Frequent loss of HLA alleles associated with copy number-neutral 6pLOH in acquired aplastic anemia

Frequent loss of HLA alleles associated with copy number-neutral 6pLOH in acquired aplastic anemia

Takamasa Katagiri[1)2)*], Aiko Sato-Otsubo[3)*], Koichi Kashiwase[4)7)], Satoko Morishima[5)], Yusuke Sato[3)], Yuka Mori[3)], Motohiro Kato[3)], Masashi Sanada[3)], Yasuo Morishima[6)], Kohei Hosokawa[2)], Yumi Sasaki[2)], Shigeki Ohtake[1)], Seishi Ogawa[3)7)†] and Shinji Nakao[2)†]

On behalf of the Japan Marrow Donor Program.

[1)] Clinical Laboratory Science, Division of Health Sciences, Kanazawa University Graduate School of Medical Science, 5-11-80 Kodatsuno, Kanazawa, Ishikawa 920-0942 Japan.
[2)] Cellular Transplantation Biology, Kanazawa University Graduate School of Medical Science, 13-1 Takaramachi, Kanazawa, Ishikawa 920-8640 Japan.
[3)] Cancer Genomics Project, Graduate School of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan.
[4)] Tokyo Metropolitan Red Cross Blood center, 2-1-67 Tatsumi, Koto-ku, Tokyo 135-0053 Japan.
[5)] Department of Hematology, Fujita Health University, 1-98 Dengakugakubo, Kutsukake-cho, Toyoake, Aichi 470-1192 Japan.
[6)] Department of Hematology and Cell Therapy, Aichi Cancer Center Hospital, 1-1 Kanokoden, Chikusa-ku, Nagoya 464-8681 Japan.
[7)] Core Research for Evolutional Science and Technology, Exploratory Research for Advanced Technology, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi-shi, Saitama 332-0012, Japan.

*† T.K. and A.S.-O., and Seishi.O. and S.N. were equally contributed to this work.

Corresponding to:

Shinji Nakao (snakao@med3.m.kanazawa-u.ac.jp)
Cellular Transplantation Biology, Kanazawa University Graduate School of Medical Science, 13-1 Takaramachi, Kanazawa, Ishikawa 920-8640 Japan.
TEL: +81-762-65-2274
FAX: +81-762-65-2274

Short title: HLA allelic loss due to 6pLOH in aplastic anemia

Abstract

Idiopathic aplastic anemia (AA) is a common cause of acquired bone marrow failure. Although autoimmunity to hematopoietic progenitors is thought to be responsible for its pathogenesis, little is known about the molecular basis of this autoimmunity. Here we show that a substantial proportion of AA patients harbors clonal hematopoiesis characterized by the presence of acquired copy number-neutral loss of heterozygosity (CNN-LOH) of the 6p arms (6pLOH). The 6pLOH commonly involved the human leukocyte antigen (HLA) locus, leading to loss of one HLA haplotype. Loss of HLA-A expression from multiple lineages of leukocytes was confirmed by flow cytometry in all 6pLOH(+) cases. Surprisingly, the missing HLA-alleles in 6pLOH(+) clones were conspicuously biased to particular alleles, including HLA-A*02:01, A*02:06, A*31:01, and B*40:02. A large scale epidemiological study on the HLA alleles of patients with various hematologic diseases revealed that the four HLA alleles were overrepresented in the germline of AA patients. These findings indicate that the 6pLOH(+) hematopoiesis found in AA represents „escape" hematopoiesis from the autoimmunity, which is mediated by cytotoxic T-cells that target the relevant auto-antigens presented on hematopoietic progenitors through these class I HLAs. Our results provide a novel insight into the genetic basis of the pathogenesis of AA.

Introduction

Acquired aplastic anemia (AA) is a rare condition associated with bone marrow failure and pancytopenia[1]. A series of classical observations and experiments have unequivocally supported that the auto-immunity to hematopoietic stem / progenitor cells (HSPCs) critically underlies the pathogenesis of the bone marrow failure in the majority of AA cases. According to the widely accepted model of immune-mediated bone marrow failure, activated cytotoxic T cells (CTLs) that recognize an auto-antigen(s) presented on HSPCs through their class I HLA molecules have a major role in initiating the autoimmune reactions[2-4]. However, no definitive evidence exists that supports this model or the presence of such CTL repertoires. Moreover, little information is available about their target antigens or about the way by which they are recognized by effector T cells.

Another long-standing issue on AA is its close relationship with clonal hematopoiesis[5,6]. It was first suspected from an apparent overlap between AA and paroxysmal nocturnal hemoglobinuria (PNH)[7,8], and was also implicated by the frequent development of late clonal disorders in AA, such as myelodysplastic syndromes (MDS), PNH, or even acute myeloid leukemia (AML)[9-11]. Clonal hematopoiesis can be explicitly demonstrated by conventional clonality assays at presentation in a substantial proportion of newly diagnosed typical AA cases[12]. Although it has been expected that the inciting autoimmune insult somehow confers selective pressures on the evolution of clonal hematopoiesis[5], the exact mechanism for such immunological selection or escape is still unclear.

The objectives of this study, therefore, were to characterize the clonal nature of the hematopoiesis that is maintained even under the severe autoimmune insult in AA, and

to explore the genetic/immunologic mechanism that could underlie the pathogenesis of AA. To achieve these aims, we performed single nucleotide polymorphism (SNP) array-based analysis of genomic copy numbers and/or allelic imbalances in peripheral blood (PB) specimens obtained from 306 patients with AA. Initially, we found that AA patients frequently showed clonal/oligoclonal hematopoiesis that lost specific HLA alleles as a result of copy number-neutral loss of heterozygosity (CNN-LOH) of the 6p arms, which led us to further analyses of the contribution of 6pLOH(+) clones to residual hematopoiesis and a large scale epidemiological study on the HLA alleles that are overrepresented in AA, involving a total of 6,613 transplants registered in the Japan Marrow Donor Program (JMDP).

Methods

Subjects

PB specimens from a total of 306 patients with AA were analyzed for the presence of genetic alterations using SNP arrays (see below). The clinical characteristics of these patients are summarized in Table 1 and Supplementary Table 1. Among the 306 patients, 107 were newly diagnosed while 199 were previously treated. Ninety-six patients received allogeneic bone marrow transplantation from unrelated donors through the JMDP, and their HLA information was available from the JMDP. The other 210 were newly genotyped for HLA-A, -B, -C, -DRB1, -DQB1 and -DPB1 alleles as described elsewhere[13]. 103 patients had been treated with anti-thymocyte globulin (ATG) plus cyclosporine (CsA), CsA alone, or anabolic steroids at the time of sampling. All patients and healthy individuals provided their informed consent before

sampling. The study protocol was approved by the ethics committee of the Graduate School of Medical Science, Kanazawa University and also by that of the Graduate School of Medicine, University of Tokyo.

