HHS Public Access

Author manuscript

Leukemia. Author manuscript; available in PMC 2023 July 01.

Published in final edited form as:

Leukemia. 2023 January ; 37(1): 202–211. doi:10.1038/s41375-022-01723-w.

Molecular landscape of immune pressure and escape in aplastic anemia

Simona Pagliuca[1,2,*], **Carmelo Gurnari**[1,3,*], **Colin Hercus**[4], **Sébastien Hergalant**[5], **Niroshan Nadarajah**[6], **Adam Wahida**[6], **Laila Terkawi**[1], **Minako Mori**[1], **Weiyin Zhou**[7,8], **Valeria Visconte**[1], **Stephen Spellman**[9], **Shahinaz M. Gadalla**[7], **Caiying Zhu**[10], **Ping Zhu**[10], **Torsten Haferlach**[6], **Jaroslaw P. Maciejewski**[1]

[1]Translational Hematology and Oncology Research Program, Cleveland Clinic, Cleveland, OH

[2]Department of Clinical Hematology, CHRU Nancy, Nancy, France

[3]Department of Biomedicine and Prevention, PhD in Immunology, Molecular Medicine and Applied Biotechnology, University of Rome Tor Vergata, Rome, Italy

[4]Novocraft Technologies Sdn Bhd, Kuala Lumpur, Malaysia

[5]Inserm UMR_S1256 Nutrition-Genetics-Environmental Risk Exposure, University of Lorraine, 54500 Nancy, France

[6]Munich Leukemia Laboratory, MLL, Munich, Germany

[7]Division of Cancer Epidemiology & Genetics, NIH-NCI Clinical Genetics Branch, Rockville, Maryland, USA

[8]Cancer Genomics Research Laboratory, Frederick National Laboratory, Frederick, Maryland, USA

[*]Corresponding Author: Jaroslaw P. Maciejewski, MD, PhD, maciejj@ccf.org, Lerner Research Institute NE6, Cleveland Clinic, 9620 Carnegie Ave n building, NE6-314, Cleveland, OH.

[*]These authors equally contributed to this work

Authorship contributions:

SP designed the study, collected, analyzed and interpreted the data, developed the pipeline for the variant calling in HLA region, performed the bioinformatic and statistical analyses, and wrote the manuscript. CG performed NGS experiments, clinical and molecular data collection and participated in the analysis interpretation and critical manuscript revision. CH developed NovoHLA and the methodology for the calculation of copy number variation. SH helped in optimizing the bioanalytical HLA workflow and developing the associated bioinformatic pipeline. NN performed WGS experiments, AW, LT, MM helped in sample and data collection; SS and SG provided HLA and SNP array data for the TOAA/CIBMTR IAA cohort and gave helpful intellectual insights. WZ analyzed SNP-array data. VV helped in data collection and interpretation, gave important intellectual inputs and edited the manuscript; CZ and PZ provided single-cell RNA samples and gave meaningful intellectual inputs for the deployed analytical methodology. TH supervised the genotyping experiments in The Munich Leukemia Laboratory (MLL), helped in data interpretation and edited the manuscript; JM designed and conceptualized the study, supervised genomic experiments, provided funding and resources, interpreted the data analysis and edited the manuscript.

Conflict-of-interest disclosure:

This research was conducted in absence of any commercial or financial relationships that could be construed as a potential conflict of interest. Collaboration between Cleveland Clinic and Novocraft was based on a no-profit academic partnership.

Data sharing

All the data that support the findings of this study are available within the Article and Supplementary Files. The pipeline used for HLA mutational study has been deposited in the following repository: https://github.com/SMNPAG/HLA-mutations-IAA-PNH. NGS targeted samples can be requested to Cleveland Clinic by contacting the first and/or last author (smnpag@gmail.com, maciejj@ccf.org). Single-cell RNA sequencing raw data have been deposited in the Gene Expression Omnibus database under accession number GSE145669. Whole genome sequencing data can be requested to The Munich Leukemia Laboratory (torsten.haferlach@mll.com). Raw SNP array data from the TOAA study is available through dbGaP accession phs001710.v1.p1 (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001710.v1.p1)

[9]CIBMTR® (Center for International Blood and Marrow Transplant Research), National Marrow Donor Program/Be The Match, Minneapolis, Minnesota, United States

[10]State Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Institute of Hematology and Blood Diseases Hospital, No. 288 Nanjing Rd, Tianjin

Abstract

Idiopathic aplastic anemia (IAA) pathophysiology is dominated by autoreactivity of human leukocyte antigen (HLA)-restricted T-cells against antigens presented by hematopoietic stem and progenitor cells (HSPCs). Expansion of *PIGA* and HLA class I mutant HSPCs have been linked to immune evasion from T-cell mediated pressures. We hypothesized that, in analogy with anti-tumor immunity, the pathophysiological cascade of immune escape in IAA is initiated by immunoediting pressures and culminates with mechanisms of clonal evolution characterized by hits in immune recognition and response genes. To that end, we studied the genetic and transcriptomic make-up of the antigen presentation complexes in a large cohort of patients with IAA and paroxysmal nocturnal hemoglobinuria (PNH) by using single-cell RNA, high throughput DNA sequencing and single nucleotide polymorphism (SNP)-array platforms. At disease onset HSPCs displayed activation of selected HLA class I and II restricted mechanisms, without extensive inhibition of immune checkpoint apparatus. Using a newly implemented bioinformatic framework we found that not only class I but also class II genes were often impaired by acquisition of genetic aberrations. We also demonstrated the presence of novel somatic alterations in immune genes possibly contributing to the evasion from the autoimmune T-cells. In contrast, these hits were absent in myeloid neoplasia. These aberrations were not mutually exclusive with PNH and did not correlate with the accumulation of myeloid-driver hits. Our findings shed light on the mechanisms of immune activation and escape in IAA and define alternative modes of clonal hematopoiesis.

Introduction

Idiopathic aplastic anemia (IAA) results from the destruction of hematopoietic stem and progenitor cells (HSPCs) by autoreactive T-lymphocytes recognizing peptides derived from yet unknown antigenic proteins.[1] The reversion of the pathologic process by immunosuppressive therapy (IST) parallels laboratory observations indicating that aberrant T-cell responses, likely involving both CD4+ and CD8+ cells and their inflammatory cytokines, mediate damage to the HSPC compartment.[2,3,4,5,6] Moreover, aberrant oligoclonal T-cell expansion demonstrates that a limited number of dominant clones are responsible for the HSPCs' destruction either directly via HLA-restricted killing or through their inhibitory cytokines.[2,7,8,9,10]

Compelling evidence exists that the immunological pressure exerted by T-cells leads to selection of mutant escaping clones in HSPC pool, of which some may lead to adaptive (i.e., non-malignant) clonal recovery, while others initiate maladaptive (i.e., malignant) clonal evolution.[11,12] For instance, the emergence of immune-privileged paroxysmal nocturnal hemoglobinuria (PNH) clones may represent a tentative of adaptive recovery. This is reflected by secondary PNH evolution, observed in about 10% of IAA patients during their follow up.[13] Conversely, the occurrence of somatic leukemogenic mutations and

chromosomal defects (chiefly monosomy 7) may enable selective expansion of immune-resistant malignant clones initiating the progression to myelodysplastic syndrome (MDS) or acute myeloid leukemia (AML).[13,14,1516,17]

Loss of expression of HLA molecules on the surface of HSPCs via genetic lesions can enable establishment of immune-privileged hematopoiesis.[18,19,20,21] Similar to the mechanisms of immune escape in solid tumors and after allogeneic hematopoietic cell transplantation (HCT), these molecular defects may impair the presentation of immunogenic peptides, evading T-cell adaptive responses.[22,23,24,25] Our group and others have previously shown that clonal microdeletions or copy neutral loss of heterozygosity of HLA region are present in IAA and hypoplastic MDS as indicators of immune selection.[26,18] Somatic mutations in the coding region of certain class I HLA alleles were later found in IAA patients, generating the hypothesis that a risk allele targeted by these hits, especially in B locus, may be involved in the presentation of immunodominant peptides.[18,20,27]

In the context of immune evasion from cellular anti-tumor immune surveillance, dysfunction of HLA peptide/T-cell synapsis has been demonstrated to occur at multiple levels via class I and II HLA-restricted mechanisms resulting from a complex process of immunoediting of tumor genome.[28,29,30,31,29,32] We hypothesized that under a T-cell mediated autoimmunity able to generate immune pressure on specific hematopoietic compartments, both class I and II HLA genes, together with other immunological components involved in antigen presentation processes and regulation of T-cell effector functions may be targets of a broad immunogenomic dysfunction in IAA, featuring clonal immune escape. Consequently, we first analyzed single-cell RNA transcriptomic profiles of HSPCs in treatment-naïve patients to dissect early and subsequent signatures of immune pressure. We then performed a large-scale genomic study encompassing granular analysis of mutational/copy number events in HLA region and in immune genes, whose dysfunction could have a role in reducing the threshold of antigen presentation and T-cell activation as a mechanism of immune evasion from both class I and class II restricted processes. We further investigated the topography of HLA disruption with gold standard SNP-array methods in an independent multicenter cohort of IAA patients and explored the frequency of non-HLA immune somatic hits in a large cohort of primary myeloid diseases.

