Genetic mechanisms of HLA-I loss and immune escape in diffuse large B cell lymphoma

Marco Fangazio1,2,6, Erik Ladewig4,5,2,3,11, Karen Gomez1,2,6, Laura Garcia-Ibanez2, Rahul Kumar1, Julie Teruya-Feldstein1,4,5,6, Davide Rossie7,9, Ioan Filip1,5, Qiang Pan-Hammarström1, Giorgio Inghirami7, Renzo Boldorini2,1,12, German Ott4, Annette M. Staiger1,4, Björn Chapuy7, Gianluca Gaidano2, Govind Bhagat2,5,8, Katia Basso1, Raul Rabadán-Cueto2,3, Laura Pasqualucci8,13,14, and Riccardo Dalla-Favera2,5,8,11,13,14,15

1Institute for Cancer Genetics, Columbia University, New York, NY 10032; 2Department of Systems Biology, Columbia University, New York, NY 10032; 3Department of Biomedical Informatics, Columbia University, New York, NY 10032; 4Division of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY 10065; 5Laboratory of Experimental Hematology, Institute of Oncology Research, 6500 Bellinzona, Switzerland; 6Clinic of Hematology, Oncology Institute of Southern Switzerland, 6500 Bellinzona, Switzerland; 7Faculty of Biomedical Science, Università della Svizzera Italiana, 6900 Lugano, Switzerland; 8Department of Biosciences and Nutrition, Karolinska Institutet, SE14183 Huddinge, Sweden; 9Department of Pathology and Laboratory Medicine, Weill Cornell Medicine, New York, NY 10065; 10Department of Health Sciences, Division of Pathology, Amedeo Avogadro University of Eastern Piedmont, 28100 Novara, Italy; 11Department of Clinical Pathology, Robert Bosch Krankenhaus, 70376 Stuttgart, Germany; 12Institute of Clinical Pharmacology, Stuttgart and University of Tuebingen, 72074 Tuebingen, Germany; 13Department of Hematology and Oncology, University of Göttingen, 37073 Göttingen, Germany; 14Department of Translational Medicine, Division of Hematology, Amedeo Avogadro University of Eastern Piedmont, 28100 Novara, Italy; 15Department of Pathology & Cell Biology, Columbia University New York, NY 10032; 16Herbert Irving Comprehensive Cancer Center, Columbia University, New York, NY 10032; 17Program for Mathematical Genomics, Columbia University, New York, NY 10032; 18Department of Microbiology and Immunology, Columbia University, New York, NY 10027; and 19Department of Genetics and Development, Columbia University, New York, NY 10032

Contributed by Riccardo Dalla-Favera, April 20, 2021 (sent for review March 24, 2021; reviewed by Carlo M. Croce and Ralf Küppers)

Fifty percent of diffuse large B cell lymphoma (DLBCL) cases lack cell-surface expression of the class I major histocompatibility complex (MHC-I), thus escaping recognition by cytotoxic T cells. Here we show that, across B cell lymphomas, loss of MHC-I, but not MHC-II, is preferentially restricted to DLBCL. To identify the involved mechanisms, we performed whole exome and targeted HLA deep-sequencing in 74 DLBCL samples, and found somatic inactivation of B2M and the HLA-I loci in 80% (34 of 42) of MHC-INEG tumors. Furthermore, 70% (22 of 32) of MHC-1POS DLBCLs harbored monoallelic HLA-I genetic alterations (MHC-1POS/mono), indicating allele-specific inactivation. MHC-1NEG and MHC-1POS/mono cases harbored significantly higher mutational burden and inferred neoantigen load, suggesting potential coselection of HLA-I loss and sustained neoantigen production. Notably, the analysis of >500,000 individuals across different cancer types revealed common germline HLA-I homozygosity, preferentially in DLBCL. In mice, germinal-center B cells lacking HLA-I expression did not progress to lymphoma and were counterselected in the context of oncogene-driven lymphomagenesis, suggesting that additional events are needed to license immune evasion. These results suggest a multistep process of HLA-I loss in DLBCL development including both germline and somatic events, and have direct implications for the pathogenesis and immunotherapeutic targeting of this disease.

Diffuse large B cell lymphoma (DLBCL), the most common B cell lymphoma (1), is a genetically, phenotypically, and clinically heterogeneous disease that can occur de novo or upon histologic transformation of indolent lymphomas (2, 3). The updated World Health Organization classification recognizes two phenotypic subtypes of DLBCL that display common as well as subtype-specific genetic alterations (4–8) and are associated with different response to standard treatment (9–11): the germinal center (GC) B cell–like DLBCL, and the activated B cell–like (ABC) DLBCL, while ~30% of cases remain unclassified. Moreover, at least five DLBCL genetic subsets have been recently defined based on the presence of concurrent genetic alterations, which were shown to correlate with distinct prognosis (12–14).

Genetic analysis of DLBCL has also provided initial clues as to how this tumor may escape immune surveillance. The B2M gene, encoding for an invariant subunit necessary for the assembly of the class I major histocompatibility complex (MHC-I), undergoes inactivating mutations and focal deletions in up to 29% of DLBCLs, constituting one of the most common altered genes in this malignancy (12, 13, 15, 16). Furthermore, the comparison of sequential biopsies obtained from follicular lymphoma (FLs) and their transition into aggressive DLBCL (tFL) revealed the acquisition of B2M genomic aberrations with loss of B2M protein expression in 13% of tFL cases (17, 18), suggesting that disruption of the MHC-I complex plays a role in the progression from indolent disease to high-grade lymphoma.

The MHC-I complex is involved in the presentation of antigenic peptides derived from the degradation of self- and nonself-proteins, including viral- and tumor-associated antigens (19–21). The complex is a heterodimer expressed on the membrane of most nucleated cells and formed by the product of the B2M gene and one of the HLA-I heavy-chain (hHLA) molecules, with HLA-A, -B, and -C being the most common. MHC-I complexes present nonself-antigens on the cell surface, where they are recognized by the αβ receptors of cytotoxic CD8+ T lymphocytes, resulting in the destruction of the target cell (22). Consistent with these notions, this open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

PNAS 2021 Vol. 118 No. 22 e2104504118
https://doi.org/10.1073/pnas.2104504118
Published May 28, 2021.

