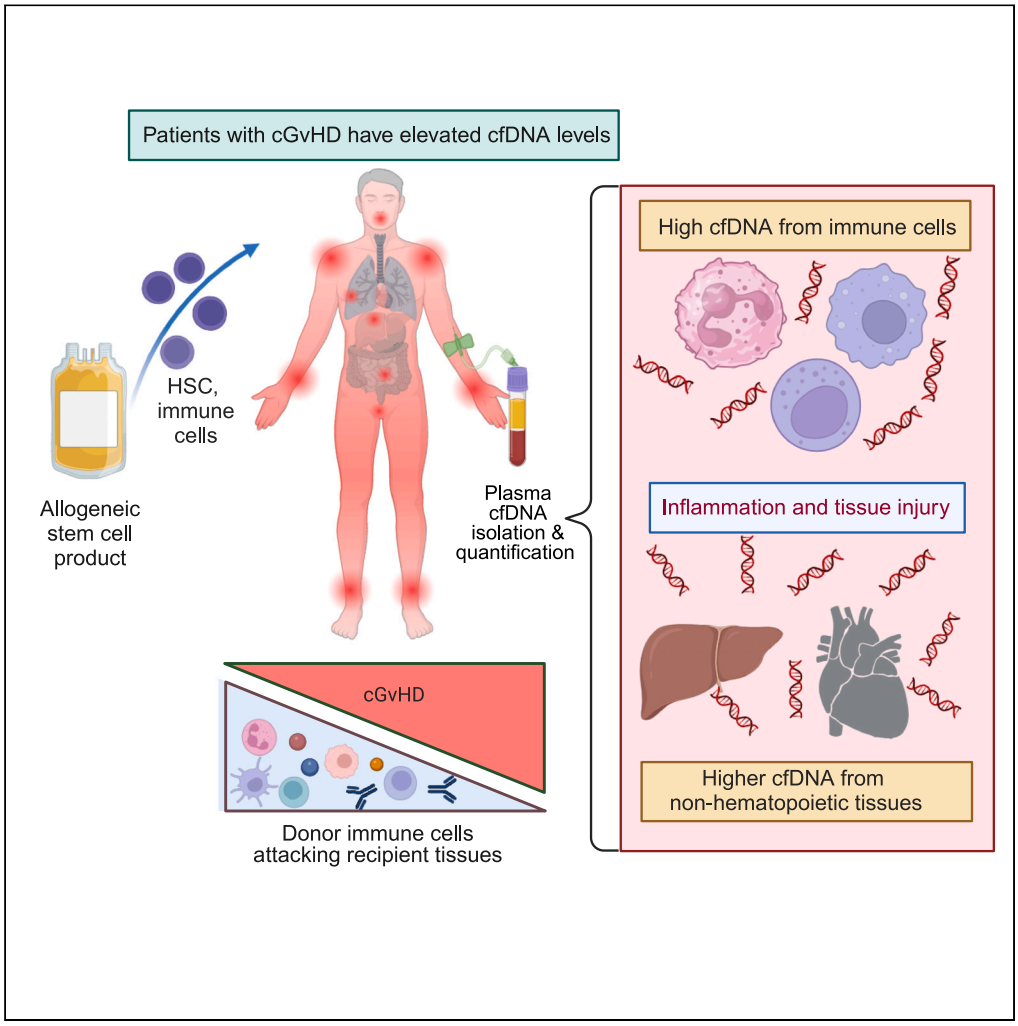
iScience

CellPress
OPEN ACCESS

Article

Chronic graft-versus-host disease is characterized by high levels and distinctive tissue-of-origin patterns of cell-free DNA



Yifan Pang, Temesgen E. Andargie, Moon Kyoo Jang, ..., Noa G. Holtzman, Steven Z. Pavletic, Sean Agbor-Enoh

sean.agbor-enoh@nih.gov

Highlights

Cell-free DNA (cfDNA) level increases with hematopoietic stem cell transplant (HSCT)

Tissue source of cfDNA differed between healthy subjects and HSCT recipients

Chronic graft-versus-host disease (cGvHD) had unique cfDNA signatures

CfDNA patterns are potentially useful biomarkers for cGvHD

Pang et al., iScience 26, 108160
November 17, 2023 © 2023
https://doi.org/10.1016/j.isci.2023.108160

iScience

CellPress
OPEN ACCESS

Article

Chronic graft-versus-host disease is characterized by high levels and distinctive tissue-of-origin patterns of cell-free DNA

Yifan Pang,[1,2] Temesgen E. Andargie,[3,4] Moon Kyoo Jang,[3] Hyesik Kong,[3] Woojin Park,[3] Thomas Hill,[5] Neelam Redekar,[5] Yi-Ping Fu,[6] Desai A. Parth,[7] Noa G. Holtzman,[8] Steven Z. Pavletic,[8] and Sean Agbor-Enoh[3,9,10,*]

SUMMARY

Chronic graft-versus-host disease (cGvHD) is a devastating complication of hematopoietic stem cell transplantation (HSCT). Effective early detection may improve the outcome of cGvHD. The potential utility of circulating cell-free DNA (cfDNA), a sensitive marker for tissue injury, in HSCT and cGvHD remains to be established. Here, cfDNA of prospectively collected plasma samples from HSCT recipients (including both cGvHD and non-cGvHD) and healthy control (HC) subjects were evaluated. Deconvolution methods utilizing tissue-specific DNA methylation signatures were used to determine cfDNA tissue-of-origin. cfDNA levels were significantly higher in HSCT recipients than HC and significantly higher in cGvHD than non-cGvHD. cGvHD was characterized by a high level of cfDNA from innate immune cells, heart, and liver. Non-hematologic tissue-derived cfDNA was significantly higher in cGvHD than non-cGvHD. cfDNA temporal dynamics and tissue-of-origin composition have distinctive features in patients with cGvHD, supporting further exploration of the utility of cfDNA in the study of cGvHD.

BACKGROUND

Annually, around 8,000 hematopoietic stem cell transplants (HSCT) are performed in the U.S. and >50,000 worldwide. By 2025, there will be 160,000 long-term survivors of HSCT in the U.S. alone.[1–3] Chronic graft-versus-host disease (cGvHD) is one of the most debilitating complications of HSCT that occurs in 30–70% of long-term survivors despite prophylactic immunosuppressants, mostly within five years of HSCT.[4,5] As a result of the transplanted donor cells (graft) attacking the recipient body (host), cGvHD is a syndrome of multiorgan damage, causing diminished quality of life and increased morbidity and mortality.[6,7] cGvHD is diagnosed using the 2014 National Institutes of Health (NIH) consensus criteria, an organ-based, severity scoring system based on patient symptoms and laboratory testing of multiple additional factors, such as transaminases.[6] Invasive tests such as colonoscopy, liver biopsy, muscle biopsy, and bronchoscopy, are frequently required to further document end-organ involvement by histopathology. Unfortunately, the invasiveness of these procedures and lack of specificity of biopsy/histopathology remains a limitation, making precise diagnosis and monitoring challenging. In cases without complete remission, these suboptimal responses may be contributing to the delayed diagnosis of cGvHD, as the current diagnostic criteria is unable to identify subclinical cGvHD prior to the development of significant and often irreversible organ damage.[8–13] One potential remedy for this situation is the use of circulating cell-free DNA (cfDNA), short DNA fragments (~160 base-pairs) released from tissue to plasma during apoptosis, necrosis, or active secretion. cfDNA is highly sensitive for organ injury and has been widely used in detecting solid organ transplant graft rejection. However, there is limited data regarding the utility of cfDNA in HSCT.[14] In this pilot study, we compared the cfDNA dynamics from HSCT recipients with and without cGvHD and healthy controls.

[1]National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 20892, USA
[2]Department of Hematologic Oncology and Blood Disorders, Levine Cancer Institute, Charlotte, NC 28204, USA
[3]Laboratory of Applied Precision Omics, National Heart Lung and Blood Institute, National Institutes of Health, Bethesda, MD 20892, USA
[4]Department of Biology, Howard University, Washington, DC 20059, USA
[5]NIAID Collaborative Bioinformatics Resource, National Institute of Allergy and Infectious Diseases, NIH, Bethesda, MD 20892, USA
[6]Office of Biostatistics Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 20892, USA
[7]Department of Hematology/Oncology, Fox Chase Cancer Center, Philadelphia, PA 19111, USA
[8]Immune Deficiency Cellular Therapy Program, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA
[9]Department of Pulmonary and Critical Care Medicine, Johns Hopkins Hospital, Baltimore, MD 21205, USA
[10]Lead contact
*Correspondence: sean.agbor-enoh@nih.gov
https://doi.org/10.1016/j.isci.2023.108160

Table 1. Clinical characteristics of the study cohort

Characteristic	cGvHD N = 63	Non-cGvHD N = 13	P (cGvHD vs. non-cGvHD)	HC N = 19	P (HSCT vs. HC)
Age at HSCT (mean ± SD)[a]	43.3 ± 14.4	44.2 ± 14.7	0.836	54.9 ± 13.3	<0.001
Age at sample collection (mean ± SD)	44.6 ± 13.2	48.5 ± 14.1	0.107	54.9 ± 13.3	0.006
Sex (female/male)	27/36	6/7	0.827	12/22	0.804
Race			0.116		0.003
AA	1	0		6	
Caucasian	56	9		24	
Hispanic	4	2		0	
Asian & others	2	2		5	
Indication			0.016		
AML, ALL, MDS	36	3			
CML, MPN	5	0			
CLL, HL, NHL	19	10			
Others	3	0			
Conditioning (MAC/RIC)	31/31[c]	2/11	0.045		
Cell source (PB/BM)	55/7[c]	13/0	0.455		
HLA (matched related/others)	33/30	6/7	0.683		
Acute GvHD	43	6	0.130		
Secondary malignancy	14	0	0.110		
Relapse/refractory disease	9	1	1.000		
Alive[b]	36	12	0.026		

cGvHD, chronic graft-versus-host disease; SD, standard deviation; AA, African-American; AML, acute myeloid leukemia; ALL, acute lymphoblastic leukemia; MDS, myelodysplastic syndrome; CML, chronic myelogenous leukemia; MPN, myeloproliferative neoplasm; CLL, chronic lymphocytic leukemia; HL, Hodgkin's lymphoma; NHL, non-Hodgkin's lymphoma; NA: not applicable; MAC, myeloablative conditioning; RIC: reduced intensity conditioning; PB, peripheral blood; BM, bone marrow; HLA, human leukocyte antigen.