Methods

Sample and patient overview, study design and IRB approval

This is a non-interventional multi-institutional prospective study exploiting biospecimens collected from multiple cohorts of bone marrow failure (BMF) patients. In total, 875 patients were included in this study: N=197 from Cleveland Clinic Foundation (CCF), with sequencing support from Munich Leukemia Laboratory (MLL) of Munich (Germany); N=15 from the Institute of Hematology and Blood Diseases Hospital (IHBDH) of Tianjin, (China);[33] and N=663 from the Center for International Blood and Marrow Transplant Research (CIBMTR). Genomic data collected from 137 healthy controls, 733 AML and 942 MDS patients were used for validation purposes.

A total 295 samples from 212 patients were analyzed (Table 1, Figure 1, Tables S1–S2), exploring: i) single HSPCs immunologic signatures before treatment (N=15)[33] compared to healthy individuals (HC; N=4), or serially in patients responding to immunosuppression (N=4); ii) granular genetic aberrations in HLA (somatic mutations and loss of heterozygosity, N=177); iii) myeloid (N=132) and iv) non-HLA immune-related genes (N=53). Finally, cohorts of 733 AML and 942 MDS patients and 137 HC were used as a disease control population for a comparative study of somatic immune alterations and confirmation of the accuracy of HLA genotyping and mutational analysis (Figure 1). The topography of HLA disruption was further studied using SNP arrays copy number variation (CNV) of HLA loci in a separate cohort of 663 patients from the CIBMTR registry.

HLA genomic data from 137 HC were used to assess the accuracy of our typing and mutational algorithm (see Supplementary methods).

This research was conducted under the Institutional Review Board (IRB) of CCF, (IRB #5024) and local ethics committees of IHBDH, MLL and all the centers involved in CIBMTR registry. All the procedures involving human subjects were carried out under the legacy and the ethical principles of the Helsinki Declaration.

Genotyping studies

HLA sequencing, myeloid targeted and whole genome sequencing (WGS) were performed on patient samples included in the biorepository of the Translational Hematology and Oncology Research Department of CCF. MLL provided support for the WGS. For IAA patients undergoing allogeneic HCT from an unrelated donor from CIBMTR registry, DNA samples were collected prior to HCT between 1989 and 2015, and SNP-Array genotyping was performed at the Cancer Genomics Research Laboratory (CGR) in the Division of Cancer Epidemiology and Genetics (DCEG) at the National Cancer Institute (NCI). Patient characteristics and genotyping methods for this last cohort were previously reported.[34,35] See Supplementary methods for further details on DNA extraction, library preparation, SNP-array, sequencing platforms and bioanalytical pipelines.

HLA mutational platform

Details of the bioinformatic approach to investigate somatic HLA mutational status are also provided within the supplementary appendix. In brief, after obtaining a confident full 8-digit typing from both targeted and WGS samples with a new implemented algorithm (NovoHLA), paired-end reads from either targeting sequencing or WGS (in this last case after MHC region isolation from aligned and sorted BAM files and conversion to FASTQ files) were directly aligned on a per-patient HLA reference using Novoalign (Novocraft Technologies Sdn Bhd). After sorting, marking duplicates and indexing, variant calling was performed through Varscan v2.4 in tumor-only mode.[36] A further filter was applied to prioritize high quality calls present in both read mates. A minimum coverage of 30 reads (>10 reads for the variant allele) was used as a threshold for the detection in the targeted platform while a minimum coverage of 10 reads (>4 for the variant allele) was used for WGS samples. This generated a list of variants with false positive events being still potentially present among true calls. To avoid this background noise, we developed a

java tool built on the multi-alignment files provided by the HLA-IPD/IMGT database,[37] to retain only non-polymorphic calls (see supplementary appendix, Figure S1). The typing and variant calling platforms were validated in a cohort of 133 healthy subjects and 109 different cell lines, with known HLA typing (Figure S2).

Single-cell RNA sequencing and analysis

Single-cell RNA libraries from bone marrow (BM) specimens of IAA patients and healthy controls were prepared and sequenced at the Institute of Hematology and Blood Diseases Hospital of Tianjin as previously described.[33] Briefly, after a flow-based single-cell sorting, HSPCs (Lin−CD34+) and T-cells were separated. A 3′-Biased single-cell tagged reverse transcription sequencing (STRT-seq)[38,39] was deployed for the samples in study, as previously published.[33] After pooling of barcoded single cells, cleaning, amplification and tagmentation, polyA-tailing, adaptor ligation and library construction, 150bp paired-end reads were generated from Illumina HiSeq Ten X platform. After quality controls and filtering processes, high quality reads were aligned to human genome (GRCh38) with HISAT2. Seurat suite was used for data processing, normalization, analysis and visualization.[40]

Statistical analysis

Median, interquartile ranges (IQR), mean and 95%CI intervals were used where appropriate. Frequency and distribution of categorical variables were expressed as percentage. For all relevant comparisons, after testing for normal distribution, comparative analyses between two groups were performed, by two-sided paired or unpaired Student's t-tests at 95% CI. In the case of not normally distributed data, the Wilcoxon matched-pair signed rank test at 95% CI was used. Fisher's exact test or Chi-square were applied for independent group comparisons. Cox regression and proportional hazard models for the competing risk sub distributions were used to assess the impact of mutational burden on cumulative incidence of progression, in univariable setting.[41] Death at any time and HCT were used as competitive events for the cumulative incidence of progression to MDS. Generalized linear models were used to evaluate the associations between HLA and myeloid or *PIGA* burden. All statistical tests were two-sided, and a P-value <0.05 was considered statistically significant.

Results

Signatures of immune pressure in hematopoietic stem cells and multipotent progenitors in IAA

First, we sought to understand whether HSPC transcriptional phenotype in IAA before treatment could reflect the effects of immune pressure. Therefore, we investigated the immune signatures in HSPC pool of bone marrows from 15 treatment-naïve patients (12 moderate and 3 severe IAA) and compared them to those of 4 HC using single cell transcriptome analysis. This particular workflow did not include infiltrating lymphocytes. HLA and immune checkpoint expression on each HSPC subtype was first explored in a subset of HC (Supplementary results, Figure S3 and S4). After quality control steps (Figure S5), filtering and normalization of raw single cell counts, uniform manifold approximation and projection (UMAP) for dimension reduction of the merged transcripts of single

cell samples confirmed that all the cells belonging to the same differentiation subtype, even if from different individuals, clustered together, therefore enabling statistically valid comparisons (Figure 2A–B).

The initial unbiased differential analysis comparing all the HSPC clusters from treatment-naïve patients with those from HCs showed important dysregulation of genes involved in signaling mediated by interleukins, cell cycle, regulation of apoptosis and myeloid differentiation (Table S3). However, in this study we focused on target cells among players of the immune attack in AA. Therefore, to avoid the bias of different cell compositions among IAA and healthy BM, we also studied the gene expression in comparable, phenotypically defined clusters of hematopoietic stem cells (HSCs/multipotent progenitors (MPPs) (Figure 2A–C) to define gene signatures resultant from/or modulating efficacy of immune response (Table S4).

Results of this differential analysis (Table S5, Figure S6, S7) highlighted the up-regulation of many immune-related genes, involved in both class I and class II HLA antigen presentation machinery, processing and regulation (B2-microglobulin, proteasome subunits, and genes involved in endoplasmic reticulum (ER)-mediated exocytosis; Figure 2D,), and of genes involved in cytokine-signaling pathways, regulation of cell adhesion and leukocyte differentiation. No significant dysregulation was seen for components of immune checkpoint genes (including VSIR, CD70, PD-L1), except for the non-conventional negative checkpoint regulator LGALS1, which was one of the top overexpressed genes in HSC/MPP of IAA patients, Figure 2C and Figure S6–S10). We then focused specifically on HLA loci and noticed that HLA-DRB1 ($p=1.8 \times 10^{-15}$), HLA-DRA (p=0.012) and HLA-B ($p=2.2 \times 10^{-9}$) were up-regulated in this cell subset, while HLA-A ($p=1.1 \times 10^{-5}$), HLA-DQB1 (p=0.029) and HLA-DQA1 ($p=2.7 \times 10^{-6}$) were slightly down-modulated (Figure 2E). When we analyzed density (intensity) of HLA transcripts distribution in HSC/MPPs in representative patients, the overexpression of HLA-DRB1 occurred in almost all the samples. Conversely, a transcript density distribution for HLA-A, HLA-B and HLA-DQB1 loci showed decrease in some but not all patients (Figure 2F) consistent with early clonal escape. DRB1, besides being one of the most upregulated genes in HPC/MPP cells of IAA patients, appeared also to be overexpressed in all other HSPC subpopulations, from multipotent to more committed stages (Figure S8–S9). When locus-specific HLA expression in HSCs/MPPs cells before and after IST was analyzed, in 4/4 responder patients, down-modulation of at least one HLA gene was detected in post-treatment samples (Figure 2G).