Analysis of genomic copy numbers and detection of 6pLOH

Genomic copy numbers, as well as allele-specific copy numbers, were analyzed by using GeneChip® 500K arrays (Affymetrix) as previously described[14,15]. Briefly, genomic DNA from AA patients and normal controls were analyzed on GeneChip® 500K arrays separately. After adjusting several biases introduced during experiments, signal ratios of the corresponding probes between test (patient) and controls were calculated across the genome to obtain genome-wide copy numbers. Genetic lesions, including copy number gains and losses, as well as CNN-LOHs, were first detected using a hidden Markov model (HMM)-based algorithm implemented in the CNAG software[14,15]. Known copy number variations were carefully excluded by referring to the Database of Genomic Variants (http://projects.tcag.ca/variation/). CNN-LOH in 6p involving the HLA locus was more specifically and sensitively detected by statistically evaluating the mean differences in allele-specific copy numbers between heterozygous SNPs on 6p (N=~1,400) that were telomeric from the 5‟end of the HLA-A locus (rs1655927) and all non-6p heterozygous SNPs (N=~10,5000) using the Mann-Whitney's U test with the R package (http://www.r-project.org/). Possible false positive findings arising from multiple testing involving the 306 samples were evaluated by maintaining the false discovery rate (FDR) under 0.01 as previously described[16], where the microarray data of 1,000 JMDP donor specimens obtained from

an ongoing whole genome association study (unpublished data) were used to calculate an empiric null distribution[17,18].

Determination of the missing HLA alleles in 6pLOH(+) clones in patients with AA

The 500K SNP data of the 1,800 JMDP donor-recipient pairs (JMDP data set), together with their HLA genotyping information, was used to generate an HLA SNP haplotype table on the GeneChip® 500K platform, which contains the consensus SNPs of the three major haplotypes (P1, P2, and P3) in Japanese subjects[18] and the SNP sequences of all observed HLA haplotypes complementary to either of P1-P3 within the JMDP set (N = 1576) (data not shown). To determine the missing HLA haplotype in each 6pLOH(+) patient, those "HLA" haplotypes were first selected from the above HLA haplotype table that were compatible with the observed HLA genotypes of that patient. Among these, a candidate haplotype was selected such that it contained the minimum number of SNPs that were incompatible with the patient's genotype. For each candidate haplotype, genomic copy numbers were inferred at the heterozygous SNPs along that haplotype using the circular binary segmentation algorithm[19,20], which divided the haplotype into one or more discrete segments with different mean copy numbers. Finally, each copy number segment was thought to be 'missing', when the alternative hypothesis $\left(\mathrm{H}a: S_i \neq \overline{S_i},\ \mathrm{for}\ ^{\forall} i\right)$ was supported against the null hypothesis $\left(\mathrm{H}_0 : S_i = \overline{S_i},\ \mathrm{for}\ ^{\forall} i\right)$ using the Wilcoxon signed rank test with a significance level of 0.05, where S_i represents the allele-specific copy number at the *i*th heterozygous SNP site within the segment of the candidate haplotype with $\overline{S_i}$ being the corresponding value for the complementary haplotype (Supplementary Figure. 1). Finally, for those

HLA types that appeared more than 8 times among 6pLOH(+) cases, their contribution to the observed allelic loss of HLA haplotypes was evaluated by multi-variate logistic regression analysis with stepwise backward selection

Flow cytometry

Heparinized PB and BM were collected from the patients at diagnosis and/or after treatment. HLA-A expression on granulocytes, monocytes, B and T cells, and BM $CD34^+$ cells was analyzed by flow cytometry using a FACS Canto II® instrument (Beckton Dickinson) with the FlowJo program (Tree Star, Inc.). The monoclonal antibodies used for this study are provided in Supplementary Table 2.

Human Androgen Receptor Assay (HUMARA)

The human androgen receptor gene was amplified from genomic DNA of 23 female patients including 3 6pLOH(+) patients, as described by Ishiyama et al.[21] with some modifications. Clonality was assessed using an "S value" as a marker of skewing in granulocytes and T lymphocytes.

Association of HLA types with AA

A total of 6,613 patients who had received allogeneic bone marrow transplantation through the JMDP between 1992 and 2008 were investigated to see whether the HLA alleles frequently missing in CNN-LOH in 6p with the development of AA could represent risk alleles for the development of AA. Thus, the frequencies of patients with each of the candidate risk alleles (HLA-A*31:01, B*40:02, A*02:01 and A*02:06) and those having none of these alleles were compared between 407 patients with AA

and those with other hematopoietic disorders (1,827 with AML, 1,606 with acute lymphocytic leukemia, 1,014 with chronic myeloid leukemia, 825 with MDS, 566 with non-Hodgkin lymphoma, and 368 with other hematopoietic neoplasms) (Supplementary Table 3) by calculating the Fisher's *P*-values in the corresponding 2 x 2 contingency tables.

Results

Genetic lesions in AA detected by SNP array analysis

After excluding known or suspected copy number variations, a total of 50 genetic lesions were identified in 46 out of the 306 (15%) PB specimens of our AA case series (Table 1 and Figure. 1). Among these by far the most conspicuous was the recurrent CNN-LOH involving the 6p arm, which was detected in 28 cases as a significant dissociation of allele-specific copy number graphs in 6p regions using a hidden Markov model (HMM)-based algorithm implemented in the CNAG software[2,14,15] (Figure. 2A and 2B). Of particular interest was that all CNN-LOH in 6p commonly affected the HLA locus, causing a haploid loss of HLA alleles and uniparental HLA expression. In some cases, the breakpoint of the 6pLOH was predicted to fall within the HLA locus (Figure. 2B). These findings strongly indicated that the HLA locus was the genetic target of these 6pLOHs. Also supporting this was the finding that in half of the cases, the dissociations in the allele-specific copy number graphs were gradually attenuated to the baseline over several mega base pair regions rather than showing a discrete breakpoint, indicating the presence of multiple 6pLOH(+) clones within a single case that had different breakpoints but still shared the same missing HLA alleles (Figure.

2C). Moreover, the 6pUPDs existing only in a minor population were more sensitively detected by statistically evaluating the size of dissociation of allele-specific copy numbers in the 6p arm. With this improved statistical test, CNN-LOH in 6p was found in a total of 40 cases (13%) (Figure. 2D and Supplementary Figure. 2), where the FDR was maintained at 0.01 to avoid too many false positive findings. In all 6pLOH(+) cases, substantial numbers of heterozygous SNP calls were retained within the affected regions, thus indicating that the CNN-LOHs in 6p were not constitutional but represented acquired genetic events only found in the affected subclones (Figure. 1). In fact, all 6pLOH(+) cases were shown to have „heterozygous" HLA alleles in high-resolution HLA typing of their PB (Table 2). Moreover, 6pLOH was not detected in the CD3-positive T cells in selected cases (cases 25 and 26, Supplementary Figure. 3). By quantitatively comparing the observed differences in allele-specific copy numbers in the 6pLOH segments with what were expected assuming 100% LOH(+) components, the 6pLOH(+) clones were estimated to account for 0.2 % to 53.9 % of the PB leukocytes (Table 2). The trend of the lower percentages of the 6pLOH(+) fraction in newly diagnosed patients compared to those in patients at remission was thought to reflect the fact that the former patients tended to have lower counts of granulocytes and monocytes, which were the predominant targets of 6pLOH (see below).

