haematologica
Journal of the Ferrata Storti Foundation

Early establishment of the naïve B-cell receptor repertoire with gradual antigen-experienced maturation after pediatric hematopoietic stem cell transplantation

by Melanie de Gier, Martijn Cordes, Ingrid Pico-Knijnenburg, Peter Blazso, Dagmar Berghuis, Jolan E. Walter, Lisa Ott de Bruin, Arjan C. Lankester and Mirjam van der Burg

Received: May 14, 2026.
Accepted: August 31, 2026.

Citation: Melanie de Gier, Martijn Cordes, Ingrid Pico-Knijnenburg, Peter Blazso, Dagmar Berghuis, Jolan E. Walter, Lisa Ott de Bruin, Arjan C. Lankester and Mirjam van der Burg. Early establishment of the naïve B-cell receptor repertoire with gradual antigen-experienced maturation after pediatric hematopoietic stem cell transplantation.
Haematologica. 2026 Sept 10. doi: 10.3324/haematol.2026.301216 [Epub ahead of print]

Publisher's Disclaimer.
E-publishing ahead of print is increasingly important for the rapid dissemination of science. Haematologica is, therefore, E-publishing PDF files of an early version of manuscripts that have completed a regular peer review and have been accepted for publication.
E-publishing of this PDF file has been approved by the authors.
After having E-published Ahead of Print, manuscripts will then undergo technical and English editing, typesetting, proof correction and be presented for the authors' final approval, the final version of the manuscript will then appear in a regular issue of the journal.
All legal disclaimers that apply to the journal also pertain to this production process.

Early establishment of the naïve B-cell receptor repertoire with gradual antigen-experienced maturation after pediatric hematopoietic stem cell transplantation

Running title: BCR repertoire analysis post-HSCT

Melanie de Gier,[1] Martijn Cordes,[1] Ingrid Pico-Knijnenburg,[1] Peter Blazso,[2,3] Dagmar Berghuis,[4] Jolan Walter,[5] Lisa Ott de Bruin,[4] Arjan C Lankester,[4] Mirjam van der Burg[1]

Affiliations

[1]Department of Pediatrics, Laboratory for Pediatric Immunology, Willem-Alexander Children's Hospital, Leiden University Medical Center, the Netherlands

[2]Bioinformatics Service Unit, Complex Molecular and Cell Biology Service Centre, HUN-REN Biological Research Centre, Szeged, Hungary

[3]Department of Pediatrics, University of Szeged, Szeged, Hungary

[4]Department of Pediatrics, Division of Pediatric Immunology and Stem Cell Transplantation, Willem-Alexander Children's Hospital, Leiden University Medical Center, the Netherlands

[5]Division of Pediatric Allergy/Immunology, University of South Florida at Johns Hopkins All Children's Hospital, St. Petersburg, USA

Corresponding author

Prof. Mirjam van der Burg, PhD
Dept. of Pediatrics, Laboratory for Pediatric Immunology, Willem-Alexander Children's Hospital, Leiden University Medical Center, the Netherlands
Leiden University Medical Center
Albinusdreef 2
2333 ZA Leiden
The Netherlands
Email: m.van_der_burg@lumc.nl
Work phone: +31 71 5264976

Author contributions

MdG, LOB, AL and MvdB wrote the manuscript. MdG and IP prepared sequencing libraries. MdG, MC, PB and JW performed data analyses. MdG, DB and AL included patients and provided clinical data. LOB, AL and MvdB supervised the project. All authors read and critically evaluated the manuscript before submission.

Disclosures

The authors declare no conflict of interest.

Data sharing agreement

The sequencing data presented in this manuscript have been deposited in the European Genome-phenome Archive (EGA) under accession number EGAXXXXXX. Processed IMGT archives are available upon request.

Word count: 4236
Tables and Figures: 7
Supplemental files: 1

Abstract

Allogeneic hematopoietic stem cell transplantation (HSCT) is a curative treatment for diverse inherited and acquired immuno-hematological disorders. While quantitative B-cell recovery is well characterized, qualitative reconstitution of the B-cell receptor (BCR) repertoire, particularly in pediatric patients, remains incompletely defined. We delineated longitudinal reconstitution of naïve and antigen-experienced immunoglobulin heavy chain (IGH) repertoires during the first two years after pediatric HSCT. Peripheral blood mononuclear cells from 10 pediatric patients were analyzed at 3, 6, 12, and 24 months post-HSCT and compared with age-matched healthy donors. High-throughput sequencing of IGH transcripts was performed to assess V(D)J gene usage, repertoire diversity, junctional characteristics, somatic hypermutation (SHM), and clonal architecture across naïve and antigen-experienced compartments.

The naïve IGHM repertoire demonstrated broad, unbiased IGHV, IGHD, and IGHJ gene usage as early as 3 months post-HSCT, with diversity and junctional features comparable to controls. In contrast, antigen-experienced repertoires (IGHG and IGHA) showed reduced SHM frequencies and diminished N-nucleotide insertions during the first year, normalizing by 24 months. Despite delayed SHM accumulation, targeting of activation-induced cytidine deaminase and polymerase η motifs and replacement-to-silent mutation ratios were preserved, indicating intact molecular affinity maturation. Clonal expansions were detectable in class-switched compartments but did not compromise overall repertoire diversity.

These findings indicate early establishment of a diverse naïve IGH repertoire after pediatric HSCT, followed by gradual maturation of antigen-experienced repertoires, resembling key features of physiological B-cell development.

Introduction

Allogeneic hematopoietic stem cell (HSC) transplantation (HSCT) is an established therapy for various inherited and acquired disorders of the hematopoietic system. Following a brief period of conditioning regimen-induced pancytopenia, innate immune cells typically demonstrate rapid recovery within weeks post-HSCT while recovery of lymphopoiesis can take up to several months.[1, 2] Post-HSCT, higher B-cell counts have been associated with improved survival but were also implicated in the development of severe acute graft-versus-host disease (GvHD).[3-5] Conversely, slow B-cell recovery counts leave patients vulnerable to infections and impact vaccination responses.[1, 6-8] In previous studies, quantitative recovery of naïve and memory B-cell subsets was shown to be most impacted by the age of donors and recipients[9-11]. Recent consensus recommendations have emphasized monitoring of quantitative immune recovery parameters after transplantation[12], yet optimal adaptive humoral defense and immune homeostasis is not only determined by B-cell quantities and their reached stages of maturation but also by their diversity and antigen-binding profile of the antibody pool they produce. Previous work has highlighted persistent defects in the memory B-cell compartment following HSCT, including studies by Storek et al. demonstrating impaired humoral immunity and delayed recovery of antigen-experienced B cells.[1] Repertoire analyses have already proven valuable in evaluating immune fitness in immunodeficiencies like RAG deficiency, CVID and others as well as during fetal development.[13-19] However, relatively few studies have comprehensively characterized qualitative B-cell receptor (BCR) repertoire reconstitution after HSCT using high-throughput sequencing approaches, and many studies have instead focused on quantitative immune recovery or the T-cell receptor repertoire.[20]

The naïve B-cell receptor repertoire is formed during B-cell development in the bone marrow and fetal liver via recombination of the variable (V), diversity (D) and joining (J) genes in the immunoglobulin heavy chain locus (IGH), and V:J gene-recombination of the Ig kappa and lambda light chain loci (IGK and IGL). Recombination and usage of various V, (D,) and J genes of an Ig heavy and light chain together with exonuclease trimming and random nucleotide addition by terminal deoxynucleotidyl transferase (TdT) at the V:D and D:J junctions result in a highly diverse BCR repertoire.[21, 22] The highest degree of diversity is reflected in the *IGH* complementarity-determining region 3 (CDR3), where the V, D and J genes join. The CDR3 region is commonly used to study BCR repertoire diversity. Intrinsic biases in gene usage, processing and selection mechanisms of the naïve BCR repertoire remove potentially autoreactive B cells.[23] When emerging from the bone marrow, the first naïve B cells will migrate to the secondary lymphoid organs and upon antigen recognition undergo additional affinity

maturation and specification in the germinal centers, where somatic hypermutation (SHM) in the variable domains and class-switch recombination from IgM, to IgG, IgA or IgE takes place, together shaping the antigen-experienced BCR repertoire. This developmental sequence provides a physiological framework against which post-transplant B-cell repertoire reconstitution can be evaluated.

Previous studies investigating BCR repertoire recovery after HSCT have consistently reported delayed maturation of the antigen-experienced compartment, although differences have been described regarding the timing of repertoire normalization.[24-26] Studies have shown that diversity of the antigen-experienced BCR repertoire remains reduced during the first two years post-HSCT, accompanied by persistence of B cells carrying somatically hypermutated B-cell receptors.[26, 27] In contrast, others report that the antigen-experienced repertoire in children reaches a polyclonal state similar to healthy donors approximately 6 months after HSCT, although the timing varies depending on transplant type and immunosuppressive therapy[25]. Although these studies provided important insights into post-transplant B-cell reconstitution, they predominantly employed CDR3 spectratyping and Sanger sequencing. An in-depth longitudinal evaluation of BCR repertoire diversity, clonal composition and affinity maturation using next-generation sequencing remains lacking. Taken together, the published data suggest that overall BCR repertoire diversity post-HSCT is initially reduced. However, the timeframe required for normalization of repertoire diversity remains incompletely defined, particularly with respect to potential differences between naïve and antigen-experienced compartments. To address this, we performed high-throughput targeted-messenger RNA sequencing of rearranged immunoglobulin heavy (IGH) gene transcripts, shortly BCR repertoire sequencing (NGS) on peripheral blood mononuclear cells (PBMCs) collected from pediatric patients at 3-, 6-, 12- and 24-months post-HSCT. We find that the naïve BCR repertoire is already highly diverse at 3 months, but observe notable differences in memory BCR repertoire diversity, clonality and affinity maturation. These data provide valuable insight into two-year longitudinal naïve and antigen-experienced BCR repertoire reconstitution after transplant and establish a framework for future studies evaluating qualitative aspects of humoral immune recovery in patients undergoing HSCT or gene therapy.

Methods

Patients and healthy donors

Peripheral blood samples were collected from ten pediatric patients who underwent allogeneic HSCT between January 2004 and December 2018. Patients with B-cell counts <100 cells*10^6/L at 3 months post-HSCT were

excluded from the study to ensure sufficient material for reliable repertoire analysis. Samples were obtained during routine follow-up visits with written informed consent, as part of the LUMC-WAKZ biobank, and in accordance with the protocol approved by said biobank and the LUMC ethics committee (P.01.028). Ten age-matched healthy donors served as a benchmark for physiological repertoire architecture and maturation rather than as a clinically matched control group.

Next-generation sequencing library

Peripheral blood mononuclear cells (PBMCs) were isolated using standard Ficoll-Isopaque density separation. Total RNA was isolated using the GenElute™ Total RNA Purification Kit (Sigma-Aldrich, St. Louis, MO, USA) according to the manufacturer's instructions, including the DNase I step (Sigma-Aldrich). Full-length cDNA libraries were generated using approximately 2µg of purified RNA with the SuperScript™ II Reverse Transcriptase kit (ThermoFisher Scientific, Walham, MA, USA). For amplification of immunoglobulin heavy chain (*IGH*) transcripts, a multiplex PCR was performed using forward IGHV1-6 FR1 (BIOMED-2) primers and *CgCH1*, *IGHCG* or *IGHCA* reverse primers using equal cDNA input.[28, 29] Agarose gel electrophoresis with MinElute gel extraction (Qiagen, Valencia, CA, USA) was used for product purification. Subsequently, all samples were pooled equimolarly and purified with Agencourt AMPure XP beads (Beckman Coulter, Fullerton, CA, USA). Sequencing adapters were ligated with the KAPA HyperPrep Kit (Roche Sequencing, Pleasonton, CA, USA). Sequencing was performed on the Illumina MiSeq with the MiSeq Reagent Kit v3 (Illumina, San Diego, USA).