[a]The comparison of age was performed by t-test, the rest of the analyses were performed by chi-square test.

[b]As of April 2021.

[c]missing data in 1 subject.

RESULTS

Patient characteristics

This pilot study included 95 subjects, 76 HSCT subjects and 19 healthy controls (Table 1). The 76 HSCT subjects comprised of 63 subjects with cGvHD and 13 subjects without cGvHD (non-cGvHD). The 19 healthy controls subjects (HC) were older than the HSCT subjects at sample collection (54.9 ± 13.3 vs. 46.2 ± 13.9, p = 0.010) and had significantly more African-American subjects (15.8% vs. 1.3%, p = 0.019). The 63 cGvHD subjects consisted of 5 patients with moderate and 58 with severe cGvHD. Median time from HSCT to the diagnosis of cGvHD was 7.1 months [interquartile range (IQR) = 5.1–10.5]. The cGvHD and non-cGvHD groups were matched in age, sex, race, and source of stem cells at HSCT. In total, 106 plasma samples from 61 cGvHD subjects and 46 samples from 12 non-cGvHD subjects were sequenced (Figure S1). For the subset of patients with longitudinal samples (n = 20), samples collected on the day of transplant (day 0) and around day 30, 60, 100, 180, and year 1, 2, 3, and 4 after transplant were included to define temporal cfDNA trends. For the rest of the HSCT cohort (n = 56), single time points were collected at a median of 1,103 (range, 433–4321) days post-HSCT in the cGvHD group and 1,828 days post-HSCT for the one non-cGvHD subject. We find no statistical differences in rates of acute GvHD (68.3% vs. 46.1%, p = 0.130) and secondary malignancy between cGvHD and non-cGvHD (22.0% vs. 0.0%, p = 0.110). However, acute leukemia/myelodysplastic syndrome was the most common indication for HSCT in the cGvHD group (57.1%), while lymphoid malignancies were the most common indication in the non-cGvHD group (76.9%) myeloablative conditioning was more common in the cGvHD group (49.2% vs. 15.4%, p = 0.045). Median post-HSCT follow-up (months) for the cGvHD and non-cGvHD were similar (105.0, IQR = 65.0–146.0 vs. 113.0, IQR = 96.5–141.0, p = 0.542). Over this time, 92.3% of the non-cGvHD patients were still alive, compared to 57.1% of the cGvHD group (p = 0.026). Median time from HSCT to cGvHD diagnosis was 214 days (IQR, 152–314). cGvHD involvement included the liver (n = 36, 57.1%), gastrointestinal tract (n = 25, 39.7%), skin (n = 49, 77.8%), and joint/fascia (n = 42, 66.7%) (Table S1). The mean forced expiratory volume (FEV1) of subjects with cGvHD was 47.6 ± 18.08% predicted at the end of study follow-up, significantly lower than the mean FEV1 of the non-GvHD group of 102.3 ± 15.2% predicted (p < 0.001). Patients with cGvHD received a median of five lines (IQR = 4–7) of systemic immunosuppressive therapies (Table S1).

Corresponding chimerism testing by short tandem repeat (STR) was available for 7 of the 9 cGvHD subjects (34 available results) with longitudinal samples and 10 of the 11 non-cGvHD subjects (31 available results) with longitudinal samples. Five of the 17 subjects achieved 100% chimerism in total, T cell and myeloid lineages by post-HSCT day 100, and 16 achieved 100% chimerism by 1-year post-HSCT. One patient had chronic mixed chimerism without evidence of leukemia relapse.

HSCT recipients were characterized by persistently elevated cfDNA

HSCT recipients had higher cfDNA levels than HC

cfDNA was isolated from plasma samples (cGvHD, 106; non-cGvHD, 46; HC, 19) and quantified by quantitative polymerase chain reaction (qPCR) using primers targeting Alu115. Lambda DNA was spiked into the starting plasma sample to correct for extraction loss. The levels of cfDNA in HSCT recipients were significantly higher than HC (52.8, 31.5–92.8 vs. 12.3, 9.0–20.2 ng/mL, $p < 0.001$), while patients with cGvHD had higher level of cfDNA than patients without cGvHD (160.2, 37.3–106.2 vs. 37.0, 22.6–67.8 ng/mL, $p < 0.001$).

Temporal cfDNA dynamics of cGvHD differed from non-cGvHD by higher cfDNA level post-HSCT

To further elucidate the temporal trend of cfDNA and the difference between cGvHD and non-cGvHD, we analyzed cfDNA trends from early post-HSCT period before diagnosis of cGvHD (day0 to day100), mid post-HSCT (day101 to 2 years), and late post-HSCT (>2 years). Median total cfDNA level remained persistently higher post-HSCT compared to healthy controls ($p < 0.001$ in all three intervals, Table S2); with HSCT subjects showing cfDNA levels of 39.2 ng/mL (23.9–79.2) between day0–day100, 61.5 ng/mL (39.1–94.8) between day101 and 2 years, and 53.9 ng/mL (33.4–95.5) beyond 2 years post-HSCT. There was no significant temporal variance in cfDNA in the HSCT group ($p = 0.302$). Similarly, subjects with cGvHD showed higher cfDNA than the non-cGvHD for all time intervals considered but only reaching statistical significance in the late post-HSCT period (57.7, 34.3–101.5 vs. 25.3, 22.6–54.9 ng/mL, $p = 0.032$) (Figures 1A–1D; Table S2). There were no significant temporal variances in total cfDNA levels between the cGvHD and non-GvHD groups (Figures 1E and 1F).

cfDNA tissue-of-origin methylation pattern

Blood cells are the major source of cfDNA in both HSCT recipients and HC

Bisulfite sequencing and the cfDNAme deconvolution algorithm were applied to determine the cell or tissue source of cfDNA as previously described. The approach has both technical and biological validity and has been used to define biologically plausibility tissue injury of clinical relevance.[15,16] The distribution of major tissue cfDNA sources for each group is shown in Figure 2 and Table S3. Consistent with prior work in HC,[17] cells of hematologic origin were the major sources of cfDNA. Components of the innate immunity, including neutrophils, eosinophils, monocytes, NK-cells, and macrophages, comprised greater than half of all cfDNA. The main sources of non-hematopoietic cfDNA included vascular endothelium (4.6%), heart (3.6%), liver (0.7%), and lung (0.5%). In HSCT recipients, cells of the innate immunity were also the major contributors to plasma cfDNA, however, the percentage of neutrophil-derived cfDNA was higher (39.2%) and eosinophil-derived cfDNA was lower (6.0%) in HSCT recipients (Figures 2A–2C). cGvHD demonstrates higher fractional contribution of cfDNA from macrophage, liver, and vascular endothelium compared to non-GvHD (Table S3). Through the course of HSCT, starting on day 0, cfDNA compositions were different between HC and HSCT recipients, indicating altered tissue remodeling and heightened tissue damage and inflammation in HSCT recipients (Figure 2). In the post-transplant course, percentage of innate immunity derived cfDNA gradually increased, while stem and progenitor cell-derived cfDNA gradually decreased HSCT recipients without cGvHD (Figures 2D–2H). Patients with cGvHD followed a similar pattern but percentage of stem and progenitor cell-derived cfDNA rose again in the late post-transplant period (Figures 2I–2M).

Absolute tissue-specific cfDNA levels were computed by multiplying the tissue fraction from deconvolution to the total cfDNA levels. Monocyte, neutrophil, stem and progenitor cells, liver, lung, pancreas, and vascular endothelium-derived cfDNA were significantly lower in HC compared with HSCT recipients (all $p < 0.001$, Table S4). Patients with cGvHD had higher cfDNA levels than the non-cGvHD group across all major tissue types, except for the lung and vascular endothelium (Figures 3A–3H). There was a trend of higher adaptive immunity (B cells and T cells)-derived cfDNA in the cGvHD group (0.3, 0.0–2.2 vs. 0.0, 0.0–0.7 ng/mL, $p = 0.051$). In HSCT recipients with ≥ 5 longitudinal samples, temporal trend cfDNA level and methylome pattern had clinical implications (Figures 3N–3U). There was significant variation among subjects. Generally, cfDNA peaked around the time of engraftment and rose again with active infection, new cGvHD symptoms, or occurrence of other complications (Figures 3N–3U; Figure S2). In the non-cGvHD group, most subjects had stable or down-trending, low cfDNA levels (Figure S3F, S3H, S3J, S3N, S3P, and S3R), but some had rising cfDNA without clinical etiology (Figure S3B, S3D, and S3L).

There was no significant correlation between tissue-derived cfDNA level from blood cell-derived cfDNA level and the respective blood cell count. There were weak correlations between liver-derived cfDNA level and alanine aminotransferase ($R^2 = 0.279$, $p < 0.001$) and aspartate aminotransferase ($R^2 = 0.230$, $p < 0.001$) but not with total bilirubin (Figure S4). Within the cGvHD group, subjects with liver cGvHD had significantly higher levels of liver-derived cfDNA than those without (4.5, 1.1–8.5 vs. 0.6, 0.2–1.8 ng/mL, $p < 0.001$). Skin-derived cfDNA did not differ significantly between those with or without skin cGvHD ($p = 0.265$) nor did lung-derived cfDNA differ between those with or without pulmonary cGvHD ($p = 0.110$).