The disease status of the 40 patients at the sampling was before treatment in 16 cases, during remission for 1-16 years after therapies in 15, and before bone marrow transplantation for refractory disease in 9. All evaluable 6pLOH(+) AA cases responded to immunosuppressive therapy (IST) (23/23), whereas 101 of 126 evaluable cases with 6pLOH(−) responded ($P = 0.014$) (Table 3).

Uniparental expression of HLA-A in multi-lineage hematopoietic cells

The genetic loss of one HLA haplotype in SNP array analysis was further confirmed by expression analysis of HLA-A in PB leukocytes using flow cytometry in 19 eligible cases with 6pLOH(+), in which the HLA-A alleles were heterozygous and fresh PB samples were available. Loss of expression of one HLA-A antigen was confirmed in all 19 6pLOH(+) cases (Figure. 3A and Supplementary Figure. 4). The HLA-A missing cells in the PB were shown to have appeared shortly after the onset or before the initiation of treatments in 2 cases, and were confirmed to persist for 1 – 16 (median 6) months in 14 patients (Supplementary Table 1 and Supplementary Figure. 5). The percentage of granulocytes lacking HLA-A antigens in the 2 patients who were responsive to IST remained almost the same during the convalescent period of 2-3 months (Supplementary Figure. 6). Importantly, uniparental expression of HLA-A alleles was detected in multiple cell lineages, including granulocytes, monocytes, B cells, and to a lesser extent, in T cells. Moreover, uniparental HLA-A expression was demonstrated in bone marrow (BM) $CD34^+$ cells in 5 patients whose BM samples were available for flow cytometry. All 5 patients possessed varying proportions of BM $CD34^+$ cells (49.7 to 71.3%), which had lost the expression of one HLA-A antigen, and in each case, the missing HLA-A allele was identical to that in the PB leukocytes (Figure. 3B). The uniparental expression of HLA-A in case 13 was also observed in the $CD34^+$ compartment of the archived BM specimen obtained 2 years before analysis (Supplementary Figure. 7). Together, these findings suggested that the 6pLOH involved early HSPCs, and that the 6pLOH occurred at the level of long-term repopulating stem cells.

Clonality of the HLA-missing granulocytes

The HUMARA (Human Androgen Receptor)-based clonality assays in granulocytes were performed in 3 6pLOH(+) and 20 6pLOH(−) patients, in which all three 6pLOH(+) and 4 (20%) of the 6pLOH(−) patients showed evidence of clonality in granulocyte populations (Supplementary Figure. 8).

Missing HLA alleles in 6pLOH

Given that the HLA is the genetic target of 6pLOH in AA, the missing HLA alleles in 6pLOH are of particular interest, because in this context, they are thought to be directly involved in the presentation of the target auto-antigens to CTLs, and therefore, to be critically important in the pathogenesis of AA. We determined the missing HLA alleles in each 6pLOH(+) AA patient by the haplotype imputation of HLA alleles based on the large data of HLA haplotypes observed in the JMDP set, followed by statistical evaluation of allele-specific copy numbers along the imputed haplotypes (Figure. 4). The imputed haplotypes were confirmed in 4 cases by the family studies on the HLA. The allelic status was imputed at least partially in 39 out of the 40 6pLOH(+) cases. The imputed results were consistent with the patterns of uniparental expression of HLA-A in flow cytometry in 18 cases with 6pLOH (Table 2) (for details, see Figure. 4) except for those in case 26, in which no valid SNP haplotype around the HLA-A locus was identified and the status of HLA-A was determined by flow cytometry. The missing HLA alleles in 6pLOH(+) AA showed a conspicuous deviation to some selected HLA alleles, including HLA-A*31:01, B*40:02, C*03:04, and to a lesser extent HLA-A*02:01 and A*02:06. After the effects of linkage disequilibrium

between individual HLA alleles were taken into consideration by multi-variate analysis, 4 HLA alleles were shown to remain as the principal determinants of the missing haplotypes, HLA-A*31:01, B*40:02, A*02:01 and A*02:06 (Supplementary Table 4).

Overrepresentation of frequently missing HLAs in AA populations

Because these missing HLA alleles in 6pLOH could be involved in the pathogenesis of AA, we next tested whether these relevant HLA alleles are associated with the risk of the development of AA among the 6,613 JMDP registrants. As shown in Table 4, the 4 major missing HLA alleles, HLA-A*31:01, B*40:02, A*02:01 and A*02:06 were more frequently observed in AA cases compared to non-significant HLA alleles (i.e. all HLA alleles other than these 4 alleles), where the odds ratios for the risk of the development of AA between each of these alleles and non-significant alleles were 1.87 (95%CI; 1.43 - 2.43) for A*02:01, 2.22 (95%CI; 1.70 - 2.90) for A*02:06, 1.37 (95%CI; 1.00 - 1.88) for A*31:01, and 1.95 (1.48 - 2.58) for B*40:02 (Table 4). The combined relative risk for all these alleles was 1.75 (1.42 – 2.17) ($P = 1.3 \times 10^{-7}$).

Discussion

The origin of clonal hematopoiesis in AA is a focus of long-standing disputes, in which a profoundly reduced hematopoietic stem cell pool and/or escape from the autoimmune insults have been implicated in the evolution of the clonal hematopoiesis in AA[5,22,23]. Our findings on 6pLOH in AA provide an intriguing insight not only into the underlying mechanism of the clonal hematopoiesis in AA, but also into the origin of the autoimmunity that is responsible for the pathogenesis of AA. A recent study from the

United States also reported three cases with 6pLOH[24]. With a sensitive detection algorithm, the presence of the 6pLOH(+) components was demonstrated in as many as 13% of typical cases with AA, and the evidence from the subsequent studies strongly indicated that the HLA genes are the genetic targets of 6pLOH in AA patients. First, the HLA locus was commonly and critically involved in all 6pLOHs found in AA. Second, some AA patients carried multiple 6pLOH(+) subclones with different breakpoints, but in all cases, the 6pLOH involved the HLA locus and occurred in a manner that targeted the same parental HLA allele. Moreover, particular class I HLA alleles were overrepresented among 6pLOH(+) cases and consistently found in the missing haplotypes. Finally, many of these HLA alleles were shown to be tightly associated with the development of AA in Japanese patients in case-control studies using the large JMDP registry.