Author contributions: M.F., L.P., and R.D.-F. designed research; M.F. and L.G.-I. performed research; E.L., K.G., J.T.-F., D.R., I.F., O.P.-H., G.I., R.B., G.O., A.M.S., B.C., G.G., and R.R. contributed new reagents/analytic tools; M.F., E.L., K.G., R.K., B.C., K.B., G.R., R.R., L.P., and R.D.-F. analyzed data; and M.F., L.P., and R.D.-F. wrote the paper.

Reviewers: C.M.C., The Ohio State University; and R.K., University of Duisburg-Essen.

Competing interest statement: R.B. is founder of Genotwin and a member of the scientific advisory board of AimedBio. R.D.-F. is a member of the scientific advisory board of Ake-mara and NeoGenomics, and a consultant of Astra Zeneca. The work reported in this paper has no relation to the current activities in these companies.
DLBCL cases carrying biallelic genetic inactivation of B2M lack cell surface expression of the MHC-I protein (15). However, the fraction of cases that fail to express surface MHC-I (43 to 75% of de novo DLBCL) is significantly higher than what can be explained by the presence of B2M genetic lesions, suggesting the existence of additional (genetic or epigenetic) mechanisms of inactivation (15, 16). Indeed, HLA-I gene mutations have been observed in large DLBCL whole-exome sequencing (WES) studies, and recent work reported that ∼70% of EZH2 mutated cases are negative for MHC-I expression, which could be restored by the use of EZH2 inhibitors in an Ezh2 mutant mouse model (15, 16). Nonetheless, the mechanisms leading to MHC-I loss remain unknown for the majority of cases.

In this study, we addressed the role and specificity of MHC-I loss among B cell malignancies, the genetic mechanisms underlying MHC-I loss in DLBCL, and the genetics of MHC-I-positive (MHC-IPOS) tumors. Additionally, we investigated the contribution of MHC-I loss to the neoplastic transformation of GC B cells in vivo, alone, or in combination with BCL6 oncogene activation.

Results

Loss of HLA-I Protein Expression Is Preferentially Associated with DLBCL among Mature B Cell Malignancies. In order to investigate whether the loss of surface MHC-I is a common phenomenon across B cell malignancies, we performed immunohistochemistry analysis with antibodies against B2M and HLA-I in a multinstitutional panel of 657 lymphoma biopsies obtained at diagnosis. This panel was representative of the most common types of mature B cell lymphoma, comprising 422 DLBCL, 25 tFL, 43 Burkitt lymphomas (BL), 54 FL, 38 mantle cell lymphomas (MCL), 39 marginal zone lymphomas (MZL), and 36 chronic lymphocytic leukemia/small lymphocytic lymphomas (CLL/SLL).

In line with previous reports, 46.2% (195 of 422) of de novo DLBCL cases and 40.0% (10 of 25) of DLBCLs derived from FL transformation were MHC-I cell surface-negative due to the complete lack of HLA-I protein expression or to aberrant cytoplasmic localization, confirming the high incidence of MHC-I loss in this disease (Fig. 1A). MHC-I-negative (MHC-INEG) DLBCLs included both GC B cell–like (21 of 33; 66.6%) and ABC-like (12 of 33; 36.4%) cases, and analogous distribution was observed in an independent series (GC B cell–like: 86 of 127; 67.7%; ABC-like: 41 of 127; 32.3%) (16). In contrast, loss of MHC-I expression was significantly less common in BL (n = 12 of 43 cases, 27.9%, P = 0.024) and FL (11 of 54, 20.4%, P < 0.001), and virtually absent in MCL (1 of 38 cases, 2.6%, P < 0.001), CLL/SLL (1 of 36, 2.8%, P < 0.001), and MZL (0 of 39, 0.0%, P < 0.001) (Fig. 1A). Within each sample, the pattern of MHC-I expression was highly uniform across the tumor cell population, with >70% of tumor cells displaying B2M and HLA-I membrane staining in cases scored as MHC-IPOS, and >90% of tumor cells lacking expression of these two proteins in cases scored as MHC-INEG.

The preferential association of MHC-I loss with DLBCL and tFL was not paralleled by the loss of MHC-II expression, observed in 30 to 50% of cases independent of lymphoma entity (81 of 140 AB

![Fig. 1](https://doi.org/10.1073/pnas.2104504118)
Appendix

0.001), 1 of 38 MCL (2.6%), and 9 of 43 BL (20.9%), biopsies analyzed, compared to 5 of 25 tFL (20%, not significant) and 36 CLL/SLL (36.1%, P = 0.27), 21 of 39 MZL (53.8%, P < 0.001), 12 of 25 tFL (48.0%), 22 of 43 BL (51.2%, P < 0.001), and none of the 39 MZL and 36 CLL/SLL (P < 0.001) were negative for both proteins, largely reflecting the preferential association of MHC-I loss with DLBCL and tFL (Fig. 1C).

Together, these data point to a selective role for loss of MHC-I and combined MHC-I/MHC-II in the pathogenesis of DLBCL, but not of more indolent or non-GC-derived mature B cell neoplasms.

Both B2M and HLA-I Gene Inactivation Contribute to Loss of MHC-I Expression in DLBCL. To comprehensively investigate the genetic mechanisms underlying the loss of MHC-I membrane expression in DLBCL, we analyzed a panel of 74 previously untreated DLBCL samples with matched normal DNA (n = 32 MHC-IPOS and 42 MHC-INEG) by integrating WES or whole-genome sequencing (WGS) with targeted deep sequencing of the hcHLA-I locus, performed in a subset of cases (Dataset S1). For the identification of lesions in the highly polymorphic HLA loci, we first genotyped the 74 patients by applying the Polysolver algorithm to both normal and tumor-derived DNA sequences, and then compared the tumor alleles to the corresponding germline hcHLA-I alleles (two each for HLA-A, -B, and -C). The HLALOH computational tool, which allows monitoring allele-specific changes in HLA gene copy number, was used to uncover allelic imbalances due to genetic deletion of one allele or copy-neutral loss of heterozygosity (cnLOH).