Characterization of HLA somatic mutational landscape

We then studied whether immune pressure led to the acquisition of immune escape hits in HLA in target cells, deploying high throughput HLA genotyping of 177 patients (31 at diagnosis, 137 after treatment, 20 at the moment of evolution). We developed a high-fidelity bioanalytic pipeline enabling the capture of HLA point mutations, microdeletions, insertions and unbalanced CNV from sequencing data (see supplementary methods, Figure S1–S2). Overall, this analysis yielded 122 somatic HLA aberrations in 64 patients (36%, Table S6). At least one HLA aberration was found at diagnosis in 13/31 (42%), at follow-up in 47/137 (34%) and at the time of progression in 4/20 (20%) patients. PNH clones were

present in almost half of the cases (N=29, 45%) with HLA abnormalities (Figure 3A–C). HLA mutations involved all HLA loci, both in coding and non-coding regions. However, the most affected genomic locations were exon 1, exon 4, and the 5' and 3' untranslated regions (UTR, Figure S11A). Most mutated class I alleles were A*02:01 and B*14:02, while the most affected class II locus was DQB1 (Figure 3A). We confirmed a previously identified hotspot in exon 1 involving class I alleles (particularly B*14:02), characterized by a truncating mutation in codon 19 (c.19C>T, p.R7X or p.R7Efs13) (Figure 3B, Figure S11A, Table S6). This hotspot was mutated in 8 patients (12.5% of all mutant cases) at a median variant allele frequency (VAF) of 9.5% (range: 4.3–20%). The most commonly affected alleles were DQB1*06:02, DPA1*01:03, B*14:02, DQB1*06:03, A*02:01 and DRB1*04:01. Particularly, lesions in DQB1 and DRB1 alleles were more commonly losses. While no recurrent hotspots were identified in class II genes, point mutations or small indels in this group of alleles tended to occur mostly in UTR regions, possibly altering transcriptional regulation. Median VAF of all the somatic events was 5.0% (range 2–97%) and no significant difference in clonal size was noticed according to the timing of sampling (4.3% [range 2–28%] at diagnosis, 6% [range 2–97%] at follow-up, and 3.2% [range 2.0–6%] at MDS/AML evolution, Figure 3D). Allelic loss events (intended as deletions or microdeletions affecting part of specific alleles and detected through a hypersensitive genomic alignment- see supplementary methods) were more frequently seen in class II loci (Figure 3A–C), chiefly DRB1 and DQB1. Specifically, allelic loss was found in 14 patients carrying also HLA mutations, indicating a non-mutual exclusivity (at least at a patient level, Figure 3E). We did not find any differences in HLA expression or frequency of somatic HLA mutations and disease severity at diagnosis (Figure S11B). For validation purposes using a different platform (SNP-arrays), we also studied the localization of copy neutral and unbalanced genomic losses of HLA locus in a historical series of IAA patients (Transplant Outcomes in Aplastic Anemia (TOAA) project).[,35] In 102/663 patients (15%) 115 chr6 events of large CN-LOH or deletions (involving>=2Mb) covering HLA region were found with a majority of aberrations (101/115, 88%) covering large portions of 6p and mostly involving both class I and II genes (on average 39.3 Mb in size) and in a minor fraction (14/115) including interstitial portion of HLA region (average of 15.5 Mb), principally as secondary events (Figure S12). This last type of aberrations affected almost uniquely class II genes (chiefly DRB1 and DQB1 loci). Similarly to the epidemiology of HLA disruption seen in our original cohort, carriers of these CNV events frequently harbored alleles such as A*02:01, B*14:02, C*07:02, DRB1*15:01, DQB1*06:02, DPB1*04:01, DPA1*03:01 (Figure S12).

Clonal evolution to MDS and PNH and HLA aberrations

Malignant clonal progression of IAA with evolution to MDS poses a question as to whether this complication is due to or associated with the acquisition of immune escape hits. Similarly, the speed of PNH expansion may be related to the acquisition of subclonal HLA or other immune mutations in *PIGA* mutant clones. When patients with and without HLA aberrations at diagnosis and during the course of the disease were analyzed, no significant difference was found in terms of cumulative incidence (CI) of malignant progression (HR: 0.45 [95%CI 0.15–1.34], p=0.1350; Figure S11C). When we examined the myeloid mutational landscape (on 132 samples, Table S7), 15% (N=7/47) of patients

with HLA aberrations had ≥1 myeloid driver lesions as opposed to 32% (N=27/85) of cases without somatic HLA alterations (Figure 4A, p=.0331). HLA aberrant cases were generally associated with lower myeloid burden (number of mutations in myeloid genes, OR=0.44 [95%CI 0.2–0.8], p=0.0262, Figure 4B) and a trend for a lower VAF of myeloid mutations compared to HLA wild-type patients (median 33% [IQR: 20–48%] vs. 43% [IQR: 16–51%], Figure S13A). Of note is that only 3% of the patients in our cohort harbored >2 myeloid hits (Figure 4E).

Median granulocytic PNH clone size was 1.6% (IQR: 0.4–6.3%) in AA patients and 81.5% (IQR: 50.5–93.7%) in PNH cases. Although both HLA mutant and wild type groups had similar frequencies of PNH clones of size ≥1% at diagnosis, a higher burden of *PIGA* mutations (mosaicism) was observed in the first group (Figure 4B, OR: 1.55 [95%CI 1 −4], p=0.0201), underling that myeloid and *PIGA* mutational acquisition could be functionally linked to the type or strength of immune selection pressure. Furthermore, analysis of available sequential samples at the time of malignant progression showed a disappearance or decrease of HLA mutant clones present prior to evolution. For instance, in an illustrative case (UPN 227) we longitudinally tracked the disappearance of HLA hits (3 HLA mutations, with a dominant clone in B*14:02) at malignant progression 4 years later (when the HLA dominant clone decreased below the sensitivity threshold, <2%), and subsequently after transplant (Figure S13B).

Genomic dysfunction of antigen presentation and processing machinery

Next, we investigated whether T-cell mediated pressure can also, in addition to HLA, affect other immune response elements involved in the antigen presenting machinery. We thus analyzed patients' specimens for possible pathogenic somatic hits in a curated selection of immune genes (Table S4) in 53 patients through WGS (Figure 1, Table S1). While characterizing the somatic HLA landscape (Table S8), we identified 18 pathogenic or likely pathogenic somatic hits in 14 patients (26%), involving 8 genes (Figure 4F median VAF of ~10.22% [range 4.9–49.4%]). The pathogenicity of these events was predicted using 10 different bioinformatics tools (Table S9 and Supplementary methods). The most frequently mutated family of immune genes included proteasome machinery (*PSMC5*) and vesicle transportation (*KLC2*), potentially affecting HLA class I antigen assembly, together with class II HLA transcriptional regulators (i.e., *CIITA*, Figure 4F). Somatic hits in these genes were not mutually exclusive with HLA mutations (35% - 5/14 of the patients had co-occurring mutations) or *PIGA* aberrations (detected in 3 out of 14 cases with somatic hits in immune genes, Figure 4F, Figure S12C). About 26% of patients presented at least one myeloid driver mutation in absence of any PIGA, HLA or immune hit. As occurring with HLA mutations, again *PIGA* clonal burden was significantly higher comparing to the additional immune mutational events, possibly suggesting a major fitness advantage of PNH clones over those harboring immune-regulatory gene aberrations (Figure 4F). Interestingly, when we explored the presence of these aberrations in an independent cohort of primary AML (N=733) and MDS (N=942), we observed that these events were extremely rare as compared to IAA patients (p<0.0001, Figure 4G).

Discussion

In tumor biology, T-cell mediated recognition and elimination of cancer cells trigger a dynamic process of immunoediting, which may finally result in the immune escape of less immunogenic clones.[42] This process may involve transcriptional down-modulation or genetic selection characterized by acquisition of specific features of resistance to immune attack. Here, we suggest that a similar mechanism may be part of the pathophysiology of IAA, wherein immune-mediated selection of HSPCs leads to a heterogeneous assortment of more or less immune-resistant clones consistent with the theoretical notion that T-cell mediated destruction operates at the base of clonal selection.

To highlight the molecular consequences of this pressure on HSCs and progenitors, we deployed single-cell transcriptomics, which demonstrated upregulation of transcripts involved in antigen presentation including HLA DRB1. Our recent findings showing the role of specific HLA DRB1 allelic configurations in initiation of IAA[7] is in agreement with this result. A possible hypothesis is that the described upregulation of effector proteins on target cells creates the impetus for the clonal selection operating in the opposite direction e.g., acquisition of somatic class II mutations described in our study.

