulfo-TAG (Mabtech, Nacka Strand, Sweden) at 0.6 µg/mL in PBS-BSA 1%-Tween 0.05% was added. After 1-hour incubation at 22°C with agitation at 600 rpm, plates were washed and 150 µL of MSD read buffer T added per well. Plates were read on the MSD Meso Sector® S600 and analyzed using the MSD Discovery Workbench 4.0 software. The above methodology was also employed for detection of sIL-2Rα, CXCL9 and CXCL10 with the exception of sIL-2Rα. Plates were coated with 25 µL of biotinylated-polyclonal goat IgG anti-human CD25/IL-2Rα (R&D systems, Minneapolis, MN) at a concentration of 0.5 µg/mL in PBS-BSA 1%. The calibration standard (recombinant human sIL-2Rα, Reprokine, Tampa, FL) ranging from 20 to 100,000 pg/mL was prepared in FBS (Serum, Sigma, Saint Louis, MO). Samples were tested neat and diluted 1:2, 1:5, 1:10, 1:20 1:50 in FBS. 25 µL of the detection antibody, anti-human CD25/IL-2Rα Sulfo-TAG (R&D systems, Minneapolis, MN), were added at a concentration of 1.2 µg/mL in PBS-BSA 1%. CXCL9: plates were coated with 25 µL of mouse anti-human CXCL9 biotinylated antibody (MSD, 501 Rockville, MD) at 0.25 µg/mL in PBS-BSA 1%. The calibration standard (recombinant human CXCL9, R&D systems, Minneapolis, MN), ranging from 30 to 100,000 pg/mL was prepared in FBS. Samples were tested neat and diluted 1:2, 1:5, 1:10, 1:20 1:50 in FBS. Standard and samples were diluted 1:10 in PBS-BSA 1%. 25 µL of goat anti-human

CXCL9 antibody Sulfo-TAG (MSD, Rockville, MD) at 0.25X in PBS-BSA 1% was used for detection. CXCL10: plates were coated with 25 µL of mouse anti human CXCL10 biotinylated antibody (MSD, 501 Rockville, MD) at 0.25µg/mL in PBS-BSA 1%. The calibration standard (recombinant human CXCL10/IP10, R&D systems, Minneapolis, MN) ranging from 20 to 50,000 pg/mL was prepared in FBS (Sigma, Saint Louis, MO). Samples were tested neat or diluted 1:2, 1:5, 1:10, 1:20 1:50 in FBS. Standard and samples were diluted 1:10 in PBS-BSA 1%. 25 µL of the anti-human CXCL10 antibody Sulfo-TAG (MSD, Rockville, MD) at 0.5X in PBS-BSA 1% was used as the detection antibody. Plates were read on the MSD Meso Sector® S600 and analyzed using the MSD Discovery Workbench 4.0 software.

Histopathology analysis

BM specimens were fixed in buffered zinc formalin, for at least 12 hours (12-18 hours), decalcified in EDTA for at least 6 hours (6-8 hours), and embedded in paraffin in Leica Tissue Embedded (Pathos Delta, Milestone). Sections were cut at a thickness of 2.5 µm, deparaffined and stained according to common histological procedures using Hematoxylin and Eosin (H&E) and May-Grunwald Giemsa (MGG) for morphological evaluation, reticulin stain for evaluation of fibrosis and Perls stain for evaluation of BM iron content. Immunohistochemical staining with an extensive panel of antibodies was used to better characterize specific hematopoietic cell compartments. In detail, Myeloperoxidase (MPO) stain (polyclonal, Dako) was used for the identification of myeloid colonies, Glycoforin stain (ab clone Jc159, Dako) for erythroid colonies, CD61 (ab clone Y25, Dako) for megacaryocytes, and CD20 (ab clone L26, Dako), CD79a (ab clone JCB117, Dako), CD3 (polyclonal, Dako), CD5 (ab clone 4C7, Novocastra), CD4 (ab clone 4B12, Leica), CD8 (ab clone C8/144B, Dako), TIA1 (ab clone 2G9A10F5, Novus Biological), Granzyme B (ab clone 11F1, Leica), Perforin (ab clone PJM50 Leica), CD163 (ab clone 10D6, Novocastra), CD68 (ab clone PGM1, Dako), for the characterization of inflammatory infiltrate (lymphocytes and histiocytes). Immunohistochemistry was performed with Autostainer link 48, according to Dako automated procedures, with pretreatment with heat-induced epitope retrieval in Dako PT link. Streptavidin Biotin +Horse Radish Peroxidase and Avidin Biotin Alkaline Phosphatase were used as detection systems. Persistence of megakaryocytic, erythroid or granulocytic lineages from precursors to differentiated cells defined persistent normal hematopoiesis. All samples were examined at least on five different magnifications (4x, 10×, 20x, 40x and 63x), and reviewed by the same pathologist. The total number of positive cell for each marker was counted in five fields per patient sample under 20-fold magnification and reported as mean ± standard deviation.(1)

Immunofluorescence

Paraffin sections of BM biopsies were deparaffined and incubated with 5% bovine serum albumin (Sigma-Aldrich) to block non-specific binding. Slides were then stained with the primary Ab against human CD3 (Dako) (1:75 dilution at 4°C for an overnight incubation), Granzyme B (Leica Microsystems, Germany) (1:200 dilution at 4°C for an overnight incubation), Ki67 (Dako) (ready for use dilution at 4°C for overnight incubation) and terminal deoxynucleotiditil transferase-mediated dUTP Nick End Labeling (Millipore, MA) according to the instructions provided by the manufacturer. The specimens were then probed with Alexa Fluor 488, 555 and 647 goat anti-rabbit or mouse secondary antibody (Cell Signaling Technology, MA) (1:500 dilution at 25°C for 1 hour). Hoechst dye (Invitrogen, CA) was used for nuclear staining. Fluorescent signals were detected using a Leica TCS-SP8X laser-scanning confocal microscope (Leica Microsystems) equipped with a tunable white light laser source, 405 nm diode laser, 3 (PMT) and 2 (HyD) internal spectral detector channels.