Consistent with previous reports (16), 17 of 42 MHC-INEG DLBCLs harbored biallelic (n = 11) or monoallelic (n = 6) mutations and deletions inactivating B2M (Fig. 2A and Datasets S2 and S3). Four additional cases (9.4%) showed biallelic disruption of one or more of the main hcHLA-I genes (Datasets S3–S5). Analogously, two B2M-WT DLBCL cell lines with aberrant MHC-I cytoplasmic localization were found to carry biallelic truncating mutations in both HLA-A and HLA-B (SI Appendix, Fig. S1). Biallelic genetic alterations of hcHLA-I were mutually exclusive with biallelic B2M disruption, together accounting for 35.7% (15 of 42) of MHC-IPOS DLBCLs (Fig. 2A).

In addition to tumors harboring biallelic hcHLA-I lesions, 18 cases showed somatic monoallelic loss of one or more of the main hcHLA-I genes in the absence (n = 11) or presence (n = 7) of B2M lesions, due to a variety of genetic mechanisms that included truncating mutations (n = 5 cases), heterozygous deletions (n = 3), and cnLOH/allelic imbalance (n = 10) (Fig. 2A and Datasets S3 and S5); missense mutations were not considered, as their functional impact is currently unclear (see reconstitution experiments in SI Appendix, Fig. S2).

Analysis of other genes implicated in antigen presentation through MHC-I, among which those encoding the known MHC-I transactivator NLRC5 (23) and for the transporter associated with antigen-processing complex (TAP1/TAP2), uncovered copy number losses in 9 of 42 cases (Fig. 2A). With one exception, these lesions were heterozygous and were observed both in the presence or absence of B2M and/or hcHLA-I genetic alterations, with no specific distribution.

Together, these results suggest that, in a subset of B2M-WT cases, loss of surface MHC-I expression could be explained by direct biallelic genetic disruption of the hcHLA-I genes. Overall, 76.2% (n = 32 of 42) of MHC-INEG DLBCLs harbor genetic alterations in B2M and/or hcHLA-I genes.

Common Somatic Monoallelic HLA-I Loss in MHC-IPOS Cases. To explore possible mechanisms of immune escape in DLBCL retaining cell-surface MHC-I expression, we investigated the genetics of B2M and hcHLA-I in the subset of 32 MHC-IPOS cases. As expected, biallelic inactivation of B2M or HC-I-I genes was never found in these tumors, and only three cases harbored monoallelic B2M deletions (9.4%; P = 0.002) (Fig. 2B). Conversely, HLA-I allelic imbalance, defined as the monoallelic loss of at least one hcHLA-I gene, was detected in 22 of 32 cases (68.8%), of which 5 harbored genetic deletions encompassing one or more hcHLA-I loci. 8 showed truncating mutations that are predicted to eliminate the protein antigen binding domains, and 10 were cnLOH, with or without a concurrent point mutation (Fig. 2B and Datasets S3–S5).

Thus, monoallelic hcHLA-I loss is a common event in a substantial fraction of MHC-IPOS DLBCL (hereafter referred to as MHC-IPOS/mono), raising the hypothesis that disruption of a single HLA-I allele could interfere with the presentation of specific antigens by the MHC-I complex (Discussion).

MHC-INEG and MHC-IPOS/mono DLBCLs Show Higher Mutation Load and Increased Activation-Induced Cytidine Deaminase-Mediated Ablant Somatic Hypermutation. In several cancer types, the loss of MHC-I expression has been associated with increased mutational load,
suggesting that the accumulation of DNA damage and the consequent generation of neoantigens are coselected with mechanisms allowing escape from immune recognition (24). To investigate whether MHC-I-defective DLBCLs accumulate a higher number of somatic mutations, we first assessed the nonsilent mutation load of the 74 DLBCL samples by interrogating WES and, in a subset of cases, WGS data.

When focusing on genes expressed in normal or transformed GC B cells, MHC-I^NEG DLBCLs showed a significantly higher nonsilent mutation load compared with MHC-I^POS cases carrying WT hcHLA-I alleles (MHC-I^POS/WT; average: 73.7 ± 48.0 vs. 33.4 ± 22.3; Mann–Whitney U = 94.5, P = 0.008) (Fig. 3A). Notably, a significantly higher mutational burden was also detected in MHC-I^POS/mono tumors (average 60.1 ± 29.5; Mann–Whitney U = 46.5, P = 0.010), which were in turn indistinguishable from the MHC-I^NEG group (Mann–Whitney U = 392.0, P = 0.326) (Fig. 3A).

We then examined whether the different mutational load across DLBCL subgroups could be driven in part by aberrant somatic hypermutation (ASHM), a mechanism described in over 50% of de novo DLBCL as well as in tFL (13, 18, 25). To this end, we analyzed 126 previously identified target genes (13) for the presence of variants targeting ~2 kb from the transcription initiation site (the hypermutable domain) (Materials and Methods). We found mutations displaying typical features of activation-induced cytidine deaminase (AID)-mediated ASHM in 30 of 42 (71.4%) MHC-I^NEG cases, but only 2 of 10 (20.0%) MHC-I^POS/WT samples (P = 0.004) (Fig. 3B). A twofold higher prevalence of ASHM-targeted cases was also detected among MHC-I^POS/mono DLBCL (n = 11 of 22, 50.0%), although the small size of this panel prevents the assessment of statistical significance. Moreover, both MHC-I^NEG and MHC-I^POS/mono cases carried a larger number of ASHM-mutated genes per case (average, 5.1 and 4.5 vs. 2.3 in MHC-I^POS/WT; P = 0.037 and 0.201, respectively; Mann–Whitney U test), which was paralleled by a higher number of mutations per case (average, 10.9 and 8.6 vs. 4.1; P = 0.019 and 0.109; Mann–Whitney U test) (Fig. 3 C and D).