Previous work stressed the role of class I HLA somatic mutations as a consequence of the immunological pressure exerted by the dysfunctional cytotoxic T-cell responses.[18,20,27] The present study not only confirms this evidence, but further expands the current knowledge pinpointing the occurrence of a broader immunogenomic dysfunction of the antigen presentation machinery, encompassing most of the expressed classical HLA loci as well as genes regulating T-cell activation. Specifically, we have identified multiple HLA-A hits and a much wider assortment of HLA-B beyond B14 allele, further expanding the class I mutational repertoire.[18,19,20,21] Moreover, we demonstrated for the first time that mutational selection events can occur also in HLA class II genes. This is consistent with the pressure generated by their up-modulation at disease onset possibly related to the effects of the BM inflammatory milieu and exaggerated in carriers of specific class II alleles e.g., DRB1*15:01.[7,43] The fact that HLA mutations are spread across different loci is also in line with prior findings describing multiple T-cell clones with different HLA restriction elements.[44,45,8] The subsequent down-modulation of HLA molecules in HSPCs may be due to genetic events (mutations or losses), transcriptional downregulation (e.g. deactivation of HLA transactivators, such as CIITA), or epigenetic modifications.[46] Finally, it is also possible that clones expressing high levels of certain HLA alleles are eliminated resulting in overall decreased expression. HLA genomic and transcriptomic deficiency may share similarities with the loss/downregulation of targets after treatment with monoclonal antibodies or immunotherapies (e.g. loss of CD19 or CD20 after blinatumomab, rituximab and CAR-T,[47] or loss of CD52 after alemtuzumab-based regimens).[48,49] Higher frequency of allelic losses seen in DQB1 and DRB1 genes appears consistent with our observations of DRB1 overexpression in treatment-naïve patients, while low expression could be due to various early mechanism described above.

Mutational landscape analysis of HLA region is challenging. So far, no tools capable of simultaneous detection of mutations and unbalanced CNV in class I and II HLA genes

have been available. For the purpose of this study, we developed a new bioanalytic strategy enabling the analysis of such highly polymorphic alleles from various sequencing platforms, theoretically deployable as mutational calling algorithm also for other highly polymorphic immune loci (e.g., KIR, TAP1, MICA, MICB, etc.).

Another concordant finding of our study was the identification of other possible targets of immune escape, such as *PSMC5, CIITA, KLC2*, most of them already described in solid or hematological malignancies as associated with tumor surveillance reactions but to date not found in IAA.[50,51,52] One could speculate that such an immune clonal mosaicism involving multiple modes of immune escape, may be similar to that operative in other conditions including for instance cancer immunotherapy or allogeneic HCT. In both clinical scenarios, cells harboring somatic hits conferring immune resistance are selected under evolutionary forces. Indeed, in IAA we also noticed the co-occurrence of HLA-mutated and PNH clones rather than their mutual exclusivity. To that end, *PIGA* mutated immune-privileged HSPCs could also be reflection of such an immunoediting process, albeit *PIGA* mutant-facilitated escape is so far unique to IAA.

Our study underscores the complex relationship existing between HLA-mediated escape and clonal evolution. At the time of the preparation of this manuscript, two studies, respectively reporting on outcomes of a Japanese and a North American population of IAA and PNH patients, showed two opposite trends in terms of HLA correlations.[53,54] Hosokawa et al. demonstrated in a cohort of 633 patients that the presence of HLA lacking clones was an independent negative predictor of AML/MDS evolution.[53] However, in the second study from the NIH,[54] specific class I HLA abnormalities and the presence of B*14:02 genotype were associated with malignant progression as previously suggested by Babushok et a,[18] although the initially proposed correlation was not found in a long term updated study focused on clonal evolution.[55] The composition of our cohort precluded a conclusive determination of the role of HLA aberrations as to the risk of MDS/AML evolution, that was instead object of a separate study from our group.[13] Nevertheless, our analysis of clonal architecture and dynamics was in line with the initial Japanese findings. Indeed, in our cohort HLA aberrant clones showed a weaker transformation propensity and a lower leukemogenic mutational burden. It is also noteworthy that in a recent pan-cancer study the prevalence of class I HLA losses in solid cancers has been reported to follow a “Goldilocks” pattern, whereby tumors with the highest and lowest mutational burden (TMB) were characterized by the lowest number of HLA-related events. Conversely, cancers with an intermediate TMB harbored the highest number of HLA aberrations.[56] Translating these considerations in IAA, one could speculate that HSPCs are inherently different from intrinsically genetically unstable malignant cells. Under the immune attack, HLA mutant cells are less effectively eliminated. This behavior, in a way similar to PNH escape, may denote a feature of immune resistance uniquely associated with IAA rather than a propensity to accumulate leukemia-associated mutations.

Finally, our study sheds light onto the types of HLA aberrations, their genomic regions and affected alleles. The non-sense mutation commonly found in class I genes confirms the previously suggested selective pressure exerted by specific CD8 T-cells. The evidence of a hotspot mutation in exon 1 and the recurrence of other mutational events outside

the region encoding for the antigen binding groove (exon 2 and 3 for class I and exon 2 for class II) represents an interesting element dictating the HLA-loss escape. Indeed, dysfunctions in HLA transcripts may occur more frequently via aberrations of exon 1, altering the translation of the entire protein, or sites of TCR binding (exon 4) or deregulation of splice, transcription factor or microRNA binding sites (intronic or UTR regions), rather than via mutations of the more polymorphic antigen binding sites. This finding is in accordance with the pattern of somatic mutations of class I HLA genes found in solid tumors, highlighting how such a mechanism could encompass benign and malignant pathophysiology.[57] Intriguingly, complete allelic losses were more prevalent in class II alleles (especially in DRB1 and DQB1 loci), perhaps as a result of the pressure exerted by aberrant T-helper responses. The occurrence of class II related losses was independently validated in a large cohort of IAA patients genotyped with SNP-array platforms. In this cohort, we were able to describe an epidemiology of HLA aberrations similar to the one showed at higher resolution with high throughput sequencing experiments.

Our study expands the previous understanding of immunoediting in IAA and PNH, unveils a more complete landscape of molecular events underpinning the immune escape and provides the first comprehensive bioinformatic pipeline for the detection of somatic events in HLA locus. Our results are consistent with the theory that following an initial immune insult, the clonal architecture of residual hematopoiesis can be dominated by multiple modes of immune escape, agonistically participating to a mechanism of “semi-adaptive” clonal recovery (HLA/*PIGA* mutational acquisition) versus “maladaptive”, malignant progression.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by US National Institute of Health (NIH) grants R35 HL135795, R01HL123904, R01 380HL118281, R01 HL128425, R01 HL132071, Edward P. Evans Foundation (to J.P.M. and C.G.), Aplastic Anemia and MDS International Foundation, Italian Society of Hematology, Fondation ARC pour la Recherche sur le Cancer, Association HPN France – Aplasie medullaire and Foundation For Rare Diseases (FFRD) (to S.P.), VeloSano Pilot Award and Vera and Joseph Dresner Foundation–MDS (to V.V.). We greatly thank Lucia D’Aprano, for her IT assistance in the development of the HLA mutational pipeline. The CIBMTR is supported primarily by Public Health Service U24CA076518 from the National Cancer Institute (NCI), the National Heart, Lung and Blood Institute (NHLBI) and the National Institute of Allergy and Infectious Diseases (NIAID); HHSH250201700006C from the Health Resources and Services Administration (HRSA); and N00014-20-1-2705 and N00014-20-1-2832 from the Office of Naval Research.

We deeply thank all the reviewers who constructively contributed to improve this work.