Isolation of mononuclear cells from BM aspirate and immune-phenotypic analysis

Mononuclear cells were isolated from BM aspirates using lymphocyte separation medium (Eurobio; France). Expression of cell surface markers was assessed by flow-cytometry using standard methodology. T-cell receptor (TCR)-Vβ repertoire of T lymphocytes was evaluated using a panel of 24 different TCR Vβ-specific mAbs (IO TEST Beta Mark TCR-Vβ repertoire kit, Beckman Coulter, CA in association with isotype control and CD3 mAb (BD Biosciences). Samples were acquired with a BD LSRFortessa X-20 and analyzed using the FACSDiva software (BD Biosciences). For each sample, a minimum of 20,000 events was analyzed.

In vivo murine model of HSCT rejection

C57BL/6 Ifngr1-/- mice, (expressing the Ly5.2 congenic marker), used as recipient mice, were donated by Prof. Roland Liblau (INSERM, Toulouse) and used at an age between 7 to 10 weeks. C57BL/6 Ifngr1+/+, (expressing the Ly5.1 congenic marker) used as donor mice, were purchased at Charles River Laboratories and used at an age between 7 and 10 weeks.

All animal experiments were performed in accordance with the Swiss animal protection law. Ifngr1-/- mice were intravenously (i.v.) infected with 1.106 CFU of Bacillus Calmette-Guérin (BCG) (strain

Pasteur 1173P2, gift from Irène Garcia-Gabay, Geneva University). After 14, 20, 28, 35 and 42 days, mice were given i.v. injections of 100 mg/kg of the rat anti-mouse (m) IFNγ-neutralizing mAb (clone XMG1.2, BioXCell) or an irrelevant rat isotype control mAb (Novimmune). At day 21, these Ifngr1-/- recipient mice were given mild doses (550 rads) of whole body irradiation, using a Cs137 radioactive source (Gammacell 40 Extractor, Geneva University). BM transfer (BMT) was then initiated by the i.v. administration, of two million total bone marrow cells freshly isolated from Ifngr1+/+ donor mice. Briefly, BM cells were flushed into PBS from the femurs and tibias of donor mice. Red blood cells were lysed and the remaining cells were counted before BM cell transfer. After engraftment, mice were kept in a specific-pathogen-free animal facility, in filter-topped cages, and treated with trimethoprim/sulfamethoxazole. All caging procedures and manipulations were carried out under a laminar flow hood.

The degree of donor hematopoietic reconstitution was assessed at different time points after BMT. Immunofluorescence analyses were carried out on whole blood. Briefly, donor BM cells expressed the Ly5.1 marker, whereas recipient leukocytes expressed the Ly5.2 marker. Chimerism was, therefore, assessed as the percentage of cells expressing the Ly5.1 isotype (engraftment) amongst the total lymphocytes. Antibodies against the following surface antigens (all from BD Bioscience Pharmingen) were used: FITC-CD3ε, PE-Ly5.2, PerCp-CD19 and APC-Ly5.1. Samples were, then, acquired on FACS Calibur (Becton Dickinson) and analyzed using CellQuest Pro software.

Serum was recovered from whole blood samples and frozen at -80°C for further analysis. IFNγ and CXCL9 were quantified using Millipore's MILLIPLEX® MAP Mouse Cytokine Kit (#MCYTOMAG-70K-06) according to the manufacturer's recommendations.

Emapalumab administration in CU to HLH patients experiencing GF

In primary HLH patients, emapalumab is administered at an initial dose of 1 mg/kg. According to the compassionate use treatment protocol in use at the time, the dose of the antibody could be modified based on clinical and laboratory criteria and on the pharmacokinetic levels of the antibody as indicated by the information derived from modelling and simulation data. Based on the data accumulated so far, it has been possible to demonstrate that the clearance of emapalumab is affected by the presence of target mediated drug disposition, namely the higher the concentration of IFNγ, the faster the clearance of the drug.

Results

Activation of macrophages and T lymphocytes characterizes GF in allogeneic HSCT

In GF patients, at the initial stage, the myeloid compartment did not show a marked difference as compared to the control group (supplementary Figure 3A and supplementary Figure 4A). Generally, maturation of the myeloid lineage until late maturation stages was conserved, with a regular distribution among the different myeloid lineages, as well as the typical paratrabecular distribution of myeloid precursors/progenitors.