These data indicate that loss of MHC-I is associated with higher mutational load, reflecting in part the aberrant activity of the somatic hypermutation mechanism. Notably, the similar mutational load of MHC-I^NEG and MHC-I^POS/mono DLBCL supports the hypothesis that loss of a single and possibly specific hcHLA-I allele might be sufficient to blunt the immunogenic potential of the tumor cells (Discussion).

MHC-I^NEG and MHC-I^POS/mono DLBCL Exhibit Higher Predicted Neoantigen Load. Studies conducted on epithelial cancers have shown that the overall mutational load directly correlates with the neoantigen load of a tumor (26). Tumors with higher neoantigen load may thus be under selective pressure to lose MHC-I in order to escape immune surveillance, which could explain their overrepresentation in MHC-I–defective cases. In order to test this hypothesis in DLBCL, we interrogated the WES/WGS data with three well-established algorithms for the prediction of tumor-associated neoantigens (NetMHC, NetMHCpan, and PickPocket) (Materials and Methods) (27–29), and compared the overall neoantigen load in MHC-I^NEG, MHC-I^POS/WT, and MHC-I^POS/mono cases.

Of 5,128 somatic coding mutations identified across the 74 DLBCL samples, 394 (7.7%) were categorized as predicted tumor neoantigens (pTNA) upon filtering for affinity (≤500 nM), mutant affinity specificity (mutant affinity > WT affinity), neoantigen size (9- to 10mers), expression in DLBCL and normal B cells, and nonhomology to human peptides (Materials and Methods and SI Materials and Methods). We then examined whether the different mutational load across DLBCL subgroups could be driven in part by aberrant somatic hypermutation (ASHM), a mechanism described in over 50% of de novo DLBCL as well as in tFL (13, 18, 25). To this end, we analyzed 126 previously identified target genes (13) for the presence of variants targeting ~2 kb from the transcription initiation site (the hypermutable domain) (Materials and Methods). We found mutations displaying typical features of activation-induced cytidine deaminase (AID)-mediated ASHM in 30 of 42 (71.4%) MHC-I^NEG cases, but only 2 of 10 (20.0%) MHC-I^POS/WT samples (P = 0.004) (Fig. 3B). A twofold higher prevalence of ASHM-targeted cases was also detected among MHC-I^POS/mono DLBCL (n = 11 of 22, 50.0%), although the small size of this panel prevents the assessment of statistical significance. Moreover, both MHC-I^NEG and MHC-I^POS/mono cases carried a larger number of ASHM-mutated genes per case (average, 5.1 and 4.5 vs. 2.3 in MHC-I^POS/WT; P = 0.037 and 0.201, respectively; Mann–Whitney U test), which was paralleled by a higher number of mutations per case (average, 10.9 and 8.6 vs. 4.1; P = 0.019 and 0.109; Mann–Whitney U test) (Fig. 3 C and D).

These data indicate that loss of MHC-I is associated with higher mutational load, reflecting in part the aberrant activity of the somatic hypermutation mechanism. Notably, the similar mutational load of MHC-I^NEG and MHC-I^POS/mono DLBCL supports the hypothesis that loss of a single and possibly specific hcHLA-I allele might be sufficient to blunt the immunogenic potential of the tumor cells (Discussion).

MHC-I^NEG and MHC-I^POS/mono DLBCL Exhibit Higher Predicted Neoantigen Load. Studies conducted on epithelial cancers have shown that the overall mutational load directly correlates with the neoantigen load of a tumor (26). Tumors with higher neoantigen load may thus be under selective pressure to lose MHC-I in order to escape immune surveillance, which could explain their overrepresentation in MHC-I–defective cases. In order to test this hypothesis in DLBCL, we interrogated the WES/WGS data with three well-established algorithms for the prediction of tumor-associated neoantigens (NetMHC, NetMHCpan, and PickPocket) (Materials and Methods) (27–29), and compared the overall neoantigen load in MHC-I^NEG, MHC-I^POS/WT, and MHC-I^POS/mono cases.

Of 5,128 somatic coding mutations identified across the 74 DLBCL samples, 394 (7.7%) were categorized as predicted tumor neoantigens (pTNA) upon filtering for affinity (≤500 nM), mutant affinity specificity (mutant affinity > WT affinity), neoantigen size (9- to 10mers), expression in DLBCL and normal B cells, and nonhomology to human peptides (Materials and Methods and SI Materials and Methods).
Consequences of MHC-I Loss on Lymphomagenesis. In order to investigate the effect of MHC-I loss in the malignant transformation process in vivo, we analyzed the GC response and lymphomagenesis in mice genetically engineered to lack B2m and thus MHC-I expression specifically in GC B cells, either as a single lesion or together with oncogenic Bcl6 expression via the IṣHA-B2m allele (30), which recapitulates the cooccurrence of BCL6 translocations reported in 30 to 50% of B2M mutated human DLBCL (12, 13).

To this end, we engineered the murine ch2qE5 locus with loxP sites flanking the B2m gene exons 2 and 3 (Fig. 6a), and crossed the resulting mice with the GC-specific C1r-Cre deleter strain (31). Southern blot analysis confirmed the correct targeting of the locus (SI Appendix, Fig. S6a), and FACS analysis using antibodies against the H-2Kk haplotype documented the loss of MHC-I surface expression in >80% of B220⁺PNA⁺CD95⁺ GC B cells isolated from B2m⁺/⁺/C1r-Cre⁺ (B2m-KO [knockout]) mice 10 d after sheep red blood cell immunization (Fig. 6b). Despite the absence of surface MHC-I, the percentage of GC B cells was comparable in both single and compound B2m-KO, B2m⁺/⁺/C1r-Cre⁻ (B2m-HET), and B2m⁺/⁺/C1r-Cre⁺ (B2m-WT) littermates (Fig. 6 B and C and SI Appendix, Fig. S6 B and C). The GC structures appeared normal also in terms of dark-zone/light-zone ratio, indicating that acute loss of B2m in GC B cells is compatible with cell proliferation and survival.