Temporal dynamic of cfDNA from hematologic and non-hematologic tissues

cfDNA from hematologic tissues, including neutrophils, monocytes and macrophages, eosinophils, dendritic cells, lymphocytes, stem and progenitor cells, were combined into heme-cfDNA, while the rest from non-hematologic tissues were combined into

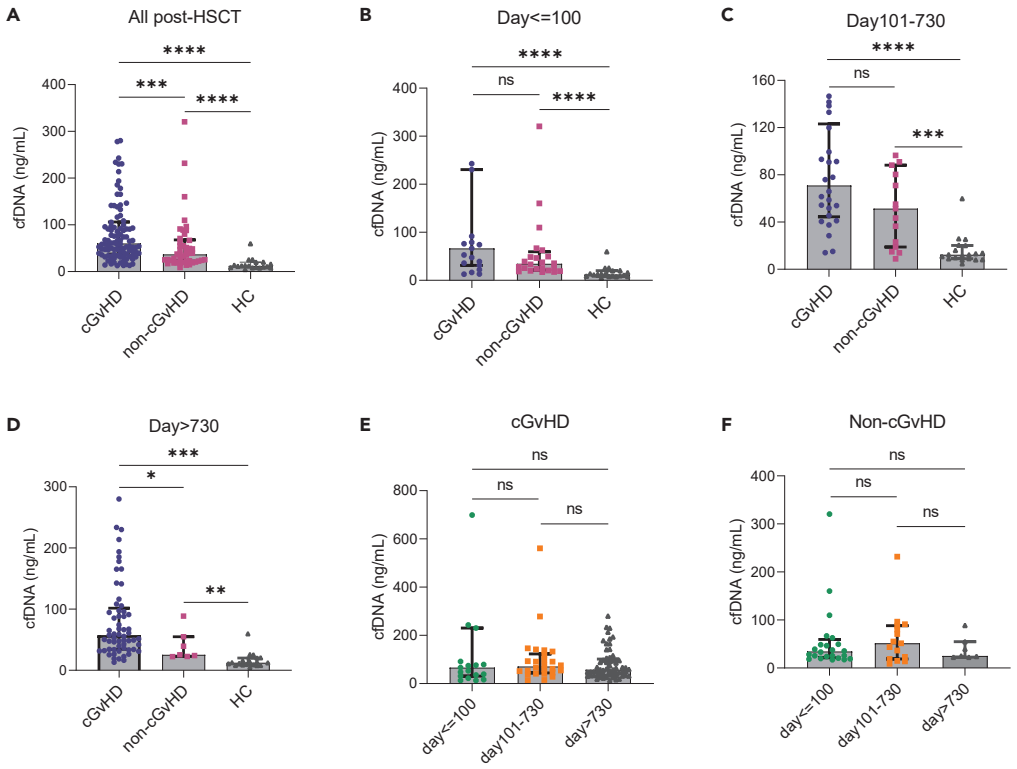


Figure 1. Sustained elevation of cfDNA levels (ng/mL) after transplant, especially in patients with cGvHD

(A) In all samples, cfDNA levels were significantly higher in patients with cGvHD compared with non-cGvHD and HC, cfDNA levels were higher in non-cGvHD than HC.

(B and C) Within day100 (B) and day101–730 post-transplant (C), cfDNA levels were higher in transplant recipients than HC, but there was no significant difference between cGvHD and non-cGvHD.

(D) Beyond day730, patients with cGvHD had higher cfDNA than non-cGvHD and HC.

(E and F) cfDNA did not have statistically significant temporal variance in cGvHD (E) or non-cGvHD (F). Mann-Whitney U test. Column: median. Error bar: interquartile range.

non-heme cfDNA (Table 2). While heme and non-heme cfDNA percentage contribution did not differ between cGvHD and non-cGvHD ($p = 0.639$ and $p = 0.550$, respectively), plasma concentration of heme cfDNA was higher in cGvHD than non-cGvHD (48.9, 29.6–91.6 vs. 26.2, 15.2–62.6 ng/mL, $p < 0.001$) (Figure 4A), as was plasma concentration of non-heme cfDNA (8.5, 4.5–16.6 vs. 5.4, 2.7–12.0 ng/mL, $p = 0.022$) (Figure 4B). The differences in heme and non-heme cfDNA between cGvHD and non-cGvHD were most significant in the late post-HSCT period, with higher non-heme derived cfDNA percentage (12.3, 9.4–18.0 vs. 7.9, 6.8–11.1%, $p = 0.022$) and quantity (8.6, 4.1–16.5 vs. 2.7, 1.6–2.7 ng/mL, $p = 0.001$) in the cGvHD group, indicating greater tissue injury than those without cGvHD (Figure 4H).

In the cGvHD group, we found no significant temporal change of heme and non-heme cfDNA after HSCT (Figures 4I and 4J). Injury from non-heme tissue types remained elevated over the early, mid, and late post-HSCT intervals (percentage, 15.1, 8.4–23.0 vs. 11.9, 6.0–19.0 vs. 12.3, 9.4–18.0%, $p = 0.445$; concentration 12.0, 6.4–21.2 vs. 7.7, 3.9–16.3 vs. 8.6, 4.1–16.5 ng/mL, $p = 0.575$). On the contrary, in the non-cGvHD group, the percentage of heme-derived cfDNA significantly increased over time (77.5, 41.0–89.2 vs. 89.1, 86.2–91.8 vs. 92.1, 89.3–93.2%, $p < 0.001$), while for non-heme tissues, the cfDNA percentage (22.3, 11.1–40.3 vs. 11.0, 8.2–13.1 vs. 7.9, 6.8–11.1%, $p < 0.001$) and concentration (8.3, 4.6–16.6 vs. 4.5, 2.5–8.8 vs. 2.7, 1.6–2.7 ng/mL, $p = 0.001$) decreased with time (Figures 4K and 4L).

Reproducibility of different analytic platforms: Meth_Atlas and cfDNAme

To demonstrate reproducibility of our findings across different analysis algorithm, we analyzed bisulfite sequence reads using another tissue-specific DNA methylation libraries, Meth_Atlas.[17] The Meth-Atlas algorithm employs array-based reference methylomes while cfDNAme is sequencing based and utilizes tissue-specific DNA methylation region. Some tissue type representation differs between the two libraries. Despite the differences, our principal observations were similar between the two approaches. In both HC and HSCT recipients, hematopoietic cells remained the major contributors of cfDNA. The absolute tissue-specific cfDNA levels of monocyte, neutrophil, adipose tissue, hepatocyte,

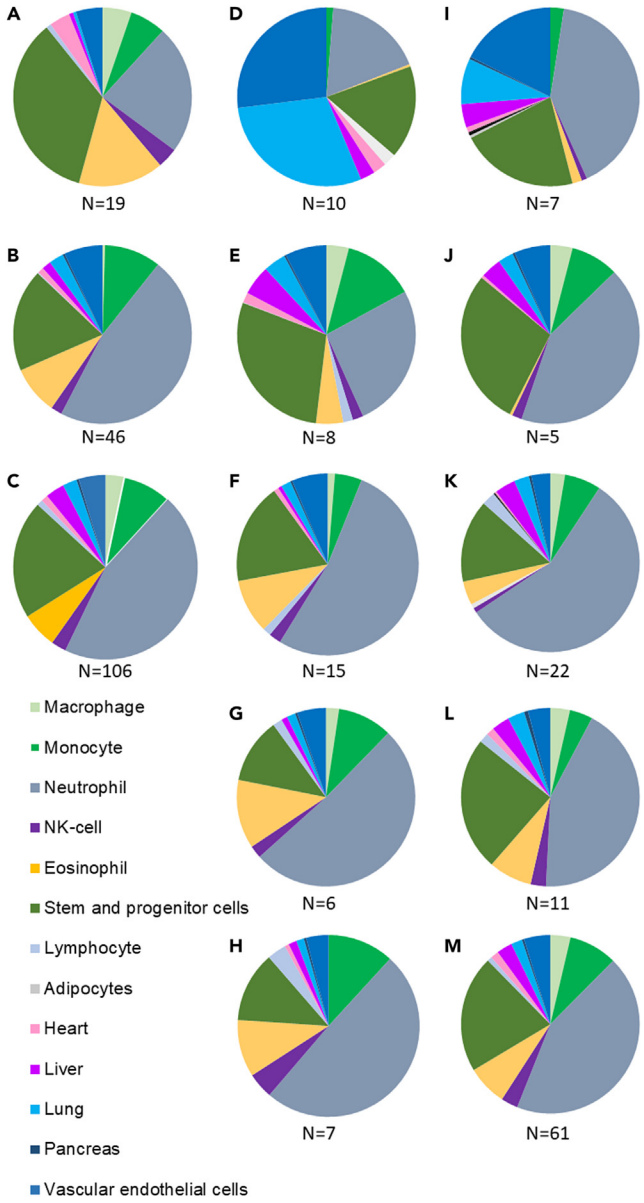


Figure 2. cfDNA methylome of HC, non-cGvHD group and cGvHD group based the median percentage of each tissue

A–C. methylome of all samples in the HC (A), non-cGvHD (B) and cGvHD (C) groups. Compared with healthy volunteers, HSCT recipients had higher percentage of neutrophil-, hepatocyte- and vascular endothelium-derived cfDNA, lower erythroid precursor-derived cfDNA (Mann-Whitney U-test, all $p<0.05$). The cGvHD group had a trend of higher respiratory tract-derived cfDNA and lower GI tract-derived cfDNA percentage than the non-cGvHD group (Mann-Whitney U-test, $p=0.068$, $p=0.051$).

D–H: Temporal changes in cfDNA methylome in the non-cGvHD group. D. Around day 0. E. Around 30-day post-HSCT, i.e., around engraftment. F. Two-month to one-year post-HSCT. G. one to two years post-HSCT. H. beyond two years post-HSCT.