BCR repertoire analysis

Raw *IGH* sequences were demultiplexed and end-trimmed to remove primer sequences. Paired-end reads were merged with a minimum overlap of 10bp, minimum paired-read length of 250bp and minimum Phred score of 25, to omit short or low-quality reads from the *IGH* repertoires. *IGH* sequences were subsequently annotated using IMGT/HighV-QUEST using release 202343-3[16]. Further data processing was performed in Antigen Receptor Galaxy (ARGalaxy), where IMGT annotated data from *IGHM*, *IGHG* and *IGHA* were concatenated before further analysis. Unproductive sequences—defined as out-of-frame rearrangements or those containing stop codons—and those labeled ‘unknown’ or ‘no results’ were excluded from the analysis. To control for amplification bias in *IGH* V(D)J gene usage and somatic hypermutation-analysis, duplicate reads were removed based on clonal definition 85% *IGHV* region identity, top *IGHV*, top *IGHJ* and nucleotide identity of the complementarity determining region 3 (CDR3) before analysis of *IGHV*, *IGHD* and *IGHJ* gene usage. For estimation of diversity

and clonal composition, clonal assignments of the full set of productive *IGH* sequences were generated by the AIRRNAT pipeline[30] using default settings including down sampling to equalize read quantities, followed by visualization in network diagrams or further downstream analysis of clonal overlap using the scRepertoire package.[31]

Statistical analysis

Statistical analyses were performed using Graphpad Prism V10. To evaluate differences between post-HSCT patients and HDs, including subgroup comparisons per timepoint, ANOVA or mixed models were used. For single-way comparisons, Kruskall-Wallis tests were used. Multiplicity-adjusted significance indicated as * $p \leq 0.0332$, ** $p \leq 0.0021$, *** $p \leq 0.0002$ and **** $p \leq 0.00001$.

Results

B cells rapidly repopulate the peripheral B-cell compartment in the first 3 months post-HSCT.

In this study, we included ten patients who underwent allogeneic HSCT and experienced limited post-transplant complications during the first two years (Table 1). Notably, none had moderate or severe acute or chronic graft-versus-host disease (GvHD). Patients were transplanted for a range of indications, including inborn errors of immunity and malignant and non-malignant hematological disorders, with a median age at transplantation of 7.2 years (range: 1.8–17.8 years). Conditioning regimens were busulfan-based in 8 of the 10 of patients. All patients received unmanipulated bone marrow grafts, and 9 out of 10 patients were 10/10 matched, of whom 3 HLA identical siblings. Three out of 10 patients developed acute GvHD (aGvHD) limited to the skin without the need for systemic steroids. Patient 4 developed limited chronic GvHD (cGvHD) affecting the skin and eyes. Most patients reached full donor chimerism in PBMCs and granulocytes at 3 months, except P05, P06 and P10, whose PBMC compartment reached full donor chimerism at 9 months (Figure 1A; Table 1). In the absence of lineage-specific B-cell chimerism (e.g. B-cell chimerism), PBMC chimerism is included as a general indicator of donor engraftment. Sixty-percent of patients received short-term intravenous immunoglobulins (IVIG) within the first 6 months post-transplant as part of standard supportive care post-HSCT (Figure 1B; Table 1).

In our cohort, lymphocytes recovered to age-matched reference levels at 3 months post-HSCT (Supplemental Figure 1A). All patients had good CD4+ T-cell recovery at 3 months (400-2100 cells*10^6/L; Supplemental Figure 1B). B-cell reconstitution showed inter-individual variability, with B cells becoming detectable at 32.7 days (range

10-68 days); the age-matched reference range of 200-600 cells*10^6/L was reached by 141.8 days (range 48-323 days; Supplemental Figure 1C). Patients receiving IVIG exhibited slightly lower circulating B-cell counts. To assess whether total B-cell recovery in our cohort aligned with previous reported patterns, we compared the B-cell counts to a univariate generalized predictive model (GAM) previously described for predicting B-cell reconstitution (Figure 1C, black line indicates GAM).[32] Total B-cell counts in our cohort followed the predictive model, stabilizing around one-year post-transplant. Memory B cells increased gradually over the two years although memory B cell counts peaked at 6 months for P04 and P05 (Supplemental Figure 1D). All patients had normal to borderline-normal IgM levels by 3 months post-transplant, except for P02 (Supplemental Figure 1E). IgM levels of P08 and P09 dropped after 3 months and remained consistently low, whereas P10 experienced an unexplained strong increase in IgM levels and received imatinib for maintenance treatment of CML. Serum IgG was in range for all patients by 6 months, although P01, P05 and P08-10 received IVIGs briefly after transplant (Supplemental Figure 1F; Figure 1B). Serum IgA levels followed the overall pattern of IgM, with low levels for P08 and P09 and a periodic increase in IgA for P10.

Early locus-wide *IGH* gene usage in the naïve BCR repertoire post-HSCT.

To investigate the reconstitution kinetics and diversity of naïve and antigen-experienced B-cell receptor (BCR) repertoire during the first two years post-HSCT, we performed bulk sequencing of *IGH* transcripts in cryopreserved PBMC from 10 patients collected at 3-, 6-, 12- and 24-months post-transplant. These samples were compared to 10 age-matched healthy donors (HD; median age 6.1 years, range 1.17-11.7). First, we examined gene usage patterns of naïve repertoires to get insight into the composition and diversity. To approximate the naïve compartment, we retained unique naïve *IGHM* sequences based on *IGHV*, *IGHJ* and CDR3 composition at the nucleotide level with low (<2%) somatic hypermutations (SHM). Locus-wide *IGHV, IGHD* and *IGHJ* gene usage was already established at 3 months post-HSCT and remained largely comparable to HDs throughout follow-up (Figure 2A-C). Although several individual genes differed significantly in frequency, including reduced *IGHV5-51, IGHV1-69, IGHJ5* and *IGHJ6* and increased *IGHV3-23, IGHV3-74* and *IGHJ3*, these differences did not alter the overall repertoire architecture. Accordingly, *IGHV-IGHJ* combinatorial diversity closely resembled that of HDs as early as 3 months post-HSCT (Figure 2D).

Junctional diversity was likewise preserved, with no differences in nucleotide deletions, N-nucleotide additions or CDR3 amino acid composition compared with HDs (Supplemental Figure 2A-C). Patients with delayed full donor chimerism (P05, P06 and P10) demonstrated comparable naïve repertoire diversity, junctional

characteristics and B-cell recovery kinetics to those achieving early full donor chimerism (data not shown). Apart from a transient reduction in *IGHV1* family usage at 6 months, no persistent differences in repertoire architecture were observed (data not shown). Together, these findings demonstrate that a broadly diverse, physiologically organized naïve BCR repertoire is established by the time of initial B-cell reconstitution after pediatric HSCT.

Gradual increase in somatic hypermutations in antigen-experienced BCR repertoires after HSCT.

We next focused on how the antigen-experienced repertoire is reinstated post-HSCT. Here, we consider *IGHG* and *IGHA* BCRs to be antigen-experienced, having undergone SHM and/or class-switch recombination (CSR) in the secondary lymphoid organs. Duplicate sequences were removed according to top *IGHV*, top *IGHJ* and CDR3 at the nucleotide level. We found that median SHM was strongly reduced in *IGHG* at 3 and 6 months post-HSCT compared to HDs ($p = 0.0061$ and $p = 0.0014$, resp.) and in *IGHA* across all post-HSCT timepoints ($p \leq 0.0233$; Figure 3A). Median SHM in *IgG* reached normal levels at approximately 24 months post-HSCT. Of those mutations, we found that SHM targeting of activation-induced cytidine deaminase (AID; RGYW/WRCY) and polymerase η (TW/WA) in both *IgG* and *IgA* were in line with those of HDs, although mutations in the WA motif were slightly reduced in both *IGHG* and *IGHA* (Figure 3B). Given the low median SHM early post-HSCT, we investigated whether this also affected the ratio of replacement and silent (R/S) mutations in the framework (FR) and the CDR regions. A higher R/S ratio in the CDR is typically associated with increased antigen specificity, whereas in the FR regions, it may compromise BCR structural stability.[33] Despite an absolute reduction of overall mutations, we found no difference in the R/S ratio in the FR and CDR regions compared to HDs in both *IGHG* and *IGHA* (Figure 3C), indicating that there was no perturbation of the position and nature of mutations acquired in the post-HSCT antigen-experienced repertoire. Taken together, our data show that SHM of the antigen-experienced repertoire is reduced post-transplant – most notably at 3 months – but gradually increases, reaching levels comparable to age-matched healthy donors by 2 years post-HSCT in *IGHG* with a slight delay observed in *IGHA*. Given that a subset of patients received short-term IVIG, we performed an exploratory comparison of IVIG-exposed and non-exposed patients and found no differences in Shannon entropy or median SHM frequency within *IGHG* or *IGHA* (data not shown), suggesting that IVIG exposure did not substantially confound antigen-experienced repertoire maturation.

Preserved molecular signatures of antigenic selection post-transplant.

Next, we investigated the characteristics of antigenic selection post-HSCT. The CDR3 composition of the antigen-experienced repertoire revealed a statistically significant increase in median deletions in *IGHG* and *IGHA*

compared to the naïve repertoire, a pattern consistently observed across all post-HSCT timepoints as well as in HDs (Figure 3D). Conversely, there were limited non-templated N-additions (N-nucleotides) in patient CDR3 regions, while these were strongly increased in age-matched HDs (Figure 3E). Despite this, there was no difference in median CDR3 length upon antigenic selection (Figure 3F). In contrast to a previous report in healthy children and adults[34], CDR3 length distribution was similar across naïve *IGHM*, *IGHG* and *IGHA* in our cohort (Figure 3F). Additionally, H-CDR3 composition of the naïve repertoire was polyclonal and length was normally distributed, whereas antigen-experienced repertoires showed more oligoclonal patterns post-HSCT (Figure 4A, Supplemental Figure 2D). Human and murine CDR3 regions are notably enriched for the amino acid (AA) tyrosine (Y) at the *IGHD-IGHJ* junction[35], which predominantly originates from *IGHJ6* in humans[36] and is crucial for antigen specificity[37]. Consistent with previous reports[34], tyrosine usage was significantly reduced in *IGHG* and *IGHA* compared to naïve *IGHM* at 6, 12 and 24 months post-HSCT and in HDs, but not yet at 3 months post-transplant (Figure 4B). Concomitantly, *IGHJ6* usage was also reduced in HDs, yet remained more readily utilized in post-transplant patients (Figure 4C). We found that the selection against *IGHV4-34*, an *IGHV* gene associated with several autoimmune diseases [38-40], was already evident in the antigen-experienced repertoire as early as three months post-HSCT, mirroring as observed in our HDs (Figure 4D). Overall, CDR3 characteristics of the post-HSCT antigen-experienced *IGH* repertoire were in line with those of HDs, with intact selection mechanisms against autoreactivity.

Reduced antigen-experienced repertoire diversity coincides with retained clonal architecture post-HSCT.