From a morphological point of view, BM infiltrating lymphocytes of GF patients appeared activated and pear-shaped. However, in order to evaluate their activation and cytotoxic activity, the expression of Granzyme B, TIA-1 and Perforin were evaluated as well, showing different levels of expression of the activation markers analyzed (supplementary Figure 3H-J, Figure 4A and supplementary Figure 5C-D). In details, mean numbers of Granzyme B, TIA-1 and Perforin positive cells in samples of GF patients and controls were 36.7 (±16.7) vs 0 ($p<0.001$), 46.5 (±26.5) vs 4.5 (±4.3) ($p<0.001$) and 6.4 (±3.4) vs 0 ($p<0.001$), respectively (Figure 4A). These data were confirmed by immunofluorescence through co-staining of T lymphocytes with both Ki67 (supplementary Figure 5A-B) and Granzyme B (supplementary Figure 5C-D), showing positivity of 66.2% (16±2.5 CD3+/Ki67+ and 8±1.5 CD3+/Ki67- cells/field) and 47.9% (26.3±1.2 CD3+/Granzyme B+ and 28.7±2.7 CD3+/Granzyme B- cells/field), respectively. By staining with CD79 and CD20 Ab, the presence of B cells, as well as of leukemia cells, was excluded (supplementary Figure 3E-F).

Polyclonal T-cell pattern with predominant CD8 effector memory phenotype

BM mononuclear cells derived from patients with GF or from controls were stained with a panel of 24 different TCR Vβ-specific mAbs. TCR repertoire profile based on the predominantly polyclonal distribution of spectratypes across all TCR Vβ gene families of controls: Vβ1 5.2%±1.0%, Vβ2 6.0%±0.8%, Vβ3 7.4%±2.4%, Vβ4 1.9%±0.2%, Vβ5.1 5.4%±0.7%, Vβ5.2 1.0%±0.2%, Vβ5.3 1.1%±0.2%, Vβ7.1 3.9%±0.6%, Vβ7.2 1.8%±0.6%, Vβ8 7.7%±1.7%, Vβ9 3.9%±0.9%, Vβ11 0.8%±0.1%, Vβ12 3.4%±1.0%, Vβ13.1 3.2%±0.25, Vβ13.2 4.2%±1.3%, Vβ13.6 1.9%±0.3%, Vβ14 6.3%±1.4%, Vβ16 2.6%±0.8%, Vβ17 4.8%±0.4%, Vβ18 1.3%±0.3%, Vβ20 5.2%±1.1%, Vβ21.3 2.7%±0.7%, Vβ22 3.5%±0.3% and Vβ23 1.9%±0.5%). Vβ TCR repertoire in patients with GF: Vβ1 6.3%±2.4%, Vβ2 6.7%±1.8%, Vβ3 8.6%±2.3%, Vβ4 5.3%±2.4%, Vβ5.1 7.0%±1.5%, Vβ5.2 2.4%±0.4%, Vβ5.3 2.1%±0.5%, Vβ7.1 5.8%±1.7%, Vβ7.2 2.9%±0.8%, Vβ8 5.4%±1.0%, Vβ9

4.4%±1.1%, Vβ11 1.5%±0.2%, Vβ12 2.7%±0.6%, Vβ13.1 3.4%±0.6%, Vβ13.2 7.2%±2.2%, Vβ13.6 4.1%±1.0%, Vβ14 9.9%±1.5%, Vβ16 5.7%±2.5%, Vβ17 8.6%±3.0%, Vβ18 1.9%±0.5%, Vβ20 5.5%±0.6%, Vβ21.3 4.1%±0.6%, Vβ22 7.8%±2.8% and Vβ23 4.3%±1.1%. The analysis of the Vβ TCR repertoire in patients with GF showed a significant difference in subclass distribution compared to the control group (Two-way ANOVA, p=0.0082), with a generally higher relative expression of the Vβ families. A further analysis between the two groups identified a significant difference in some Vβ families; in particular, a higher expansion of the Vβ families 5.2, 5.3, 11, 13.6 and 23 was observed in GF patients (p=0.001, p=0.033, p=0.005, p=0.023 and p=0.042 respectively, see also supplementary Figure 6).

Regarding lymphocytes infiltrating the BM of GF patients, the natural killer T-cell population in GF patients showed a significant difference in comparison with the controls, namely a significant increase in the CD4+/CD56+ (17.8%±15.7% *versus* 2.8%±2.5%, GF patients *versus* controls, p=0.043) and a reduction in the CD8+/CD56+ (4.5%±2.0% *versus* 16.8%±10.5%, GF vs CTRL patients, p=0.012) being observed (Figure 5B-C). These findings suggest a possible role of this population in GF pathophysiology.

Increasing expression of activation and exhaustion markers on T cells during GF

The reduced number of quiescent cells in BM also confirmed the presence of activation. Indeed, the number of double negative CD27-/CD28- in the CD8+ cell subset was significantly reduced (41.4%±18.7% *versus* 11.9%±7.6%, controls *versus* GF patients, respectively) (p=0.003) (supplementary Figure 7A).

In both CD4+ and CD8+ cell subsets, up-regulation of CD57 and its co-expression with CD95 was observed (CD57+/CD95+: 9.2%±11.0% *versus* 37.4%±12.4% and 28.2%±17.9% *versus* 66.1%±17.9% controls *versus* GF patients in CD4+ and CD8+ cell subsets, respectively) (p=0.002 and p=0.005) (Figure 5F and supplementary Figure 7B). The analysis of the activation and exhaustion marker CD279 (PD1) revealed the same trend observed on the expression of CD57 (20.3%±18.0% *versus* 65.5%±29.7% and 12.0%±8.1% *versus* 68.8%±29.7% controls *versus* GF patients in CD4+ and CD8+ cell subsets, respectively) (p=0.013 and p=0.002) (supplementary Figure 7C). We further investigated the expression of other specific exhaustion markers, namely CD223 and CD366 (reported to be markers for exhaustion on the CD4+ and CD8+ cell subsets, respectively). Supplementary Figure 7D-E shows that both CD4+ and CD8+ cells in BM of patients with GF showed up-regulation of these markers (CD4+/CD223+: 3.0%±1.8% *versus* 27.5%±9.1%;