References

1. Young NS. Current concepts in the pathophysiology and treatment of aplastic anemia. Hematology 2013; 2013: 76–81. [PubMed: 24319166]
2. Risitano AM. Oligoclonal and polyclonal CD4 and CD8 lymphocytes in aplastic anemia and paroxysmal nocturnal hemoglobinuria measured by Vbeta CDR3 spectratyping and flow cytometry. Blood 2002; 100: 178–183. [PubMed: 12070025]
3. Sloand E Intracellular interferon-gamma in circulating and marrow T cells detected by flow cytometry and the response to immunosuppressive therapy in patients with aplastic anemia. Blood 2002; 100: 1185–1191. [PubMed: 12149196]

4. Kordasti S, Marsh J, Al-Khan S, Jiang J, Smith A, Mohamedali A et al. Functional characterization of CD4+ T cells in aplastic anemia. Blood 2012; 119: 2033–2043. [PubMed: 22138514]
5. de Latour RP, Visconte V, Takaku T, Wu C, Erie AJ, Sarcon AK et al. Th17 immune responses contribute to the pathophysiology of aplastic anemia. Blood 2010; 116: 4175–4184. [PubMed: 20733158]
6. Giudice V, Selleri C. Aplastic anemia: Pathophysiology. Semin Hematol 2022; 59: 13–20. [PubMed: 35491054]
7. Pagliuca S, Gurnari C, Awada H, Kishtagari A, Kongkiatkamon S, Terkawi L et al. The Similarity of Class II HLA Genotypes Defines Patterns of Autoreactivity in Idiopathic Bone Marrow Failure Disorders. Blood 2021; : blood.2021012900.
8. Kook H, Risitano AM, Zeng W, Wlodarski M, Lottemann C, Nakamura R et al. Changes in T-cell receptor VB repertoire in aplastic anemia: effects of different immunosuppressive regimens. Blood 2002; 99: 3668–3675. [PubMed: 11986222]
9. Giudice V, Feng X, Lin Z, Hu W, Zhang F, Qiao W et al. Deep sequencing and flow cytometric characterization of expanded effector memory CD8+CD57+ T cells frequently reveals T-cell receptor Vβ oligoclonality and CDR3 homology in acquired aplastic anemia. Haematologica 2018; 103: 759–769. [PubMed: 29419434]
10. Patel BA, Giudice V, Young NS. Immunologic effects on the haematopoietic stem cell in marrow failure. Best Pract Res Clin Haematol 2021; 34: 101276. [PubMed: 34404528]
11. Yoshizato T, Dumitriu B, Hosokawa K, Makishima H, Yoshida K, Townsley D et al. Somatic Mutations and Clonal Hematopoiesis in Aplastic Anemia. New England Journal of Medicine 2015; 373: 35–47. [PubMed: 26132940]
12. Kulasekararaj AG, Jiang J, Smith AE, Mohamedali AM, Mian S, Gandhi S et al. Somatic mutations identify a subgroup of aplastic anemia patients who progress to myelodysplastic syndrome. Blood 2014; 124: 2698–2704. [PubMed: 25139356]
13. Gurnari C, Pagliuca S, Prata PH, Galimard J-E, Catto LFB, Larcher L et al. Clinical and Molecular Determinants of Clonal Evolution in Aplastic Anemia and Paroxysmal Nocturnal Hemoglobinuria. JCO 2022; : JCO.22.00710.
14. Babushok DV. A brief, but comprehensive, guide to clonal evolution in aplastic anemia. Hematology Am Soc Hematol Educ Program 2018; 2018: 457–466. [PubMed: 30504346]
15. Sun L, Babushok DV. Secondary myelodysplastic syndrome and leukemia in acquired aplastic anemia and paroxysmal nocturnal hemoglobinuria. Blood 2020; 136: 36–49. [PubMed: 32430502]
16. Tiu R, Gondek L, O'Keefe C, Maciejewski JP. Clonality of the stem cell compartment during evolution of myelodysplastic syndromes and other bone marrow failure syndromes. Leukemia 2007; 21: 1648–1657. [PubMed: 17554386]
17. Young NS, Maciejewski JP. Genetic and environmental effects in paroxysmal nocturnal hemoglobinuria: this little PIG-A goes 'Why? Why? Why?' J Clin Invest 2000; 106: 637–641. [PubMed: 10974016]
18. Babushok DV, Duke JL, Xie HM, Stanley N, Atienza J, Perdigones N et al. Somatic HLA Mutations Expose the Role of Class I-Mediated Autoimmunity in Aplastic Anemia and its Clonal Complications. Blood Adv 2017; 1: 1900–1910. [PubMed: 28971166]
19. Katagiri T, Sato-Otsubo A, Kashiwase K, Morishima S, Sato Y, Mori Y et al. Frequent loss of HLA alleles associated with copy number-neutral 6pLOH in acquired aplastic anemia. Blood 2011; 118: 6601–6609. [PubMed: 21963603]
20. Mizumaki H, Hosomichi K, Hosokawa K, Yoroidaka T, Imi T, Zaimoku Y et al. A frequent nonsense mutation in exon 1 across certain HLA-A and -B alleles in leukocytes of patients with acquired aplastic anemia. Haematologica 2020. doi:10.3324/haematol.2020.247809.
21. Imi T, Katagiri T, Hosomichi K, Zaimoku Y, Hoang Nguyen V, Nakagawa N et al. Sustained clonal hematopoiesis by HLA-lacking hematopoietic stem cells without driver mutations in aplastic anemia. Blood Adv 2018; 2: 1000–1012. [PubMed: 29720492]
22. McGranahan N, Rosenthal R, Hiley CT, Rowan AJ, Watkins TBK, Wilson GA et al. Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. Cell 2017; 171: 1259–1271.e11. [PubMed: 29107330]

23. Christopher MJ, Petti AA, Rettig MP, Miller CA, Chendamarai E, Duncavage EJ et al. Immune Escape of Relapsed AML Cells after Allogeneic Transplantation. New England Journal of Medicine 2018; 379: 2330–2341. [PubMed: 30380364]
24. Jiménez P, Cantón J, Collado A, Cabrera T, Serrano A, Real LM et al. Chromosome loss is the most frequent mechanism contributing to HLA haplotype loss in human tumors. Int J Cancer 1999; 83: 91–97. [PubMed: 10449614]
25. Vago L, Perna SK, Zanussi M, Mazzi B, Barlassina C, Stanghellini MTL et al. Loss of Mismatched HLA in Leukemia after Stem-Cell Transplantation. New England Journal of Medicine 2009; 361: 478–488. [PubMed: 19641204]
26. Afable MG, Wlodarski M, Makishima H, Shaik M, Sekeres MA, Tiu RV et al. SNP array-based karyotyping: differences and similarities between aplastic anemia and hypocellular myelodysplastic syndromes. Blood 2011; 117: 6876–6884. [PubMed: 21527527]
27. Zaimoku Y, Takamatsu H, Hosomichi K, Ozawa T, Nakagawa N, Imi T et al. Identification of an HLA class I allele closely involved in the autoantigen presentation in acquired aplastic anemia. Blood 2017; 129: 2908–2916. [PubMed: 28232583]
28. Seliger B, Maeurer MJ, Ferrone S. Antigen-processing machinery breakdown and tumor growth. Immunol Today 2000; 21: 455–464. [PubMed: 10953098]
29. Toffalori C, Zito L, Gambacorta V, Riba M, Oliveira G, Bucci G et al. Immune signature drives leukemia escape and relapse after hematopoietic cell transplantation. Nature Medicine 2019; 25: 603–611.
30. Tomasi TB, Magner WJ, Khan ANH. Epigenetic regulation of immune escape genes in cancer. Cancer Immunology, Immunotherapy 2006; 55: 1159–1184. [PubMed: 16680460]
31. Norde WJ, Maas F, Hobo W, Korman A, Quigley M, Kester MGD et al. PD-1/PD-L1 Interactions Contribute to Functional T-Cell Impairment in Patients Who Relapse with Cancer After Allogeneic Stem Cell Transplantation. Cancer Research 2011; 71: 5111–5122. [PubMed: 21659460]
32. O'Sullivan T, Saddawi-Konefka R, Vermi W, Koebel CM, Arthur C, White JM et al. Cancer immunoediting by the innate immune system in the absence of adaptive immunity. J Exp Med 2012; 209: 1869–1882. [PubMed: 22927549]
33. Zhu C, Lian Y, Wang C, Wu P, Li X, Gao Y et al. Single-cell transcriptomics dissects hematopoietic cell destruction and T-cell engagement in aplastic anemia. Blood 2021; 138: 23–33. [PubMed: 33763704]
34. Alsaggaf R, Katta S, Wang T, Hicks BD, Zhu B, Spellman SR et al. Epigenetic Aging and Hematopoietic Cell Transplantation in Patients With Severe Aplastic Anemia. Transplant Cell Ther 2021; 27: 313.e1–313.e8.
35. Savage SA, Viard M, O'hUigin C, Zhou W, Yeager M, Li SA et al. Genome-wide Association Study Identifies HLA-DPB1 as a Significant Risk Factor for Severe Aplastic Anemia. Am J Hum Genet 2020; 106: 264–271. [PubMed: 32004448]
36. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L et al. VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 2012; 22: 568–576. [PubMed: 22300766]
37. Robinson J, Barker DJ, Georgiou X, Cooper MA, Flicek P, Marsh SGE. IPD-IMGT/HLA Database. Nucleic Acids Research 2019; : gkz950.
38. Li L, Dong J, Yan L, Yong J, Liu X, Hu Y et al. Single-Cell RNA-Seq Analysis Maps Development of Human Germline Cells and Gonadal Niche Interactions. Cell Stem Cell 2017; 20: 858–873.e4. [PubMed: 28457750]
39. Gao S, Yan L, Wang R, Li J, Yong J, Zhou X et al. Tracing the temporal-spatial transcriptome landscapes of the human fetal digestive tract using single-cell RNA-sequencing. Nat Cell Biol 2018; 20: 721–734. [PubMed: 29802404]
40. Alessandrì L, Arigoni M, Calogero R. Differential Expression Analysis in Single-Cell Transcriptomics. In: Proserpio V (ed). Single Cell Methods Springer New York: New York, NY, 2019, pp 425–432.
41. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. Journal of the American Statistical Association 1999; 94: 496–509.