When monitored over 15 mo, chronically immunized B2m-KO/ IṣHA-B2m6 mice did not show significant differences in event-free survival correlating with the genotype (SI Appendix, Fig. S6d), except for the reported increased mortality of old IṣHA-Bcl6 mice, independent of B2m gene status (30). B2m loss had no significant impact on the overall incidence of lymphoproliferative disorders driven by deregulated BCL6 expression, which were detected in 6 of 26 (23.1%) IṣHA-Bcl6/B2m-WT mice, 13 of 29 (44.8%) IṣHA- Bcl6/B2m-HET mice, and 9 of 26 (34%) IṣHA-Bcl6/B2m-KO mice, although the B2m-KO compound animals showed a higher proportion of oligoclonal, B cell lymphoproliferative diseases (LPD) (Fig. 6 D and E). Notably, however, all seven LPDs and both DLBCLs diagnosed in IṣHA- Bcl6/B2m-KO mice showed expression of surface MHC-I by FACS analysis of H-2Kk- and B2m immunofluorescence, which revealed only sparse negative cells (Fig. 6 c and SI Appendix, Fig. S6e). Thus, the expanded B cell population in these animals must have originated from cells that had escaped B2m deletion. Taken together, these data strongly indicate that B2m-null (MHC-I-null) GC B cells are counterselected over time. These observations are consistent with the role of natural killer (NK) cells in eliminating MHC-I-null cells (32, 33), and suggest the requirement for additional alterations in order to foster immune escape (Discussion).

Discussion

Loss of surface MHC-I expression has been observed as a common phenomenon across malignancies of different cellular origin, including epithelial and hematologic cancers (34, 35). The first finding of our study is that, in the context of mature B lymphoid malignancies, and in line with previous studies in Hodgkin lymphoma and primary mediastinal large B cell lymphoma (5, 7, 15, 17, 18, 36–41), down-regulation of MHC-I expression represents a recurrent and specific event in more aggressive diseases, such as DLBCL and FL, but not in MCL (1%), CLL, and MZL. Within DLBCL, MHC-I loss was observed in both GC B cell (~60%) and ABC (~40%) molecular subtypes, with some nonsignificant differences among genetically defined classes (SI Appendix, Fig. S7) (12, 13). These data suggest a specific role for escape from MHC-I-mediated immune-surveillance mechanisms in the pathogenesis of DLBCL, consistent with the notion that immune evasion is linked to the higher mutational burden of these diseases.

We found that, in addition to B2m genetic lesions, bi-allelic disruption of hcHLA-I genes represents an alternative mechanism.

Appendix, Fig. S3). Peptide binding-affinity assays of 12 representative pTNAs validated the in silico prediction, confirming the robustness of the approach (SI Appendix, Fig. S4). We found that the average number of pTNAs was significantly higher in the MHC-I-NEG cases ($n = 6$ per case; range: 0 to 23) compared to MHC-I-POS/WT DLBCL ($n = 3$ per case; range: 0 to 8; Mann-Whitney $U = 104.0, P = 0.014$). Of note, the pTNA load of MHC-I-POS/mono DLBCLs was also significantly higher compared to that of MHC-I-POS/WT cases ($6$ per case; range: 0 to 13; Mann-Whitney $U = 52.0, P = 0.019$) (Fig. 4 and Dataset S6), further supporting a mechanistic analogy in terms of immune-escape between MHC-I NEG and MHC-I POS/mono DLBCL.

DLBCL Patients Show Significantly Higher Rate of HLA-I Germline Homozygosity. The similarity between MHC-I-NEG and MHC-I-POS/mono DLBCL suggested that decreased HLA-I gene diversity may impair the ability of cells to present endogenous and exogenous antigens by lowering the repertoire of functional MHC-I molecules (Discussion). To test whether reduced HLA-I gene diversity may initiate in the germline, we evaluated the percentage of cases harboring one, two, and/or all three classic hcHLA-I (A, B, C) genes in homozygosis in a cohort of 9,623 patients with 30 different types of cancer from The Cancer Genome Atlas (TCGA). We found that 18 of 48 (38%) TCGA DLBCL patients had at least one homozygous HLA-I germline locus, representing the tumor type with the highest frequency of homozygosity across all cancers (Fig. S4 and SI Appendix, Fig. S5a). This frequency was also significantly higher than that observed in a comparable healthy population obtained from the Genotype-Tissue Expression (GTEx) database ($21\%, P = 0.0117$, binomial test). The association of germline HLA-I homozygosity with increased risk of DLBCL was confirmed in a larger panel of cancer patients from the UK BioBank cohort (in total, 502,506 individuals), where DLBCL showed as one of the tumors most strongly associated with germline HLA-I homozygosity ($26\%, 258$ of $1,004$), significantly higher than the normal rate in individuals without a cancer diagnosis in the same population ($23\%, P = 0.0236$, binomial test) (Fig. Sb and SI Appendix, Fig. S5b). Finally, we calculated the odds ratio (OR) of DLBCL diagnosis as related to the number of homozygous HLA-I genes in the UK BioBank cohort. We found that the OR of DLBCL increases with the number of homozygous HLA-I genes (1.11, 1.27, and 1.47 for one, two, and three homozygous genes) (Fig. 5c). Collectively, these observations suggest a multistep process of HLA-I loss including both germline and somatic events in DLBCL pathogenesis.