CD8+/CD3663+: 22.7%±22.2% *versus* 60.6%±21.7% in controls *versus* GF patients, respectively) ($p<0.001$ and $p=0.014$). For extensively defining the exhaustion profile, we analyzed the co-expression of CD4/CD57/CD95/CD223/CD279 and CD8/CD57/CD95/CD279/CD366. As documented in supplementary Figure 7F, both CD4+ and CD8+ cell subsets were shown to be significantly more exhausted as compared to their counterpart in the control group (CD4: 0.3%±0.3% *versus* 5.2%±4.4%; CD8: 1.1%±0.9% *versus* 28.3%±15.1% in controls *versus* GF patients, respectively) ($p=0.022$ and $p=0.003$). The phenotype analysis revealed that the CD4-/CD8- γδ T lymphocytes also showed an exhausted phenotype (0.0% vs 1.6%±1.5% in controls *versus* GF patients, respectively) ($p=0.04$) (supplementary Figure 8).

References

1. Fedchenko N, Reifenrath J. Different approaches for interpretation and reporting of immunohistochemistry analysis results in the bone tissue - a review. Diagnostic pathology. 2014 Nov 29;9:221

Table S1. Data on patients receiving emapalumab after HSCT.

UPN	Age at (first) HSCT (months)	Sex	Disease	Mutation	Donor	*Ex-vivo* TCD	Number of emapalumab doses (after second HSCT)	Neutrophil recovery	Days	Platelet recovery	Days	Outcome	Complications	Notes
1	11	M	HLH	UNC13D	Mother	Yes	8 (6 mg/kg)	Yes	12	Yes	14	A/W	None	
2	12	F	HLH	unknown (FLH1)	Mother	Yes	7 (3 mg/kg)	Yes	13	Yes	13	A/W	None	
3	9	F	HLH	UNC13D	Mother	Yes	6 (1 mg/kg)	No		No		Alive and well after a third successful HSCT from unrelated cord blood unit	none	No HLH flare after rejection