42. Pagliuca S, Gurnari C, Rubio MT, Visconte V, Lenz TL. Individual HLA heterogeneity and its implications for cellular immune evasion in cancer and beyond. Front Immunol 2022; 13: 944872. [PubMed: 36131910]
43. Maciejewski JP, Follmann D, Nakamura R, Saunthararajah Y, Rivera CE, Simonis T et al. Increased frequency of HLA-DR2 in patients with paroxysmal nocturnal hemoglobinuria and the PNH/aplastic anemia syndrome. Blood 2001; 98: 3513–3519. [PubMed: 11739151]
44. Zeng W, Nakao S, Takamatsu H, Yachie A, Takami A, Kondo Y et al. Characterization of T-cell repertoire of the bone marrow in immune-mediated aplastic anemia: evidence for the involvement of antigen-driven T-cell response in cyclosporine-dependent aplastic anemia. Blood 1999; 93: 3008–3016. [PubMed: 10216097]
45. Nakao S, Takamatsu H, Yachie A, Itoh T, Yamaguchi M, Ueda M et al. Establishment of a CD4+ T cell clone recognizing autologous hematopoietic progenitor cells from a patient with immune-mediated aplastic anemia. Exp Hematol 1995; 23: 433–438. [PubMed: 7720814]
46. Tsuji N, Hosokawa K, Urushihara R, Tanabe M, Takamatsu H, Ishiyama K et al. Epigenetic Loss of the HLA-DR15 Expression on Hematopoietic Stem Progenitor Cells in Patients with Acquired Aplastic Anemia Characterized By Cyclosporine Dependency: A Novel Mechanism Underlying the Immune Escape of Hematopoietic Stem Progenitor Cells. Blood 2020; 136: 23–24.
47. He MY, Kridel R. Treatment resistance in diffuse large B-cell lymphoma. Leukemia 2021; 35: 2151–2165. [PubMed: 34017074]
48. Johansson P, Klein-Hitpass L, Röth A, Möllmann M, Reinhardt HC, Dührsen U et al. Mutations in PIGA cause a CD52-/GPI-anchor-deficient phenotype complicating alemtuzumab treatment in T-cell prolymphocytic leukemia. Eur J Haematol 2020; 105: 786–796. [PubMed: 32875608]
49. Loeff FC, Falkenburg JHF, Hageman L, Huisman W, Veld SAJ, van Egmond HME et al. High Mutation Frequency of the PIGA Gene in T Cells Results in Reconstitution of GPI Anchor − /CD52 − T Cells That Can Give Early Immune Protection after Alemtuzumab-Based T Cell–Depleted Allogeneic Stem Cell Transplantation. JI 2018; 200: 2199–2208.
50. Jia D, Augert A, Kim D-W, Eastwood E, Wu N, Ibrahim AH et al. Crebbp Loss Drives Small Cell Lung Cancer and Increases Sensitivity to HDAC Inhibition. Cancer Discov 2018; 8: 1422–1437. [PubMed: 30181244]
51. Morin RD, Mendez-Lago M, Mungall AJ, Goya R, Mungall KL, Corbett RD et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. Nature 2011; 476: 298–303. [PubMed: 21796119]
52. Robinson DR, Wu Y-M, Lonigro RJ, Vats P, Cobain E, Everett J et al. Integrative clinical genomics of metastatic cancer. Nature 2017; 548: 297–303. [PubMed: 28783718]
53. Hosokawa K, Mizumaki H, Yoroidaka T, Maruyama H, Imi T, Tsuji N et al. HLA class I allele-lacking leukocytes predict rare clonal evolution to MDS/AML in patients with acquired aplastic anemia. Blood 2021; 137: 3576–3580. [PubMed: 33754630]
54. Zaimoku Y, Patel BA, Adams SD, Shalhoub R, Groarke EM, Lee AAC et al. HLA associations, somatic loss of HLA expression, and clinical outcomes in immune aplastic anemia. Blood 2021; 138: 2799–2809. [PubMed: 34724566]
55. Groarke EM, Patel BA, Shalhoub R, Gutierrez-Rodrigues F, Desai P, Leuva H et al. Predictors of clonal evolution and myeloid neoplasia following immunosuppressive therapy in severe aplastic anemia. Leukemia 2022. doi:10.1038/s41375-022-01636-8.
56. Montesion M, Murugesan K, Jin DX, Sharaf R, Sanchez N, Guria A et al. Somatic HLA Class I Loss Is a Widespread Mechanism of Immune Evasion Which Refines the Use of Tumor Mutational Burden as a Biomarker of Checkpoint Inhibitor Response. Cancer Discov 2021; 11: 282–292. [PubMed: 33127846]
57. Shukla SA, Rooney MS, Rajasagi M, Tiao G, Dixon PM, Lawrence MS et al. Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. Nature Biotechnology 2015; 33: 1152–1158.

Key points:

- In idiopathic bone marrow failure, as in tumor biology, both CD8+ and CD4+ T-cell selective pressures modify the patterns of antigen presentation and promote the emergence of immune escape phenotypes.
- Hematopoietic stem cell evasion from immune selection may dictate the trajectories of clonal hematopoiesis and evolution.

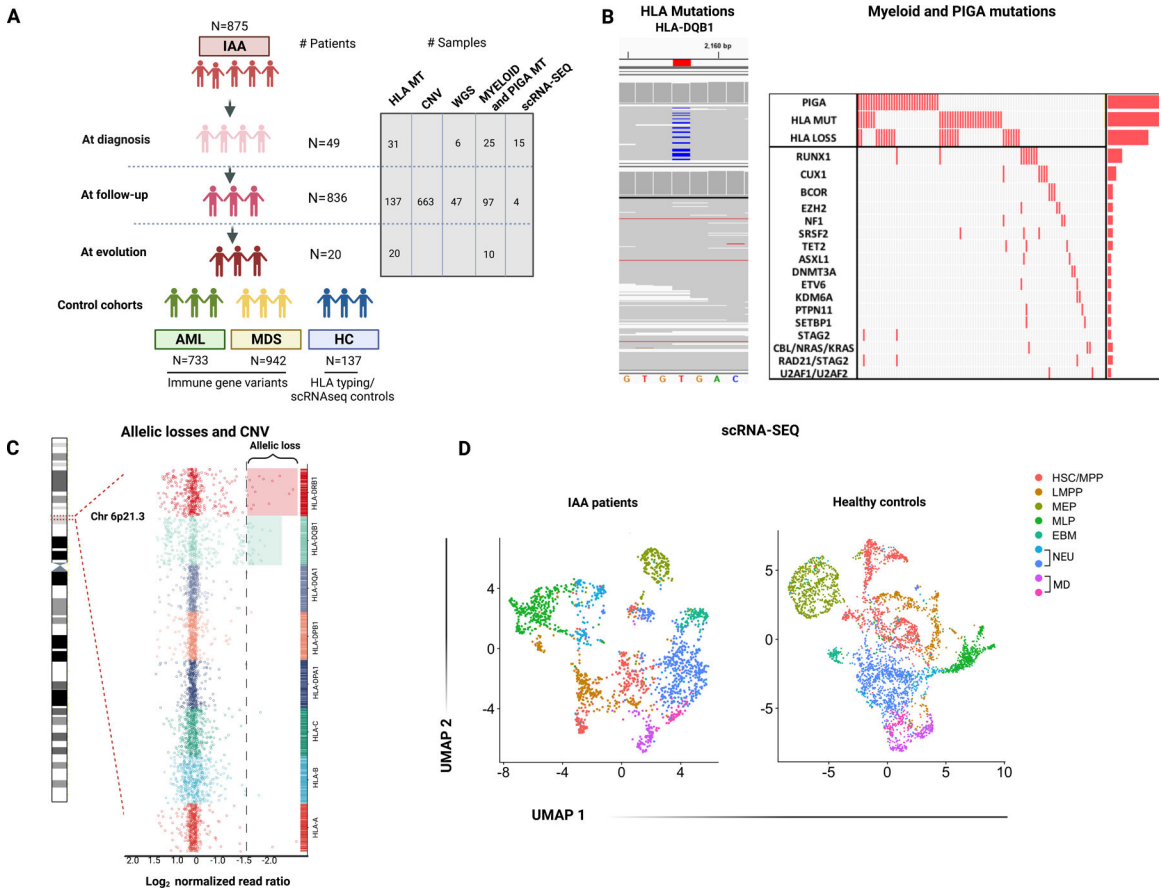

Figure 1: Study design.