Having established that *IGH* repertoire composition of pediatric patients post-HSCT was largely in line with age-matched HDs, we next explored the clonality of the patients' reconstituting repertoire. To examine this, we selected all productive *IGH* sequences and computed clonal diversity and dominance by rarefaction of the Shannon entropy and Gini-Simpson index using Hill numbers, respectively. In our cohort, the *IGHG* repertoire diversity was reduced at 3-6 months post-HSCT yet can still be considered diverse and increased gradually throughout the first two years post-transplant (Figure 5A). There was no difference in overall *IGHA* repertoire diversity, and nor a large contribution of clonally dominant sequences in both *IGHG* and *IGHA* (Figure 5B). Still, we could identify several large clones in the antigen selected (*IGHG* and *IGHA*) repertoire of patients, predominantly in the first 6 months post-transplant, although expanded clones were also identified in the repertoire of HDs (Figure 5C). Of the top 50 BCR clonotypes, *IGHG* and *IGHA* displayed strong clonal convergence post-HSCT, although we observed a switch in predominant clonotypes between 12-24 months (Figure 6A; Supplemental Figure S3). Several clones persisted for up to two years post-HSCT (Figure 6B). When comparing

the naïve repertoire between patients and HDs, we found that clones were shared between patients over time, as well as with the repertoire of HDs (Figure 7). Concomitant with previous findings[27], *IGHG* and *IGHA* BCR clones were seldom shared between individual patients. Overall, naïve repertoires showed greater overlap between patients than antigen-experienced repertoires, while longitudinal clonal overlap within individual patients remained high for *IGHG* and *IGHA*.

Discussion

To advance our understanding on BCR repertoire reconstitution dynamics after HSCT, we provide an in-depth longitudinal analysis in ten pediatric patients receiving HSCT for malignant and non-malignant hematological disorders with largely uncomplicated post-transplant courses. Here, we demonstrate that at the time of initial peripheral B-cell reconstitution (3 months post-HSCT), the naïve repertoire is already diverse, with broad *IGHV*, *IGHD* and *IGHJ* gene usage. The antigen-experienced repertoire shows low SHM early after HSCT, followed by gradual normalization over the first two years with slightly delayed recovery in *IgA* compared to *IgG*. Although diversity in the antigen-experienced compartment was only modestly reduced, dynamic clonal expansions and shifts in dominant clonotypes were observed, reflecting ongoing antigen-driven maturation. These longitudinal data provide insight into the kinetics of repertoire reconstitution that cannot be captured by cross-sectional analyses, demonstrating recovery according to B-cell ontogeny rapid, locus-wide diversification during early B-lymphopoiesis, as classically described for naïve repertoire generation in the bone marrow, followed by antigen- and germinal-center-dependent maturation of the memory compartment.[41-43] This sequence closely mirrors the recapitulation of normal B-cell ontogeny previously described after adult bone marrow transplantation using isotype-based approaches[44], and extends this concept to the pediatric setting using high-resolution repertoire sequencing. The high degree of overlap between naïve repertories across repertoires further supports this concept, suggesting that early post-transplant B-cell development follows a conserved developmental program despite heterogenous underlying diseases and transplant characteristics.

Early studies using low-resolution techniques such as CDR3 spectratyping suggested that naïve B-cell receptor (BCR) repertoire reconstitution remained oligoclonal and abnormal at 6 months post-HSCT.[24, 25, 45] Furthermore, several studies indicated that reconstitution is accompanied by increased usage of IGHV3, while IGHV6 usage is reported to be limited or transient.[45, 46] These studies were performed before the advent of Next-Generation Sequencing, and even in more recent work, molecular insights remain limited, particularly regarding

V(D)J usage and combinatorial diversity. For example, Sethi et al., 2017[27] reported only fragmentary data on IGHV1 gene usage. Here we show that the entire *IGHV* gene locus is used as early as 3 months post-HSCT, suggesting earlier reconstitution of the naïve compartment than previously assumed. Consistent with previous reports, *IGHV3* was predominantly used, while IGHV6-1 usage – previously described as limited or transient – was maintained over time. We also report a marked reduction in *IGHV5-51* and *IGHV*1-69, which might be impacted by antigenic exposure as has been suggested before.[32, 46, 47] *IGHJ3* gene use was also strongly increased post-HSCT. Importantly, we describe that the naïve H-CDR3 length is polyclonal and remains diverse throughout the first two years post-HSCT. While biased usage of specific *IGHV* genes has been reported previously, we observe preservation of overall combinatorial diversity despite subtle shifts in gene frequencies, suggesting that global repertoire is maintained and, leaving a broad naïve BCR repertoire ready to undergo affinity maturation at the periphery. These findings suggest that early numerical B-cell recovery after HSCT is accompanied by rapid restoration of the substrate required for subsequent adaptive humoral responses, although functional antigen-specific immunity still depends on later germinal-center maturation.

Empirically, one might expect that the antigen-experienced BCR repertoire would show perturbations, specifically a reduction in affinity maturation as newly emerged B cells have not yet passed through germinal centers to obtain SHM and undergo CSR. Consistent with previous findings, we confirm that the antigen-experienced compartment and especially SHM frequency recovers more slowly than the naïve pool.[27, 48, 49] This parallels normal development of the antigen-experienced repertoire in healthy children.[34] Despite reduced SHM early post-transplant, targeting of AID and polymerase η motifs, as well as selection against autoreactive clones such as IGHV4-34, remained intact, indicating intact peripheral tolerance and affinity-driven selection. Thus, B-cell intrinsic processes such as AID induction, affinity-based selection and tolerance checkpoints appear intact but are functionally $CD4^+$ T-cell help-dependent within germinal centers, which recover gradually during the first years after HSCT.[2, 50] This is consistent with the established requirement for cognate T-cell help within the germinal center light zone for B-cell affinity-based selection and class-switch recombination.[51, 52] Accordingly, restoration of quantitative B-cell numbers should not be interpreted as equivalent to complete humoral immune recovery during the first post-transplant year.

The composition of the memory B-cell compartment is highly dynamic and continually subject to change in response to antigenic exposure. Previous studies determining repertoire diversity through CDR3 spectratyping and/or quantification of clones of *IGHV1*, *IGHV3* and/or *IGHV6* concluded that the BCR repertoire is oligoclonal in the first 6 months post-HSCT and has limited inter-individual clonal overlap.[24, 25, 27] In parallel, seminal studies

have demonstrated that recovery of the memory B-cell compartment after HSCT is delayed and often functionally impaired, contributing to prolonged immune deficiency and suboptimal vaccine responses.[1, 53-55] More recent work supports persistent qualitative defects in memory B cells despite numerical recovery, underscoring the important of assessing immune competence beyond B-cell counts.[56, 57] Through diversity indices and analyses of clonal dynamics, we now show that the antigen-experienced repertoire of pediatric patients is highly diverse post-HSCT. Furthermore, we demonstrate that clonally expanded BCR clonotypes are present and persist post-transplant, yet memory B cells appear to undergo a shift in the top clonally expanded sequences; the majority of the top clonotypes is replaced during the second year post-HSCT. This could, at least in part, be explained by antigenic exposure and revaccination at approximately 6 to 12 months post-HSCT, although this could not be directly assessed in our cohort. Notably, *IGHG* and *IGHA* displayed intra- and inter-individual sharedness. Notably, overlap between antigen-experienced repertoires was limited across patients, whereas substantial clonal overlap was observed within individual patients over time. This suggests that, once established, antigen-experienced immunity becomes increasingly individualized through antigen-driven selection while maintaining longitudinal stability. Clinically, these observations support the concept that restoration of durable humoral immunity after HSCT depends not only on recovery of B-cell numbers but also on progressive maturation and maintenance of patient-specific memory B-cell responses.

Our cohort intentionally comprised patients with successful early B-cell recovery, setting a benchmark for BCR repertoire development post-HSCT. Whether similar repertoire dynamics occur in patients with delayed or impaired B-cell reconstitution remains to be determined. Exploratory analyses of patients with delayed full donor chimerism did not reveal consistent alterations in repertoire architecture but were limited by small subgroup size. IVIG-exposure in our cohort depended on varying clinical indications, such as low serum IgG, immune reconstitution parameters and individual considerations such as infection risk troubling the assessment of a causal relationship to B-cell repertoire recovery, The cohort was heterogeneous with respect to underlying diagnosis, including hematological and non-hematological disorders, and predominantly comprised younger children with only two adolescent patients. Larger, prospectively designed cohorts will be required to determine whether repertoire reconstitution differs by underlying diagnosis, age at transplantation, or degree of B-cell recovery at 3 months.

Taken together, our data show that pediatric HSCT recipients rapidly establish a diverse naïve IGH repertoire, reflective of intact *de novo* B-lymphopoiesis. A portion of the naïve clonotypes of patients in our cohort was also shared with that of age-matched HDs. Maturation of the antigen-experienced compartment proceeded

gradually over the first two years post-transplant, with gradual accumulation of SHM up to one-year post-transplant. Additional studies will be needed to confirm that IVIG exposure and/or delayed full donor chimerism indeed do not affect BCR repertoire development after HSCT. Our findings suggest that numerical B-cell recovery alone may not fully reflect restoration of humoral immune competence. While de novo generation of a diverse naïve repertoire occurs rapidly after HSCT, maturation of antigen-experienced immunity continues well beyond the first post-transplant year. These observations support current approaches to post-transplant immune monitoring that incorporate qualitative assessment of memory B-cell recovery and vaccine responsiveness in addition to quantitative lymphocyte recovery. Future studies integrating repertoire analysis with functional immune readouts such as autoreactivity and antigen specificity and comparisons across clinical subgroups like GvHD and post-transplant autoimmunity will give further insights into aberrations in establishments of adaptive humoral defense and immune homeostasis post-HSCT.

References

1. Storek J, Geddes M, Khan F, et al. Reconstitution of the immune system after hematopoietic stem cell transplantation in humans. Semin Immunopathol. 2008;30(4):425-437.
2. Abdel-Azim H, Elshoury A, Mahadeo KM, Parkman R, Kapoor N. Humoral immune reconstitution kinetics after allogeneic hematopoietic stem cell transplantation in children: a maturation block of igM memory B cells may lead to impaired antibody immune reconstitution. Biol Blood Marrow Transplant. 2017;23(9):1437-1446.
3. Ando T, Tachibana T, Tanaka M, et al. Impact of graft sources on immune reconstitution and survival outcomes following allogeneic stem cell transplantation. Blood Adv. 2020;4(2):408-419.
4. Yao Y, Li B, Li J, et al. The frequencies of lymphocyte subsets on "day 30″ correlate with the clinical outcome of pediatric hematopoietic stem cell transplantation. Immunol Lett. 2023;254:21-29.
5. Troullioud Lucas AG, Lindemans CA, Bhoopalan SV, et al. Early immune reconstitution as predictor for outcomes after allogeneic hematopoietic cell transplant; a tri-institutional analysis. Cytotherapy. 2023;25(9):977-985.
6. Avanzini MA, Locatelli F, Dos Santos C, et al. B lymphocyte reconstitution after hematopoietic stem cell transplantation: functional immaturity and slow recovery of memory CD27+ B cells. Exp Hematol. 2005;33(4):480-486.
7. D'Orsogna LJ, Wright MP, Krueger RG, et al. Allogeneic hematopoietic stem cell transplantation recipients have defects of both switched and igm memory B cells. Biol Blood Marrow Transplant. 2009;15(7):795-803.
8. Corre E, Carmagnat M, Busson M, et al. Long-term immune deficiency after allogeneic stem cell transplantation: B-cell deficiency is associated with late infections. Haematologica. 2010;95(6):1025-1029.
9. Savage WJ, Bleesing JJ, Douek D, et al. Lymphocyte reconstitution following non-myeloablative hematopoietic stem cell transplantation follows two patterns depending on age and donor/recipient chimerism. Bone Marrow Transplant. 2001;28(5):463-471.
10. van der Maas NG, Berghuis D, van der Burg M, Lankester AC. B cell reconstitution and influencing factors after hematopoietic stem cell transplantation in children. Front Immunol. 2019;10:782.
11. Zhou G, Zhan Q, Huang L, et al. The dynamics of B-cell reconstitution post allogeneic hematopoietic stem cell transplantation: a real-world study. J Intern Med. 2024;295(5):634-650.
12. Hammoudi T, Nucera S, Troullioud Lucas AG, et al. Harmonized immune recovery monitoring after HCT: evidence and practical guidance from the Westhafen Intercontinental Group. Blood Adv. 2025;9(23):6141-6157.
13. Lee YN, Frugoni F, Dobbs K, et al. Characterization of T and B cell repertoire diversity in patients with RAG deficiency. Sci Immunol. 2016;1(6):eaah6109.
14. Ott de Bruin LM, Bosticardo M, Barbieri A, et al. Hypomorphic Rag1 mutations alter the preimmune repertoire at early stages of lymphoid development. Blood. 2018;132(3):281-292.
15. van Schouwenburg PA, Ijspeert H, Pico-Knijnenburg I, et al. Identification of CVID patients with defects in immune repertoire formation or specification. Front Immunol. 2018;9:2545.
16. Ijspeert H, van Schouwenburg PA, Pico-Knijnenburg I, et al. Repertoire sequencing of B cells elucidates the role of UNG and mismatch repair proteins in somatic hypermutation in humans. Front Immunol. 2019;10:1913.
17. van Schouwenburg PA, van der Burg M, Ijspeert H. NGS-based b-cell receptor repertoire analysisrepertoire analyses in the context of inborn errors of immunity. Methods Mol Biol. 2022;2453:169-190.
18. Ijspeert H, Wentink M, van Zessen D, et al. Strategies for B-cell receptor repertoire analysis in primary immunodeficiencies: from severe combined immunodeficiency to common variable immunodeficiency. Front Immunol. 2015;6:157.
19. Rechavi E, Lev A, Lee YN, et al. Timely and spatially regulated maturation of B and T cell repertoire during human fetal development. Sci Transl Med. 2015;7(276):276ra225.
20. Cowan MJ, Yu J, Facchino J, et al. Lentiviral gene therapy for artemis-deficient SCID. N Engl J Med. 2022;387(25):2344-2355.
21. Jackson KJ, Kidd MJ, Wang Y, Collins AM. The shape of the lymphocyte receptor repertoire: lessons from the B cell receptor. Front Immunol. 2013;4:263.
22. Elhanati Y, Sethna Z, Marcou Q, Callan CG Jr., Mora T, Walczak AM. Inferring processes underlying B-cell repertoire diversity. Philos Trans R Soc Lond B Biol Sci. 2015;370(1676):20140243.
23. Neehus A-L, Romberg N, Sankaran VG. Human genetic influences on early B cell development. J Hum Immun. 2025;1(3):e20250042.