Supplementary Figures

Figure S1

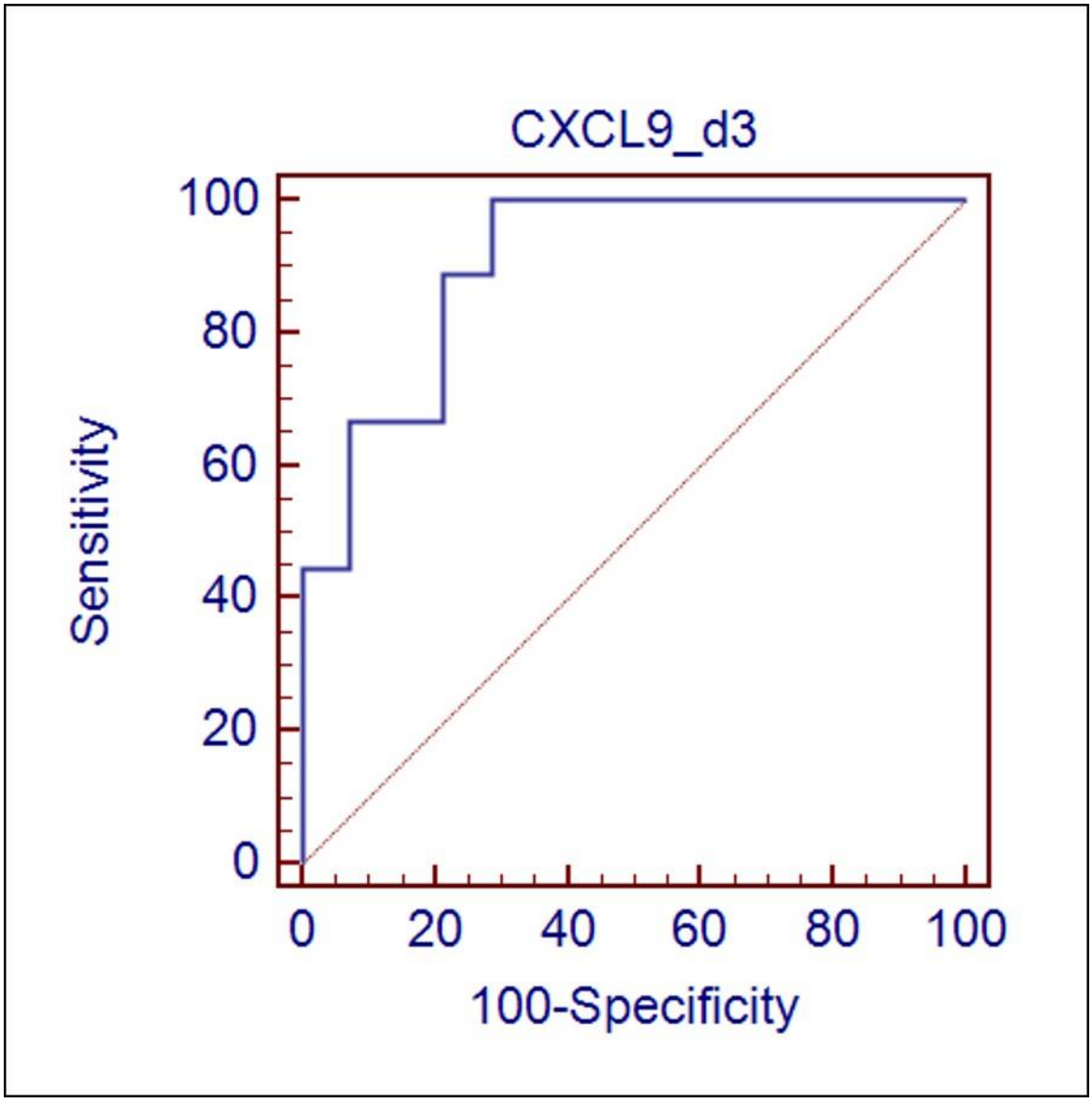

Fig. S1. **ROC curve for CXCL9 at day +3 after HSCT.**

Figure S2

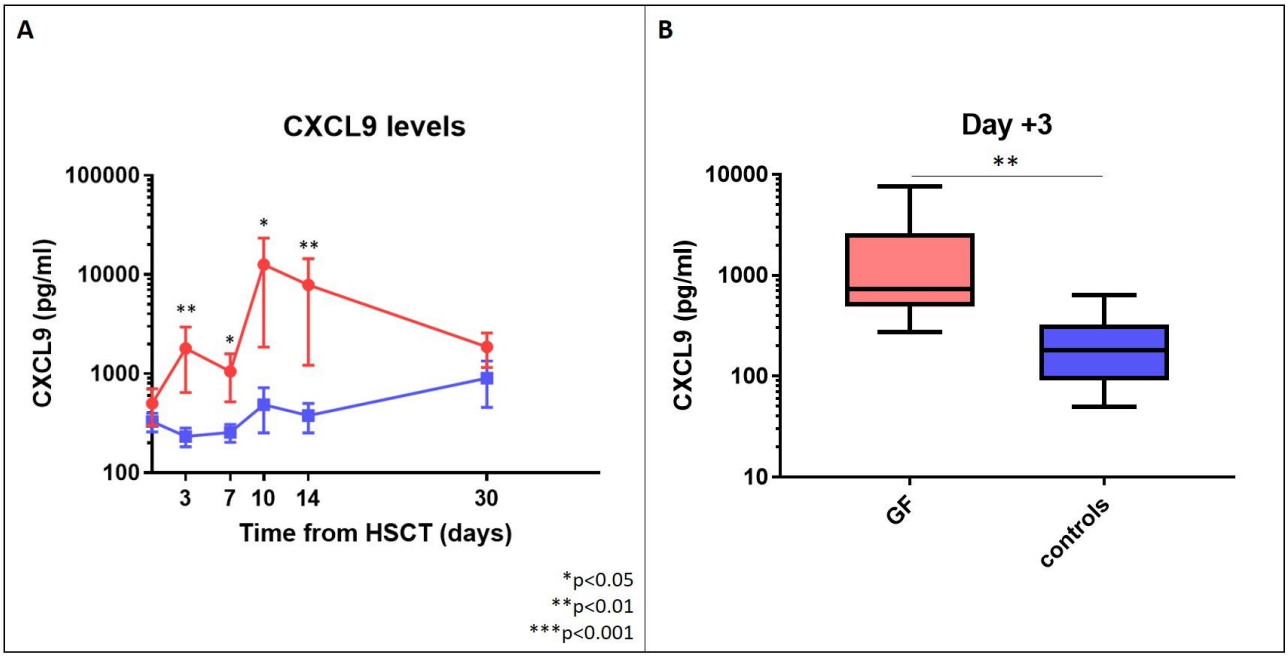

Fig. S2. **CXCL9 levels of GF patients, without HLH patients, compared to controls**. A. CXCL9 levels during time (mean and SEM). B. CXCL9 levels at day +3 after HSCT (median, 25°-75° centiles, range).