The conspicuous bias of the missing HLA alleles in 6pLOH to particular HLA types, and the significant association of AA with those HLA types strongly suggest that the recurrent 6pLOH in AA is a phenomenon tightly related to the pathogenesis of AA rather than mere secondary event during the course of AA. Based on these observations, it is well reasoned that in 6pLOH(+) AA cases, the auto-immunity to HSPCs is mediated by the CTLs that target the antigens presented via specific class I HLA molecules, and that the 6pLOH(+) cells found in AA could be explained as escape hematopoiesis that survives the auto-immune insult by genetically deleting the relevant HLA species that are required for antigen presentation (Figure. 5). These scenarios are further supported by the recent reports showing that the CNN-LOH in 6p provides a common mechanism of leukemic relapse after HLA haplo-identical stem cell transplantations, in which leukemic cells that lost the mismatched HLA haplotype

through CNN-LOH in 6p are thought to escape the immunological surveillance of the engrafted donor-T cells[25,26]. Importantly, it was experimentally demonstrated by immunological assays that the 6pLOH(+) leukemic cells actually escaped GVL by CTLs, while 6pLOH(−) leukemic cells were effectively killed by the same CTLs. Although the immunological targets of CTLs are different between relapse after haplo-identical transplants (mismatched HLAs themselves) and AA (still unknown autoantigens presented on missing HLAs), the prominent similarities found in both cases further supports that CNN-LOH in 6p confers an escape mechanism from auto-reactive CTLs in AA.

In light of the above considerations, the chronological behavior of the 6pLOH(+) components in PB is also interesting and worth discussing. Despite the assumption that 6pLOH is an effective escape mechanism from CTLs, the 6pLOH(+) stem cells were unable to repopulate the BM to cure AA, unless effective IST was applied (Supplementary Figure. 6). This is most likely explained by the presence of inflammatory cytokines, such as IFN-γ and TNF-α, which have also been shown to play an important role in the bone marrow failure in AA, and are thought to be responsible for the continued prevention of the 6pLOH(+) stem cells from fully expanding and reconstituting the BM (Supplementary Figure. 9A and 9B)[27,28].

When the autoimmune insults are removed after IST, no further injury of normal stem cells would occur. However, this does not necessarily mean the surviving normal stem cells can eventually outnumber the 6pLOH(+) stem cells over time. Note that once the autoimmune insults disappear, nothing could biologically or immunologically discriminate a 6pLOH(+) stem cell from a 6pLOH(−) stem cell (Supplementary Figure. 9A). In particular, a 6pLOH(+) stem cell and a 6pLOH(−)

stem cell will produce the same number of progeny on average, and feed the same number of mature blood cells. As a consequence, once established, the predominance of 6pLOH(+) stem cells over 6pLOH(−) stem cells should be maintained, after the severely reduced hematopoietic stem cell pool has been re-expanded with removal of the inciting autoimmunity. It is also of note that the recovery of myeloid components after IST, which are affected more strongly by 6pLOH than lymphoid cells, contributes to an apparent increase in 6pLOH components in the SNP array analysis in PB (Supplementary Figure. 6A).

One of the most significant findings in the current study is the identification of the HLA alleles that are overrepresented in the Japanese AA populations, including HLA-A*31:01, B*40:02, A*02:01, and A*02:06. All of these HLA alleles belong to class I major histocompatibility complexes (MHCs) and thus are thought to be involved in the antigen presentation to CTLs. This provides another prominent example, in which specific HLA-types play a critical role in the development of a human disease, and the information about these particular HLA types provides a solid basis on which we can ultimately isolate the relevant antigens responsible for the development of AA. Of particular note, there was a previous report that indicated that HLA-B*40:02 and A*02:06 were overrepresented in PNH as well as AA, although the study size was much smaller than the current study[29]. Combined with our study, these findings support the hypothesis that AA and PNH are the different outcomes of the same immunologic insult[5,30], and may also provide the genetic basis of the high prevalence of AA and PNH in East Asia[31,32]

In some AA cases, hematopoiesis could be maintained over years by the progenitors that escaped and survived the inciting autoimmune insult by deleting the target HLA

through CNN-LOH in 6p. Given that the 6pLOH was detected in only 13 % of our series, it is likely that other escape mechanisms may also operate to maintain hematopoiesis in AA. In fact, clonality was clearly demonstrated in 20% of the 6pLOH(−) cases in the HUMARA study (Supplementary Figure. 8). In addition, our SNP array analysis also revealed a variety of clonal abnormalities in AA cases (Figure. 1), although it is still open to question whether these abnormalities actually represent the mechanism of escape hematopoiesis or were related to some neoplastic process. Further studies on the genetic basis of the escape mechanisms would contribute to our understanding of the molecular pathogenesis of AA.

Acknowledgements

This study was supported in part by the Core Research for Evolutional Science and Technology (CREST), the Japan Science and Technology Agency, Grant-in-Aids for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and Grant-in-Aids from the Ministry of Health, Labor and Welfare of Japan. We are deeply indebted to the patients and donors and their physicians including K. Kawakami of Suzuka General Hospital, A Okamoto of Nagoya Daini Red Cross Hospital for contributing to this study.

Authorship Contributions

T.K. and A.S.-O., Seishi.O. and S.N. were equally contributed to this work.

Shigeki.O., Seishi.O. and S.N. developed the concept of the study and supervised the project. T.K., Shigeki.O. and S.N. designed the experiments. T.K., A.S.-O., Y.S., Y.M., M.K., M.S., K.H., and Y.S. performed the experiments and analyzed the data. K.K. performed high-resolution HLA typing. S.M. and Y.M. provided the information of JMDP donor-recipient pairs (JMDP data set). T.K., A.S.-O., Seishi.O. and S.N. wrote the paper. All authors approved the final version of this paper.

Author Information

Affiliations

Clinical Laboratory Science, Division of Health Sciences, Kanazawa University Graduate School of Medical Science, 5-11-80 Kodatsuno, Kanazawa, Ishikawa 920-0942 Japan.

Takamasa Katagiri and Shigeki Ohtake.

Cellular Transplantation Biology, Kanazawa University Graduate School of Medical Science, 13-1 Takaramachi, Kanazawa, Ishikawa 920-8640 Japan.

Takamasa Katagiri, Kohei Hosokawa, Yumi Sasaki and Shinji Nakao

Cancer Genomics Project, Graduate School of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan.

Aiko Sato-Otsubo, Yusuke Sato, Yuka Mori, Motohiro Kato, Masashi Sanada and Seishi Ogawa.

Tokyo Metropolitan Red Cross Blood center, 2-1-67 Tatsumi, Koto-ku, Tokyo 135-0053 Japan.

Koichi Kashiwase

Department of Hematology, Fujita Health University, 1-98 Dengakugakubo, Kutsukake-cho, Toyoake, Aichi 470-1192 Japan.

Satoko Morishima

Department of Hematology and Cell Therapy, Aichi Cancer Center Hospital, 1-1 Kanokoden, Chikusa-ku, Nagoya 464-8681 Japan.