![Fig. 4. MHC-I NEG and MHC-I POS/mono DLBCL show higher neoantigen load. Number of predicted true neoantigens identified in different HLA-defined subsets of DLBCL using NETMHC, NETMHCpan, and PickNPocket. * P < 0.05.](https://doi.org/10.1073/pnas.2105040118)
to explain the loss of MHC-I expression in B2M unmutated tumors. Of note, only one of our samples showed complete simultaneous loss of all three major HLA-I alleles, suggesting a role for allele dosage, and consistent with the idea that HLA-I genes are not functionally redundant. Conversely, no clear underlying genetic cause was identified for the remaining large fraction of MHC-I\(^{-}\)NEG DLBCLs, which retained either one (43%) or both (26%) intact B2M and \(hc\)HLA-I alleles. This negative result is unlikely due to technical issues (e.g., low coverage depth) because analogous results were obtained by using WES or the more robust targeted \(hc\)HLA-I DNA deep-sequencing approach. A search for mutations in other genes implicated in MHC-I expression revealed recurrent heterozygous deletions of \(NLC3\) and \(TAPI/TAP2\), but their distribution was independent of surface MHC-I expression and of genetic alterations in B2M and \(hc\)HLA-I (Fig. 2) (16). Thus, additional nongenetic mechanisms of (allele-specific) repression, such as epigenetic silencing/DNA hypermethylation or signals from the tumor microenvironment, are likely responsible for the down-regulation of MHC-I expression in this malignancy, as recently reported in solid tumors (42). Among these mechanisms, \(EZH2\) activating mutations were found significantly associated with lack of MHC-I expression, consistent with the increased levels of the repressive H3K27me3 mark observed at the promoter of \(Nlc5\) upon overexpression of \(Ezh2\) in mouse B cells (16). Indeed, all three \(EZH2\) mutated DLBCL samples in our study were MHC-I\(^{-}\)NEG, although two of them concurrently harbored monoallelic mutations in B2M and \(hc\)HLA-I. Thus, future investigations based on comprehensive genetic and epigenetic analysis will be necessary to clarify these issues.

A notable finding of this study was the identification of monoallelic \(hc\)HLA-I inactivation (due to genetic lesions or \(cn\)LOH) in as many as 69% of MHC-I\(^{POS}\) DLBCL, where these lesions are predicted to cause allele-specific loss of expression. These findings suggest that HLA-I LOH may represent a pervasive mechanism of immune evasion also in MHC-I\(^{POS}\) DLBCL, analogous to what was observed at lower frequency in lung cancer (15, 43). One explanation for this result could be that the loss of specific HLA haplotypes bearing the highest affinity for relevant tumor (neo)antigens is sufficient to escape T cell recognition. The first notable feature of such a model is that these cells would remain MHC-I\(^{POS}\) and therefore evade NK cell-mediated attack otherwise triggered by the loss of MHC-I inhibitory signals (34). Additionally, this model would suggest that attempts to epigenetically reactivate silent HLA-I alleles (44) may not lead to the restoration of tumor immunogenicity, since the most relevant neoantigen presentation may have already been lost via genetic inactivation of the corresponding presenting allele. A confirmation of these notions requires the combined analysis of the genetic HLA-I\(^{-}\) makeup and the neoantigen allele-specific load of the tumor vis à vis the T cell receptor specificities of autologous T cells.

Our analysis revealed an increased risk of DLBCL in individuals with germline homozygosity for the HLA-I genes (45). The potential relevance of this finding is further emphasized by the significantly higher frequency of DLBCL versus other tumor types, suggesting a specific role in this disease. Consistent with the model suggested above for somatic alterations, these observations lend support to a model of multistep restriction of HLA alleles that would start at the germline level and then progressively reduce the repertoire of antigen-presenting molecules during DLBCL pathogenesis via somatic gene alterations involving HLA-I genes. The specific predisposition for this cancer type could be explained by its intrinsic mutagenic feature due to AID-mediated ASHM (25), which would require progressively increasing protection from immune recognition.

Similar to results in lung cancer (24), MHC-I\(^{-}\)NEG and MHC-I\(^{POS\text{mono}}\) DLBCLs were associated with significantly higher load of non-silent mutations compared to MHC-I\(^{POS\text{WT}}\) lymphomas, paralleled by an increased number of predicted neoantigens (Fig. 5). These data suggest that the aberrant activity of AID could be a major contributor to the selection of cells capable of evading immune
recognition stimulated by increased neoantigen load. Supporting this concept, both ASHM and HLA-I loss are generally absent in the dominant FL clone, while they are commonly acquired upon histologic transformation (18).

In vivo, conditional loss of \textit{B2m} was not associated with increased penetrance of BCL6-driven lymphomas. However, the LPDs that develop in compound \textit{I\textsubscript{μ}HA-Bcl6/\textit{B2m}\textsuperscript{fl/fl}} mice all retained expression of MHC-I, pointing to a strong selective pressure against the \textit{B2m}-deficient GC B cell population over time. This antitumor response could be mediated by the activation of NK cell cytotoxicity that is innately elicited by the down-regulation of surface MHC-I, with consequent lack of engagement of NK cell inhibitory receptors (32, 33, 46). Thus, additional pathways may need to be disrupted in order to confer survival advantage to the precursor tumor cell. Indeed, as many as 61% of human DLBCLs concurrently lack the expression of MHC-I and CD58 (15), a ligand for the CD2 receptor required for NK cell–mediated recognition (47). Unfortunately, the role of CD58-mediated NK cell escape could not be addressed in vivo, as no mouse homolog has been identified for this gene.

In conclusion, the results herein broaden the role of (complete or haplotype-specific) MHC-I inactivation in the escape from antitumor immune surveillance during the evolution of DLBCL.
with direct clinical implications for the development of thera-
peutic approaches based on immunomodulatory molecules.

Material and Methods

Study Panel. A multiinstitutional panel of 657 formalin-fixed paraffin-embedded (FFPE) biopsies obtained at diagnosis from various B cell lymphoma types was used for the analysis of B2M and HLA-I protein expression (422 DLBCL, 25 TFL, 43 BL, 54 FL, 38 MCL, 39 MZL, and 36 CLL/SLL). A subset of 74 DLBCL samples was then selected for molecular studies based on availability of 1) paired normal DNA, and 2) FFPE sections for immunohistochemistry analysis of MHC-I expression (n = 42 MHC-IPOS and 32 MHC-IPPOD (Dataset S1). The study was approved by the Columbia University Institutional Review Board as exempt research of ano-
nymized/dé-identified existing pathological specimens, under regulatory guide-
line 45 CFR 46.101(b) (4). Thirty-nine of these cases have been used in previously published genomic analyses of DLBCL (12, 48, 49).