24. Omazic B, Lundkvist I, Mattsson J, Permert J, Nasman-Bjork I. Memory B lymphocytes determine repertoire oligoclonality early after haematopoietic stem cell transplantation. Clin Exp Immunol. 2003;134(1):159-166.
25. Di Martino D, Terranova MP, Scuderi F, et al. V(H)3 and V(H)6 immunoglobulin M repertoire reconstitution after hematopoietic stem-cell transplantation in children. Transplantation. 2005;79(1):98-107.
26. Omazic B, Hentschke P, Näsman-Björk I, et al. Reconstitution of the Ig heavy chain CDR3 repertoire after allogeneic haematopoietic stem cell transplantation with myeloablative or reduced-intensity conditioning regimens. Scand J Immunol. 2005;61(1):72-81.
27. Sethi MK, Thol F, Stadler M, et al. VH1 Family immunoglobulin repertoire sequencing after allogeneic hematopoietic stem cell transplantation. PLoS One. 2017;12(1):e0168096.
28. Langerak AW, Groenen PJ, Brüggemann M, et al. EuroClonality/BIOMED-2 guidelines for interpretation and reporting of Ig/TCR clonality testing in suspected lymphoproliferations. Leukemia. 2012;26(10):2159-2171.
29. van Dongen JJ, Langerak AW, Brüggemann M, et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. Leukemia. 2003;17(12):2257-2317.
30. Blazso P, Csomos K, Tipton CM, Ujhazi B, Walter JE. Lineage reconstruction of in vitro identified antigen-specific autoreactive B cells from adaptive immune receptor repertoires. Int J Mol Sci. 2022;24(1):225.
31. Borcherding N, Bormann NL, Kraus G. scRepertoire: an R-based toolkit for single-cell immune receptor analysis. F1000Res. 2020;9:47.
32. van der Maas NG, von Asmuth EGJ, Berghuis D, et al. Modeling influencing factors in B-cell reconstitution after hematopoietic stem cell transplantation in children. Front Immunol. 2021;12:684147.
33. Julian MC, Li L, Garde S, Wilen R, Tessier PM. Efficient affinity maturation of antibody variable domains requires co-selection of compensatory mutations to maintain thermodynamic stability. Sci Rep. 2017;7:45259.
34. IJspeert H, van Schouwenburg PA, van Zessen D, et al. Evaluation of the antigen-experienced B-cell receptor repertoire in healthy children and adults. Front Immunol. 2016;7:410.
35. Ivanov, II, Schelonka RL, Zhuang Y, Gartland GL, Zemlin M, Schroeder HW Jr. Development of the expressed Ig CDR-H3 repertoire is marked by focusing of constraints in length, amino acid use, and charge that are first established in early B cell progenitors. J Immunol. 2005;174(12):7773-7780.
36. Wang H, Yan K, Wang R, et al. Antibody heavy chain CDR3 length-dependent usage of human IGHJ4 and IGHJ6 germline genes. Antib Ther. 2021;4(2):101-108.
37. Birtalan S, Zhang Y, Fellouse FA, Shao L, Schaefer G, Sidhu SS. The intrinsic contributions of tyrosine, serine, glycine and arginine to the affinity and specificity of antibodies. J Mol Biol. 2008;377(5):1518-1528.
38. Tipton CM, Fucile CF, Darce J, et al. Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. Nat Immunol. 2015;16(7):755-765.
39. Yuuki H, Itamiya T, Nagafuchi Y, Ota M, Fujio K. B cell receptor repertoire abnormalities in autoimmune disease. Front Immunol. 2024;15:1326823.
40. Bashford-Rogers RJM, Smith KGC, Thomas DC. Antibody repertoire analysis in polygenic autoimmune diseases. Immunology. 2018;155(1):3-17.
41. Hardy RR, Hayakawa K. B cell development pathways. Annu Rev Immunol. 2001;19:595-621.
42. LeBien TW, Tedder TF. B lymphocytes: how they develop and function. Blood. 2008;112(5):1570-1580.
43. Pieper K, Grimbacher B, Eibel H. B-cell biology and development. J Allergy Clin Immunol. 2013;131(4):959-971.
44. Velardi A, Cucciaioni S, Terenzi A, et al. Acquisition of Ig isotype diversity after bone marrow transplantation in adults. A recapitulation of normal B cell ontogeny. J Immunol. 1988;141(3):815-820.
45. Gokmen E, Raaphorst FM, Boldt DH, Teale JM. Ig heavy chain third complementarity determining regions (H CDR3s) after stem cell transplantation do not resemble the developing human fetal H CDR3s in size distribution and Ig gene utilization. Blood. 1998;92(8):2802-2814.
46. Fumoux F, Guigou V, Blaise D, Maraninchi D, Fougereau M, Schiff C. Reconstitution of human immunoglobulin VH repertoire after bone marrow transplantation mimics B-cell ontogeny. Blood. 1993;81(11):3153-3157.
47. Di Martino D, Dallorso S, Terranova P, Scarso L, Morreale G, Valetto A. B-cell repertoire reconstitution after hematopoietic stem cell transplantation in children evaluated by immunoglobulin heavy chain third complementarity determining region fingerprinting. Haematologica. 2004;89(4):506-508.

48. Suzuki I, Milner EC, Glas AM, et al. Immunoglobulin heavy chain variable region gene usage in bone marrow transplant recipients: lack of somatic mutation indicates a maturational arrest. Blood. 1996;87(5):1873-1880.
49. Glas AM, Nottenburg C, Milner EC. Analysis of rearranged immunoglobulin heavy chain variable region genes obtained from a bone marrow transplant (BMT) recipient. Clin Exp Immunol. 1997;107(2):372-380.
50. Ogonek J, Kralj Juric M, Ghimire S, et al. Immune reconstitution after allogeneic hematopoietic stem cell transplantation. Front Immunol. 2016;7:507.
51. Victora GD, Nussenzweig MC. Germinal centers. Annu Rev Immunol. 2012;30:429-457.
52. Victora GD, Nussenzweig MC. Germinal Centers. Annu Rev Immunol. 2022;40:413-442.
53. Storek J, Zhao Z, Lin E, et al. Recovery from and consequences of severe iatrogenic lymphopenia (induced to treat autoimmune diseases). Clin Immunol. 2004;113(3):285-298.
54. Trück J, Ramasamy MN, Galson JD, et al. Identification of antigen-specific B cell receptor sequences using public repertoire analysis. J Immunol. 2015;194(1):252-261.
55. Podgorny PJ, Liu Y, Dharmani-Khan P, et al. Immune cell subset counts associated with graft-versus-host disease. Biol Blood Marrow Transplant. 2014;20(4):450-462.
56. Nielsen SCA, Roskin KM, Jackson KJL, et al. Shaping of infant B cell receptor repertoires by environmental factors and infectious disease. Sci Transl Med. 2019;11(481):eaat2004.
57. Heck C, Steiner S, Kaebisch EM, et al. CD4+ T cell dependent B cell recovery and function after autologous hematopoietic stem cell transplantation. Front Immunol. 2021;12:736137.

Tables

Table 1.

ID	P01	P02	P03	P04	P05	P06	P07	P08	P09	P10
Diagnosis	β-TM	SAA	CAMT	SCN	HLH	SCN	BMF	CML blast crisis	T-ALL late relapse	CML
Genetic defect (where applicable)	*HBB*		*MPO*	*SRP54*	*STX11*	*ELA2*	*ADA2*			
Age at HSCT (years)	1.77	3.15	2.47	4.37	6.22	3.57	7.24	9.92	17.83	15.69
Conditioning	Bu Mel Cy ATG	Flu Cy ATG	Bu Cy	Bu Flu ATG	Bu Flu ATG	Bu Flu ATG	Treo Flu ATG	Bu Flu ATG	Bu Flu Clo	Bu Flu Clo
GvHD prophylaxis	CsA MTX	CsA MTX	CsA MTX	CsA MTX	CsA MTX	CsA MTX	CsA MTX	CsA MTX	CsA MTX	CsA/Tac* MTX
Graft source	MUD BM	MUD BM	IRD BM	MUD BM	IRD BM	IRD BM	MUD BM	MUD BM	MUD BM	MUD BM
Age donor (years)	34.72	47.52	5.36	29.14	28.5	6.63	41.61	30.72	22.37	45.24
HLA matching	10/10	10/10	10/10	10/10	10/10	10/10	9/10 *HLA-C*	10/10	10/10	10/10
Donor chimerism	Full	Full	Full	Full	Full	Full	Full	Full	Full	Full
aGvHD grade	I *skin*	-	-	-	-	-	-	II *skin*	-	I *skin*
cGvHD grade	-	-	-	Limited *skin-eyes*	-	-	-	-	-	-
First B cells detected (days)	64	11	10	11	43	24	22	68	53	21
B cells in normal range (>100 cells*10^6/L; days)	166	323	48	50	300	181	72	156	68	54
IVIG	Yes	No	No	No	Yes	No	Yes	Yes	Yes	Yes

Abbreviations: β-TM, β-thalassemia major; SAA, severe aplastic anemia; CAMT, congenital megakaryocytic thrombocytopenia; SCN, severe congenital neutropenia; HLH, hemophagocytic lymphohistiocytosis; BMF, bone marrow failure; AML, acute myeloid leukemia; CML, chronic myeloid leukemia; T-ALL, T-cell acute lymphoblastic leukemia; Bu, Bulsulfan; Mel, Melphalan; Cy, cyclophosphamide; ATG, anti-thymocyte globulin; Flu, Fludarabine; Clo, Clofarabine; CsA, Cyclosporin A; MTX, methotrexate; Tac, Tacrolimus; MUD, matched unrelated donor; IRD, identical related donor; aGvHD, acute graft-versus-host-disease; cGvHD, chronic graft-versus-host-disease; BO, bronchiolitis obliterans. **switched to Tacrolimus on day+1 due to intolerance to Cyclosporin A.*

Figure legends

Figure 1. Donor Chimerism and Immune Reconstitution During the First Two Years After HSCT. **A** Donor chimerism of PBMCs and granulocytes in the first 2 years post-HSCT. **B** Administration of post-transplant IVIGs in P01 and P05-10. **C** Total $CD19^+$ B-cell counts per patient compared to the univariate generalized predictive model of van der Maas et al. 2021 (black line).