Figure S3

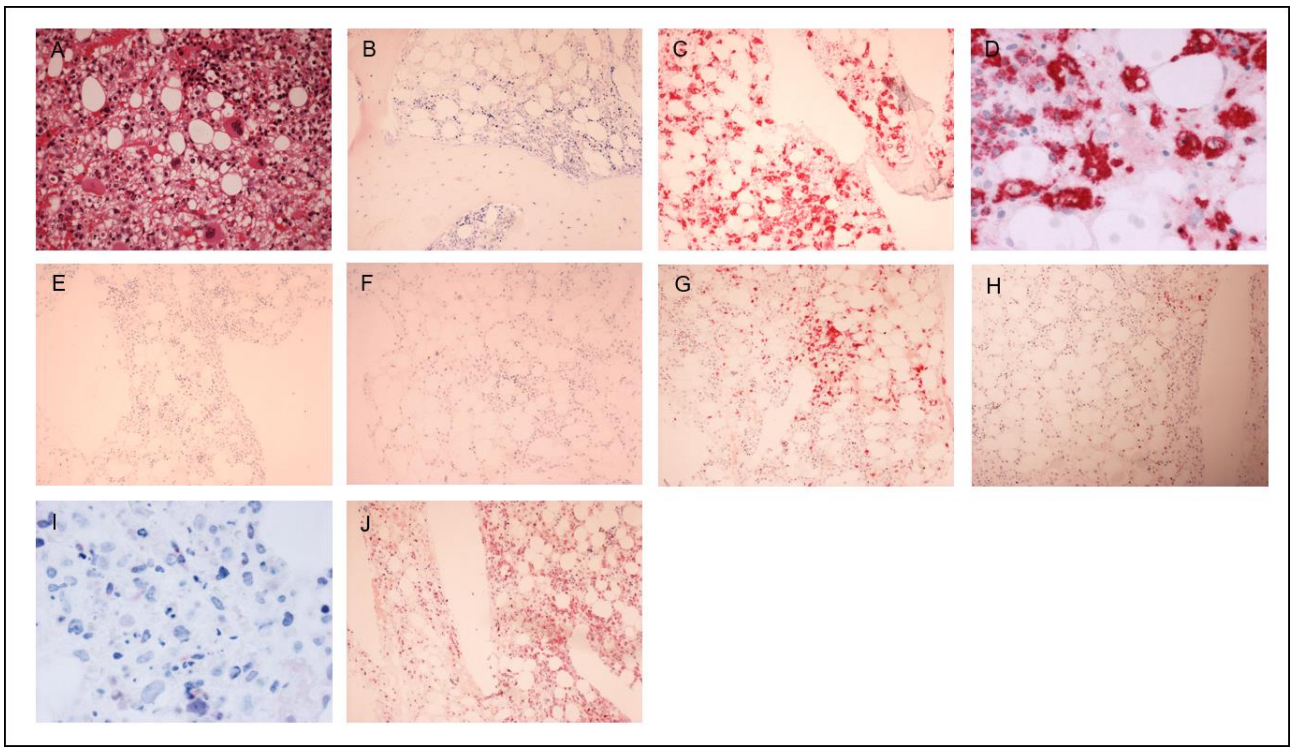

Fig. S3. **Immunohistochemistry evaluation of bone marrow specimens in a patients with graft failure (Pt#4 and #6).** Panel A. H&E staining of BM specimen at 20X magnification. Panel B. May Grünwald-Giemsa staining for the evaluation of the hemopoietin colonies (10X). Panel C-D. Characterization of macrophage population by CD68 (10X and 40X). Pane E-F. Evaluation of B cell compartments by the expression of CD79 and CD20 markers (10X). T lymphocytes was confirmed by the expression of CD5 (Panel G). Cytotoxic activity of T lymphocytes was evaluated by the expression of granzyme B (Panel H),Perforin (Panel I) and TIA-1 (Panel J) (10X).

Figure S4

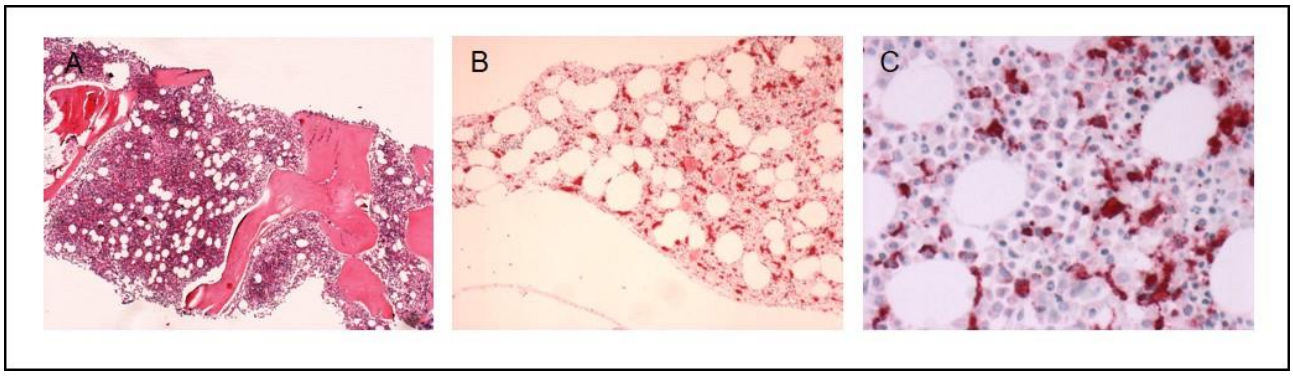

Fig. S4. **Immunohistochemistry evaluation of bone marrow specimens in a patient with sustained donor cell engraftment after HSCT (Pt#9)**. A. H&E staining of BM specimen at 4X magnification. B-C. Characterization of macrophage population by CD68 (10X and 40X).