I–M: Temporal changes in cfDNA methylome in the cGvHD group. I. Around day 0. J. Around 30-day post-HSCT, i.e., around engraftment. K. Two-month to one-year post-HSCT. L. one to two years post-HSCT. M. beyond two years post-HSCT.

and kidney cfDNA were significantly lower in HC compared with HSCT recipients ($p < 0.001$, <0.001, <0.001, <0.001, and = 0.022, respectively). cGvHD patients had higher neutrophil (28.2, 14.8–57.9 vs. 15.6, 6.1–40.0 ng/mL, $p = 0.005$), erythroid/platelet precursor (6.1, 3.4–12.3 vs. 3.9, 1.5–6.3 ng/mL, $p < 0.001$), monocyte (7.9, 3.5–13.9 vs. 4.4, 2.1–8.6 ng/mL, $p = 0.012$), hepatocyte (1.2, 0.1–6.2 vs. 0.4, 0.0–1.0 ng/mL, $p = 0.005$), and airway stratified epithelium cfDNA (0.6, 0.0–1.6 vs. 0.0, 0.0–0.7 ng/mL, $p = 0.005$) (Table S6). Finally, Meth_Atlas analysis found the same temporal trend of heme and non-heme cfDNA in HSCT recipients with or without cGvHD as cfDNAme analysis. Subjects with cGvHD again had higher plasma concentrations of heme cfDNA and non-heme cfDNA than those without cGvHD (47.8, 29.4–89.9 vs. 26.1, 15.5–62.6 ng/mL,

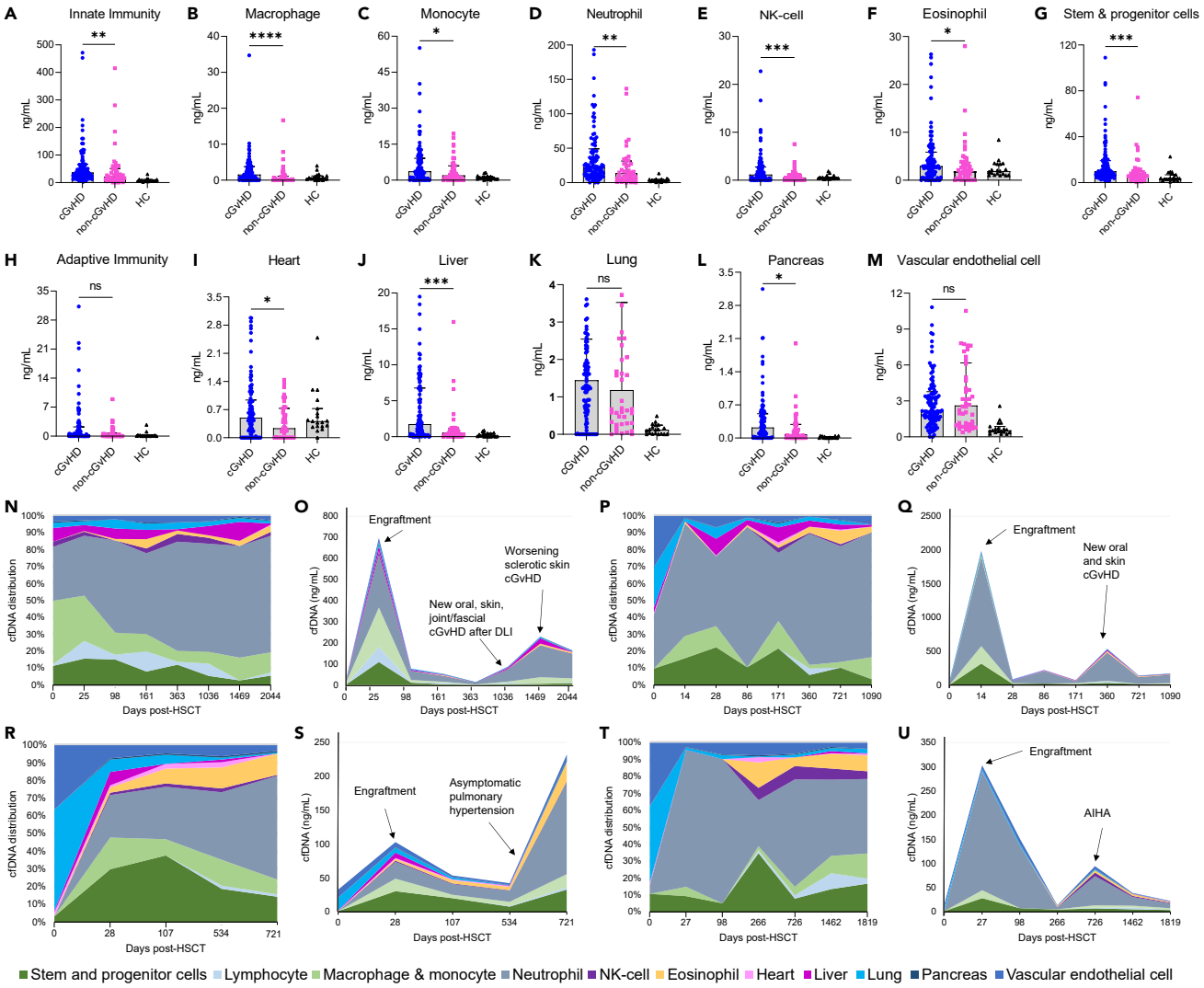


Figure 3. Distinctive cfDNA methylome patterns among cGvHD, non-cGvHD, and HC

A–M: cfDNA from different tissue types of the three groups of samples. Significant differences in tissue-specific cfDNA levels between cGvHD and non-cGvHD were demonstrated by Mann-Whitney U-test in cells in innate immunity (A), specifically macrophage (B), monocytes (C), neutrophils (D), NK-cells (E) and eosinophils (F). Significant differences were also observed in stem/progenitor cells (G), heart (I), liver (J), pancreas (K) and vascular endothelium-derived cfDNA (M). No significant differences were observed in adaptive immunity (H) and lung-derived cfDNA (K). N-U: Temporal trend of cfDNA distribution and level in prototype subjects with ≥ 5 longitudinal samples.

N–Q. subjects with cGvHD. R-U. non-cGvHD subjects. cfDNA trend corelated with major clinical events such as engraftment, active infection, the occurrence of new cGvHD or other post-HSCT complications. cGvHD: chronic graft-versus-host disease. DLI: donor lymphocyte infusion. AIHA: autoimmune hemolytic anemia.

p = 0.008; 8.2, 4.6–18.4 vs. 6.0, 2.5–11.4 ng/mL, p = 0.005). The late post-HSCT period was also the time when the differences in heme and non-heme cfDNA between cGvHD and non-cGvHD were the most pronounced. Higher percentage (13.1, 10.0–20.4 vs. 7.4, 4.8–11.6%, p = 0.003) and quantity (7.8, 4.4–18.0 vs. 1.9, 1.7–3.2 ng/mL, p = 0.015) of non-heme derived cfDNA in the cGvHD group was seen, indicating greater tissue injury than those without cGvHD (Figure S5).

DISCUSSION

The diagnosis and treatment of cGvHD has improved in the past two decades owing to the development of the cGvHD NIH Diagnostic and Staging Consensus Criteria, which was first introduced in 2005 and revised in 2014, as well as better therapies.[6,10–12] However, outcomes of patients with cGvHD remain limited due to the inability to diagnose subclinical cGvHD with sensitive and specific non-invasive biomarkers prior to advanced clinical organ damage. In this pilot study, cfDNA demonstrates distinctive tissue sources and temporal trends between

Table 2. Comparison of hematologic tissue and non-hematologic tissue-derived cfDNA between cGvHD and non-cGvHD

	cGvHD (median, IQR)	non-cGvHD (median, IQR)	p (Mann-Whitney U)	HSCT (median, IQR)	HC[a] (median, IQR)	p (Mann-Whitney U)
All						
Heme, %	87.4, 80.8–91.5	87.3, 76.9–91.9	0.639	87.4, 80.2–91.6	88.4, 86.6–90.0	0.300
Non-Heme, %	12.6, 8.5–18.8	12.6, 8.9–23.0	0.550	12.6, 8.6–20.0	11.6, 10.0–13.4	0.291
Heme, ng/mL	48.9, 29.6–91.6	26.2, 15.2–62.6	**<0.001**	45.8, 22.8–81.0	11.1, 7.9–17.5	**<0.001**
Non-Heme, ng/mL	8.5, 4.5–16.6	5.4, 2.7–12.0	0.022	7.7, 4.0–15.9	1.4, 0.9–2.4	**<0.001**
day ≤ 100						
Heme, %	84.9, 77.0–91.6	77.5, 41.0–89.2	0.149	81.2, 68.4–90.4	88.4, 86.6–90.0	**0.010**
Non-Heme, %	15.1, 8.4–23.0	22.3, 11.1–40.3	0.129	20.4, 9.6–28.9	11.6, 10.0–13.4	**0.007**
Heme, ng/mL	46.4, 26.7–212.3	21.4, 13.7–46.0	0.056	29.9, 14.0–69.6	11.1, 7.9–17.5	**<0.001**
Non-Heme, ng/mL	12.0, 6.4–21.2	8.3, 4.6–16.6	0.478	10.8 5.5–16.8	1.4, 0.9–2.4	**<0.001**
day 101–730						
Heme, %	88.0, 80.1–94.0	89.1, 86.2–91.8	0.516	89.0, 84.3–93.3	88.4, 86.6–90.0	0.968
Non-Heme, %	11.9, 6.0–19.0	11.0, 8.2–13.1	0.645	11.2, 6.7–15.6	11.6, 10.0–13.4	0.905
Heme, ng/mL	60.5, 34.9–99.8	46.7, 16.3–77.4	0.151	56.9, 28.7–86.7	11.1, 7.9–17.5	**<0.001**
Non-Heme, ng/mL	7.7, 3.9–16.3	4.5, 2.5–8.8	**0.037**	6.3, 3.6–13.6	1.4, 0.9–2.4	**<0.001**
day>730						
Heme, %	87.6, 81.3–90.7	92.1, 89.3–93.2	0.013	88.9, 82.6–91.6	88.4, 86.6–90.0	0.689
Non-Heme, %	12.3, 9.4–18.0	7.9, 6.8–11.1	0.022	11.2, 8.6–17.2	11.6, 10.0–13.4	0.637
Heme, ng/mL	46.1, 29.7–84.7	22.5, 20.9–53.8	0.074	45.0, 29.4–82.2	11.1, 7.9–17.5	**<0.001**
Non-Heme, ng/mL	8.6, 4.1–16.5	2.7, 1.6–2.7	**0.001**	6.7, 3.6–16.0	1.4, 0.9–2.4	**<0.001**

heme, hematologic tissue-derived; non-heme, non-hematologic tissue-derived; cGvHD, chronic graft-versus-host disease; HC, healthy control.
[a]Temporal change was not applicable to HC, all comparison was used with the same data.

cGvHD and non-GvHD. The differences in cfDNA trends start early before the diagnosis of cGvHD. Over time, the cfDNA trends also correlate with clinical manifestations with levels increasing at times of other clinical complications such as infection. Taken together, this pilot study highlights a potential utility of cfDNA for diagnosis of cGvHD as well as risk stratification to identify end-organ involvement.