A) Sample and patient overview. HC: Healthy controls; lAA: idiopathic aplastic anemia; AML: acute myeloid leukemia; scRNAseq: single-cell RNA sequencing; MT:mutations; CNV: copy number variation. WGS: whole genome sequencing. Number of patients are indicated on the left of the gray box and number of specimens within the gray box. **B)** Integrative genomic viewer (IGV) example of a HLA mutation in a representative IAA patient and example of oncoplot showing the results of the myeloid, PIGA and HLA analysis. **C)** Scatter plot describing the normalized read number ratio of each allele in IAA cohort (data from N=203 specimens analyzed with HLA targeted sequencing). Each dot represents one allele per patient and its position captures the Log2 normalized read ratio (- NRR - number of reads explaining the coverage divided by the mean of the coverage for both alleles of a given locus -see supplementary methods–). The dotted gray line indicates the threshold set up for the allelic loss (Log2 NRR < −1.5). The rug represents the density of alleles reported for a given locus. **D)** Uniform Manifold Approximation and Projection (UMAP) distribution of hematopoietic stem cells and progenitors (HSPC) from IAA treatment-naïve patients and healthy controls. The colors of the clusters indicate all the CD34+ cell populations identified according the following annotation (see [33]). Hematopoietic stem cells and multipotent progenitors (HSCs/MPPs); lymphoid-primed multipotent progenitors (LMPPs); megakaryocyte and erythroid progenitors (MEPs); multipotent lymphoid progenitors (MLPs); eosinophil, basophil, and mast cell progenitors (EBMs); Neutrophil progenitors (Neu); monocyte/dendritic progenitors (MD).

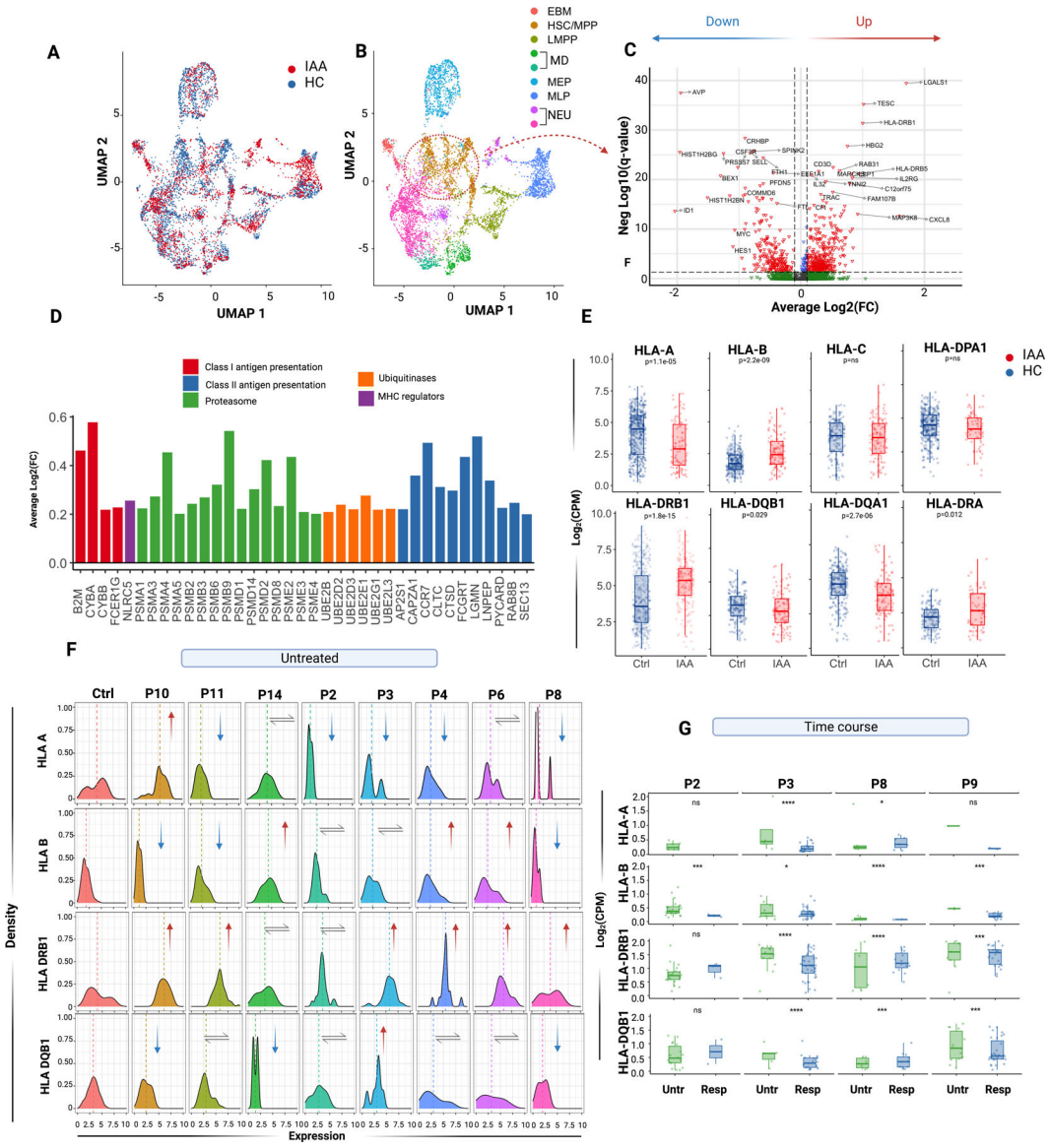

Figure 2: Single-cell RNAseq analysis of antigen presentation machinery genes in HSPCs in aplastic anemia patients and healthy controls.

A) Spatial distribution based on the Uniform Manifold Approximation and Projection (UMAP) algorithm of bone marrow hematopoietic stem cells and progenitors (HSPCs) in idiopathic aplastic anemia (IAA) patients and healthy controls (HC). **B)** Cell population annotation based on HSPC clusters (see [33]). Hematopoietic stem cells and multipotent progenitors (HSCs/MPPs); lymphoid-primed multipotent progenitors (LMPPs); megakaryocyte and erythroid progenitors (MEPs); multipotent lymphoid progenitors (MLPs); eosinophil, basophil, and mast cell progenitors (EBMs); Neutrophil progenitors (Neu); monocyte/dendritic progenitors (MD). **C)** Volcano plot showing the results of the differential analysis IAA and HC in HSC/MPP subset. X-axis depicts the single-cell average logarithmic fold change [Average log2(FC)] for each gene, while Y-axis indicates the negative logarithm of the adjusted p-value (q-value - calculated applying the Bonferroni correction). The labels highlight the genes differentially expressed at statistical significant

levels. **D)** Bar plot indicating the average log (FC) derived from the differential analysis IAA vs HC in HSC/MPP subset (only average log FC> 0.2 are displayed). **E)** Boxplots showing at single-cell level the differences in expression of HLA genes in HSC/MPP subset in controls (Blue) and IAA patients (red). **F**) Density plots indicating the density of expression of A, B, DRB1 and DQB1 loci in HSC/MPP cell population across controls and untreated patients. **G**) Boxplots showing pre- and post-treatment distribution of expression of A, B, DRB1 and DQB1 loci in representative cases for which a sample at the moment of response after treatment was available.

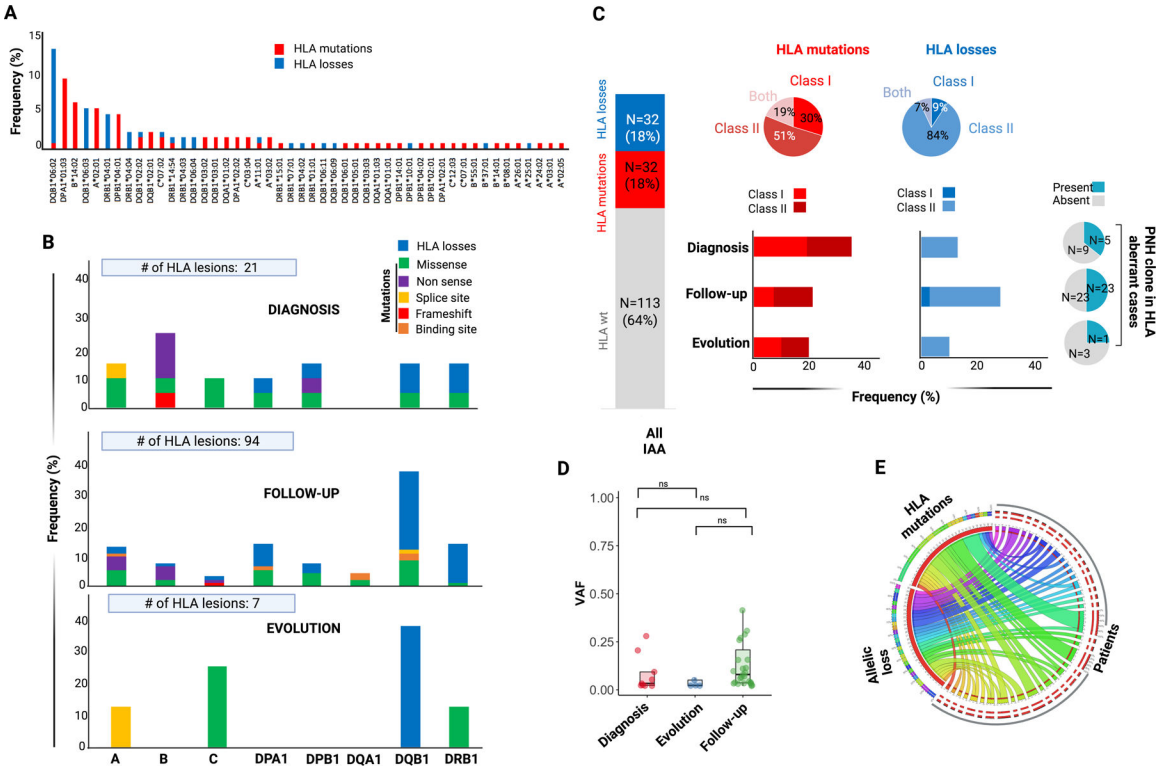

Figure 3: Landscape of HLA aberrations.