Figure 2. Early, Broad IGH V(D)J Gene Usage in the Naïve BCR Repertoire After HSCT. Distribution of **A** *IGHV*, **B** *IGHD* and **C** *IGHJ* genes of the unique *IGHM* naïve BCR repertoire at 3, 6, 12 and 24 months post-HSCT compared to HDs. **D** Heatmap depicting the IGHV and IGHJ combinations observed in the BCR repertoire. Cell values represent the mean percentage of VJ-combination usage per group. Multiplicity-adjusted significance indicated as * $p \leq 0.0332$, ** $p \leq 0.0021$, *** $p \leq 0.0002$ and **** $p \leq 0.00001$.

Figure 3. Somatic Hypermutation and Antigenic Selection in Antigen-experienced BCR Repertoires After HSCT. A Median SHM in the *IGHV* gene and CDR3 region of *IGHG* and *IGHA* sequences. **B** The percentage of SHM targeting activation-induced cytidine deaminase (AID; RGYW/WRCY) and polymerase η (TW/WA) in *IGHG* and *IGHA*. **C** The ratio of replacement and silent mutations (R/S) in the framework (FR) and CDR regions. Median **D** deletions, **E** N-nucleotide additions and **F** CDR3 length (amino acid; aa) in the antigen-experienced repertoire. Multiplicity-adjusted significance indicated as * $p \leq 0.0332$, ** $p \leq 0.0021$, *** $p \leq 0.0002$ and **** $p \leq 0.00001$.

Figure 4. Structural and Germline Characteristics of Reconstituting BCR Repertoires After HSCT. A Distribution of CDR3 length in *IGHM* naïve, *IGHG* and *IGHA* in post-HSCT patients at 3, 6, 12 and 24 months post-HSCT compared to HDs. Relative proportion of the use of **B** the amino acid tyrosine (Y) in the CDR3 region, **C** *IGHJ6* and **D** *IGHV4-34* genes in the naïve and antigen-experienced repertoire of post-HSCT patients and HDs. Multiplicity-adjusted significance indicated as * $p \leq 0.0332$, ** $p \leq 0.0021$, *** $p \leq 0.0002$ and **** $p \leq 0.00001$.

Figure 5. Diversity and Network Architecture of Antigen-experienced BCR Repertoires After HSCT. For calculations of BCR diversity and clonality, all high-quality and productive *IGHM* naïve, *IGHG* and *IGHA* sequences were used. **A** Shannon entropy and **B** Gini-Simpson index of the *IGHG* and *IGHA* repertoire were calculated in AIRRNAT using standard settings. **C** Network diagrams from the AIRRNAT analysis pipeline of the naïve and antigen-experienced repertoires of two representative HDs and two representative patients at 3, 6, 12 and 24 months post-HSCT. Multiplicity-adjusted significance indicated as * $p \leq 0.0332$, ** $p \leq 0.0021$, *** $p \leq 0.0002$ and **** $p \leq 0.00001$.

Figure 6. Longitudinal Clonal Dynamics and Repertoire Overlap After HSCT. A Alluvial plots of top 50 BCR clones based on clonotype definition Top V, Top D, Top J gene and CDR3 at nucleotide level and **B** Circos of all overlapping sequences between timepoints of two representative patients showing post-transplant longitudinal and cross-sectional overlap dynamics of the naïve, *IGHG* and *IGHA* repertoire. Circos only depict clones overlapping between timepoints.

Figure 7. Clonal Overlap Between Patients and HDs. Clonal overlap is calculated at each timepoint using Morisita's overlap index after downsampling to equalize read quantities.

A

chimerism

PBMC

granulocyte

donor chimerism (%)