Yasuo Morishima

Core Research for Evolutional Science and Technology, Exploratory Research for Advanced Technology, Japan Science and Technology Agency, 4-1-8 Honcho,

Kawaguchi-shi, Saitama 332-0012, Japan.

Koichi Kashiwase and Seishi Ogawa

Disclosure of Competing Financial Interests

All authors have no financial or personal relationships with other people or organizations that could inappropriately influence this study. The authors declare no competing financial interest.

Corresponding Author

Readers are welcome to comment on the online version of this article at www.nature.com/nm/index.html. Correspondence and requests for materials should be addressed to S. Ogawa (sogawa-tky@umin.ac.jp), S. Nakao (snakao@med3.m.kanazawa-u.ac.jp).

References

1. Young NS, Calado RT, Scheinberg P. Current concepts in the pathophysiology and treatment of aplastic anemia. Blood. 2006;108(8):2509-2519.
2. Nakao S, Takami A, Takamatsu H, et al. Isolation of a T-cell clone showing HLA-DRB1*0405-restricted cytotoxicity for hematopoietic cells in a patient with aplastic anemia. Blood. 1997;89(10):3691-3699.
3. Chen J, Ellison FM, Eckhaus MA, et al. Minor antigen h60-mediated aplastic anemia is ameliorated by immunosuppression and the infusion of regulatory T cells. J Immunol. 2007;178(7):4159-4168.
4. Risitano AM, Maciejewski JP, Green S, Plasilova M, Zeng W, Young NS. In-vivo dominant immune responses in aplastic anaemia: molecular tracking of putatively pathogenetic T-cell clones by TCR beta-CDR3 sequencing. Lancet. 2004;364(9431):355-364.
5. Young NS. The problem of clonality in aplastic anemia: Dr Dameshek's riddle, restated. Blood. 1992;79(6):1385-1392.
6. 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(8):1648-1657.
7. Lewis SM, Dacie JV. The aplastic anaemia--paroxysmal nocturnal haemoglobinuria syndrome. Br J Haematol. 1967;13(2):236-251.
8. Dameshek W. Riddle: what do aplastic anemia, paroxysmal nocturnal hemoglobinuria (PNH) and "hypoplastic" leukemia have in common? Blood. 1967;30(2):251-254.
9. Socie G, Rosenfeld S, Frickhofen N, Gluckman E, Tichelli A. Late clonal diseases of treated aplastic anemia. Semin Hematol. 2000;37(1):91-101.
10. Tichelli A, Gratwohl A, Wursch A, Nissen C, Speck B. Secondary leukemia after severe aplastic anemia. Blut. 1988;56(2):79-81.
11. de Planque MM, Kluin-Nelemans HC, van Krieken HJ, et al. Evolution of acquired severe aplastic anaemia to myelodysplasia and subsequent leukaemia in adults. Br J Haematol. 1988;70(1):55-62.
12. van Kamp H, Landegent JE, Jansen RP, Willemze R, Fibbe WE. Clonal hematopoiesis in patients with acquired aplastic anemia. Blood. 1991;78(12):3209-3214.
13. Kawase T, Morishima Y, Matsuo K, et al. High-risk HLA allele mismatch combinations responsible for severe acute graft-versus-host disease and implication for its molecular mechanism. Blood. 2007;110(7):2235-2241.
14. Nannya Y, Sanada M, Nakazaki K, et al. A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. Cancer Res. 2005;65(14):6071-6079.
15. Yamamoto G, Nannya Y, Kato M, et al. Highly sensitive method for genomewide detection of allelic composition in nonpaired, primary tumor specimens by use of affymetrix single-nucleotide-polymorphism genotyping

microarrays. Am J Hum Genet. 2007;81(1):114-126.
16. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proc Natl Acad Sci U S A. 2003;100(16):9440-9445.
17. Ogawa S, Matsubara A, Onizuka M, et al. Exploration of the genetic basis of GVHD by genetic association studies. Biol Blood Marrow Transplant. 2009;15(1 Suppl):39-41.
18. Morishima S, Ogawa S, Matsubara A, et al. Impact of highly conserved HLA haplotype on acute graft-versus-host disease. Blood. 2010;115(23):4664-4670.
19. Olshen AB, Venkatraman ES, Lucito R, Wigler M. Circular binary segmentation for the analysis of array-based DNA copy number data. Biostatistics. 2004;5(4):557-572.
20. Venkatraman ES, Olshen AB. A faster circular binary segmentation algorithm for the analysis of array CGH data. Bioinformatics. 2007;23(6):657-663.
21. Ishiyama K, Chuhjo T, Wang H, Yachie A, Omine M, Nakao S. Polyclonal hematopoiesis maintained in patients with bone marrow failure harboring a minor population of paroxysmal nocturnal hemoglobinuria-type cells. Blood. 2003;102(4):1211-1216.
22. Murakami Y, Kosaka H, Maeda Y, et al. Inefficient response of T lymphocytes to glycosylphosphatidylinositol anchor-negative cells: implications for paroxysmal nocturnal hemoglobinuria. Blood. 2002;100(12):4116-4122.
23. Bessler M, Mason PJ, Hillmen P, et al. Paroxysmal nocturnal haemoglobinuria (PNH) is caused by somatic mutations in the PIG-A gene. Embo J. 1994;13(1):110-117.
24. Afable MG, 2nd, Wlodarski M, Makishima H, et al. SNP array-based karyotyping: differences and similarities between aplastic anemia and hypocellular myelodysplastic syndromes. Blood. 2011;117(25):6876-6884.
25. Vago L, Perna SK, Zanussi M, et al. Loss of mismatched HLA in leukemia after stem-cell transplantation. N Engl J Med. 2009;361(5):478-488.
26. Villalobos IB, Takahashi Y, Akatsuka Y, et al. Relapse of leukemia with loss of mismatched HLA resulting from uniparental disomy after haploidentical hematopoietic stem cell transplantation. Blood. 2010;115(15):3158-3161.
27. Zoumbos NC, Gascon P, Djeu JY, Trost SR, Young NS. Circulating activated suppressor T lymphocytes in aplastic anemia. N Engl J Med. 1985;312(5):257-265.
28. Hinterberger W, Adolf G, Aichinger G, et al. Further evidence for lymphokine overproduction in severe aplastic anemia. Blood. 1988;72(1):266-272.
29. Shichishima T, Noji H, Ikeda K, Akutsu K, Maruyama Y. The frequency of HLA class I alleles in Japanese patients with bone marrow failure. Haematologica. 2006;91(6):856-857.

30. Karadimitris A, Manavalan JS, Thaler HT, et al. Abnormal T-cell repertoire is consistent with immune process underlying the pathogenesis of paroxysmal nocturnal hemoglobinuria. Blood. 2000;96(7):2613-2620.
31. Issaragrisil S, Kaufman DW, Anderson T, et al. The epidemiology of aplastic anemia in Thailand. Blood. 2006;107(4):1299-1307.
32. Montane E, Ibanez L, Vidal X, et al. Epidemiology of aplastic anemia: a prospective multicenter study. Haematologica. 2008;93(4):518-523.