Early detection and risk stratification of patients for cGvHD is important for appropriate management. cGvHD biomarkers may improve the risk stratification, prediction of response to treatment, and prognostication of cGvHD. To develop a clinically applicable biomarker, several important steps need to be undertaken by candidates, including discovery, verification, and qualification.[18] Several biomarkers such as B cell activating factor (BAFF), chemokine (C-X-C motif) ligand 9 (CXCL9) and CXCL10, are in discovery and/or verification phase, but to date, none have been qualified for clinical use.[19] Higher levels of BAFF and CXCL9 have been observed in patients with active cGvHD, the former is associated with non-relapse mortality while the latter is not.[20–22] Elevated serum CXCL10 level at 3-month post-HSCT is predictive of cGvHD.[23] Certain biomarkers may have organ-specific properties, e.g., high plasma matrix metalloproteinase-9 (MMP9) concentration is associated with lung cGvHD, high tear MMP9 concentration with ocular cGvHD, and high plasma regenerating islet-derived protein 3-α with GI cGvHD.[24–26]

Previous studies of circulating cfDNA in HSCT focused on detection of residual or recurrent disease using cfDNA chimerism or identification of disease-specific mutations in cfDNA.[2,27,28] cfDNA has also shown utility in detecting infections by measuring pathogen-specific DNA in the cfDNA compartment.[29–31] The dynamics of cfDNA in patients with cGvHD, as compared to those without, is not clearly defined, which limits the ability to utilize cfDNA approaches in the early diagnosis of cGvHD. Here, we presented a pilot study aiming to define the temporal dynamics of cfDNA post-HSCT and identify important trends that could be further explored to optimize cGvHD diagnosis and management. Our study highlighted the persistently elevated cfDNA levels post-HSCT and distinctive cfDNA tissue-of-origin methylation patterns in HSCT recipients compared with HC. Even in HSCT recipients without any complications and who were clinically well months to years post-HSCT, their cfDNA level and composition did not return to a baseline level close to that of HC. Within HSCT recipients, patients with cGvHD differ from patients without cGvHD by the higher cfDNA levels and different cfDNA tissue-of-origin in the former. Elevated cfDNA levels from multiple organs outside the usual cGvHD target organs, such as the heart and pancreas, potentially suggested globally increased inflammation, tissue injury, and regeneration in HSCT recipients. The current study identified that plasma cfDNA assays could become a useful tool in studying hematopoiesis, immune reconstitution, and tissue regeneration post-HSCT, and could be a promising biomarker for cGvHD. A prospectively designed study, ideally in conjunction with clinical trials, is required to verify the value of cfDNA assays as a cGvHD biomarker and the biomarker category according to the FDA-NIH definition.[19]

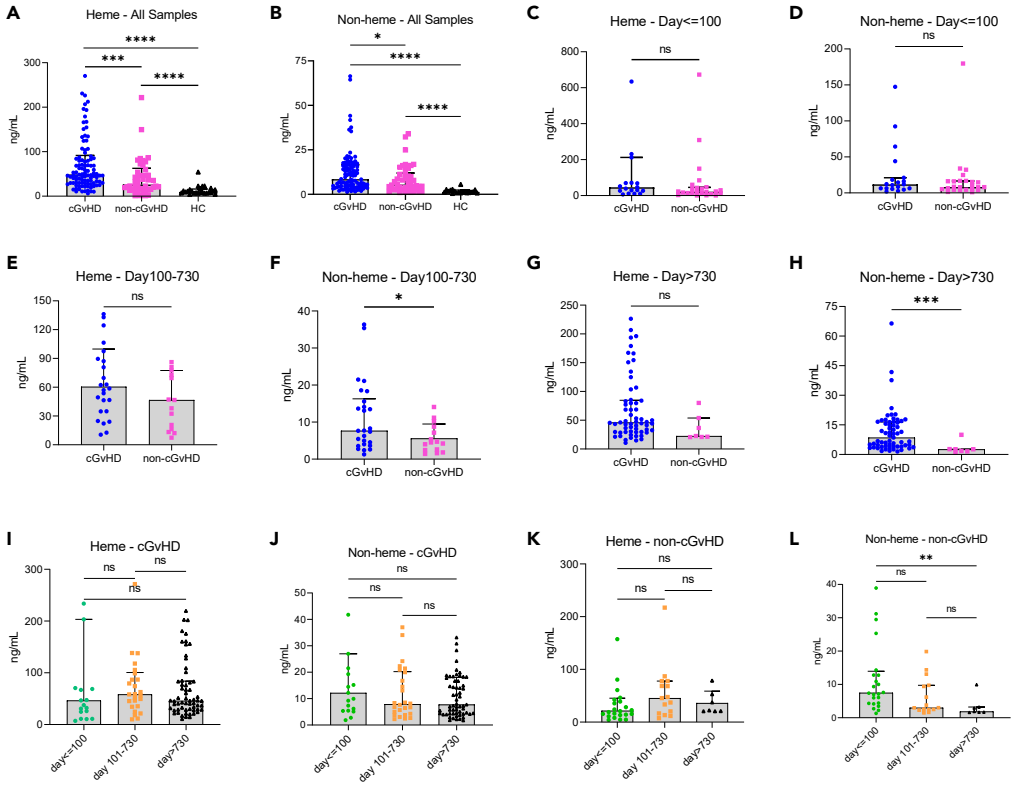


Figure 4. cfDNA (ng/mL) derived from hematologic tissue (heme) and non-hematologic tissue (non-heme) distinguished patients with cGvHD from patients without cGvHD

(A and B) In all samples, heme-derived cfDNA level (A) and non-heme-derived cfDNA level (B) were higher in cGvHD (n = 106) than non-cGvHD (n = 46).

(C and D) Within day100 post-HSCT, level of heme- (C) and non-heme-derived cfDNA (D) did not differ significantly between cGvHD (n = 19) and non-cGvHD (n = 24).

(E and F) Between day 101 and 730, there were no significant difference in heme-derived cfDNA (E), while non-heme-derived cfDNA (F) was higher in cGvHD (n = 26) than non-cGvHD (n = 15).

(G and H) After day 730, there were no significant difference in heme-derived cfDNA (G), while non-heme-derived cfDNA (H) was higher in cGvHD (n = 61) than non-cGvHD (n = 7).

(I and J) In patients with cGvHD, the temporal variance of heme-derived (I) and non-heme-derived cfDNA (J) levels had no statistical significance.

(K and L) In patients without cGvHD, the temporal variance of heme-derived cfDNA (K) had no statistical significance, while non-heme-derived cfDNA significantly decreased over time (L). Mann-Whitney U test. Column: median. Error bar: interquartile range.

Cheng et al., demonstrated the diagnostic value of solid-organ-derived cfDNA in the diagnosis of acute GvHD, however, there was limited data in cGvHD due to the limited duration of follow-up.[16] Duque-Afonso et al., showed that the percentage of recipient-derived cfDNA was higher in patients with cGvHD than without but lower than for patients with relapsed disease.[32] Our results are consistent with previous studies, cfDNA remained higher in HSCT survivors than in HC. With longer follow-up time, cfDNA levels are higher in patients with cGvHD than non-cGvHD. Studies have identified cfDNA as not only a marker of injury but also a damage-associated molecular patterns (DAMP)—immune-regulatory molecule that may directly trigger tissue injury through intracellular signaling pathways.[14] The persistent elevation of cfDNA in HSCT survivors, even in those without clinical evidence of GvHD or other complications, may indicate a long-lasting high level of cell turnover and/or inflammation post-HSCT. Given that extracellular cfDNA is a DAMP, a pathogenic role of cfDNA in cGvHD deserves further investigation.

The tissue sources of cfDNA showed differences between HC and HSCT subjects. The elevated level of neutrophil and monocyte-derived cfDNA in HSCT recipients may suggest highly active immune dysregulation mediated by the innate immune system. Not surprisingly, hepatocyte-derived cfDNA was elevated in HSCT recipients, especially those with cGvHD, correlating with the frequency of liver injury post-HSCT due to cGvHD, drug toxicity, and other etiologies.[33] Within the HSCT subjects, those who develop cGvHD showed different temporal dynamics of hematopoietic and non-hematopoietic sources; the cGvHD group demonstrated persistently higher cfDNA from non-hematopoietic tissue sources, consistent with the underlying chronic inflammation and tissue injury associated with cGvHD (Figure 1). Indeed, cfDNA from certain tissues such as monocyte, neutrophil, and eosinophils could help differentiate cGvHD and non-cGvHD. The higher level of erythroblast-derived cfDNA could reflect active destruction and maturation blockage of the erythroid precursors which has been observed in cGvHD.[17,34,35] The elevation of liver-derived cfDNA among those with cGvHD of the liver suggested that it could be a tool to differentiate

cGvHD-liver injury vs. other types of liver injury. However, there was no clear evidence that lung-derived cfDNA correlated with clinical pulmonary cGvHD or skin-derived cfDNA with skin cGvHD. The reason may be the following: (1) cfDNA has a very short half-life and may only be elevated during active inflammation. Many of the subjects in this study were already diagnosed with cGvHD based on advanced clinical symptoms or PFT findings. There may not be much active inflammation in the lung or skin at the time of sample collection, therefore the respective tissue-derived cfDNA was not elevated. (2) It is likely that in HSCT recipients, the majority of the cfDNA would be derived from the hematologic system because the hematologic/immune system is constantly being remodeled. In the experience from our group and others, even in patients with significant lung etiology such as acute lung allograft rejection, COVID-19 pneumonia, and pulmonary hypertension, lung-derived cfDNA only made up 0.65–1.2% of total cfDNA.[36,37] In our study lung-derived cfDNA made up about 2% of total cfDNA (Table S3), which suggested potentially higher level of lung tissue inflammation in HSCT recipients than subjects with primary lung diseases. (3) The algorithm we used may not accurately represent the small/distal airways, airway smooth muscles, and other target tissues of pulmonary or skin cGvHD. Most current methods for cell deconvolution of WGBS data use a non-negative least square algorithm to bin bisulfite-sequenced reads by the similarity of their methylation profile to previously determined databases of methylation calls for different tissues. The sensitivity of this approach needs to be improved. For example, the reference methylome in Meth_Atlas captures less than 10% of the 28 million CpG sites of the human genome and might not be sufficient to detect low-fraction cell types accurately. The cfDNAme uses cell-type-specific methylation windows deduced from WGBS data but might have higher noise than array data. Nevertheless, both methods can accurately determine the tissue make up of samples at higher coverages, with decreasing accuracy as coverage decreases.