Immunohistochemistry and Immunofluorescence Analysis. Analysis of human and mouse samples was performed on FFPE tissue sections according to standard protocols, with the antibodies reported in SI Appendix, Supplementary Materials and Methods. Samples were independently scored by two investi-
gators, using the standard cutoff of 20% to discriminate negative vs. positive cases; however, >90% of cases that were scored as MHC-IPOS showed mem-
brane staining in >70% of tumor cells, and all cases scored as MHC-IPPOD lacked membrane signal in >90% of tumor cells. Only samples with concordant calls were included in the study.

Genomic Analyses. Purified genomic DNA from matched tumor and normal tissues was used for WES, WGS, targeted sequencing of the HLA regions, and ASHM analysis, as reported in details in SI Appendix, Supplementary Materials and Methods (see also Datasets S7–S9 for details of the sequencing performance). Somatic variants were identified with the Statistical Algorithm for Variant Identification (SAVI) (50); HLA genotyping and mutation calling were performed by applying the Polysolver computational tool, according to published methods (51), followed by Sanger-sequencing validation. The presence of copy number aberrations was determined by Sequenza (52) and of SwaI/HpaI (53) and SNP6 array or GATK analysis; LOHHLA was used to identify haplotype-specific copy number changes of the HLA locus, as described previously (24).

Assessment of Germline Homozygosity at HLA-I Loci in TCGA and UK Biobank. Germline homozygosity at the HLA-I locus was assessed in 9,623 patients with 30 different types of cancer from the TCGA project (https://www.cancer.gov/tcga), using previously published HLA-I genotypes (53). In addition, HLA-I genotypes were obtained for 488,265 individuals from the UK Biobank (https://www.ukbiobank.ac.uk); HLA alleles were imputed as the pair of alleles with maxi-
mum posterior probability from HLA*IMP:02, as described previously (54). For both cohorts, individuals whose HLA genotypes showed two identical alleles in at least one HLA-I gene were classified as being homozygous. The rate of homozygosity in the general population was estimated using normal samples from the GTEX project (https://gtexportal.org/home) (see SI Appendix, Supplementary Materials and Methods for additional details).

Targeting Vector Construction and Generation of B2m Conditional KO Mice. The B2m targeting vector was constructed by sequential subcloning of PCR-generated fragments into the pEMC1-Neo vector. An ftr-flanked neomycin-resistance cas-
sette (Neo) was cloned upstream of two loxP-site flanking exon 2 and 3 of the B2m gene (Fig. 6A). The targeting vector was electroporated in the murine em-
byronic stem (ES) cell line Sv129, and clones were selected with Neomycin (1 μg/mL). Resistant clones were screened for homologous recombination by Southern blot analysis of BamHI-digested DNA, using a 5’ external probe, and of SwaI/HpaI double-digested DNA using a 3’ external probe (Fig. 6A). Homologous recombin-
ant ES cell clones were injected into blastocysts derived from C57BL6 mice. Chimeric mice able to transmit the B2m conditional allele through the germline were crossed with a mouse expressing the flipase (Flp) recombinase in order to elimi-
nate the Neomycin-resistance cassette, and then backcrossed onto a C57BL6 background.

Mice. Deletion of B2m was directed to GC B cells by breeding the mice with the C57-Cre deleter strain (31). The offspring were bred with I-Ag7−/−B6C3 knockin mice, which carry a BCLS transgene downstream of the endogenous immuno-
globulin Iγ promoter (30), to generate compound mice. Analysis of T cell-
dependent immune responses was performed on animals intraperitoneally in-
jected with SRBC (Cocalico Biologicals) (n = 500 million per mouse in PBS) and analyzed 10 d postimmunization at 3 and 6 mo of age. Both genders were included in the experiments. Tumor watch studies were conducted for a mini-
um of 15 mo on 26 to 29 animals per genotype, which were killed upon evidence of illness or at endpoint. Mice were housed in a dedicated pathogen-
free environment, and all animal work was performed according to Institutional Animal Care and Use Committee. Genotyping was performed by PCR analysis, and the protocol is available upon request. Mice were killed according to the regulations of the Department of Comparative Medicine, Columbia University. Details on the flow cytometric and immunohistochemistry analyses of mouse cohorts are reported in SI Appendix, Supplementary Materials and Methods.

Detailed procedures and methods for the in silico neoantigen prediction, neoantigen peptides synthesis, HLA-I binding affinity assays, flow cytometric analysis, and statistical analysis can be found in SI Appendix.

Data Availability. The WES and WGS data for the 74 DLBCL patients are available in the European Genome-phenome Archive (EGA) (accession no. EGA500001005054) and the National Center for Biotechnology Information (accession no. phv000450.v3.p).

ACKNOWLEDGMENTS. We thank David Dominguez-Sola (Icahn School of Medicine) for expert advice on the immunohistochemical analysis of mouse and human tissues, and Hongyan Tang and Tongwei Mo for technical assistance with the mouse husbandry and analysis. Data from the UK Biobank were accessed under Application Number 47884. This study was supported by NIH Grants R35CA-210105 to R.D.-F. and U01CA243073 and U54CA193313 (to R.R.); the Stand Up To Cancer/Lustgarden Foundation (R.R.); and Associazione Italiana per la Ricerca sul Cancro Grant 5 x 1000 No. 21198 “Molecular bases of disease dissemination in lymphoid malignancies to optimize curative therapeutic strategies” (to G.G.). E.L. was supported by National Cancer Institute (NCI) Grant K00CA124748. K.G. was supported by the Medical Scientist Training Program Grant 2T32GM007367. The study was funded in part through the NIH/NCI Can-
cer Center Support Grant P30CA136396 and used the resources of the Cancer Center Flow Core Facility (Columbia University), the Genomics and High Throughput Screening Shared Resource (Columbia University), and the Molecular Pathology Shared Resource at Columbia University Irving Medical Center.