100
80
60
40
20
0

3 6 9 12 18 24

months post-HSCT

P01
P02
P03
P04
P05
P06
P07
P08
P09
P10

B

IVIG use

P01
P02
P03
P04
P05
P06
P07
P08
P09
P10

3 6 9 12

months post-HSCT

C

Total B cells

cells per µl

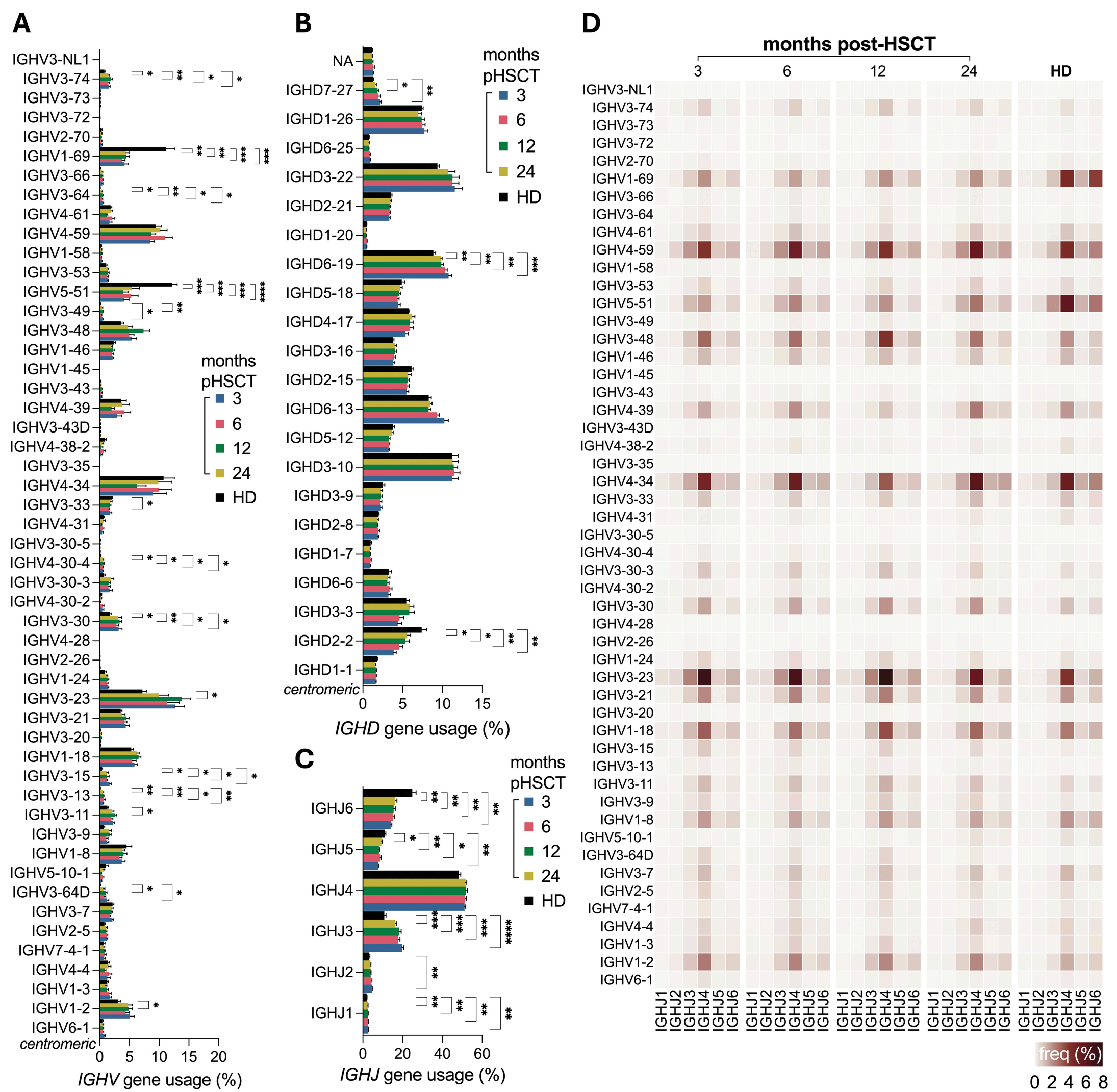
1000
100
10
1

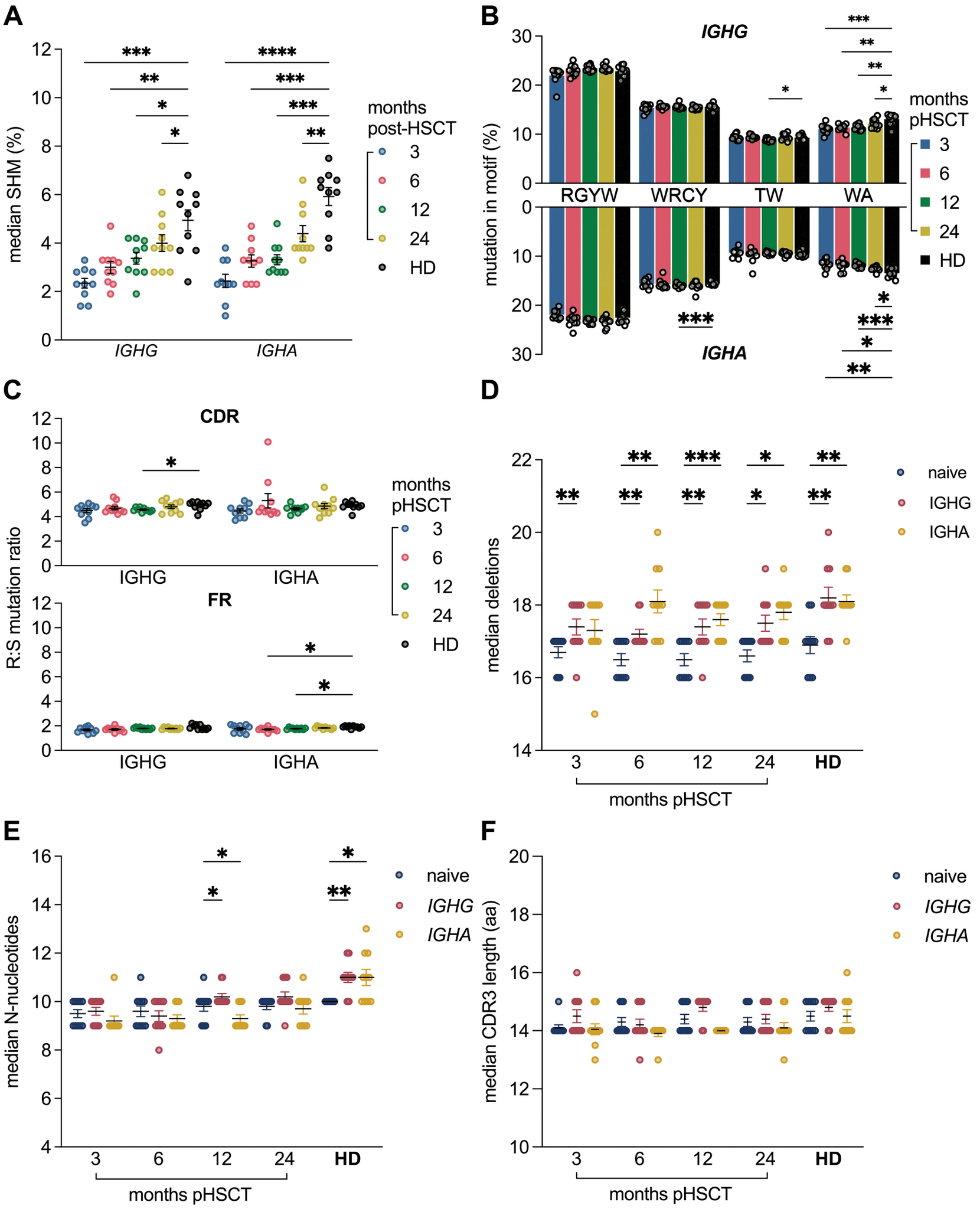
6 12 18 24

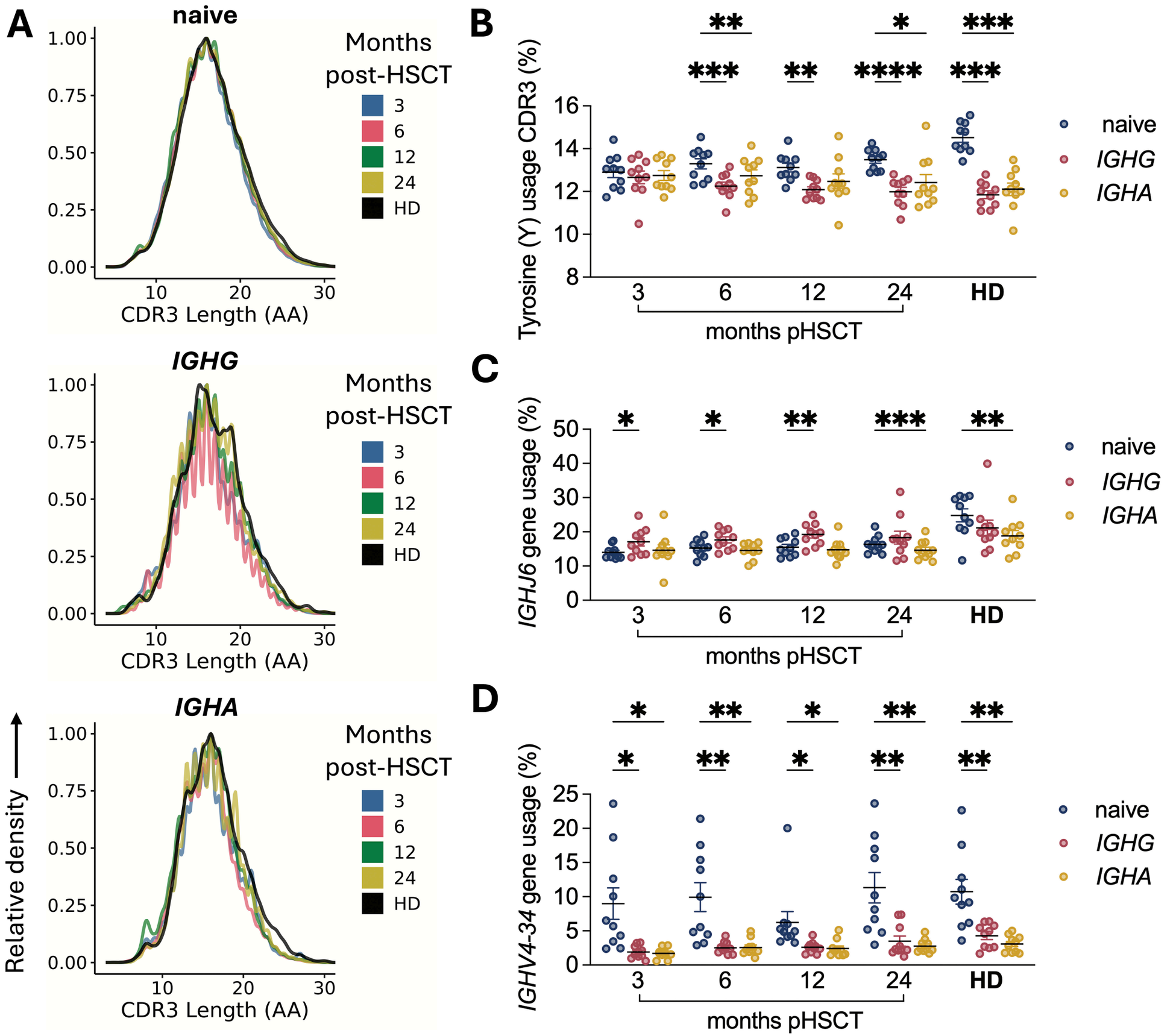
months post-HSCT

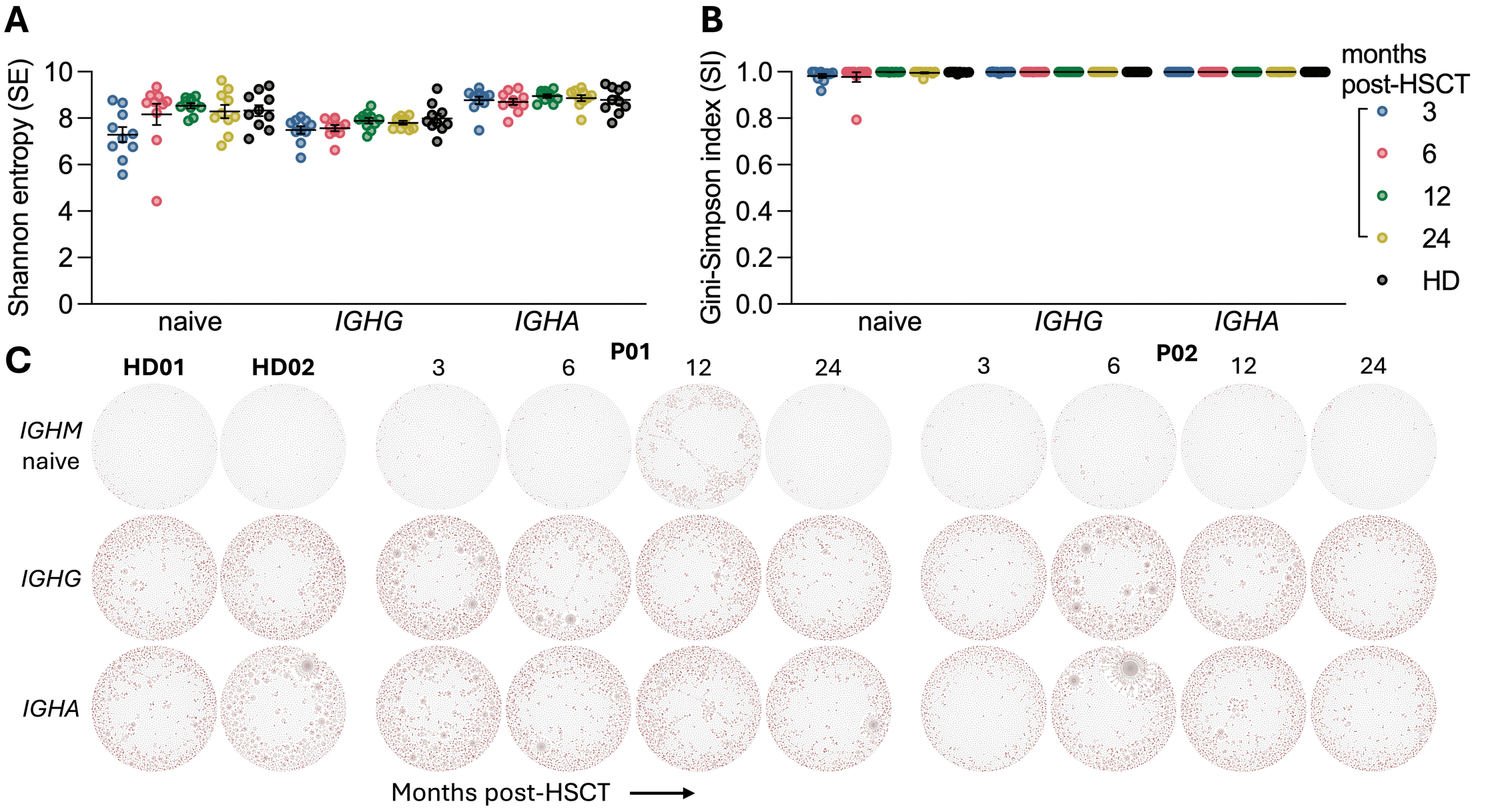
Patient

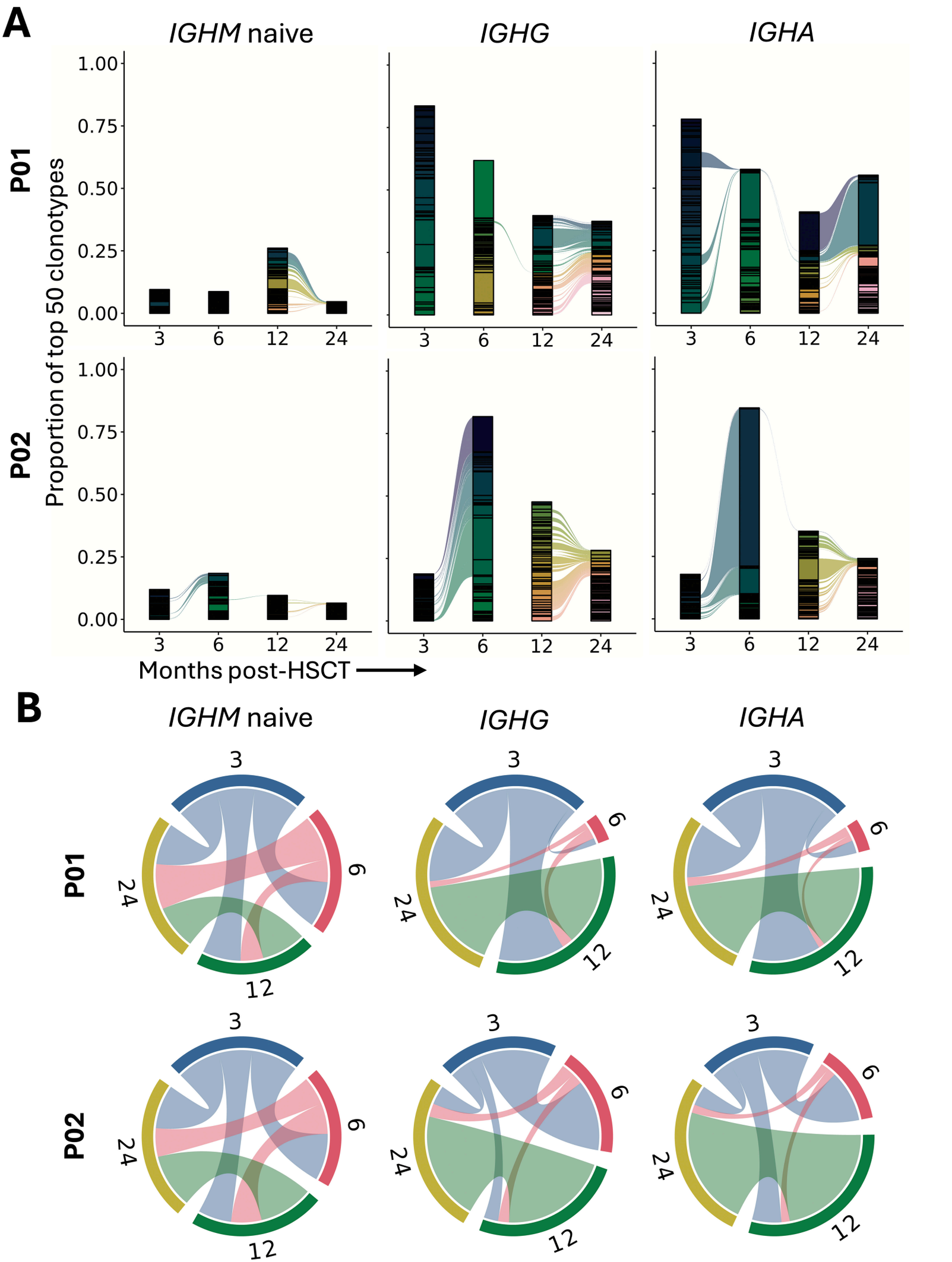
P01
P02
P03
P04
P05
P06
P07
P08
P09
P10

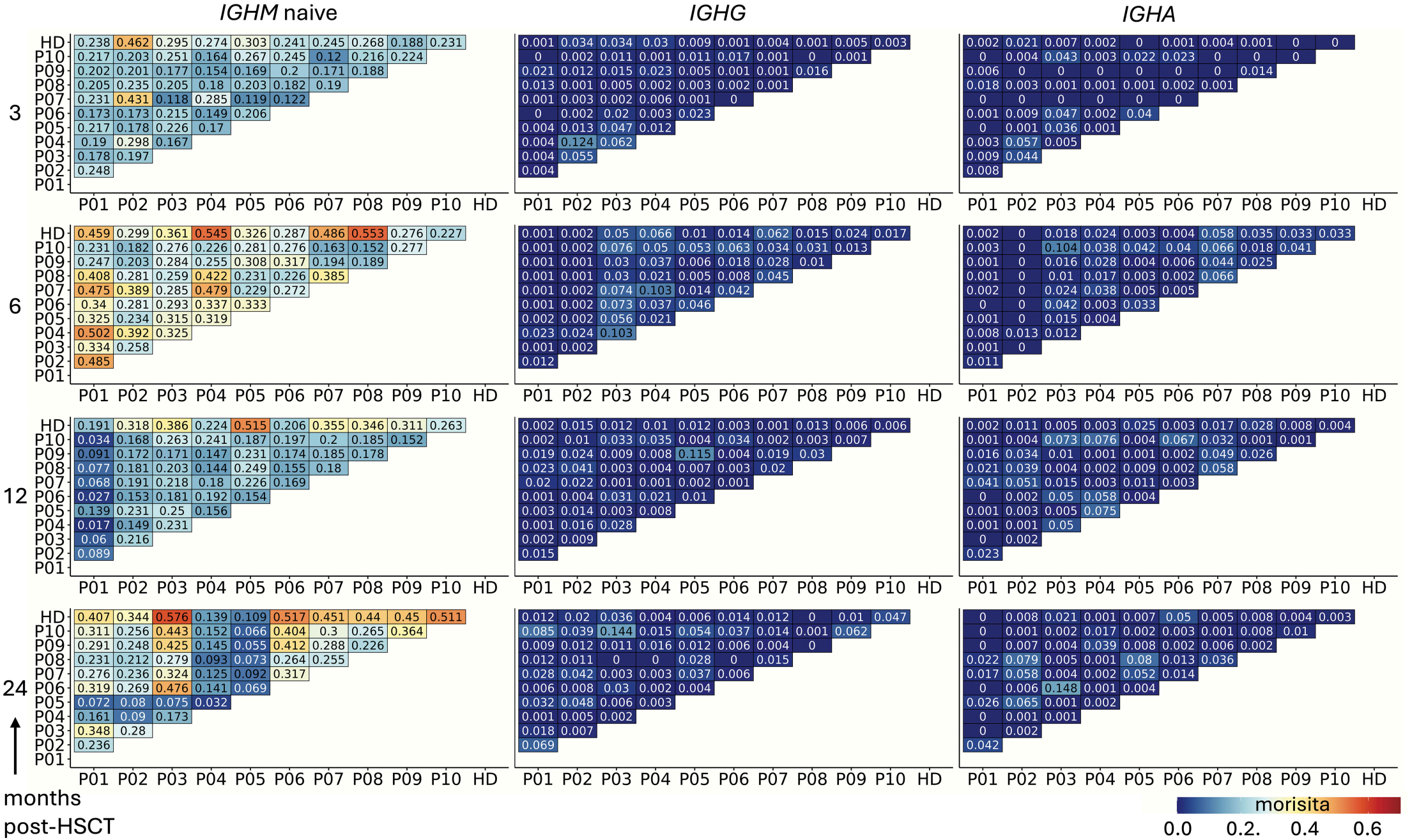

A
IGHV3-NL1
IGHV3-74
IGHV3-73
IGHV3-72
IGHV2-70
IGHV1-69
IGHV3-66
IGHV3-64
IGHV4-61
IGHV4-59
IGHV1-58
IGHV3-53
IGHV5-51
IGHV3-49
IGHV3-48
IGHV1-46
IGHV1-45
IGHV3-43
IGHV4-39
IGHV3-43D
IGHV4-38-2
IGHV3-35
IGHV4-34
IGHV3-33
IGHV4-31
IGHV3-30-5
IGHV4-30-4
IGHV3-30-3
IGHV4-30-2
IGHV3-30
IGHV4-28
IGHV2-26
IGHV1-24
IGHV3-23
IGHV3-21
IGHV3-20
IGHV1-18
IGHV3-15
IGHV3-13
IGHV3-11
IGHV3-9
IGHV1-8
IGHV5-10-1
IGHV3-64D
IGHV3-7
IGHV2-5
IGHV7-4-1
IGHV4-4
IGHV1-3
IGHV1-2
IGHV6-1
centromeric
0 5 10 15 20
IGHV gene usage (%)
months pHSCT
3
6
12
24
HD
B
NA
IGHD7-27
IGHD1-26
IGHD6-25
IGHD3-22
IGHD2-21
IGHD1-20
IGHD6-19
IGHD5-18
IGHD4-17
IGHD3-16
IGHD2-15
IGHD6-13
IGHD5-12
IGHD3-10
IGHD3-9
IGHD2-8
IGHD1-7
IGHD6-6
IGHD3-3
IGHD2-2
IGHD1-1
centromeric
0 5 10 15
IGHD gene usage (%)
months pHSCT
3
6
12
24
HD
C
IGHJ6
IGHJ5
IGHJ4
IGHJ3
IGHJ2
IGHJ1
0 20 40 60
IGHJ gene usage (%)
months pHSCT
3
6
12
24
HD
D
months post-HSCT
3
6
12
24
HD
IGHV3-NL1
IGHV3-74
IGHV3-73
IGHV3-72
IGHV2-70
IGHV1-69
IGHV3-66
IGHV3-64
IGHV4-61
IGHV4-59
IGHV1-58
IGHV3-53
IGHV5-51
IGHV3-49
IGHV3-48
IGHV1-46
IGHV1-45
IGHV3-43
IGHV4-39
IGHV3-43D
IGHV4-38-2
IGHV3-35
IGHV4-34
IGHV3-33