In summary, this pilot study reports cfDNA temporal dynamics after HCST, with HSCT patients showing consistent higher cfDNA than healthy controls. Further, HSCT patients who developed cGvHD show higher cfDNA than those without cGvHD. The pattern of cfDNA also differs over time post-HSCT; cGvHD patients show persistently high cfDNA levels from non-heme tissue types, and non-cGvHD patients show reducing levels of non-heme-derived cfDNA (Figure 4). The specific tissue sources of cfDNA also show difference between cGvHD and non-cGvHD. Our results, together with others, provide compelling evidence for the continued research to depict the landscape of cfDNA post-HSCT in patients with or without post-HSCT complications and explore the utility of cfDNA as a potential biomarker for cGvHD.

Limitations of the study

There were several limitations to the study. First, it was limited by a small sample size, especially the number of patients with longitudinal sample collection. The post-transplant time point of sample collection schedule was not standardized between groups. Second, the patients had a diverse background of primary disease, HSCT preparative therapy and cGvHD treatment, potentially introducing confounders in cfDNA analysis. Third, there was a lack of genotype data to reliably assign cfDNA as donor-derived versus recipient-derived. Genomic chimerism data are not available for all patients. It remains to be elucidated whether cfDNA chimerism corresponds to genomic DNA chimerism. An on-going computation model in our group, and larger, prospective sample collection that includes pre-transplant genomic DNA data should address this limitation. Fourth, several tissue types that might be affected by HSCT and involve in the pathogenesis of cGvHD such as megakaryocytes, were not adequately separated or included in the tissue-specific DNA methylation library. Ongoing assay developments are poised to include these additional tissue types and assess the role of other cfDNA approaches such as fragment analysis or chromatin immunoprecipitation in HSCT. Lastly, correlative data such as cytokine levels and other inflammatory markers were not available with each sample, reducing the power of the study to confirm the correlation between high cfDNA level and the high level of inflammation and tissue damage. A current ongoing multicenter prospective cohort study should address these limitations.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Plasma sample processing and cfDNA isolation
  - cfDNA quantitation
  - Library preparation and sequencing
  - cfDNA sequence analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
- ADDITIONAL RESOURCES

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.108160.

ACKNOWLEDGMENTS

This research is supported by the NHLBI and/or NIH intramural research program. Pavletic: HRSA Health Systems Bureau's Division of Transplantation-National Institutes of Health Inter-agency Agreement Late Effects Initiative. Center for Cancer Research: Research Funding; National Cancer Institute: Research Funding; Celgene: Research Funding; Actelion: Research Funding; Eli Lilly: Research Funding; Pharmacyclics: Research Funding; Kadmon: Research Funding.

AUTHOR CONTRIBUTIONS

Conceptualization, Y.P., S.Z.P., and S.A.-E.; Methodology, Y.P., T.E.A., M.K.J., S.Z.P., and S.A.-E.; Investigation: Y.P., T.E.A., M.K.J., H.K., W.P., T.H., N.R., Y.-P.F., D.A.P., N.G.H., S.Z.P., and S.A.-E; Writing – Original Draft: Y.P.; Writing – Review & Editing: Y.P., T.E.A., T.H., S.Z.P.,N.G.H., and S.A.-E.; Resources: S.Z.P., and S.A.-E.

DECLARATION OF INTERESTS

The authors declare no competing interest.

Received: January 31, 2023
Revised: August 21, 2023
Accepted: October 5, 2023
Published: October 9, 2023

REFERENCES

1. D'Souza, A., Lee, S., Zhu, X., and Pasquini, M. (2017). Current Use and Trends in Hematopoietic Cell Transplantation in the United States. Biol. Blood Marrow Transplant. *23*, 1417–1421. https://doi.org/10.1016/j.bbmt.2017.05.035.
2. Aljurf, M., Weisdorf, D., Alfraih, F., Szer, J., Müller, C., Confer, D., Hashmi, S., Kröger, N., Shaw, B.E., Greinix, H., et al. (2019). Worldwide Network for Blood & Marrow Transplantation (WBMT) special article, challenges facing emerging alternate donor registries. Bone Marrow Transplant. *54*, 1179–1188. https://doi.org/10.1038/s41409-019-0476-6.
3. Majhail, N.S., Tao, L., Bredeson, C., Davies, S., Dehn, J., Gajewski, J.L., Hahn, T., Jakubowski, A., Joffe, S., Lazarus, H.M., et al. (2013). Prevalence of hematopoietic cell transplant survivors in the United States. Biol. Blood Marrow Transplant. *19*, 1498–1501. https://doi.org/10.1016/j.bbmt.2013.07.020.
4. Zeiser, R., and Blazar, B.R. (2017). Pathophysiology of Chronic Graft-versus-Host Disease and Therapeutic Targets. N. Engl. J. Med. *377*, 2565–2579. https://doi.org/10.1056/NEJMra1703472.
5. Le, R.Q., Bevans, M., Savani, B.N., Mitchell, S.A., Stringaris, K., Koklanaris, E., and Barrett, A.J. (2010). Favorable outcomes in patients surviving 5 or more years after allogeneic hematopoietic stem cell transplantation for hematologic malignancies. Biol. Blood Marrow Transplant. *16*, 1162–1170. https://doi.org/10.1016/j.bbmt.2010.03.005.
6. Jagasia, M.H., Greinix, H.T., Arora, M., Williams, K.M., Wolff, D., Cowen, E.W., Palmer, J., Weisdorf, D., Treister, N.S., Cheng, G.-S., et al. (2015). National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: I. The 2014 Diagnosis and Staging Working Group Report. Biol. Blood Marrow Transplant. *21*, 389–401.e1. https://doi.org/10.1016/j.bbmt.2014.12.001.
7. Boyiadzis, M., Arora, M., Klein, J.P., Hassebroek, A., Hemmer, M., Urbano-Ispizua, A., Antin, J.H., Bolwell, B.J., Cahn, J.Y.Y., Cairo, M.S., et al. (2015). Impact of Chronic Graft-versus-Host Disease on Late Relapse and Survival on 7,489 Patients after Myeloablative Allogeneic Hematopoietic Cell Transplantation for Leukemia. Clin. Cancer Res. *21*, 2020–2028. https://doi.org/10.1158/1078-0432.CCR-14-0586.
8. Yalniz, F.F., Murad, M.H., Lee, S.J., Pavletic, S.Z., Khera, N., Shah, N.D., and Hashmi, S.K. (2018). Steroid Refractory Chronic Graft-Versus-Host Disease: Cost-Effectiveness Analysis. Biol. Blood Marrow Transplant. *24*, 1920–1927. https://doi.org/10.1016/j.bbmt.2018.03.008.
9. Lee, S.J., Wolff, D., Kitko, C., Koreth, J., Inamoto, Y., Jagasia, M., Pidala, J., Olivieri, A., Martin, P.J., Przepiorka, D., et al. (2015). Measuring Therapeutic Response in Chronic Graft-versus-Host Disease. National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: IV. The 2014 Response Criteria Working Group Report. Biol. Blood Marrow Transplant. *21*, 984–999. https://doi.org/10.1016/j.bbmt.2015.02.025.
10. Miklos, D., Cutler, C.S., Arora, M., Waller, E.K., Jagasia, M., Pusic, I., Flowers, M.E., Logan, A.C., Nakamura, R., Blazar, B.R., et al. (2017). Ibrutinib for chronic graft-versus-host disease after failure of prior therapy. Blood *130*, 2243–2250. https://doi.org/10.1182/blood-2017-07-793786.
11. Zeiser, R., Polverelli, N., Ram, R., Hashmi, S.K., Chakraverty, R., Middeke, J.M., Musso, M., Giebel, S., Uzay, A., Langmuir, P., et al. (2021). Ruxolitinib for Glucocorticoid-Refractory Chronic Graft-versus-Host Disease. N. Engl. J. Med. *385*, 228–238. https://doi.org/10.1056/nejmoa2033122.
12. Cutler, C., Lee, S.J., Arai, S., Rotta, M., Zoghi, B., Lazaryan, A., Ramakrishnan, A., DeFilipp, Z., Salhotra, A., Chai-Ho, W., et al. (2021). Belumosudil for Chronic Graft-versus-Host Disease (cGVHD) After 2 or More Prior Lines of Therapy: The ROCKstar Study. Blood *138*, 2278–2289. https://doi.org/10.1182/blood.2021012021.
13. Kitko, C.L., Pidala, J., Schoemans, H.M., Lawitschka, A., Flowers, M.E., Cowen, E.W., Tkaczyk, E., Farhadfar, N., Jain, S., Steven, P., et al. (2021). National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: IIa. The 2020 Clinical Implementation and Early Diagnosis Working Group Report. Transplant. Cell. Ther. *27*, 545–557. https://doi.org/10.1016/j.jtct.2021.03.033.
14. Tsuji, N., and Agbor-Enoh, S. (2021). Cell-free DNA beyond a biomarker for rejection: Biological trigger of tissue injury and potential therapeutics. J. Heart Lung Transplant. *40*, 405–413. https://doi.org/10.1016/j.healun.2021.03.007.
15. Cheng, A.P., Cheng, M.P., Loy, C.J., Lenz, J.S., Chen, K., Smalling, S., Burnham, P., Timblin, K.M., Orejas, J.L., Silverman, E., et al. (2021). Cell-free DNA Profiling Informs Major Complications of Hematopoietic Cell Transplantation. Preprint at bioRxiv. https://doi.org/10.1101/2020.04.25.061580.
16. Cheng, A.P., Cheng, M.P., Loy, C.J., Lenz, J.S., Chen, K., Smalling, S., Burnham, P., Timblin, K.M., Orejas, J.L., Silverman, E., et al. (2022). Cell-free DNA profiling informs all major complications of hematopoietic cell transplantation. Proc. Natl. Acad. Sci. USA *119*, e2113476118. https://doi.org/10.1073/pnas.2113476118.
17. Moss, J., Magenheim, J., Neiman, D., Zemmour, H., Loyfer, N., Korach, A., Samet, Y., Maoz, M., Druid, H., Arner, P., et al. (2018). Comprehensive human cell-type methylation atlas reveals origins of circulating cell-free DNA in health and disease. Nat. Commun. *9*, 5068. https://doi.org/10.1038/s41467-018-07466-6.
18. Paczesny, S., Hakim, F.T., Pidala, J., Cooke, K.R., Lathrop, J., Griffith, L.M., Hansen, J., Jagasia, M., Miklos, D., Pavletic, S., et al. (2015). National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: III. The 2014 Biomarker Working Group Report. Biol. Blood Marrow

Transplant. *21*, 780–792. https://doi.org/10.1016/j.bbmt.2015.01.003.