Figure Legends

Figure 1. Copy number changes and allelic imbalances in 46 out of the 306 AA cases.

The copy number changes and allelic imbalances (or CNN-LOHs) in each case are summarized in the chromosomal order vertically for 46 AA cases with copy number abnormalities. Gains and losses, as well as CNN-LOHs, are shown in the indicated colors.

Figure 2. Acquired 6pLOHs in AA patients that target the HLA locus.

Panel A shows typical CNAG outputs in SNP array analysis showing CNN-LOH (purple line) that appears as significant dissociation in allele-specific copy number graphs (red and green lines) from the baseline with normal total copy numbers (tCN) (upper panel). As a result of an allelic conversion, the affected segment causes LOH, where asterisks indicate one (lower panel). The "acquired" origin of these lesions is indicated by the retention of substantial numbers of heterozygous SNP calls (green bars below the chromatogram) that would otherwise mostly disappear. Panel B summarizes the breakpoints of 6pLOHs found in a total of 28 AA cases, all involving the HLA locus in common. In more than half of cases (indicated by arrowheads in Panel B), the exact location of the breakpoint was difficult to uniquely determine, where dissociation of the allele-specific copy number graphs continuously tapered along the 6p arm, indicating the presence of multiple 6pLOH(+) clones with common missing alleles (Panel C). In fact, the breakpoint containing regions are separated into multiple segments having significantly different copy numbers in the circular binary

segmentation model, as indicated by solid lines with *P*-values. Note that the most telomeric breakpoint is located within (case 24) or centromeric to (case 23) the HLA locus in each case. Panel D shows a skewed distribution of the logarithm of *p*-values in AA cases compared with normal individuals. The *p*-values were calculated in the Mann-Whitney's U test, with which the difference in the mean allele-specific copy numbers between 6p and other chromosomal regions was evaluated (see Methods section). >250 values are plotted as 250.

Figure 3. Uniparental expression of HLA in AA cases with CNN-LOH in 6p.

Allele-specific expression of HLA-A antigens in AA specimens was examined by flow cytometry using monoclonal antibodies that specifically recognize the indicated HLA types (red lines), where leukocytes from healthy individuals were used as a control (blue lines). Panels A and B show the uniparental expression of HLA-A antigens in PB leukocytes and BM $CD34^+$ cells obtained from 3 AA cases with CNN-LOH in 6p. Different leukocyte compartments were separately examined, including granulocytes (G), monocytes (M), B-lymphocytes (B), and T-lymphocytes (T).

Figure 4. Imputation of missing HLA haplotypes.

The observed allelic copy numbers at heterozygous SNP sites along each candidate SNP haplotype are color-coded as indicated at the bottom. Green bars showed the SNPs that are incompatible with the patient's genotype. Case IDs and haplotype ID (HT_ID) are indicated on the left. The locations of the 500K SNPs and HLA-A, C, B, DRB1, DQB1, and DPB1 are indicated in the figure. For each allele, genomic copy numbers were imputed using the circular binary segmentation algorithm. This divided each haplotype into one or more segments having discrete mean allelic copy numbers

(blue arrows on the right). The positions of breakpoints are indicated by arrowheads. Finally, the mean allelic copy number of each segment was statistically compared to that of corresponding segment on the other haplotype using the Wilcoxon signed rank test. Missing HLA haplotypes were determined based on the result of the statistic tests. Purple and Blue lines indicated the retained and missing segments, respectively, while the allelic status was not determined statistically for those segments shown by green lines.

Figure 5. A proposed mechanism for escape hematopoiesis in 6pLOH(+) AA.

In AA, the targets of CTLs are the hematopoietic stem/progenitor (HSPC) cells that present some auto-antigen through particular Class I HLA molecules, including HLA-A*02:01, A*02:06, A*31:01, and B*40:02. In the presence of these auto-immune insults, the HSPCs that lose their expression of the antigen-presenting HLA molecule as a result of CNN-LOH in 6p would acquire a growth advantage over other HSPCs expressing the relevant HLA, leading to clonal outgrowth of the 6pLOH(+) progenies.

Table 1. Patients' characteristics

	Newly diagnosed (n=107)	Previously Treated (n=199)
Age at diagnosis (months), median (range)	64 (9-88)	24 (2-80)
Gender, Male/Female	58/49	110/89
Severity of AA at onset, no. of patients (%)		
severe	79 (74%)	185 (93%)
non severe	28 (26%)	14 (7%)
History (months), median (range)	19 (0.1-251)	51 (0.1-372)
Past treatment, no. of patients (%)		
ATG+CsA	-	39 (20%)
CsA alone	-	51 (26%)
Anabolic steroid alone	-	13 (7%)
Unknown[§]	-	96 (48%)

§Information regarding previous therapies of 96 cases (from Japan Marrow Donor Program) was unavailable.

Abbreviations: ATG, antithymocyte globulin; CsA, cyclosporin A; PR, partial remission; CR, complete remission; -, not applicable.