Figure S5

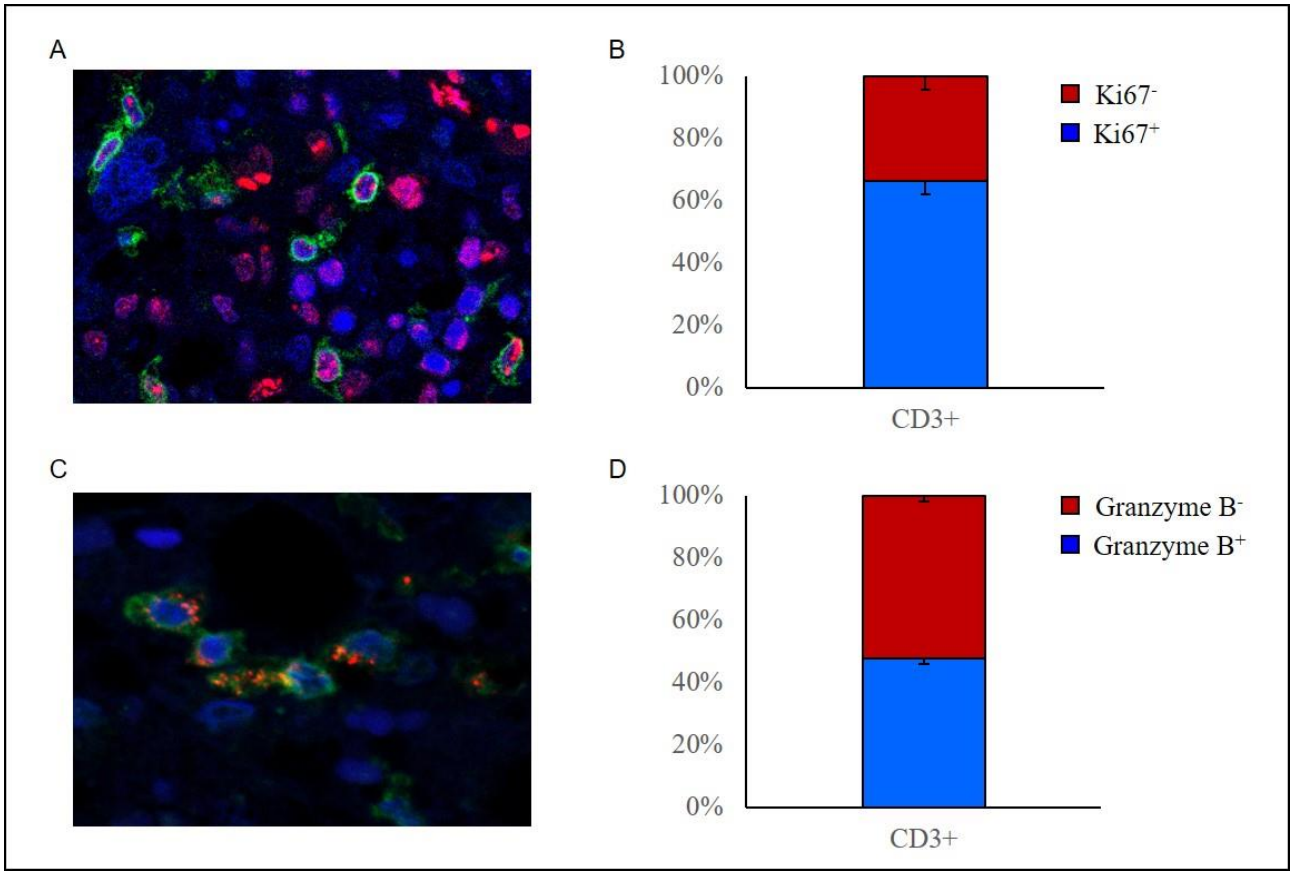

Fig. S5. **Characterization of the T-cell compartment in a bone marrow biopsy of a patient experiencing graft failure (Pt#4)**. A. Particular of 40X magnification of immunofluorescence staining of T lymphocytes (green) and Ki67 proliferation marker (red). B. The graph shows the average number ± SE of Ki67 expression in CD3+ cells per high power field (n=6, 145 CD3+ cells). C. Particular of 40X magnification of immunofluorescence staining of T lymphocytes (green) and Granzyme B as a marker for cytotoxic activity (red). D. The graph shows the average number ± SE of Granzyme B expression in CD3+ cells per high power field (n=3, 165 CD3+ cells).

Figure S6

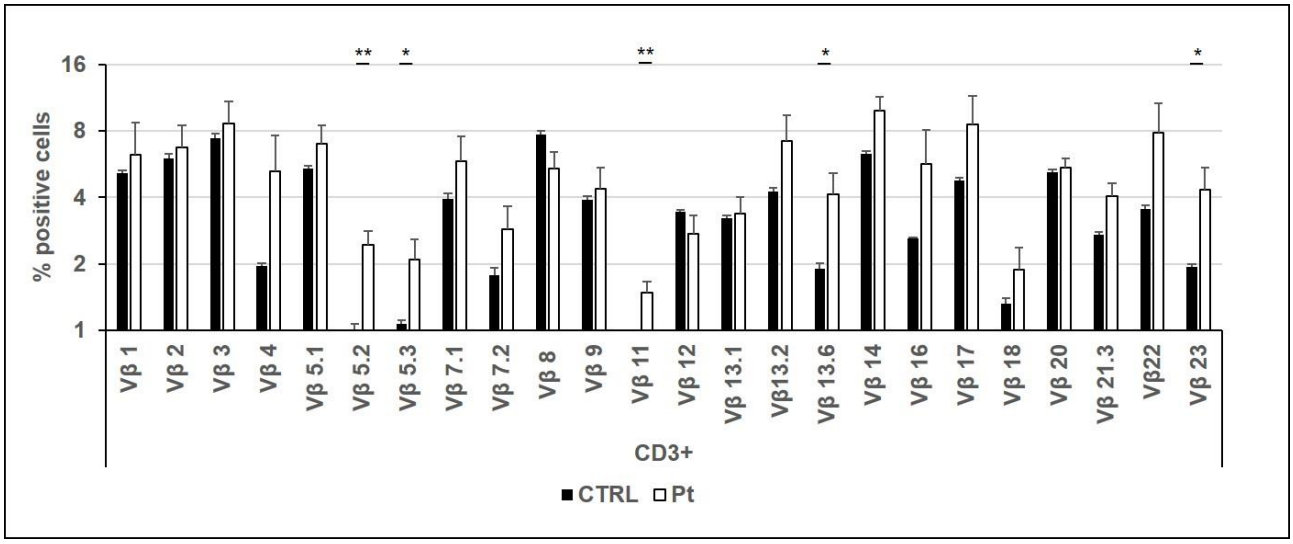

Fig. S6. **Analysis of T-cell receptor repertoire in patients who did or did not experience GF.** Panel A. Frequency of positive T lymphocytes stained with each indicated Vβ Abs among gated CD3+ cells in CTRL (black bars; n=11) and GF patients (white bars; n=7). Bars indicate the mean ± standard error. Two-way ANOVA and Student's t-test were carried out between CTRL and patients with GF on single Vβ families.