19. Bidgoli, A., DePriest, B.P., Saatloo, M.V., Jiang, H., Fu, D., and Paczesny, S. (2022). Current Definitions and Clinical Implications of Biomarkers in Graft-versus-Host Disease. Transplant. Cell. Ther. *28*, 657–666.
20. Sarantopoulos, S., Stevenson, K.E., Kim, H.T., Cutler, C.S., Bhuiya, N.S., Schowalter, M., Ho, V.T., Alyea, E.P., Koreth, J., Blazar, B.R., et al. (2009). Altered B-cell homeostasis and excess BAFF in human chronic graft-versus-host disease. Blood *113*, 3865–3874. https://doi.org/10.1182/blood-2008-09-177840.
21. Sarantopoulos, S., Stevenson, K.E., Kim, H.T., Bhuiya, N.S., Cutler, C.S., Soiffer, R.J., Antin, J.H., and Ritz, J. (2007). High Levels of B-Cell Activating Factor in Patients with Active Chronic Graft-Versus-Host Disease. Clin. Cancer Res. *13*, 6107–6114. https://doi.org/10.1158/1078-0432.ccr-07-1290.
22. Saliba, R.M., Sarantopoulos, S., Kitko, C.L., Pawarode, A., Goldstein, S.C., Magenau, J., Alousi, A.M., Churay, T., Justman, H., Paczesny, S., et al. (2017). B-cell activating factor (BAFF) plasma level at the time of chronic GvHD diagnosis is a potential predictor of non-relapse mortality. Bone Marrow Transplant. *52*, 1010–1015. https://doi.org/10.1038/bmt.2017.73.
23. Chirumbolo, G., Dicataldo, M., Barone, M., Storci, G., De Matteis, S., Laprovitera, N., Sinigaglia, B., Barbato, F., Maffini, E., Cavo, M., et al. (2023). A Multiparameter Prognostic Risk Score of Chronic Graft-versus-Host Disease Based on CXCL10 and Plasmacytoid Dendritic Cell Levels in the Peripheral Blood at 3 Months after Allogeneic Hematopoietic Stem Cell Transplantation. Transplant. Cell. Ther. *29*, 302.e1–302.e8.
24. Inamoto, Y., Martin, P.J., Onstad, L.E., Cheng, G.S., Williams, K.M., Pusic, I., Ho, V.T., Arora, M., Pidala, J., Flowers, M.E.D., et al. (2021). Relevance of Plasma Matrix Metalloproteinase-9 for Bronchiolitis Obliterans Syndrome after Allogeneic Hematopoietic Cell Transplantation. Transplant. Cell. Ther. *27*, 759.e1–759.e8. https://doi.org/10.1016/j.jtct.2021.06.006.
25. Berchicci, L., Aragona, E., Arrigo, A., Marchese, A., Miserocchi, E., Bandello, F., and Modorati, G. (2021). Conjunctival Matrix Metalloproteinase-9 Clinical Assessment in Early Ocular Graft versus Host Disease. J. Ophthalmol. *2021*, 9958713–9958717. https://doi.org/10.1155/2021/9958713.
26. DePriest, B.P., Li, H., Bidgoli, A., Onstad, L., Couriel, D., Lee, S.J., and Paczesny, S. (2022). Regenerating islet-derived protein 3-α is a prognostic biomarker for gastrointestinal chronic graft-versus-host disease. Blood Adv. *6*, 2981–2986. https://doi.org/10.1182/bloodadvances.2021005420.
27. Ramírez, M., Díaz, M.A., García-Sánchez, F., Velasco, M., Casado, F., Villa, M., Vicario, J.L., and Madero, L. (1996). Chimerism after allogeneic hematopoietic cell transplantation in childhood acute lymphoblastic leukemia. Bone Marrow Transplant. *18*, 1161–1165.
28. Albitar, M., Zhang, H., Pecora, A.L., Ip, A., Goy, A.H., Estella, J.J., De Dios, I., Ma, W., Kaur, S., Suh, H.C., et al. (2021). Reliability of Liquid Biopsy and Next Generation Sequencing in Monitoring Residual Disease Post-Hematopoietic Stem Cell Transplant. Blood *138*, 1828.
29. Fung, M., Zompi, S., Seng, H., Hollemon, D., Parham, A., Hong, D.K., Bercovici, S., Dolan, E., Lien, K., Teraoka, J., et al. (2018). Plasma Cell–Free DNA Next-Generation Sequencing to Diagnose and Monitor Infections in Allogeneic Hematopoietic Stem Cell Transplant Patients. Open Forum Infect. Dis. *5*, ofy301. https://doi.org/10.1093/ofid/ofy301.
30. Armstrong, A.E., Rossoff, J., Hollemon, D., Hong, D.K., Muller, W.J., and Chaudhury, S. (2019). Cell-free DNA next-generation sequencing successfully detects infectious pathogens in pediatric oncology and hematopoietic stem cell transplant patients at risk for invasive fungal disease. Pediatr. Blood Cancer *66*, e27734. https://doi.org/10.1002/pbc.27734.
31. Hill, J.A., Dalai, S.C., Hong, D.K., Ahmed, A.A., Ho, C., Hollemon, D., Blair, L., Maalouf, J., Keane-Candib, J., Stevens-Ayers, T., et al. (2021). Liquid Biopsy for Invasive Mold Infections in Hematopoietic Cell Transplant Recipients With Pneumonia Through Next-Generation Sequencing of Microbial Cell-Free DNA in Plasma. Clin. Infect. Dis. *73*, e3876–e3883. https://doi.org/10.1093/cid/ciaa1639.
32. Duque-Afonso, J., Waterhouse, M., Pfeifer, D., Follo, M., Duyster, J., Bertz, H., and Finke, J. (2018). Cell-free DNA characteristics and chimerism analysis in patients after allogeneic cell transplantation. Clin. Biochem. *52*, 137–141. https://doi.org/10.1016/j.clinbiochem.2017.11.015.
33. Abdelbary, H., Magdy, R., Moussa, M., and Abdelmoaty, I. (2020). Liver disease during and after hematopoietic stem cell transplantation in adults: a single-center Egyptian experience. J. Egypt. Natl. Canc. Inst. *32*, 11. https://doi.org/10.1186/s43046-020-0020-1.
34. Jung, N., Dai, B., Gentles, A.J., Majeti, R., and Feinberg, A.P. (2015). An LSC epigenetic signature is largely mutation independent and implicates the HOXA cluster in AML pathogenesis. Nat. Commun. *6*, 8489. https://doi.org/10.1038/ncomms9489.
35. Lin, Y., Hu, X., Cheng, H., Pang, Y., Wang, L., Zou, L., Xu, S., Zhuang, X., Jiang, C., Yuan, W., et al. (2014). Graft-versus-Host Disease Causes Broad Suppression of Hematopoietic Primitive Cells and Blocks Megakaryocyte Differentiation in a Murine Model. Biol. Blood Marrow Transplant. *20*, 1290–1300. https://doi.org/10.1016/j.bbmt.2014.05.009.
36. Andargie, T.E., Tsuji, N., Seifuddin, F., Jang, M.K., Yuen, P.S., Kong, H., Tunc, I., Singh, K., Charya, A., Wilkins, K., et al. (2021). Cell-free DNA maps COVID-19 tissue injury and risk of death and can cause tissue injury. JCI Insight *6*, e147610. https://doi.org/10.1172/jci.insight.147610.
37. Brusca, S.B., Elinoff, J.M., Zou, Y., Jang, M.K., Kong, H., Demirkale, C.Y., Sun, J., Seifuddin, F., Pirooznia, M., Valantine, H.A., et al. (2022). Plasma Cell-Free DNA Predicts Survival and Maps Specific Sources of Injury in Pulmonary Arterial Hypertension. Circulation *146*, 1033–1045. https://doi.org/10.1161/circulationaha.121.056719.
38. Köster, J., and Rahmann, S. (2012). Snakemake—a scalable bioinformatics workflow engine. Bioinformatics *28*, 2520–2522. https://doi.org/10.1093/bioinformatics/bts480.
39. Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics *27*, 1571–1572. https://doi.org/10.1093/bioinformatics/btr167.
40. Zhao, H., Sun, Z., Wang, J., Huang, H., Kocher, J.P., and Wang, L. (2014). CrossMap: a versatile tool for coordinate conversion between genome assemblies. Bioinformatics *30*, 1006–1007. https://doi.org/10.1093/bioinformatics/btt730.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Plasma	Human/patient	The samples are not from commercial sources or biological repositories
Chemicals, peptides, and recombinant proteins		
DNA Methylation-Gold Kits	Zymo Research	D5008
QIAsymphony DSP Circulating DNA Kit	Qiagen	937556
xGen Methyl-Seq DNA Library Prep Kit	IDT Swift (Accel-NGS Methyl-Seq DNA Library Kit)	10009824
NovaSeq 6000 S4 Reagent Kit v1.5 (200 cycles)	Illumina	20028313
Quant-iT PicoGreen dsDNA assay kit	Thermo Fisher Scientific	P11496
Lambda DNA	Promega	D1501
iTaq Universal SYBR Green Supermix	Bio-Rad	1725122
Primers for Alu115/Alu247	Customized (Available from ITD)	N/A (10009856)
Primers for Lambda DNA	Customized and synthesized by IDT Boulter N, Suarez FG, Schibeci S, Sunderland T, Tolhurst O, Hunter T, Hodge G, Handelsman D, Simanainen U, Hendriks E, Duggan K. A simple, accurate and universal method for quantification of PCR. BMC Biotechnol. 2016 Mar 9; 16:27. https://doi.org/10.1186/s12896-016-0256-y. PMID: 26956612; PMCID: PMC4784296.	N/A
Cell-free DNA ScreenTape Analysis	Agilent	5067-5630, 5067–5631, 5067-5633
High Sensitivity DNA ScreenTape Analysis	Agilent	5067-5603, 5067–5584, 5067-5585
Software and algorithms		
methyl-seek Snakemake pipeline	https://github.com/OpenOmics/methyl-seek	N/A
TrimGalore	https://github.com/FelixKrueger/TrimGalore: GitHub	N/A
Bismark version 0.23.0	Krueger F, Andrews SR. **Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications**. *Bioinformatics* 2011; 27(11):1571-1572	N/A
cfDNAme	https://github.com/alexpcheng/cfDNAme	N/A
meth_atlas	https://github.com/nloyfer/meth_atlas	N/A
CrossMap	Zhao H, Sun Z, Wang J, Huang H, Kocher JP, Wang L. **CrossMap: a versatile tool for coordinate conversion between genome assemblies**. *Bioinformatics* 2014; 30(7):1006-1007	N/A
IBM SPSS Statistics Subscription	IBM Corp, Armonk, NY	N/A
GraphPad Prism 9	Home - GraphPad	N/A
Excel	Microsoft	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Sean Agbor-Enoh (sean.agbor-enoh@nih.gov).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- All data reported in this paper will be shared by the lead contact upon request.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