Table 2. 6pLOH(+) AA cases and imputed allelic status of HLA alleles

UID	6pUPD(+) Fraction(%)※	MISSING ALLELES&						RETAINED ALLELES&					
		A	B	C	DRB1	DQB1	DPB1	A	B	C	DRB1	DQB1	DPB1
19	53.9%	**31:01***	**40:02**	**03:04**	12:01	03:01	05:01	24:02	52:01	12:02	15:02	06:01	05:01
12	51.8%	**02:01***	**40:02**	03:03	15:01	06:02	05:01	26:02	40:06	08:01	09:01	03:03	05:01
17	51.6%	24:02	13:01	**03:04**	12:02	03:01	04:02	24:02	52:01	12:02	15:02	06:01	09:01
304	49.3%	**31:01***	55:02	01:02	12:02	03:01	41:01	24:02	07:02	07:02	01:01	05:01	04:02
11	48.0%	**02:06***	**40:02**	**03:04**	15:01	06:02	ND§1	11:01	67:01	07:02	16:02	05:02	ND§1
21	46.2%	**31:01***	51:01	14:02	14:05	05:03	03:01	24:02	07:02	07:02	01:01	05:01	04:02
24	44.9%	**31:01**	**40:02**	**03:04**	11:01	03:01	02:01	24:02	40:06	08:01	09:01	03:03	05:01
26	44.3%	**31:01*♮**	40:01	**03:04**	04:05	04:01	03:01	26:03	52:01	12:02	15:02	06:01	09:01
27	43.5%	**02:06**	**40:02**	**03:04**	04:10	04:02	02:01	11:01	52:01	12:02	15:02	06:01	09:01
10	42.1%	**31:01**	**40:02**	**03:04**	08:03	06:01	02:01	24:02	51:01	14:02	09:01	03:03	02:01
8	40.8%	**02:06***	**40:02**	03:03	12:01	03:01	05:01	24:02	52:01	12:02	15:02	06:01	04:02
23	35.2%	**02:01**	**40:02**	**03:04**	09:01	03:03	02:01	24:02	54:01	01:02	04:05	04:01	04:02
25	32.1%	**02:06***	no LOH					01:01	no LOH				
9	23.5%	**02:06***	39:01	07:02	08:02	04:02	02:01	24:02	15:18	07:04	04:01	03:01	14:01
20	21.7%	26:01*	**40:02**	03:03	15:01	06:02	05:01	02:18	46:01	01:02	08:03	06:01	05:01
14	21.7%	**31:01***	51:01	14:02	09:01	03:03	05:01	24:02	52:01	12:02	15:02	06:01	09:01
22	20.6%	**02:01**	39:01	07:02	08:03	06:01	05:01	24:02	52:01	12:02	15:02	06:01	09:01
18	17.6%	**02:01***	40:06	08:01	09:01	03:03	02:01	24:02	35:01	03:03	15:01	06:02	04:02
15	17.4%	**02:06**	40:06	08:01	09:01	03:03	02:01	24:02	07:02	07:02	01:01	05:01	02:01
41	15.2%‡	**31:01***	35:01	03:03	09:01	03:03	03:01	26:01	39:01	07:02	08:03	06:01	05:01
28	12.8%	24:02	54:01	01:02	01:01	05:01	04:02	24:02	52:01	12:02	15:02	06:01	09:01
29	11.7%	**31:01**	**40:02**	**03:04**	15:01	06:02	02:01	24:02	54:01	01:02	04:05	04:01	05:01
305	10.3%	**02:06***	**40:02**	15:02	15:02	06:01	04:01	24:02	51:01	14:02	09:01	03:03	02:01
13	9.6%	24:02*	**40:02**	**03:04**	15:01	06:02	02:01	**02:01**	35:01	08:01	09:01	03:03	02:01
306	8.5%	24:02*	**40:02**	**03:04**	09:01	03:03	02:01	26:02	40:06	08:01	09:01	03:03	02:01
16	8.1%	11:01	40:06	08:01	no LOH			24:02	46:01	01:02	no LOH		
30	8.0%	**02:06**	39:01	07:02	no LOH			24:02	40:06	08:01	no LOH		
72	5.6%	**02:01**	**40:02**	**03:04**	09:01	03:03	05:01	02:07	46:01	01:02	08:03	06:01	02:02
36	4.0%	**02:01***	ND§2	ND§3	15:02	06:01	09:01	24:02	ND§2	ND§3	15:02	06:01	09:01
124	3.5%	24:02	**40:02**	**03:04**	12:01	03:01	02:01	24:02	52:01	12:02	15:02	06:01	09:01
223	2.8%	**31:01***	48:01	**03:04**	09:01	03:03	05:01	**02:06**	39:01	07:02	15:01	06:02	02:01
215	2.8%	**31:01**	51:01	14:02	08:02	04:02	04:02	03:01	44:02	05:01	13:01	06:03	05:01
181	1.3%	**02:06**	13:01	**03:04**	12:02	03:01	05:01	24:02	52:01	12:02	15:02	06:01	09:01
97	1.0%	24:02	07:02	07:02	01:01	05:01	05:01	**02:01**	39:01	07:02	15:01	06:02	02:01
252	0.9%	ND§4	**40:02**	**03:04**	09:01	03:03	05:01	ND§4	46:01	01:02	04:05	04:01	05:01
118	0.9%	**02:06***	**40:02**	**03:04**	08:02	03:02	05:01	24:02	52:01	12:02	15:02	06:01	09:01
298	0.8%	24:02	**40:02**	**03:04**	15:01	06:02	05:01	24:02	52:01	12:02	15:02	06:01	09:01
188	0.7%	24:02	52:01	12:02	15:02	06:01	09:01	**02:01**	52:01	12:02	11:01	03:01	05:01
291	0.7%	**31:01**	51:01	14:02	15:01	06:02	02:01	24:02	40:01	**03:04**	11:01	03:01	05:01
196	0.2%	ND§5 (A*02:06/24:02, B*35:01/51:01, C*03:03/15:02, DRB1*04:03/15:01, DQB1*03:02/06:02, DPB1*0:201/02:01)											

※The percentage of 6pUPD(+) fraction is derived from total peripheral blood leukocytes that include lymphoid as well as myeloid element.

* The allelic loss was confirmed by FCM.

& HLA types significantly deviated to missing alleles are indicated by shadows.

§1: DPB1*04:02/05:01, §2: B*15:18/52:01, §3: C*08:01/12:02, §4; A*02:01/02:07, §5; Missing allele was not determined because copy number changes in these segments were not statistically significant.

♮ The missing haplotype was determined by flow cytometry.

Table 3. Response rate (CR+PR) according to the Camitta criteria

	Newly diagnosed (n=107)		Previously treated (n=103)	
	6pLOH(−) (n=91)	6pLOH(+) (n=16)	6pLOH(−) (n=88)	6pLOH(+) (n=15)
Immunosuppressive therapies (all)	36/49 (73%)	11/11 (100%)	65/77 (84%)	12/12 (100%)
ATG+CsA	14/19 (74%)	7/7 (100%)	27/33 (82%)	5/5 (100%)
CsA alone	22/30 (73%)	4/4 (100%)	38/44 (86%)	7/7 (100%)
Anabolic steroid alone	0/0 (0%)	0/0 (0%)	7/11 (64%)	2/2 (100%)
Unknown/Not evaluable	42	5	0	1

Abbreviations: ATG, antithymocyte globulin; CsA, cyclosporin A; PR, partial remission; CR, complete remission.

Table 4. Association of missing HLA alleles with AA in Japanese patients

Risk allele	AA (N=407)	Other diseases (N=6206)	Total (N=6613)	*P*-value (χ^2 test)	Odds Ratio (95% CI) (vs. no risk alleles)
A*02:01	103	1173	1276	2.5×10^{-6}	1.87 (1.43 – 2.43)
A*02:06	100	957	1057	$< 1.0 \times 10^{-7}$	2.22 (1.70 – 2.90)
A*31:01	58	899	957	0.048	1.37 (1.00 – 1.88)
B*40:02	86	938	1024	1.8×10^{-6}	1.95 (1.48 – 2.58)
All risk alleles	268	3250	3518	1.3×10^{-7}	1.75 (1.42 – 2.17)
No risk alleles	139	2956	3095		