Figure S7

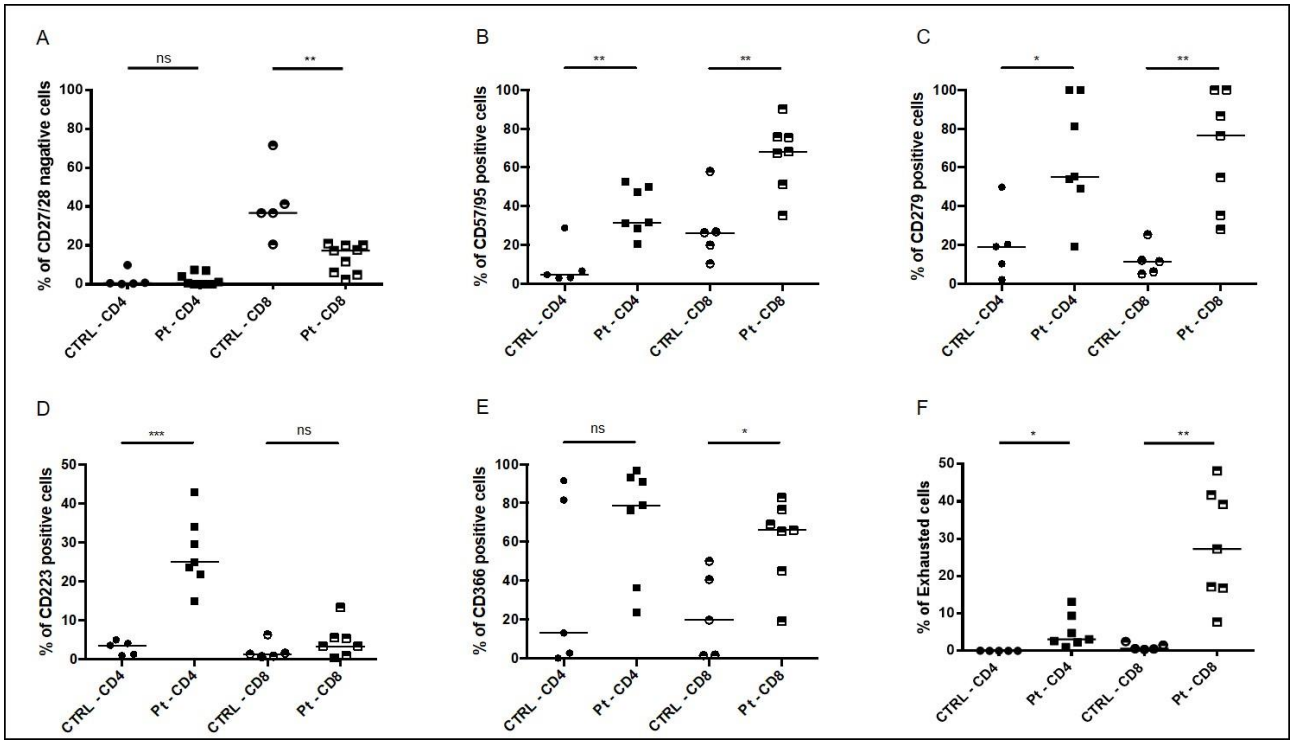

Fig. S7. **Exhaustion profile of CD4+ or CD8+. Flow-cytometry analysis of both CD4+ and CD8+ T-cell populations for the expression of activation and exhaustion markers.** CD27/CD28 (A), CD95/CD57 (B), CD279 (C), CD223 (D), CD366 (E) and the combination of CD57/CD95/CD223/CD279 for the CD4 population or CD57/CD95/CD279/CD366 for the CD8 population (F). Data showing the percentage of each patient or CTRL is represented by a symbol and the median is marked by a horizontal line.

Figure S8

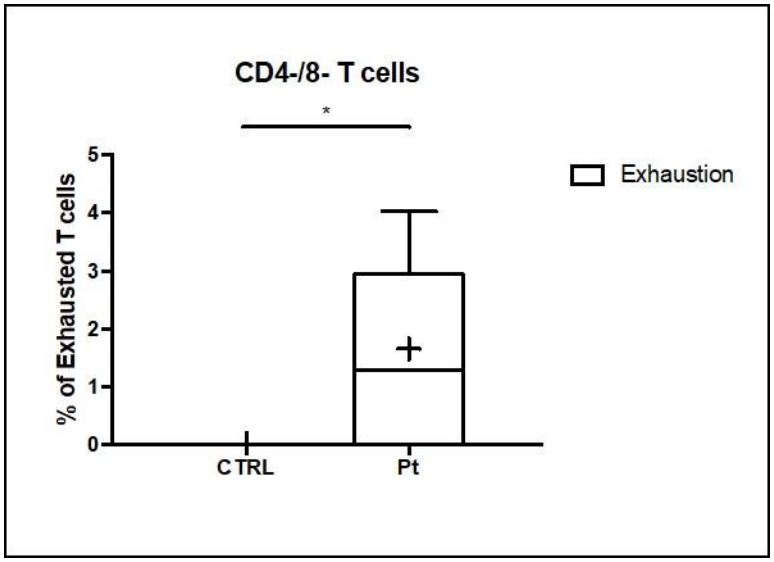

Fig. S8. **Immuno-phenotype analysis investigating exhaustion markers in γδ T lymphocytes**. Evaluation of the co-expression of CD223, CD279 and CD366 on CD3+ CD4 and CD8 double negative population in both CTRL and patients with GF. Data show average (+) and median ± SD.

.