IGHV4-31
IGHV3-30-5
IGHV4-30-4
IGHV3-30-3
IGHV4-30-2
IGHV3-30
IGHV4-28
IGHV2-26
IGHV1-24
IGHV3-23
IGHV3-21
IGHV3-20
IGHV1-18
IGHV3-15
IGHV3-13
IGHV3-11
IGHV3-9
IGHV1-8
IGHV5-10-1
IGHV3-64D
IGHV3-7
IGHV2-5
IGHV7-4-1
IGHV4-4
IGHV1-3
IGHV1-2
IGHV6-1
IGHJ1 IGHJ2 IGHJ3 IGHJ4 IGHJ5 IGHJ6
freq (%)
0 2 4 6 8


A
median SHM (%)
months post-HSCT
3
6
12
24
HD
IGHG
IGHA
B
IGHG
IGHA
mutation in motif (%)
RGYW
WRCY
TW
WA
months pHSCT
C
CDR
FR
R:S mutation ratio
IGHG
IGHA
D
median deletions
naive
IGHG
IGHA
months pHSCT
E
median N-nucleotides
F
median CDR3 length (aa)


A
naive
IGHG
IGHA
Relative density
CDR3 Length (AA)
Months post-HSCT
3
6
12
24
HD
B
Tyrosine (Y) usage CDR3 (%)
months pHSCT
C
IGHJ6 gene usage (%)
D
IGHV4-34 gene usage (%)
naive
IGHG
IGHA


A
Shannon entropy (SE)
10
8
6
4
2
0
naive
IGHG
IGHA
B
Gini-Simpson index (SI)
1.0
0.8
0.6
0.4
0.2
0.0
naive
IGHG
IGHA
months post-HSCT
3
6
12
24
HD
C
HD01
HD02
P01
3
6
12
24
P02
3
6
12
24
IGHM naive
IGHG
IGHA
Months post-HSCT


A
IGHM naive
IGHG
IGHA
P01
P02
Proportion of top 50 clonotypes
1.00
0.75
0.50
0.25
0.00
3
6
12
24
Months post-HSCT
B
IGHM naive
IGHG
IGHA
P01
P02


IGHM naive
IGHG
IGHA
3
6
12
24
months post-HSCT
HD P10 P09 P08 P07 P06 P05 P04 P03 P02 P01
P01 P02 P03 P04 P05 P06 P07 P08 P09 P10 HD
morisita
0.0 0.2. 0.4 0.6

Early establishment of the naïve B-cell receptor repertoire with gradual antigen-experienced maturation after pediatric hematopoietic stem cell transplantation

Running title: BCR repertoire analysis post-HSCT

Melanie de Gier,[1] Martijn Cordes,[1] Ingrid Pico-Knijnenburg,[1] Peter Blazso,[2] Dagmar Berghuis,[3] Jolan Walter,[4] Lisa Ott de Bruin,[3] Arjan C Lankester,[3] Mirjam van der Burg[1]