There were 63 cGvHD subjects selected from 447 enrolled patients from the "Natural History Study of Clinical and Biological Factors Determining Outcomes in Chronic Graft-Versus-Host Disease" (NCT00922235), with 108 plasma samples collected. The inclusion criteria for cGvHD subject selection were cGvHD global score = severe, involved organs ≥ 2, and FEV1 $\leq 70\%$ predicted (Figure S1). There were 13 non-cGvHD subjects with 73 samples from other HSCT protocols (NCT00074490, NCT00520130, NCT03000244, NCT02356159). Two samples from the GvHD group, and 27 samples from the non-GvHD group were of poor quality and were excluded, leaving 106 and 46 samples from 61 to 12 subjects in the cGvHD and non-cGvHD group for final analysis. Nineteen HC subjects were recruited as part of sample collection at the NIH Blood Bank. The sample collection and analysis were adherent to the requirements according to the Institutional Review Board. The age, sex and race of the study participants are detailed in Table 1. The gender, ancestry, ethnicity, and socioeconomic status of the participants are unavailable.

METHOD DETAILS

Plasma sample processing and cfDNA isolation

Plasma samples were isolated from whole blood collected in EDTA (BD) tubes by centrifugation at 1600 *g* for 10 min at 4°C, aliquoted and stored immediately at −80°C. The aliquoted plasma was thawed at room temperature and centrifuged at 16000 g for 5 min at 4°C to remove residual debris. cfDNA was extracted from 1 mL of plasma by QIAsymphony circulating DNA kit (QIAGEN). The plasma samples were spiked with 0.142 ng/mL unmethylated 160 bp lambda phage DNA (Promega), to measure the efficiency of cfDNA extraction and bisulfite conversion for genome-wide methylation sequencing, and to correct for cfDNA concentrations for extraction loss. The isolated cfDNA were eluted into 60 μL elution buffer and stored at −20°C until further use.

cfDNA quantitation

cfDNA was quantified by quantitative PCR using a primer targeting short (Alu115) sequences: forward, 5′-GACCTCTATGCCAACACAGT-3′ and reverse, 5′-AGTACTTGCGCTCAGGAGGA-3′.[36] After extraction, cfDNA quality was analyzed by assessing Integrity Score (the proportion of long (Alu247) versus short (Alu115) fragments) and by using Tapestation cfDNA ScreenTape assay. Samples with an integrity score >0.8 and/or an absence of one distinct cfDNA peak in TapeStation were eliminated for bisulfite sequencing due to risks of contamination with high molecular weight nuclear DNA. The rest of the samples were selected for bisulfite sequencing, with a median Integrity score of 0.49 (IQR, 0.36–0.65).

Library preparation and sequencing

An input of 5–50 ng of isolated cfDNA was used to perform bisulfite conversion using EZ DNA methylation-Gold kit (Zymo Research) according to the manual. Libraries were prepared using the Accel-NGS Methyl-Seq DNA Library Kit with Unique Dual Indexing (Swift Biosciences) for whole-genome bisulfite sequencing according to the manual. The quality of the constructed cfDNA library was visualized using a high-sensitivity D1000 ScreenTape and quantified using the Quant-iT PicoGreen dsDNA Assay kit (Life Technologies). The DNA libraries were then normalized in 8 equimolar concentrations and were subjected for ~200 M reads by 2X100 bp, paired-end DNA sequencing on NovaSeq 6000 (Illumina).

cfDNA sequence analysis

Sequencing yield was on an average 162 million reads/sample (min = 45 million; max = 290 million). The bisulfide sequencing data were analyzed using the methyl-seek Snakemake pipeline (https://github.com/OpenOmics/methyl-seek).[38] We first trimmed raw sequencing reads using TrimGalore (https://github.com/FelixKrueger/TrimGalore: GitHub), and then aligned to the bisulfite converted hg38 human genome reference using Bismark version 0.23.0.[39] Next, we used Bismark to remove duplicate read alignments and extract the methylation

status of each cytosine base in the aligned position. This step generated a CpG bed file with hg38 genomic coordinates, which we converted to hg19 genome coordinates to be compatible with the deconvolution method using the tool CrossMap.[40]

For cfDNAme (https://github.com/alexpcheng/cfDNAme), we first extracted the CpG windows found in the cfDNAme database from our cfDNA samples. cfDNAme then uses a non-negative least squares method to identify the best possible match for each read to one of 27 different cell/tissue types based on the methylation state of cell/tissue-specific DNA methylation markers present on each read, corresponding to the cfDNAme database. We then calculated the proportion of reads mapping to each methylation marker in the cfDNAme database, to estimate the proportion of each cell/tissue type in each CpG sample. The absolute concentration of tissue-specific cfDNA was computed by multiplying the tissue-specific cfDNA percentage with the total cfDNA concentration. cfDNA quantification and bisulfite sequencing data were available for all HC from a prior study.

For Meth_atlas (https://github.com/nloyfer/meth_atlas),[17] we first extracted the CpG sites found in the meth_atlas database from our cfDNA samples. Meth_atlas then uses a non-negative least squares method to identify the best possible match for each read to one of 25 different cell/tissue types based on the methylation state of cell/tissue-specific DNA methylation markers present on each read, corresponding to the meth_atlas database. We then calculated the proportion of reads mapping to each methylation marker in the meth_atlas database, to estimate the proportion of each cell/tissue type in each CpG sample. The absolute concentration of tissue-specific cfDNA was computed by multiplying the tissue-specific cfDNA percentage with the total cfDNA concentration. cfDNA quantification and bisulfite sequencing data were available for all HC from a prior study.[36]

To ensure the reproducibility among different batches of samples, we analyzed replicate plasma samples for 12 healthy controls, each replicate performed as a different batch by a different analyst. The replicate plasma samples were utilized for cfDNA isolation and whole genome bisulfite sequencing. Sequence reads were analyzed by cfDNAme. As shown below, fraction of each tissue type was similar between the two batches. Variance between the two batches were also not significant. For example, the variance for neutrophil cfDNA was 0.000517 between the batches. The variances of cfDNA derived from liver, lung, heart, monocyte, or macrophage were all less than 0.0001 (Figure S6).

QUANTIFICATION AND STATISTICAL ANALYSIS

We described demographics and transplant characteristics as counts and percentages for categorical variables, and as mean and standard deviation (SD) for continuous variables. We used Kruskal-Wallis and Chi-square tests to compare continuous variables and nominal variables between groups, respectively. We compared the cfDNA characteristics between HSCT recipients including cGvHD, non-cGvHD and HC using a Kruskal-Willis test. Comparisons of the cfDNA characteristics between cGvHD and non-cGvHD at different post-HSCT time-period were made by the Mann-Whitney U test. The correlations between cfDNA and other clinical markers such as liver transaminases were analyzed by the Pearson correlation coefficient. All analysis were completed by IBM SPSS Statistics Subscription plan (IBM Corp. Armonk, NY). Statistical significance was defined as a 2-tailed p value ≤ 0.05. Figures were made by GraphPad Prism Version 9.5.1 (GraphPad, San Diego, CA) and Microsoft Excel.

ADDITIONAL RESOURCES

The plasma samples of the patients with cGvHD and non-cGvHD were collected through one of the clinical trials listed below. The clinical trials are conducted solely at the National Institutes of Health. No other external sites had been generated to support discussion or use of the information/data/material created by the manuscript.

NCT00092235: https://www.clinicaltrials.gov/study/NCT00092235?term=NCT00092235&rank=1.
NCT00074490: https://www.clinicaltrials.gov/study/NCT00074490?term=NCT00074490&rank=1.
NCT00520130: https://www.clinicaltrials.gov/study/NCT00520130?term=NCT00520130&rank=1.
NCT03000244: https://www.clinicaltrials.gov/study/NCT03000244?intr=NCT03000244&rank=1.
NCT02356159: https://www.clinicaltrials.gov/study/NCT02356159?intr=NCT02356159&rank=1.