1. S. Swerdlow et al., WHO Classification of Tumours of Haematopoietic and Lym-
phoid Tissues (WHO, 2008).
2. L. Pasquale et al., Cancer Genetics of diffuse large B-cell lymphoma. Blood 131, 2307–2319 (2018).
3. K. Dunleavy, W. H. Wilson, Appropriate management of molecular subtypes of diff-
use large B-cell lymphoma. Oncology (Williston Park) 28, 326–334 (2014).
4. J. G. Lohr et al., Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. Proc. Natl. Acad. Sci. U.S.A. 109, 3879–3884 (2012).
5. R. D. Morin et al., Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. Nature 476, 298–303 (2011).
6. L. Pasquale et al., inactivating mutations of acetyltransferase genes in B-cell lym-
phoma. Nature 471, 189–195 (2011).
7. L. Pasquale et al., Analysis of the coding genome of diffuse large B-cell lymphoma. Nat. Genet. 43, 830–837 (2011).
8. R. D. Morin et al., Mutation and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. Blood 122, 1256–1265 (2013).
9. G. Wright et al., A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. Proc. Natl. Acad. Sci. U.S.A. 100, 9991–9996 (2003).
10. G. Lenz, L. M. Staudt, Aggressive lymphomas. N. Engl. J. Med. 362, 1417–1429 (2010).
11. A. A. Alizadeh et al., Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 408, 503–511 (2000).
12. B. Chapuy et al., Molecular subtypes of diffuse large B-cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. Nat. Med. 24, 679–690 (2018).
13. R. Schmitz et al., Genetics and pathogenesis of diffuse large B-cell lymphoma. N. Engl. J. Med. 378, 1396–1407 (2018).
14. G. W. Wright et al., A probabilistic classification tool for genetic subtypes of diffuse large B-cell lymphoma with therapeutic implications. Cancer Cell 37, 551–568.e14 (2020).
15. M. Challa-Malladi et al., Combined genetic inactivation of j2-microglobulin and CD58 reveals frequent escape from immune recognition in diffuse large B cell lymphoma. Blood 120, 728–740 (2012).
16. G. Ennishi et al., Molecular and genetic characterization of MHC deficiency identifies HZ2H as therapeutic target for enhancing immune recognition. Cancer Discov. 9, 546–563 (2019).
17. J. Okosun et al., Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. Nat. Genet. 46, 176–181 (2014).
18. L. Pasquale et al., Genetics of follicular lymphoma transformation. Cell Rep. 6, 130–140 (2014).
19. Fangazio et al. Genetic mechanisms of HLA-I loss and immune escape in diffuse large B cell lymphoma
A. R. Townsend, F. M. Gotch, J. Davey, Cytotoxic T cells recognize fragments of the influenza nucleoprotein. Cell 42, 457–467 (1985).

R. M. Zinkernagel, P. C. Doherty, Immunological surveillance against altered self components by sensitised T lymphocytes in lymphocytic choriomeningitis. Nature 251, 547–548 (1974).

R. M. Zinkernagel, P. C. Doherty, Restriction of in vitro T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. Nature 248, 701–702 (1974).

J. W. Yewdall, E. Reits, J. Neefjes, Making sense of mass destruction: Quantitating MHC class I antigen presentation. Nat. Rev. Immunol. 3, 952–961 (2003).

K. S. Kobayashi, P. J. van den Elsen, NRKCS: A key regulator of MHC class I-dependent immune responses. Nat. Rev. Immunol. 12, 813–820 (2012).

N. McGranahan et al., TRACERx Consortium, Allele-specific HLA loss and immune escape in lung cancer evolution. Cell 171, 1259–1271.e11 (2017).

L. Pasqualucci et al., Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. Nature 412, 341–346 (2001).

T. N. Schumacher, R. D. Schreiber, Neoantigens in cancer immunotherapy. Science 348, 69–74 (2015).

M. Andreatto, M. Nielsen, Gapped sequence alignment using artificial neural networks: Application to the MHC class I system. Bioinformatics 32, 511–517 (2016).

V. Jurtz et al., NetMHCpan-4.0: Improved peptide-MHC class I interaction predictions for non-HLA class I alleles. Immunity 54, 116–131.e10 (2021).

S. S. Wang et al., HLA class I and II diversity contributes to the etiologic heterogeneity of non-Hodgkin lymphoma subtypes. Cancer Res. 78, 4086–4096 (2018).

D. Pende et al., The susceptibility to natural killer cell-mediated lysis of HLA class I-positive melanomas reflects the expression of insufficient amounts of different HLA class I alleles. Eur. J. Immunol. 28, 2384–2394 (1998).

C. Binder et al., CD2 immunobiology. Front. Immunol. 11, 1090 (2020).

W. Ren et al., Genetic landscape of hepatitis B virus-associated diffuse large B-cell lymphoma. Blood 131, 2670–2681 (2018).

N. F. de Miranda et al., Exome sequencing reveals novel mutation targets in diffuse large B-cell lymphomas derived from Chinese patients. Blood 124, 2544–2553 (2014).

V. Trifonov, L. Pasqualucci, E. Tiacco, B. Falini, R. Rabanad, SAVI: A statistical algorithm for variant frequency identification. BMC Syst. Biol. 7 (suppl. 2), S2 (2013).

S. A. Shukla et al., Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. Nat. Biotechnol. 33, 1152–1158 (2015).

F. Favero et al., Sequenza: Allele-specific copy number and mutation profiles from tumor sequencing data. Ann. Oncol. 26, 64–70 (2015).

V. Thomson et al.; Cancer Genome Atlas Research Network, The immune landscape of cancer. Immunity 48, 812–830.e14 (2018).

C. Bycroft et al., The UK Biobank resource with deep phenotyping and genomic data. Nature 562, 203–209 (2018).