Figure 1

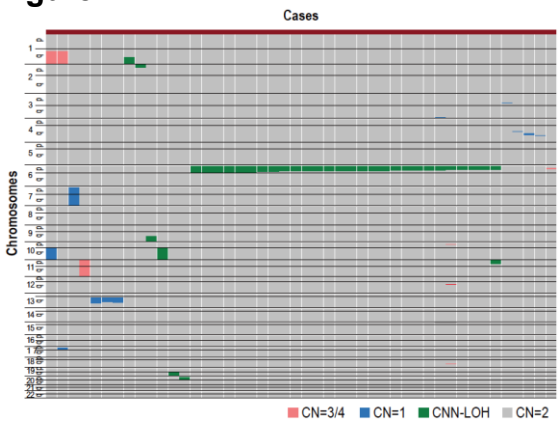


Figure 2

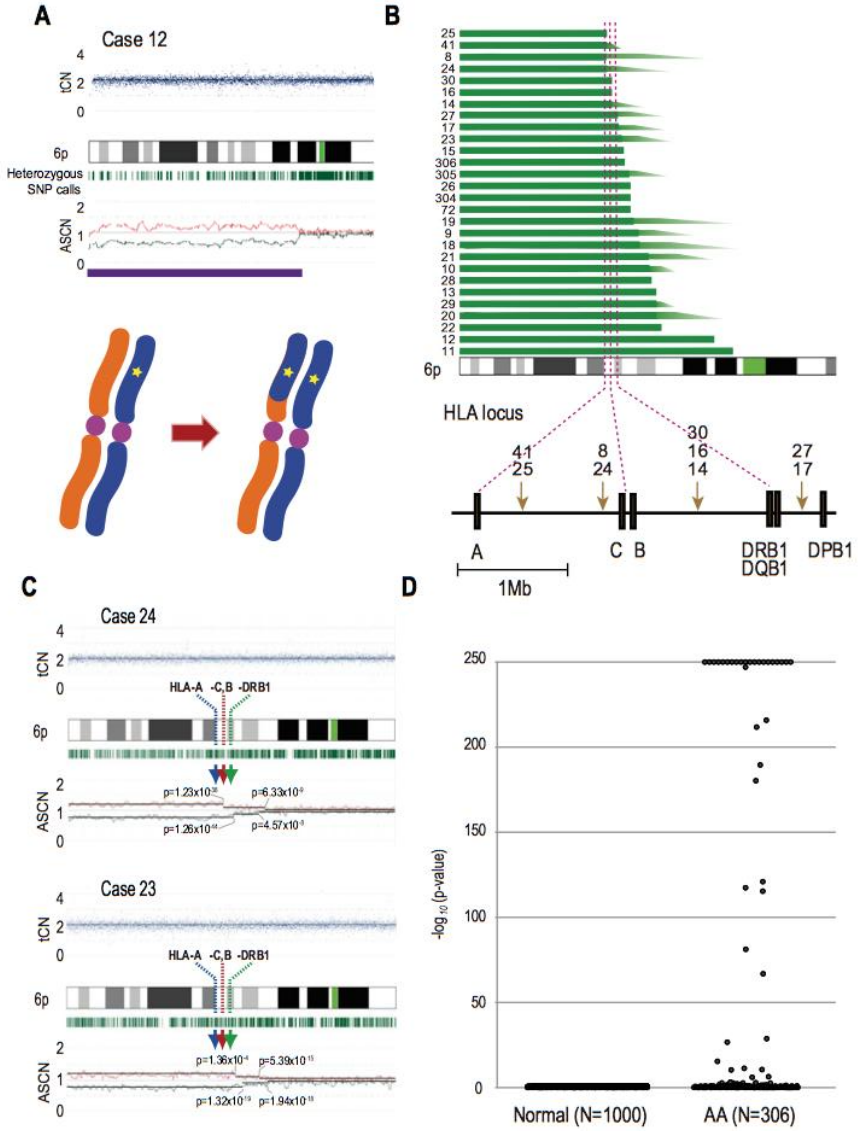


Figure 3

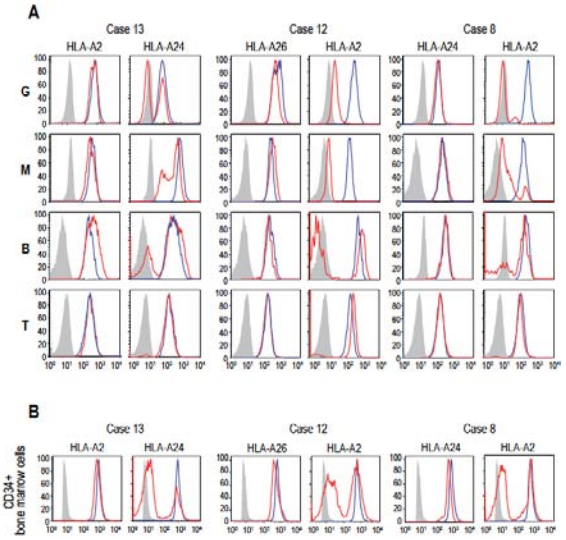


Figure 4

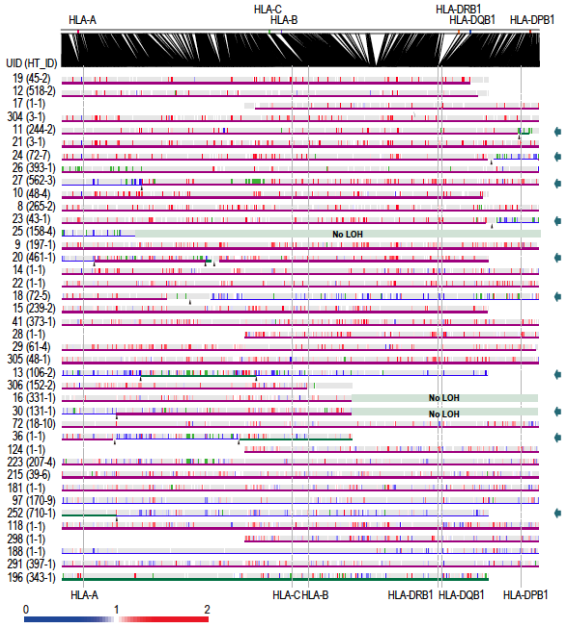
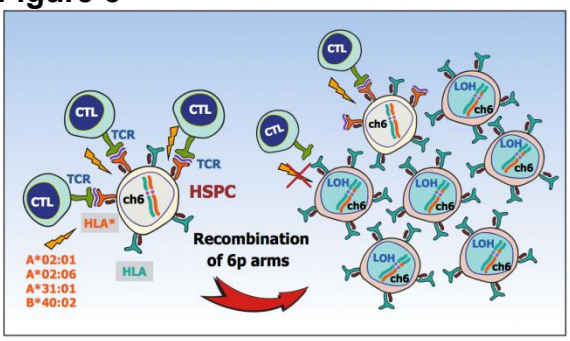


Figure 5


CTL
TCR
HSPC
ch6
HLA*
HLA
A*02:01
A*02:06
A*31:01
B*40:02
Recombination
of 6p arms
LOH