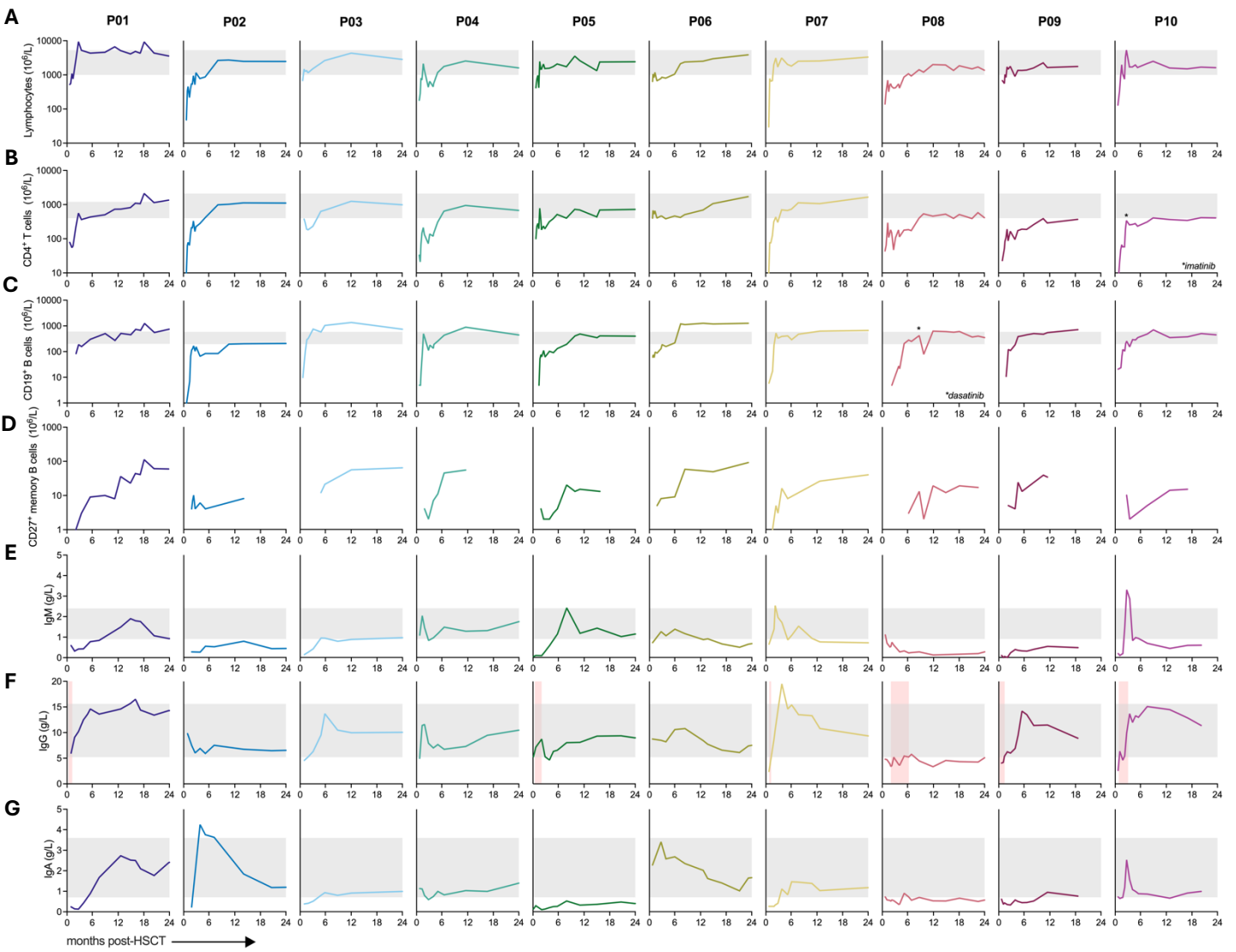
Supplemental Figure Legends

Supplemental Figure 1. Post-HSCT cellular reconstitution kinetics. Cell counts of **A** total lymphocytes, **B** $CD4^+$ T cells, **C** $CD19^+$ B cells were collected at regular intervals during routine follow-up in the first 24 months post-HSCT. Grey bars indicate reference values. **D** CD27-positive memory B-cell counts. Serum titers of **E** *IgM*, **F** *IgG* and **G** *IgA* collected during routine follow-up, with grey horizontal bars as reference values and vertical red bars indicating IVIG administration.

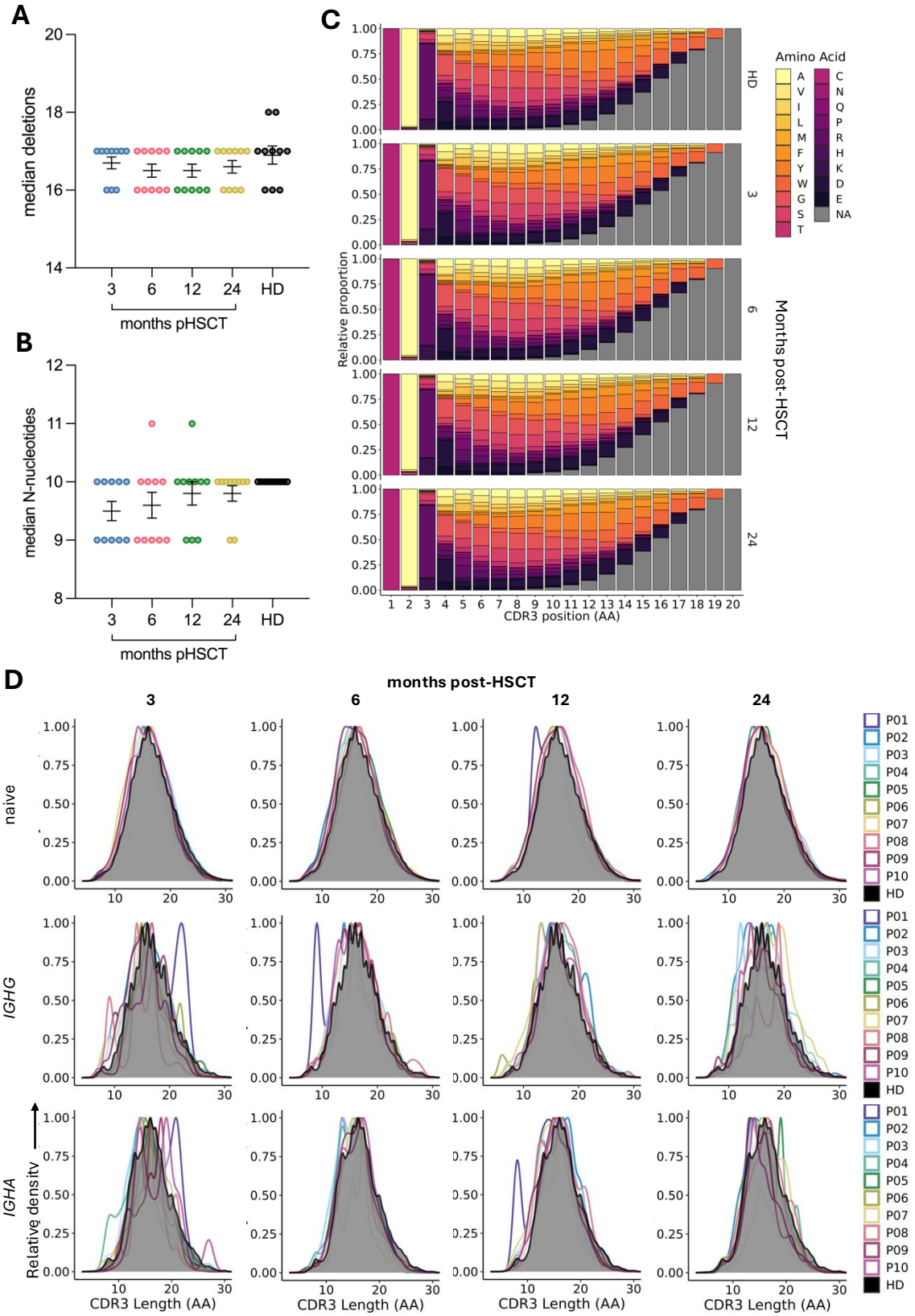
Supplemental Figure 2 CDR3 characteristics of the *IGHM* naive repertoire. Average **A** deletions and **B** N-additions of the naïve BCR repertoire at 3, 6, 12 and 24 months compared to healthy donors (HD) based on unique sequences with clonotype definition Top V, Top D, Top J and CDR3 at nucleotide level. **C** Relative proportion of amino acid use across each position of the CDR3 of all HDs and at 3, 6, 12 and 24 months of two representative patients of unique sequences. **D** Distribution of CDR3 length in *IGHM* naïve, *IGHG* and *IGHA* of individual patients post-HSCT at 3, 6, 12 and 24 months compared to the average of 10 HDs.

Supplemental Figure 3. Clonal Overlap Between Post-HSCT Timepoints. A Alluvial plots of top 50 BCR clones based on clonotype definition Top V, Top D, Top J gene and the CDR3 at nucleotide level and **B** Circos of P03-10 showing longitudinal clonal overlap dynamics of *naïve*, *IGHG* and *IGHA.* Circos only depict clones overlapping between timepoints.

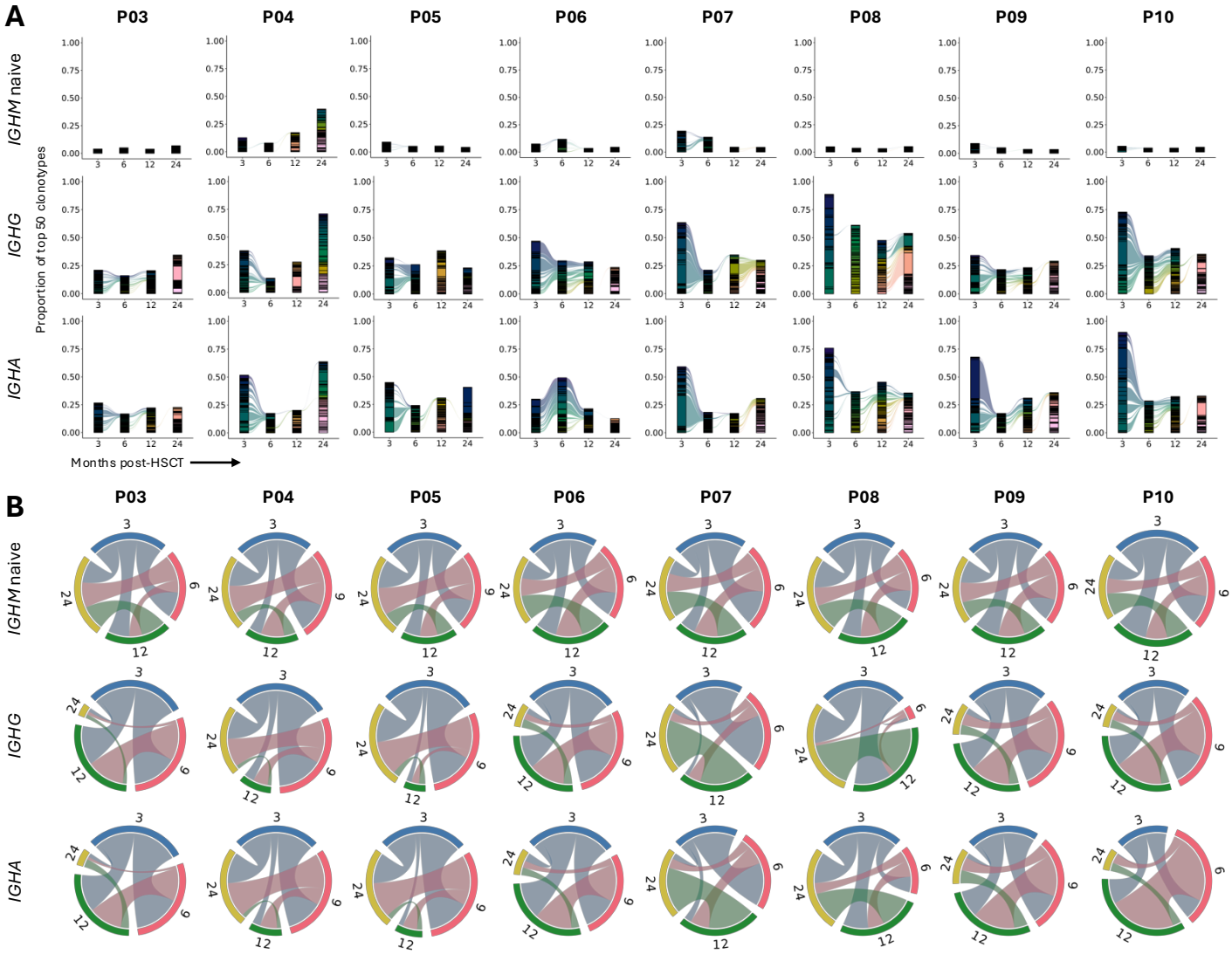
Supplemental Figures


A
P01
P02
P03
P04
P05
P06
P07
P08
P09
P10
Lymphocytes (10⁶/L)
B
CD4⁺ T cells (10⁶/L)
*imatinib
C
CD19⁺ B cells (10⁶/L)
*dasatinib
D
CD27⁺ memory B cells (10⁶/L)
E
IgM (g/L)
F
IgG (g/L)
G
IgA (g/L)
months post-HSCT

S1


A
median deletions
months pHSCT
B
median N-nucleotides
months pHSCT
C
Relative proportion
CDR3 position (AA)
Amino Acid
Months post-HSCT
D
months post-HSCT
naive
IGHG
IGHA
Relative density
CDR3 Length (AA)

S2


A
P03
P04
P05
P06
P07
P08
P09
P10
IGHM naive
IGHG
IGHA
Proportion of top 50 clonotypes
Months post-HSCT
B
P03
P04
P05
P06
P07
P08
P09
P10
IGHM naive
IGHG
IGHA

S3