A) Barplot depicting the frequency of aberrant alleles (harboring HLA mutations -red – or losses -blue -) **B)** Stacked bar plots showing the distribution of the number of aberrations per locus according to their functional consequence and disease phase. **C)** Distribution of HLA mutations and losses across the disease course. On the right, pie charts capturing the distribution and number of patients presenting with a PNH clone in each group**. D)** Boxplot showing the VAF of HLA mutations (each dot represents a mutated sample). P-value assessed with wilcoxon rank sum test**. E)** Circle plot showing the number of HLA mutations and allelic loss per patient and their co-occurrence. This mode of visualization has been produced by means of Circos table viewer (http://mkweb.bcgsc.ca/tableviewer/visualize/). **Abbreviations:** Fup: Fullow-up; PNH: paroxysmal nocturnal hemoglobinuria; ns: non-significant p-value; VAF: variant allele frequency.

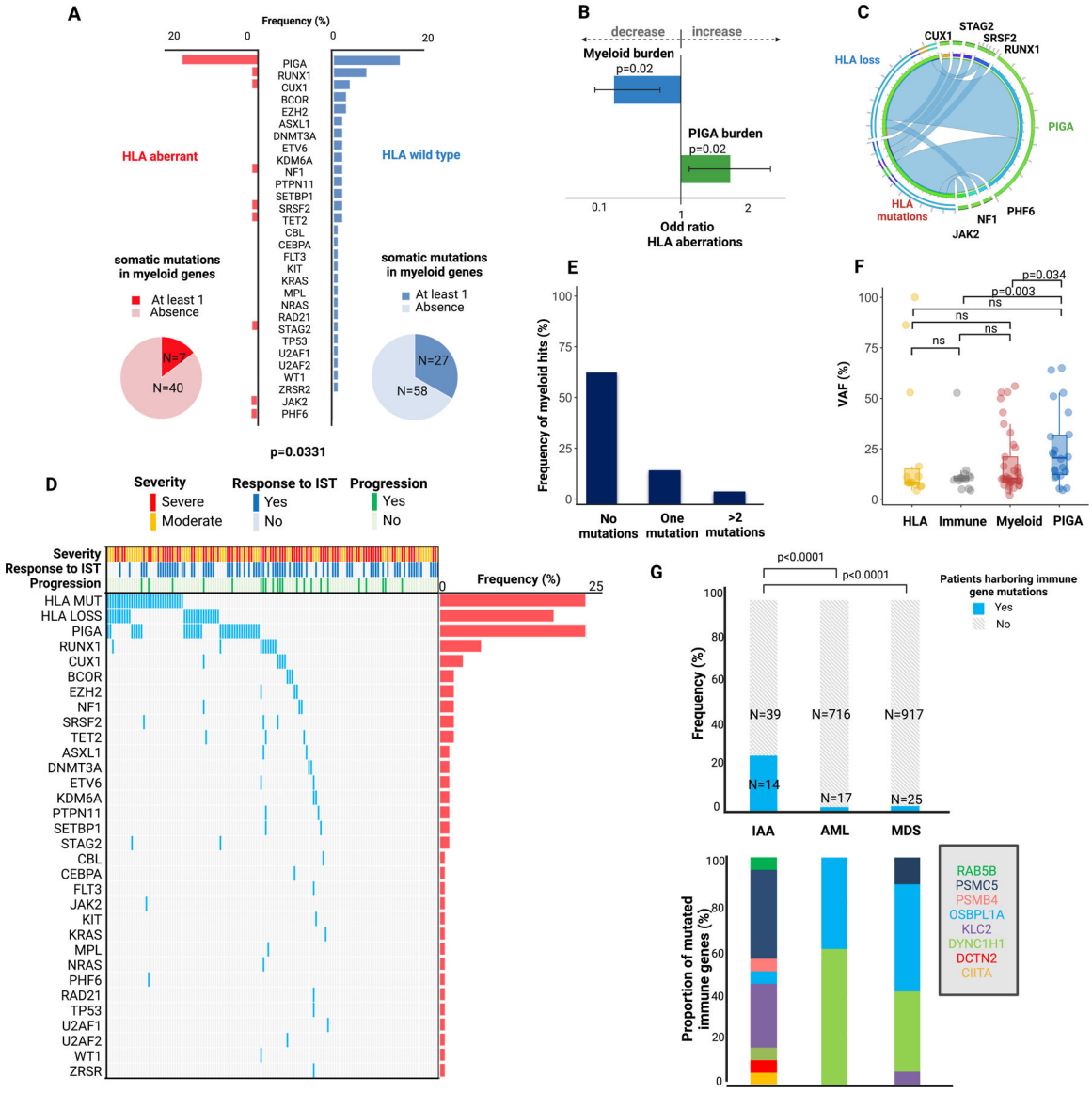

Figure 4: Landscape of HLA aberrations and myeloid molecular lesions.

A) Bar plot representing the frequency of myeloid alterations within HLA-mutated and HLA wild type group, with pie charts depicting the distribution and the total number of myeloid mutations. p-value calculated using chi-square test. **B)** Bar plot indicating the odd ratio (position of the box on x-axis) and the 95% confident interval (positions of error bars) for associations with HLA aberrations and respectively the myeloid burden (number of mutations in myeloid genes, blue box) or PIGA burden (number of mutations in PIGA gene, green box). Odd ratio calculated with a general linear regression model using Poisson distribution **C)** Circle plot showing the number and reciprocal distribution of the pattern of co-mutations within HLA aberrant samples. **D)** Waterfall plot showing the distribution of HLA and myeloid lesions (clustered for HLA mutations). Each colored spot indicates the presence of the lesion while gray spots its absence. The bar plot on the right captures the frequency of each aberrant entity. **E)** Barplots showing the rank of myeloid somatic hits

in patients undergone a targeted myeloid and PIGA gene panel sequencing (N=132). **F)** Boxplots showing the distribution of VAF in HLA, myeloid, PIGA and non-HLA immune genes. **G)** Barplots representing the distribution of mutated immune genes across different cohorts of patients identified through WGS (IAA: idiopathic aplastic anemia; AML: acute myeloid leukemia: MDS: myelodysplastic syndromes) Numbers indicate number of patients in each group (light blue: mutated; gray: non mutated).

Table 1:

Patient characteristics

Variable		N/median (%/IQR)
Cohort	**All**	212
	CCF	197
	IHBDH	15
Age (years)	median (IQR)	49.4 (28.7–64.2)
Gender	Female	119 (56%)
	Male	94 (44%)
Disease phenotype at diagnosis	IAA +/− non-hemolytic PNH clone	171 (80%)
	IAA + hemolytic PNH clone	8 (4%)
	Primary hemolytic PNH	34 (16%)
PNH clone *	Presence (%)	87 (44%)
	median size at diagnosis (IQR)	5.7 (1–43.19)
Severity IAA	Severe	106 (63%)
	Moderate	72 (37%)
Cytogenetics at diagnosis	Normal	161 (75%)
	Abnormal	2 (1%)
	No growth/Not available	49 (23%)
First-line treatment **	ATG/CSA± eltrombopag	92 (44%)
	CSA alone	21 (10%)
	Other IST	52 (25%)
	HCT	4 (2%)
	No IST/No HCT ***	39 (19%)
Anti-complement therapy	N (%)	26 (14%)

Abbreviations: IQR: Interquartile range, IAA: Idiopathic aplastic anemia ; PNH: Paroxysmal Nocturnal Hemoglobinuria; MN: Myeloid Neoplasia; IST: Immunosuppressive treatment ; HCT Hematopoietic cell transplantation; ATG: Anti-thymocyte globulin; CSA: Cyclosporine; CCF: Cleveland Clinic Foundation, Cleveland, Ohio; IHBDH: Institute of Hematology and Blood Diseases Hospital, Tianjin, China.

* Missing data on PNH clones, N=32.

** Missing data on First-line treatment, N=4.

*** This category includes also patients receiving anti-complement therapy as only treatment